

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Entanglement, Complexity, and Geometry Beyond AdS/CFT

Permalink

<https://escholarship.org/uc/item/1wb8r2hb>

ISBN

9798189278761

Author

Ritchie, Kyle James

Publication Date

2026-05-01

Peer reviewed|Thesis/dissertation

Entanglement, Complexity, and Geometry Beyond AdS/CFT

by

Kyle Ritchie

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Yasunori Nomura, Chair

Professor Luca Iliesiu

Professor Ezra Rubenstein

Spring 2026

Entanglement, Complexity, and Geometry Beyond AdS/CFT

Copyright 2026
by
Kyle Ritchie

Abstract

Entanglement, Complexity, and Geometry Beyond AdS/CFT

by

Kyle Ritchie

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Yasunori Nomura, Chair

The holographic principle posits that the physics within a region of spacetime can be encoded in degrees of freedom living on its geometric boundary. This principle is most robustly exemplified by the AdS/CFT correspondence, which provides a dictionary between quantum gravitational degrees of freedom in Anti-de Sitter (AdS) spacetime and those of the conformal field theory (CFT) at its asymptotic boundary. Beyond AdS/CFT, the precise structure of holographic encoding is not fully understood. This dissertation, divided into two parts, addresses various aspects of holography beyond AdS/CFT.

In Part I, we study how degrees of freedom outside causal horizons holographically encode the spacetime beyond them. We examine the near-horizon mode structure of Schwarzschild and de Sitter horizons and argue that, from the perspective of semiclassical observers whose ability to resolve fine-grained quantum information is inherently limited, the physics beyond the horizon emerges from near-horizon dynamics. We discuss the relationship between this framework and the standard global framework used in AdS/CFT, emphasizing the complementary mechanisms through which they incorporate unitarity and the equivalence principle.

In Part II, we investigate the relationship between the geometry of generic spacetime subregions and their quantum information content, as captured by entanglement entropy and computational complexity. We first use the relationship between random tensor networks and fixed-geometry states to motivate a gravitational path integral derivation of the quantum extremal surface prescription for bulk subregions of AdS. In particular, we show that the von Neumann entropy of a bulk subregion is equal to the generalized entropy of the bulk entanglement wedge containing it. We then propose a notion of generalized complexity for bulk subregions as the sum of their volume and the complexity of the semiclassical bulk fields they support. We demonstrate that if the semiclassical complexity is quantified by a codimension-0 measure which satisfies strong subadditivity, then timelike focusing arguments imply a quantum generalization of the strong energy condition, as well an upper bound on the semiclassical complexity of a slice by its volume.

To those whose journeys are concluded

and for whom mine continues

Chad Broumley

Joan Ritchie

Lynn Ritchie

Rita Medina

Contents

Contents	ii
List of Figures	iv
List of Tables	xi
1 Introduction	1
I The Emergence of Spacetime Beyond Causal Horizons	4
2 Black Hole and de Sitter Microstructures from a Semiclassical Perspective	5
2.1 Introduction and Summary	5
2.2 Quantum Field on Background Spacetime	14
2.3 The Absence of Spacetime below the String Length	21
2.4 Holographic Description	38
2.5 Analytic Extension of Spacetime in Quantum Gravity	53
2.6 Intrinsically Extended Spacetimes	75
2.7 Gravitational Path Integral	89
2.8 Conclusion	98
3 Complementarity for a Dynamical Black Hole	100
3.1 Introduction	100
3.2 Exterior Description	102
3.3 Emergence of Global Spacetime	103
3.4 A Puzzle in the Global Spacetime Picture and Its Resolution	111
3.5 Causal Structure of Black Hole Formation and Evaporation	120
3.6 Global Spacetime as a Coarse-Grained Description	127
3.7 Conclusion	130
II Entanglement and Complexity for Gravitating Subregions	132
4 Hollow-grams: Generalized Entanglement Wedges from the Gravitational Path Integral	133
4.1 Introduction	133

4.2	BP Proposal in RTNs	136
4.3	Holographic Rényi Entropy via Fixed-Geometry States	143
4.4	BP Proposal from the Gravitational Path Integral	148
4.5	Examples	151
4.6	Discussion	159
5	A Timelike Quantum Focusing Conjecture	163
5.1	Introduction	163
5.2	Definitions	165
5.3	Timelike Focusing and its Consequences	167
5.4	Discussion	173
A	Appendix	179
A.1	The Hartle-Hawking Wavefunction	179
A.2	Bulk Subregions Spanning the Horizon	181
	Bibliography	184

List of Figures

2.1	The effective potential V_ℓ of a Schwarzschild black hole (solid lines) for a minimally-coupled massless scalar field with $\ell = 0, 1, 2$ as a function of the tortoise coordinate r_* . The Hawking temperature $T_H = 1/4\pi r_+$ is indicated by the horizontal dashed line. The approximated curves in Eq. (2.2.20) are also plotted (dotted lines) in the region where they are relevant.	16
2.2	The effective potential V_ℓ of de Sitter spacetime for a minimally-coupled massless scalar field with $\ell = 0, 1, 2$ as a function of the tortoise coordinate r_* . The Hawking temperature $T_H = 1/2\pi r_+$ is indicated by the horizontal dashed line.	18
2.3	A schematic depiction of approximately orthogonal, independent modes having $\omega \approx O(T_H)$ and localized in the zone for a field of negligible mass. (a) For $\ell \approx O(1)$, there are $\sim \ln(r_+/l_s)$ independent ingoing and outgoing modes. Their approximate basis can be taken to be wavepackets of width $\sim r_+$ in r_* , distributed uniformly over the classically allowed region between r_{*s} and $r_{*b}^{(\ell)}$, where $r_{*b}^{(\ell)} - r_{*s} \sim r_+ \ln(r_+/l_s)$. (b) For modes with $\ell \approx O(r_+/l_s)$, the classically allowed region is near the stretched horizon with $r_{*b}^{(\ell)} - r_{*s} \sim r_+$, so all these modes must be regarded as horizon modes.	27
2.4	A schematic depiction of quanta occupying various modes. Solid arrows indicate that the occupation numbers of the modes are determined by the thermodynamic consideration as described in Section 2.3.1. For zone modes, this applies to the outgoing mode as well as the ingoing modes with $\omega \lesssim \omega_{\text{barrier}}^{(\ell)}$, except that backreaction of Hawking emission at $r \sim r_z$ leads to a deficit for some of these ingoing modes as indicated by the dashed arrow.	35
2.5	Holographic slices in a black hole spacetime in ingoing Eddington-Finkelstein coordinates (v, r) obtained (a) without and (b) with quantum effects in the bulk. These slices correspond to equal-time hypersurfaces in the bulk, which approach Schwarzschild time slices near the black hole. The figures are taken from Refs. [35] and [36].	40
2.6	A holographic theory resides on a holographic screen, which is a hypersurface foliated by quantum marginally antitrapped (or trapped) surfaces called leaves. At a given boundary time, a boundary state lives on a leaf, from which one can reconstruct a bulk equal-time hypersurface using the flow equation in Eq. (2.4.1).	42

2.7	Renormalized holographic screens, on which regularized holographic theories reside, are timelike hypersurfaces in an asymptotically AdS and flat spacetimes, which are depicted by the dotted blue lines in the Penrose diagrams in (a) and (b), respectively. In both cases, holographic slices agree with the time slices given by timelike Killing vectors in the asymptotic regions. In the non-renormalized limit, physics occurring in a finite time is described by holographic theories located at the conformal boundary ($r = \infty$) and spatial infinity (i^0) in the cases of asymptotically AdS and flat spacetimes, respectively.	43
2.8	Penrose diagram of de Sitter spacetime. A de Sitter bubble corresponds to the upper-left half of this diagram, where FLRW equal-time hypersurfaces are drawn by green lines. The holographic screen is located along the trajectory of the apparent horizon, which is depicted by the solid blue curve.	45
2.9	A holographic description of “single-sided” de Sitter spacetime. Holographic slices agree (approximately) with equal-time hypersurfaces in static coordinates. The smaller the amount of UV renormalization becomes, the closer renormalized leaves are to the bifurcation surface. The geometry in the vicinity of the past horizon (indicated by light green) and the bifurcation surface (light blue) is regularization dependent.	47
2.10	A sketch of the collection of states in holographic Hilbert space $\mathcal{H}_{\text{eff}}(\mathcal{A})$ which have the same entanglement structure at leading order in $1/\mathcal{N}$. It forms an $e^{Q\mathcal{N}}$ -dimensional subspace of $\mathcal{H}_{\text{eff}}(\mathcal{A})$ (represented by the pink plane) except that exponentially rare nongeneric states are excluded (white arrows), and that $e^{\mathcal{N}} - e^{Q\mathcal{N}}$ isolated states orthogonal to it are added (red arrow).	52
2.11	The spacetime region K described by the effective theory of the interior erected at a boundary (Schwarzschild) time $t = t_*$.	58
2.12	The spacetime regions described by effective theories erected at $t = t_*$ (blue) and $t = t_* + \Delta t$ with $\Delta t > 0$ (red). The black arrow indicates a falling object that reaches the stretched horizon at $t = t_*$.	65
2.13	(a) Radiation at $t = t_R$, $R(t_R)$, contains information about some of the horizon and zone modes at $t < t_R - t_{\text{scr}} - t_{\text{sig}}$. This allows us to reconstruct a portion $\mathcal{I}_{\text{rad}}$ of the interior spacetime corresponding to the inner wedge of the stretched horizon at $t = t_R - t_{\text{scr}}$. The amount of information that can be reconstructed, however, depends on the location of an object in ingoing time, which is indicated by the gradation of the red shade. (b) Radiation after the black hole is fully evaporated, $R(t > t_{\text{evap}})$, allows for reconstruction of the full interior region.	68
2.14	Information outside the static patch in the effective theory can be retrieved if the system leaves the de Sitter phase at a later time, for example by tunneling into a Minkowski bubble universe.	73

2.15	An object falling into the stretched horizon at $t = t_{\text{obj}}$ can try to send a physical signal to the observer at $r = 0$ who enters a Minkowski bubble nucleated at $r = 0$ at $t = t_{\text{obj}} + \Delta t$. The holographic static-patch description of de Sitter spacetime suggests that this is not possible if Δt is larger than the scrambling time t_{scr} . This expectation is consistent with the semiclassical picture, since for $\Delta t > t_{\text{scr}}$ the object would have to send the signal within proper time $\tau_{\text{obj}} \lesssim l_{\text{P}}$, which is not possible.	74
2.16	Holographic theory of the interior may be constructed by distributing the interior degrees of freedom over the horizon $H(t)$, which consists of two marginally trapped surfaces as indicated by the wedges. Time t evolves toward the future on both components of the horizon. Entanglement entropy between the two sides is computed by finding the maximin surface, obtained by minimizing the area of a surface homologous to a component of the horizon on a Cauchy surface anchored to $H(t)$ (green dots) and then maximizing it over possible Cauchy surfaces, which leads to the bifurcation surface (blue dot). The entanglement entropy is thus given by the Bekenstein-Hawking entropy of the black hole.	84
2.17	Preparation of a holographic state of the thermofield double form by Euclidean path integral. In each panel, the thick red line represents the location at which the holographic theory resides, while the shaded region corresponds to the bulk region that can be simply reconstructed, e.g., by the procedure of continuously pulling in the boundary.	86
2.18	The holographic theory based on static patches consists of two weakly interacting systems located on the stretched horizons of two static patches, with the simply reconstructed region staying within the two patches (red). In the DS/dS correspondence, each of the holographic systems is located by itself on de Sitter spacetime with one lower dimensions, leading to the simply reconstructed region called the DS/dS patch (blue). The structures of the two theories are similar at $t = 0$	88
2.19	Holographic theory placed on the de Sitter horizons of two static patches can describe the region outside the horizons (as viewed by polar observers). Entanglement entropy between the two sides is computed by finding the maximin surface, which is nothing other than one of the horizons (blue dot). The entanglement entropy, therefore, is given by the Gibbons-Hawking entropy. . .	89
2.20	The density matrix of the exterior region, whose element is specified by the field configurations ψ_+ and ψ_- , can be computed by Euclidean path integral, which can be interpreted to receive equal contributions from all the black hole microstates. Euclidean path integral can also be used to prepare a two-sided state Ψ , which is pure and unique.	91

2.21	(a) The (nonnormalized) density matrix $\hat{\rho}_R$ of radiation in nongravitational region R can be calculated by performing path integral with the boundary condition fixing the spatial field configurations above and below the cut along R . (b) The square of the density matrix, $\hat{\rho}_R^2$ is calculated using the replica method, which replicates the spacetime into two and imposes the boundary condition such that one side of the cuts along R on the two sheets are appropriately sewn together (as represented by blue arrows).	93
2.22	Two contributions to replica method calculation of $\overline{\text{Tr}} \hat{\rho}_R^2$. (a) The normal contribution in which the gravitational region is filled separately on each sheet. Integration of semiclassical matter gives $e^{S_{\text{rad}}}$ coming from Euclidean evolution relating all 4 field configurations above and below the cuts on two sheets (red arrows), while gravitational path integral gives the $e^{S_{\text{bh}}}$ factor for each sheet. (b) For the replica wormhole configuration, the gravitational regions on two sheets are geometrically connected. The contribution from gravitation path integral is thus only $e^{S_{\text{bh}}}$, while the contribution from matter integration gives $e^{2S_{\text{rad}}}$ since there are two independent cycles for evolution (red and blue arrows).	95
2.23	(a) The contribution to Z_n with the gravitational region filled separately (light blue) on each sheet. All semiclassical matter configurations above and below the cuts R (green dashed) are related by Euclidean evolution (red arrow), giving the factor of $e^{S_{\text{rad}}}$. (b) For the replica symmetric wormhole configuration, the gravitational regions on n sheets are geometrically connected. This allows for n independent cycles for evolution (red arrows), giving the contribution from matter integration of order $e^{nS_{\text{rad}}}$	96
3.1	In the exterior description, equal-time slices foliate the spacetime region outside the “inner boundary” of space, which for a (quasi-)static black hole corresponds to the stretched horizon.	104
3.2	The effective theory of the interior erected at time t gives an emergent two-sided black hole geometry. The degrees of freedom associated with the zone $\tilde{Z}_t$ of the second exterior arise from soft and far mode states coupled to the states in the zone Z_t of the first exterior in the vacuum microstate. The spacetime region described by this effective theory is the region indicated as K_t	106
3.3	The effective theory of the interior erected at an exterior time t describes the region of the black hole interior corresponding to the interior of the effective two-sided black hole composed of the zone at t and its second exterior mirror. To obtain the whole black hole interior, we need to patch the pictures obtained by effective theories erected at different exterior times.	108
3.4	A detector can be set up such that it extracts from Hawking radiation the information that is located at a spacetime point x in the interior of the black hole. Interpreting this as being caused by nonlocal interactions between x and the spacetime region r in which the radiation resides, we encounter the problem of the existence of a time slice (blue line) that goes through both the future of r and the past of x . On such a slice, the information in question is cloned, leading to a contradiction with the no-cloning theorem of quantum mechanics.	113

3.5	The physics occurring in the region described by the effective theory erected at time t_x cannot be obtained by any effective theory erected after $T_{t_x} = t_x + 2r_+ \ln(r_+/l_P)$, implying that it must be viewed as having occurred before T_{t_x} in the exterior description.	115
3.6	The physics associated with an object that fell into the horizon at time t_f cannot be described by any effective theory erected after $t_{\text{comp}} = t_f + t_{\text{scr}}$. This implies that it must be viewed as having occurred before t_{comp} in the exterior description.	117
3.7	In the exterior description, the causal flow of information is obvious after the black hole is formed. The question is if there exists an analogous picture throughout the history of the black hole evolution, including the times during its formation.	120
3.8	Consider a diary in the region near $r = 0$ far in the past. We then fire a null shell of infalling particles toward $r = 0$, creating a black hole of mass M . The unitarity of the S -matrix implies that the information about the diary must eventually be emitted from the stretched horizon in the form of Hawking radiation. An exterior description with trivial causality, the right panel of Fig. 3.7, is obtained by extending the stretched horizon toward the past using the event horizon.	121
3.9	An excitation in a semiclassical theory can be viewed as a perturbation on the network of entanglement representing a spatial slice (left). The network can be cut at a closed surface Σ surrounding the perturbation and be replaced with an equivalent network in which the interior of Σ is removed (right). An operator corresponding to the original excitation can be represented by a nonlocal operator acting on the internal legs of the new network, obtained by pushing the original operator through the network.	124
3.10	An equal-time hypersurface of the holographic boundary, called a leaf, can be pulled inward by a renormalization flow. The trajectory of each leaf, depicted in the figure, then provides an equal-time hypersurface in the bulk called a holographic slice in [35, 36]. We can terminate the flow when the renormalized leaf reaches the stretched event horizon, giving the manifestly causal exterior description.	125
3.11	The global spacetime description can be obtained by evolving the initial equal-time hypersurface in such a way that at each time, the equal-time hypersurface is a Cauchy slice. This description, however, is consistent only if we disallow a process involving an exponentially fine-tuned detector configuration accessing information about the interior; see Section 3.4.	128
3.12	Assuming that the singularity at the endpoint of black hole evaporation is sufficiently weak, the evaporation of a black hole can be viewed as the creation of a baby universe in (but only in) the effective global spacetime description.	130

4.1	A bulk input region in a tensor network. The green dots represent the vertices comprising the input region a . The yellow edges correspond to the internal edges $\mathcal{E}_a$, while the red edges represent the boundary ∂a and the turquoise legs indicate the bulk legs. The right-hand panel shows the diagrammatic representation of the hollow-graphic state used to compute the entropy associated with a	139
4.2	Schematic representation of the GHZ example. On the left, we see the tensor network analogue of the type of state we consider. The GHZ state is analogous to the state with subregions a and b both removed. The right side schematically represents the GHZ-type entanglement between respective subregions.	142
4.3	$\psi(h)$ is computed via a Euclidean gravitational path integral whose boundary conditions consist of the fixed geometry h on the bulk slice Σ and asymptotic AdS boundary conditions with sources. Gluing two copies along Σ computes the norm $ \psi(h) ^2$ and the Euclidean geometry g_h generically has a jump in extrinsic curvature across Σ . Integrating over h computes the norm of the original state $\langle \psi \psi \rangle = \int dh \psi(h) ^2$, resulting in a smooth saddle g_ψ	144
4.4	A figure representing the domain wall configurations contributing to the action when the input is a boundary subregion (left) and a bulk subregion (right).	146
4.5	The dominant contribution to the replica path integral comes from geometries in which the regions of the bulk slice Σ within the domain wall (sharing a conformal boundary with A) are glued cyclically, while the regions outside the domain wall are glued identically.	147
4.6	A schematic depiction of the Euclidean path integral with additional boundaries a^- and a^+ . When the induced metric on a^-, a^+ is chosen to agree with h on region a , the Euclidean spacetime is the original fixed-geometry saddle with an infinitesimal-width cut around the subregion a . The hollow-graphic map is norm-preserving.	149
4.7	The dominant contribution to the replica path integral comes from geometries in which the regions of the bulk slice Σ within the domain wall containing a are glued cyclically, while the regions outside the domain wall are glued identically.	150
4.8	The Hartle-Hawking wavefunction is computed using a saddle-point geometry g_{Φ_h} with an asymptotic boundary of length $\frac{\beta}{2}$ and a boundary Σ with the specified dilaton value Φ_h at the horizon. At generic values of Φ_h , there is a kink in the slice Σ	153
4.9	Left: For a bulk interval on a single side of the horizon, the entanglement wedge extends to the horizon. In the case where the value of the dilaton at the endpoints is specified, the opening angle at the horizon is given by $\frac{2\pi}{n}$ so that the parent spacetime is smooth. Right: For a bulk interval which crosses the horizon, the entanglement wedge is the region a itself.	154
4.10	Two gauge-invariant specifications of region a : On the left, we specify the value of the dilaton at its endpoints, along with whether the points lie on the same or opposite sides of the horizon from a chosen asymptotic boundary. On the right, we specify the geodesic distance of its endpoints from a chosen asymptotic boundary.	155

4.11	The region we choose is an annulus bounded by spheres with areas $\mathcal{A}_A$ and $\mathcal{A}_B$. On the left, the annulus is on the same exterior region in which case, the entanglement wedge extends out to the horizon. On the right, the annulus includes the horizon in which case, the entanglement wedge is the same as the region itself.	158
4.12	An example configuration illustrating subtleties that arise in time-dependent settings. For simplicity, we work at leading order in G and choose wedges a and b such that we have $e_{\max} = e_{\min} = e$. In this case, $e(a)$ is given by a restricted maximin procedure (see Ref. [266] for a proof of this statement). Consider a subregion a that forms a horseshoe-shaped wedge on a Cauchy slice in asymptotically flat spacetime. The portion of the boundary of $e(a)$ that does not coincide with the boundary of a is shown in red. Subregion b is chosen such that a portion of its boundary (shown in green) is null like to the red segment of $e(a)$. Generically, this green portion of ∂b will have smaller area than the red portion of $e(a)$. We also choose b such that it is normal everywhere such that $e(b) = b$. Thus a and b satisfy $a \subset e(b)'$ and $b \subset e(a)'$. Since computing the restricted entanglement wedge $e(a) _{b'}$ requires all candidate Cauchy slices to include the boundary of b , the surface $e(a)$ can no longer be a valid maximin surface in b' . Instead, a new surface—comprising the green portion of ∂b along with the null surfaces that connect it to region a —will have smaller area and be present on all Cauchy slices. This surface therefore becomes $e(a) _{b'}$	161
5.1	Panel (a) depicts a hypersurface orthogonal timelike congruence generated by a slice Σ_0 . In panel (b), deformations of the slice are defined by surfaces $\tau = F(x)$ where τ is the proper time along the timelike generator through $x \in \Sigma_0$. Panel (c) depicts a local deformation along a “pencil” of generators near a point x_0	166
5.2	Depicted are the codimension-0 regions between the chosen reference slice Σ_{ref} and the local deformations of the slice Σ_0 . The undeformed slice is defined by $\tau = F[x]$ (top left), and the deformations at x_1 and x_2 (right, bottom left) are implemented by $F_{\epsilon_1}(x)$ and $F_{\epsilon_2}(x)$. The terms appearing in the second variation correspond to the complexities associated with these codimension-0 regions.	170
5.3	If the timelike quantum expansion is everywhere negative throughout the congruence, then the complexity associated with all future slices Σ_t is bounded above by the volume of the initial slice Σ_0	172

List of Tables

2.1	Correspondence between an evaporating black hole and cosmological de Sitter spacetime.	71
-----	--	----

Acknowledgements

Firstly, I thank my advisor, Yasunori Nomura, for his willingness to support my scientific and professional endeavors, and for his encouragement to pursue my own ideas. I would also like to thank the collaborators I've had the great fortune to work with at Berkeley and beyond, including Chitraang Murdia, Violet Concepcion, Samuel Weiss, Sami Kaya, Pratik Rath, and Lucas Kovalsky.

Beyond scientific collaboration, I would like to thank the graduate students and postdocs who first welcomed me into the LITP (then BCTP) as a young graduate student after the COVID shutdown lifted. In particular, I thank Vincent Su, Liz Wildenhain, Hugo Marrochio, and Masamichi Miyaji. Their friendly presence, engaging intellectual discourse, and willingness to integrate me into the BCTP during a precarious time was invaluable.

Throughout the years, I have met many wonderful people in Berkeley's physics community. I am especially thankful to have met Hana Mir, who over the last several years has become a dear friend, and whose friendship has been a Shining Light in some of the most challenging times.

I am eternally grateful for my family, and especially my parents, for their unlimited support and encouragement in all of my endeavors, no matter how big or small.

I would also like to extend special thanks to one of my oldest friends, Casey Tessier, who first stood by me on the high school drumline, and continued to stand by me for the next decade.

Lastly, I thank the many wonderful undergraduate students I've had the privilege of teaching, whose enthusiasm for learning served as a daily reminder of why I started this journey in the first place.

Chapter 1

Introduction

One of the most profound insights to come out of the last half-century of quantum gravity is the holographic relationship between information and geometry. The first hints at this relationship appeared in the study of black hole thermodynamics, with the realization that the entropy of a black hole is equal to its area [1, 2]

$$S_{BH} = \frac{A_{BH}}{4G} \quad (1.1)$$

These early investigations naturally lead to the notion of generalized entropy as the sum of the horizon area and semiclassical matter entropy outside the horizon,

$$S_{gen} = \frac{A_{BH}}{4G} + S_{out} , \quad (1.2)$$

and produced many novel insights, like the generalized second law of thermodynamics, Bekenstein’s bound on the energy/entropy ratio in gravitating spacetimes (and its covariant generalizations), and even a derivation of Einstein’s equations from the thermodynamics of causal horizons[3–6]. By the end of the 20th century it had become clear that semiclassical quantum gravitational systems obey a general holographic principle, whereby information about a region is holographically encoded on its boundary [7–9].

The most compelling instantiation of this principle is the AdS/CFT correspondence, which provides a formal mathematical duality between quantum gravity in $d+1$ -dimensional Anti-de Sitter (AdS) spacetimes, and non-gravitating conformal field theories (CFTs) living at their d -dimensional boundaries [10]. In the last several decades, computational techniques developed within this context have allowed us to directly relate information theoretic quantities in the boundary CFT to geometric quantities in the AdS bulk, elevating the holographic principle to an exact mathematical equality.

One important example of this relationship is entanglement wedge reconstruction, by which operators in a boundary subregion R can be used to construct operators in the bulk subregion $e(R)$ (the entanglement wedge), which it encodes [11, 12]. Another important example is the quantum extremal surface (QES) formula, derived using gravitational path integral techniques, which equates the von Neumann entropy of boundary subregions R with the generalized entropy of the bulk surface $\gamma_R = \partial e(R)$ bounding its entanglement wedge

[13–16],

$$S_{vN}(R) = \min_{\gamma} \left\{ \frac{A(\gamma)}{4G} + S_{bulk} \right\} = S_{gen}(\gamma_R) \quad (1.3)$$

thus refining earlier conceptions of entropy-area duality. Such results have been used, for example, to understand the flow of information during black hole evaporation, resolving longstanding conceptual paradoxes [17–20].

In the last decade it has also become understood that information-theoretic quantities beyond entanglement entropy, such as state/circuit complexity and the complexity of bulk reconstruction, capture important features of quantum gravitational dynamics, and likewise admit a geometric description [21–24]. It has been proposed, for example, that the quantum complexity of CFT states describing AdS black holes scales with the volume of the black hole interior [25, 26]

$$C \sim \frac{V_{BH}}{G\ell} \quad (1.4)$$

and, more generally, that the complexity of a state on a boundary subregion R may scale with the volume of the maximal slice through its bulk entanglement wedge [27, 28]

$$C(R) \sim \max_{\Sigma \in e(R)} \frac{\text{Vol}(\Sigma)}{G\ell} + \text{corrections} \quad (1.5)$$

Despite the immense progress, it remains a challenge to understand precisely how the holographic principle manifests for spacetimes like our own, where the computational control afforded by AdS/CFT is no longer available. Although it is still believed that the semiclassical physics within a gravitating region is encoded holographically on its boundary, there exists no AdS/CFT-like dictionary between the semiclassical and holographic degrees of freedom. Beyond the absence of a dictionary, there are subtler issues. For example, once gravitational fluctuations are allowed, it becomes challenging to specify a subregion consistently across geometries. To avoid the latter issue, one could try to restrict their attention to holography of causal horizons (which are naturally coordinate independent), but the primary issue persists. Moreover, it is not fully understood which insights and computational techniques from AdS/CFT are adaptable to general spacetime subregions.

Understanding holography in contexts beyond AdS/CFT is the primary aim of this dissertation. In Part I, we draw on many insights from AdS/CFT, such as the relationship between entanglement structure and geometry, to address how, for limited semiclassical observers, regions beyond causal horizons emerge from the collective dynamics of near-horizon degrees of freedom, thus realizing the holographic principle. In Part II, we consider general gravitating subregions and investigate how quantum information measures such as entanglement entropy and quantum complexity relate to geometry, thereby providing generalizations of the relations in Eq. 1.3 and Eq. 1.5.

Part I is organized as follows. In Chapter 2, we consider how, for exterior observers, the interior of a black hole emerges from the collective dynamics of microscopic degrees of freedom living near the horizon. We analyze the near-horizon mode structure, and show that their collective semiclassical dynamics and entanglement structure constitute an effective holographic theory of the interior. We contrast this framework with the standard framework

of AdS/CFT, where the global geometry manifestly exists, but where unitarity (e.g. during black hole evaporation) is only emergent after including non-perturbative quantum effects. We also show that a similar analysis in the de Sitter static patch suggests that the region beyond the cosmological horizon emerges from the dynamics of near-horizon modes on the interior.

In Chapter 3, we refine the relationship between the exterior observer’s description and the global description. We investigate an apparent paradox which arises in evaporating black hole spacetimes. In the global picture, it is understood that Hawking radiation encodes the black hole interior, giving exterior observers apparent causal access to the interior, which should be forbidden. We show that in the exterior picture, this apparent paradox is manifestly resolved by the fact that the interior emerges from the effective theory near the horizon exterior, a region which always remains in causal contact with the radiation. We argue that, in the exterior picture, the interior only becomes effectively causally separated due to the semiclassical observer’s inability to perform arbitrarily complex operations on the radiation.

Part II is organized as follows. In Chapter 4, we develop a gravitational path integral approach for computing the von Neumann entropy of bulk subregions, and show that it obeys a QES-like formula. For boundary subregions, the path integral approach draws on the relationship between fixed geometry states and random tensor networks. We argue that this correspondence can be extended to bulk subregions, allowing one to relate the bulk path integral calculation to a tensor network calculation. We explicitly employ this path integral technique for bulk subregions in JT gravity. In this context, we show that different covariant specifications of the bulk subregion can lead to different Renyi entropies, but that the von Neumann entropy is robust against this choice.

In Chapter 5, we investigate to what extent the techniques used for studying the dynamics of entanglement entropy for bulk subregions can be adapted to study the dynamics of computational complexity. In order to do this, we propose a notion of generalized complexity as the sum of a region’s volume plus the semiclassical complexity of the fields supported on it. We establish a set of consistency conditions necessary for applying geodesic focusing arguments. We show that codimension-0 measures of complexity are required, and that these measures need to obey strong subadditivity. We then demonstrate that focusing conditions can be used to establish a volume upper-bound on the complexity, as well as a novel quantum strong energy condition which bounds the timelike energy-momentum flux below by local variations of the complexity.

Part I

The Emergence of Spacetime Beyond Causal Horizons

Chapter 2

Black Hole and de Sitter Microstructures from a Semiclassical Perspective

2.1 Introduction and Summary

For a long time, the puzzle of black hole information loss has caused confusion about how gravity works at the quantum level [17, 29, 30]. This confusion arises mostly due to the following features of the semiclassical description of black holes:

- (i) Despite the fact that spacetime is fundamentally quantum mechanical, it is described as a classical object.
- (ii) While the fundamental theory has a preferred class of time foliations for spacetimes with a horizon, general relativity seems to treat all the coordinates equally.

The purpose of this paper is to elucidate these points and present a coherent picture in which the results of semiclassical theory are consistently interpreted to address issues related to the information puzzle. This picture builds on recent developments of our understanding of quantum gravity.

The first hint that spacetime consists of quantum degrees of freedom came from the discovery that black holes have entropy [1, 31]. In the standard statistical mechanical interpretation, this implies that the horizon, a region which general relativity describes as empty space, contains a quantum mechanical substance. This interpretation is indeed supported by the anti-de Sitter (AdS)/conformal field theory (CFT) correspondence [10]—a concrete realization of holography [7–10, 32].

Recent progress in understanding the AdS/CFT correspondence, and holography more generally, has shown that while general relativity is diffeomorphism invariant, there appear

to be a preferred class of coordinates at the quantum level. From the viewpoint of the boundary theory, these coordinates correspond to descriptions based on quantum operators constructed by a simple procedure or operators which are not exponentially complex in fundamental degrees of freedom [23, 33–37], and they cover only a portion of spacetime when there is a horizon. In a system with a black hole, these coordinates are associated with Schwarzschild time slicing, or any other time foliation associated with an observer located outside the horizon. This conforms to the earlier idea that black hole evolution obeys the standard rules of quantum mechanics—and hence is unitary—when viewed by an external observer [38–41]. Indeed, we see that there is nothing unusual with the results of semiclassical theory if a black hole is described using only external frames.

The issue, then, is how to interpret the existence of the black hole interior which arises when the external coordinates are analytically extended in general relativity. We take the view that the picture of the black hole interior arises only effectively at the semiclassical level. In particular, we adopt the construction in Refs. [42–46] in which the interior emerges because of the special, chaotic nature of the horizon dynamics.¹ A key idea is that when the black hole is described in an external quasi-static reference frame, its large acceleration with respect to the free falling frame makes the dynamics at the horizon—more precisely the stretched horizon [40]—appear string theoretic, which is chaotic across all low energy species. This makes the black hole vacuum microstate generic in the relevant microcanonical ensemble, allowing us to erect the effective theory of the interior.

The erected effective theory is defined only up to errors of order $e^{-(S_{\text{BH}}+S_{\text{rad}})/2}$, where S_{BH} and S_{rad} are the Bekenstein-Hawking entropy and the coarse-grained entropy of the degrees of freedom entangled with the black hole, respectively. Operators describing the interior are state dependent [47, 52, 53], though only weakly in the sense of Refs. [46, 51]. Reflecting the fact that the Hilbert space associated with the black hole system is finite dimensional, the effective theory can be used only for a finite time interval; the existence of the black hole singularity is consistent with this [42]. Similarly, weak cosmic censorship can be viewed as a statement that a distant description, which corresponds to a simple boundary description in holography, can be used for an arbitrarily long time, to the extent that the theory is well defined in the infrared.

In this paper, we review the construction described above and refine it, including subtle evolutionary effects for a dynamically formed black hole. We present the whole framework in a coherent manner and address various issues associated with it, including its realization in the boundary description of holography. In doing so, we also discuss the relation of the present construction to other recent related works. In particular, we discuss the relation between the picture described here and that based on quantum extremal surfaces and entanglement wedge reconstruction [18–20, 54, 55]. We argue that the latter emerges through coarse graining necessary to describe a semiclassical black hole without specifying its microscopic structure [45]. We also see that the two are closely tied, respectively, to the canonical and path integral formulations of quantum mechanics [56].

¹ A detailed description of relations of this construction to earlier work [47–51] is given in Ref. [46] and throughout this paper.

We expect that at the macroscopic level, generic (quasi-)static horizons, including cosmic horizons, have locally the same statistical features. Indeed, the horizon of de Sitter spacetime has the same entropy per area as that of a black hole [57], and so is the temperature of the Hawking cloud at the stretched horizon. In fact, building on the analyses in Refs. [43, 58], we see that the holographic description of a static patch of de Sitter spacetime is very much an “inside-out” version of that of a black hole. We present this description in the context of more general holography for cosmological spacetimes [59, 60]. We also discuss the relationship of the present description with other recent proposals [61–66].

2.1.1 Overall picture and outline of the paper

In the rest of this section, we present an overview of the picture presented in this paper, pointing to where the details of each subject are covered. The assumptions about the setup which we adopt throughout the paper will be summarized at the end of this section.

No puzzle for a black hole when viewed from the exterior

Consider a non-rotating, non-charged black hole in a 4-dimensional asymptotically flat spacetime. At the classical level, the black hole is uniquely specified by one continuous parameter: its mass M . We view this system from a distance, i.e., we describe it using Schwarzschild time slicing or something related to it in a simple manner.

When quantum effects are included, the black hole has a finite entropy

$$S_{\text{BH}}(M) = 4\pi \frac{G_{\text{N}} M^2}{\hbar c}, \quad (2.1.1)$$

where G_{N} is Newton’s constant. This means that at the quantum level the black hole is characterized by discrete—albeit exponentially dense—states, rather than a continuous, classical number.² Specifically, the number of independent states in the energy interval between $M - \delta M/2$ and $M + \delta M/2$ is given by

$$\mathcal{N}(M) \sim e^{S_{\text{BH}}(M)} \frac{\delta M}{M}. \quad (2.1.2)$$

Note that this is a general phenomenon in quantum mechanics. Like other physical systems, the entropy in Eq. (2.1.1) diverges in the limit $\hbar \rightarrow 0$.

The semiclassical theory treats the black hole as a classical object while retaining the finite nature of the entropy. This is not inconsistent: the concept of entropy can be defined at the level of thermodynamics without explicitly taking into account the fundamental discreteness. The disadvantage of this treatment, however, is that one can no longer resolve each microstate, hence requiring a statistical, or thermodynamic, treatment of the system [1–3, 31].

² Of course, quantum mechanics allows for a superposition of these independent states, so that the expectation value of the energy, or M , can take continuous values.

At the semiclassical level, a black hole is described as having a definite mass M but with the Hawking cloud around it, which is in a thermal mixed state of temperature

$$T_{\text{H}} = \frac{\hbar c^3}{8\pi k G_{\text{N}} M}, \quad (2.1.3)$$

where k is the Boltzmann constant. This is a proxy for an ensemble of black hole microstates with energies between $M - \delta M/2$ and $M + \delta M/2$, which dominates the corresponding microcanonical ensemble of states associated with the spacetime region near the black hole. Below, we adopt natural units $\hbar = c = k = 1$.

The formation and evaporation of a black hole described in an external frame is a process in which quantum information in the initial collapsing matter is dispersed among spacetime degrees of freedom, represented by the Bekenstein-Hawking entropy; Hawking emission then transfers it back to matter degrees of freedom in the semiclassical description. For an external observer, this entire process is unitary if all the microscopic degrees of freedom are accounted for. An important point is that with quantum effects, the instantaneously-defined apparent horizon is stretched, at which the local (Tolman) Hawking temperature is the string scale [40].³ The trajectory of this stretched horizon is timelike, so an object falling into a black hole reaches there within finite time. This object is then absorbed by the black hole, whose information will eventually be sent back to ambient space as Hawking radiation.

This implies that in the external frame description, the stretched horizon behaves as a regular material surface as far as the flow of quantum information is concerned. The semiclassical theory, however, treats the horizon degrees of freedom appearing in the middle of the process to be classical, or at best an ensemble of quantum states represented by the Hawking cloud. Therefore, in any semiclassical calculation, the quantum information in the initial matter *must* appear to be lost in the final state; there is no way to completely describe the microscopic information on the horizon while staying in the semiclassical regime. In other words, Hawking's calculation [31] must have led to information loss, and indeed it did [17].

Emergence of the interior

A puzzling feature of the picture described above is that general relativity seems to imply that the black hole horizon is smooth. Namely, when an object freely falls into the horizon, it does not experience anything special there. This is obviously not the case when an object falls into a regular material surface.

As discussed in Refs. [42–46], the stretched horizon is distinguished from other, regular material surfaces by its chaotic [67], fast-scrambling [68, 69] dynamics across *all* low energy species. Here, by all low energy species, we mean all quantum fields appearing in the low energy effective theory below the string scale. These features arise because of the large relative acceleration, of order the string scale, between the quasi-static frame (a natural

³ By the string scale, we mean the scale at which the low energy effective field theory description breaks down. The appearance of this scale is associated with the nonzero Newton's constant, which controls quantum corrections to the system since it always appears with $\hbar$.

frame in holography) and the free falling frame at the stretched horizon. It is this aspect of the horizon dynamics that allows us to erect a description in which an object falls freely through the stretched horizon. This is done by making the state take a fully generic form in the relevant microcanonical ensemble.⁴

Specifically, when described in a quasi-static reference frame, quantum degrees of freedom of a black hole consist of modes in a spatial region near the stretched horizon (called the zone) as well as those at the stretched horizon. We call them zone and horizon modes, respectively. While the dynamics of the former is described by the low energy effective field theory, that of the latter is not. Note that these modes are defined for each time interval in which the black hole can be viewed as quasi-static.

Let us now focus on a (small) subset of the zone modes which is relevant for describing an infalling physical object. In particular, the object will be described by excitations of these modes over the black hole vacuum. Let us call these modes hard modes and all other black hole (zone and horizon) modes soft modes. From the assumption that the horizon degrees of freedom obey chaotic dynamics, we can then conclude that a black hole vacuum microstate takes the form

$$|\Psi(M)\rangle = \sum_n e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} c_{ni_n} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle, \quad (2.1.4)$$

where $|\{n_\alpha\}\rangle$ represent states of the hard modes, specified by the occupation numbers n_α for each mode α , and $|\psi_{i_n}^{(n)}\rangle$ are the states of the soft modes that have energy $M - E_n$, where E_n is the energy carried by $|\{n_\alpha\}\rangle$. Note that since the total energy of the black hole system is constrained to be M , the number of independent horizon-mode states $|\psi_{i_n}^{(n)}\rangle$ coupling to $|\{n_\alpha\}\rangle$ in Eq. (2.1.4) is given by $S_{\text{bh}}(M - E_n)$, the density of states at energy $M - E_n$. Here,

$$S_{\text{bh}}(E) = 4\pi E^2 l_{\text{P}}^2 \quad (2.1.5)$$

is the Bekenstein-Hawking entropy density at energy E with l_{P} being the Planck length. (The contribution from the hard modes to the black hole entropy is negligible.) Also note that because of the chaotic nature of the horizon dynamics, the coefficients c_{ni_n} take random values across all low energy species. In fact, the quasi-static form of Eq. (2.1.4) is achieved quickly after any disturbance, i.e. within the scrambling timescale of order $M l_{\text{P}}^2 \ln(M l_{\text{P}})$ [68, 69].

The universality of the form of the state in Eq. (2.1.4) allows us to erect the effective theory of the interior. Specifically, we can define the normalized state coupling to $|\{n_\alpha\}\rangle$ in Eq. (2.1.4)

$$\|\{n_\alpha\}\rangle = \varsigma_n e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} c_{ni_n} |\psi_{i_n}^{(n)}\rangle, \quad (2.1.6)$$

where

$$\varsigma_n = \frac{1}{\sqrt{\sum_{i_n=1} e^{S_{\text{bh}}(M-E_n)} |c_{ni_n}|^2}} = e^{\frac{E_n}{2T_{\text{H}}}} \sqrt{\sum_m e^{-\frac{E_m}{T_{\text{H}}}} \left[1 + O\left(e^{-\frac{1}{2}S_{\text{bh}}(M)}\right)\right]}. \quad (2.1.7)$$

⁴ For a charged or rotating black hole, the relevant ensemble consists of microstates with charge or angular momentum constrained to lie within a small window dictated by quantum uncertainties.

Plugging this into Eq. (2.1.4), we obtain the standard thermofield double form [70, 71]

$$|\Psi(M)\rangle = \frac{1}{\sqrt{\sum_m e^{-\frac{E_m}{T_H}}}} \sum_n e^{-\frac{E_n}{2T_H}} |\{n_\alpha\}\rangle |\{n_\alpha\}\rangle, \quad (2.1.8)$$

up to exponentially small corrections of order $e^{-S_{\text{bh}}/2}$. We can thus evolve an object in the zone (generated by acting creation/annihilation operators on $|\{n_\alpha\}\rangle$'s) using the time evolution operator associated with the proper time of the falling object, which is different from the original time evolution operator in the boundary theory but still local in bulk spacetime [42–46]. This implies the existence of a description in which the falling object passes the horizon and enters the black hole interior smoothly.

We emphasize that the criterion for a state to take the universal form in Eq. (2.1.4) is much stronger than that for regular thermalization; in particular, it must exhibit universal thermalization across all low energy species. It is reasonable to expect that such strong thermalization is achieved within a reasonable timescale only by the string scale dynamics, which singles out the stretched horizon. We also stress that the prescription of obtaining interior spacetime described here does not require a detailed knowledge about microscopic dynamics of quantum gravity; only some basic assumptions are sufficient.

As the evaporation of a black hole progresses, information stored in the horizon degrees of freedom is gradually transferred to Hawking radiation. After the Page time [41], entanglement between hard modes and early Hawking radiation becomes nonnegligible, so that the interior description must involve this Hawking radiation [50]. The way this occurs is that the black hole interior, more precisely $|\{n_\alpha\}\rangle$'s in Eq. (2.1.8), must involve early Hawking radiation *in addition to* the black hole degrees of freedom.

This last statement may seem to contradict the recent claim about entanglement wedge reconstruction that the interior of an old black hole can be reconstructed using only early Hawking radiation [18, 54, 55]. This is, however, not the case [45]. A key point is that such entanglement wedge reconstruction makes crucial use of boundary time evolution, which allows for reconstructing states of the black hole and radiation modes at an earlier time. On the other hand, the effective theory of the interior uses these modes directly at the time relevant for describing an infalling object. This understanding resolves an apparent puzzle associated with causality in entanglement wedge reconstruction.

In this paper, we first perform a detailed analysis of the near horizon mode structure in Sections 2.2 and 2.3, and study its relation to the holographic boundary picture in Section 2.4. The construction of the effective theory of the interior is presented in Sections 2.3 and 2.5, where we also discuss the refinement of the construction needed to accommodate the effects of evaporation on the instantaneous state of the zone modes. The relationship of this picture to entanglement wedge reconstruction is discussed in Section 2.5.

de Sitter holography

The similarity between the static patch description of de Sitter spacetime and the external description of black hole spacetime has been studied for a long time [57]. Based on this similarity, it was suggested that de Sitter spacetime may admit a holographic description

with finite-dimensional Hilbert space [72, 73]. In this paper, building on the earlier analysis in Refs. [43, 58], we develop a holographic description of de Sitter spacetime based on the (quasi-)static picture.

First, we note that the analogue of a collapse formed, single-sided black hole is cosmological de Sitter spacetime, which arises approximately in the middle of a cosmological history, e.g. at late times in a bubble universe with positive cosmological constant. In this case, there is a single static patch as viewed by an observer (timelike geodesic), which is the inside-out analogue of the exterior of the black hole. We define the zone and horizon modes as those inside (on the observer side of) the stretched horizon and those on the stretched horizon, respectively. As in the case of a black hole, a microstate for the de Sitter vacuum then takes the form

$$|\Psi(E)\rangle = \sum_n e^{S_{\text{dS}}(E-E_n)} \sum_{i_n=1} c_{ni_n} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle, \quad (2.1.9)$$

where $|\{n_\alpha\}\rangle$ represent states of the hard modes and $|\psi_{i_n}^{(n)}\rangle$ are the states of the soft modes that have energy $E - E_n$, where E_n is the energy of $|\{n_\alpha\}\rangle$ measured at the location of the observer. Here,

$$S_{\text{dS}}(E) = \pi E^2 l_{\text{P}}^2 \quad (2.1.10)$$

represents the de Sitter entropy, with E being the “energy” of the de Sitter vacuum related to the Hubble radius α by

$$E = \frac{\alpha}{l_{\text{P}}^2}. \quad (2.1.11)$$

Again, the coefficients c_{ni_n} take random values across all low energy species because of universally chaotic dynamics at the horizon. This randomization is achieved rather quickly, within the scrambling timescale of order $\alpha \ln(\alpha/l_{\text{P}})$ [74].

Using a microstate in Eq. (2.1.9), the effective theory describing a region outside the horizon (as viewed from the observer) can be constructed analogously to the black hole case. By identifying the normalized state $\|\{n_\alpha\}\rangle\rangle$ using Eq. (2.1.9), we can rewrite the full state as

$$|\Psi(E)\rangle = \frac{1}{\sqrt{\sum_m e^{-\frac{E_m}{T_{\text{H}}}}}} \sum_n e^{-\frac{E_n}{2T_{\text{H}}}} |\{n_\alpha\}\rangle \|\{n_\alpha\}\rangle\rangle, \quad (2.1.12)$$

where T_{H} is the de Sitter temperature. This state represents the semiclassical vacuum state of global de Sitter spacetime at time when the effective theory is erected.

We thus find that the effective global de Sitter picture emerges from a cosmological de Sitter spacetime as a collective phenomenon involving horizon modes. Like the black hole case, the effective theory of global de Sitter spacetime is intrinsically semiclassical in that there is an intrinsic ambiguity of order $e^{-S_{\text{GH}}/2}$ in the definition of the theory, where S_{GH} is the Gibbons-Hawking entropy. This, therefore, addresses the issue identified as a puzzling feature in Ref. [75] that symmetries of classical (global) de Sitter spacetime cannot be implemented exactly in a finite-dimensional Hilbert space.

The holographic theory of de Sitter spacetime described here is presented in Sections 2.3 and 2.5. Its relation to holography in more general spacetimes is discussed in Section 2.4, where we will see how static patch de Sitter holography arises naturally from holography in

more general cosmological spacetimes. In Section 2.5, we will describe the construction of the effective theory and see that this theory is sufficient to provide a semiclassical description of future measurements of the observer.

Intrinsically two-sided systems

While we find that the analytically extended Schwarzschild and global de Sitter spacetimes emerge from the more physical, single-sided spacetimes via collective dynamics involving horizon modes, there is nothing theoretically wrong with considering an “intrinsically two-sided” system, which comprises two copies of the holographic theory for the single-sided system.

In Section 2.6, we analyze such two-sided systems using our framework. We will see that these two-sided systems lead to physics similar to the corresponding single-sided systems at the semiclassical level, despite the fact that the microscopic structures of relevant states are significantly different. We also comment on possible relations of our framework to other proposals for holographic theories developed in the context of intrinsically two-sided systems. In particular, we discuss the relationship of our description of de Sitter spacetime with the DS/dS correspondence [61, 62] and the Shaghoulian-Susskind proposal [63–66].

Gravitational path integral: ensemble from coarse graining

The approach described so far is formulated most naturally in the canonical formalism of quantum mechanics. For example, we have introduced the concept of horizon modes and their associated states. However, quantum mechanics must also be formulable using path integral. What does the picture look like in this case?

The path integral formalism has a very different starting point. In the context of quantum gravity, it is the collection of classical field configurations on classical geometries, which are then integrated over to obtain physical results. This implies, however, that a black hole (or de Sitter) spacetime appearing in the path integral must be treated as a (semi)classical object, so its detailed microscopic structure cannot be discriminated. In fact, since the energy differences between different black hole microstates are suppressed by $e^{-S_{\text{BH}}(M)}$, these states cannot be discriminated by direct measurement.

From the microscopic point of view, this means that a black hole appearing in gravitational path integral is a coarse-grained object. This mandatory coarse graining introduces the concept of ensemble averaging in interpreting the result of gravitational path integral. This interpretation, in fact, reproduces many features which are attributed to the ensemble nature of holographic theories in lower dimensional quantum gravity [19] using an ensemble of microscopic states in a single theory [45]. However, if the classical black hole spacetime necessarily represents an ensemble of microstates, how can the underlying unitarity of a black hole evolution be captured by the quantum extremal surface method [18, 54, 55] which uses such classical spacetime?

A key observation is that while gravitational path integral can only calculate ensemble averages, it can still do so for many different quantities. In particular, adopting the replica

method, the gravitational path integral can calculate the traces of powers of the density matrix of Hawking radiation $\overline{\text{Tr} \rho_R^n}$. This is the replica wormhole calculation of Refs. [19, 20]. From this, the ensemble average of the von Neumann entropy of the radiation can be obtained

$$\overline{S_R} = -\lim_{n \rightarrow 1} \frac{\partial}{\partial n} \overline{\text{Tr} \rho_R^n} \sim \min\{S_{\text{rad}}, S_{\text{bh}}\}, \quad (2.1.13)$$

which follows the Page curve. The reason why we could obtain the Page curve here is because we have calculated the ensemble average of the microscopic von Neumann entropy, which obeys the Page curve for all members of the ensemble. This is unlike Hawking’s calculation which gives the von Neumann entropy of the averaged state.

This replica method prescription involving gravitational path integrals is virtually equivalent to the entanglement island prescription for calculating entropies, adopted in Refs. [18, 54, 55]. This, therefore, provides the following interpretation of the results in Refs. [18, 54, 55]: while any formalism treating a black hole as a classical object, including the gravitational path integral, must involve an ensemble average over microstates, microscopic information about some quantities can still be deduced by computing such ensemble averages. The entanglement island prescription (implicitly) adopts this for von Neumann entropies.

The issue described here is discussed in Section 2.7. It is based on the picture outlined in Refs. [45, 56]. The idea that the semiclassical description involves an ensemble of microstates was also discussed in Refs. [76–80], and the understanding of the Page curve presented there is based on the developments in Refs. [81–87].

Quantum mechanics vs general relativity in quantum gravity

It is important that the extremization procedure involved in the calculation of an entanglement island is performed on global spacetime of general relativity. In particular, for a black hole spacetime, it must be performed on the whole spacetime including the interior of the black hole.

This picture, therefore, is complementary to that based on the external view of the black hole. Here the existence of the black hole interior is manifest, while understanding the unitarity of black hole evolution requires a method incorporating nonperturbative effects of quantum gravity, such as replica wormholes. On the other hand, in the framework based on the external view, the unitarity of the evolution is built-in, and the interior emerges only effectively as a collective phenomenon involving horizon (and possibly other, entangled) degrees of freedom.

Despite the fact that the two pictures appear very different, they give the same physical conclusions. In particular, a black hole evolves unitarily and has a smooth horizon. The origin of historical confusions about black hole physics come from the fact that only one of these features is manifest in a given low energy description; the other appears in a highly nontrivial manner. It is interesting that the description in which unitarity of quantum mechanics is manifest naturally comes with the canonical/Hamiltonian formulation of quantum mechanics, while the one in which the interior predicted by general relativity is manifest is naturally associated with the path integral/Lagrangian formulation. It is an interesting question if a

microscopic formulation of quantum gravity can make both these features manifest.

2.1.2 List of assumptions

Here we list out the assumptions that we use throughout the paper. We focus on spacetimes which are spherically symmetric at the semiclassical level, although excitations on them are not restricted to be spherically symmetric. We take the number of spacetime dimensions to be $3 + 1$, although our arguments can be generalized straightforwardly to $d + 1$ dimensions with $d \geq 3$.

When discussing black holes, we mostly consider a spherically symmetric, non-charged black hole in an asymptotically flat spacetime (or an analogous object such as a spherically symmetric, non-charged small black hole in an asymptotically AdS spacetime). It is straightforward to include the effect of a charge or rotation unless the black hole is extremal or near extremal. An extension to (near) extremal and lower dimensional black holes is expected to be nontrivial, since these black holes have different structures for the densities of states than “generic” black holes considered in this paper; see, e.g., Refs. [88, 89].

2.2 Quantum Field on Background Spacetime

In this section, we discuss the behavior of quantum fields in spherically symmetric background spacetimes. The calculation presented here is elementary. The results, however, are used in later sections, so we include them for completeness.

We consider a 4-dimensional spacetime with the metric

$$ds^2 = -f(r)dt^2 + \frac{1}{f(r)}dr^2 + r^2d\Omega^2, \quad (2.2.1)$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. For simplicity, we consider a minimally-coupled real scalar field Φ in this spacetime, whose action is given by

$$I = \frac{1}{2} \int d^4x \sqrt{-g} (-g^{\mu\nu} \partial_\mu \Phi \partial_\nu \Phi - m_\Phi^2 \Phi^2). \quad (2.2.2)$$

After changing the radial coordinate to the tortoise coordinate r_* defined by

$$dr_* = \frac{dr}{f(r)}, \quad (2.2.3)$$

the metric becomes

$$ds^2 = f(r) (-dt^2 + dr_*^2) + r^2 d\Omega^2. \quad (2.2.4)$$

Here, r should be regarded as a function of r_* . The equation of motion derived from Eq. (2.2.2) is then given by

$$\left(-\partial_t^2 + \partial_{r_*}^2 - \frac{f(r)f'(r)}{r} + \frac{f(r)}{r^2} \partial_\Omega^2 - m_\Phi^2 f(r) \right) (r\Phi) = 0, \quad (2.2.5)$$

where ∂_Ω^2 is defined by $\partial_\Omega^2 \chi = (1/\sin \theta) \partial_\theta (\sin \theta \partial_\theta \chi) + (1/\sin^2 \theta) \partial_\phi^2 \chi$.

We look for positive frequency solutions of the form

$$\Phi(t, r_*, \theta, \phi) = e^{-i\omega t} \frac{\varphi_{\ell m}(r_*)}{r} Y_{\ell m}(\theta, \phi) \quad (2.2.6)$$

with $\omega \geq 0$, where $Y_{\ell m}(\theta, \phi)$ represent real spherical harmonics, satisfying $\partial_\Omega^2 Y_{\ell m} = -\ell(\ell + 1)Y_{\ell m}$. This results in a linear, second-order differential equation for $\varphi_{\ell m}(r)$

$$-\frac{d^2}{(dr_*)^2} \varphi_{\ell m}(r_*) + \{V_\ell(r_*) - \omega^2\} \varphi_{\ell m}(r_*) = 0, \quad (2.2.7)$$

where the effective potential $V_\ell(r_*)$ is given by

$$V_\ell(r) = f(r) \left(\frac{f'(r)}{r} + \frac{\ell(\ell + 1)}{r^2} + m_\Phi^2 \right). \quad (2.2.8)$$

The corresponding equations for higher spin fields are generally more complicated, but they can also be derived using a semiclassical method [90–93].

We are interested in spacetime that has a horizon at radius r_+ determined by

$$f(r_+) = 0. \quad (2.2.9)$$

The location, r_s , of the stretched horizon [40] is determined by the condition that the proper distance between $r = r_+$ and r_s is the string length l_s :

$$\left| \int_{r_+}^{r_s} \frac{dr}{\sqrt{f(r)}} \right| \approx l_s, \quad (2.2.10)$$

giving

$$|r_s - r_+| \approx \frac{|f'(r_+)| l_s^2}{4}. \quad (2.2.11)$$

Here, we have assumed that $|f'(r_+)|$ is not much suppressed compared with the natural size determined by dimensional analysis, $|f'(r_+)| \sim 1/r_+$, which is indeed the case for the spacetimes that we consider in this paper. The Hawking temperature, as measured at r satisfying $f(r) = 1$, is given by

$$T_H = \frac{|f'(r_+)|}{4\pi}, \quad (2.2.12)$$

so that the location of the stretched horizon also coincides with the place where the local (Tolman) Hawking temperature $T_{\text{loc}}(r) = T_H / \sqrt{f(r)}$ becomes the string scale, $\approx 1/2\pi l_s$ [45].

2.2.1 Schwarzschild black hole

Let us consider a Schwarzschild black hole of mass M . The metric is given by

$$f(r) = 1 - \frac{r_+}{r}, \quad (2.2.13)$$

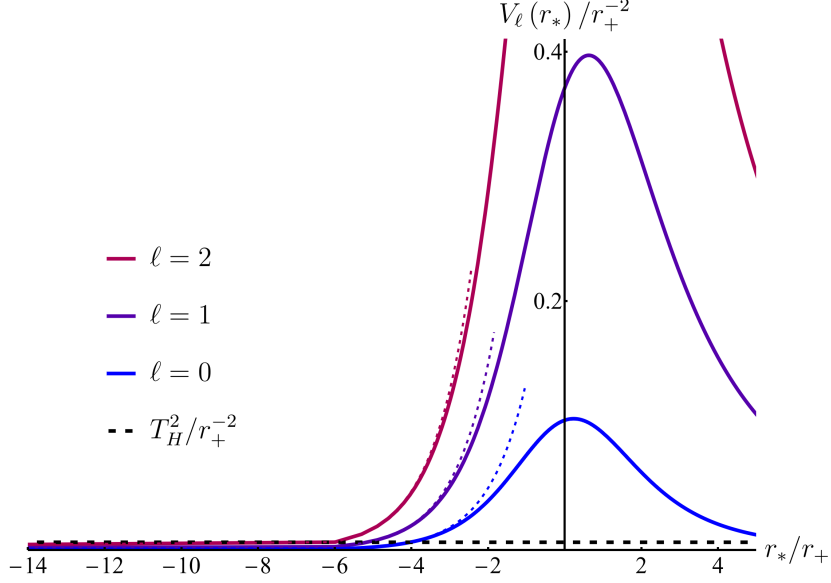


Figure 2.1: The effective potential V_ℓ of a Schwarzschild black hole (solid lines) for a minimally-coupled massless scalar field with $\ell = 0, 1, 2$ as a function of the tortoise coordinate r_* . The Hawking temperature $T_H = 1/4\pi r_+$ is indicated by the horizontal dashed line. The approximated curves in Eq. (2.2.20) are also plotted (dotted lines) in the region where they are relevant.

where $r_+ = 2Ml_p^2$. The tortoise coordinate is given by

$$r_* = r + r_+ \ln \frac{r - r_+}{r_+}, \quad (2.2.14)$$

which maps $r : (r_+, \infty)$ to $r_* : (-\infty, \infty)$, and the effective potential is

$$V_\ell(r_*) = \left(1 - \frac{r_+}{r}\right) \left(\frac{r_+}{r^3} + \frac{\ell(\ell+1)}{r^2} + m_\Phi^2\right), \quad (2.2.15)$$

which is plotted in Fig. 2.1 for $m_\Phi = 0$.⁵ The stretched horizon is located at

$$r_s - r_+ \approx \frac{l_s^2}{4r_+} \iff r_{*s} - r_+ \approx -2r_+ \ln \frac{2r_+}{l_s}, \quad (2.2.17)$$

⁵ For a general, not necessarily a scalar, field of $m_\Phi = 0$, we have

$$V_\ell(r_*) = \left(1 - \frac{r_+}{r}\right) \left(\frac{(1-s^2)r_+}{r^3} + \frac{\ell(\ell+1)}{r^2}\right), \quad (2.2.16)$$

where s is the spin-weight parameter and $\ell \geq |s|$. In the near horizon limit, this leads to the approximate potential of Eq. (2.2.20) but with $\lambda_\ell = (\ell^2 + \ell + 1 - s^2)/r_+^2$, giving a higher potential barrier for larger $|s|$ for a fixed value of $\ell - |s|$.

and the Hawking temperature is given by

$$T_H = \frac{1}{4\pi r_+}. \quad (2.2.18)$$

While we cannot solve Eq. (2.2.7) analytically, we can study the behavior of the field in the near horizon and far regions by making appropriate approximations. We are mostly interested in the near horizon “zone” region

$$r_s < r < r_z, \quad r_z \approx \frac{3}{2}r_+, \quad (2.2.19)$$

where r_z indicates the location of the potential barrier. The near horizon limit corresponds to $r - r_+ \ll O(r_+)$, where $r_* \approx r_+ + r_+ \ln[(r - r_+)/r_+] \iff r \approx r_+ + r_+ e^{(r_* - r_+)/r_+}$. The effective potential in this region is given by

$$V_\ell(r_*) \approx \lambda_\ell e^{\frac{r_* - r_+}{r_+}}, \quad \lambda_\ell = \frac{\ell^2 + \ell + 1}{r_+^2} + m_\Phi^2. \quad (2.2.20)$$

This approximation is valid if r_* is negative with $|r_*|$ sufficiently larger than r_+ ; see Fig. 2.1. This implies that we can trust solutions obtained using Eq. (2.2.20) if $\omega^2 \ll V_\ell(r_* = -r_+) \sim \lambda_\ell$.

Two independent real solutions of Eq. (2.2.7) with Eq. (2.2.20) can be taken as

$$\text{Re} \left[I_{2ir_+ \omega} \left(2\sqrt{\lambda_\ell} r_+ e^{\frac{r_* - r_+}{2r_+}} \right) \right], \quad \text{Im} \left[I_{2ir_+ \omega} \left(2\sqrt{\lambda_\ell} r_+ e^{\frac{r_* - r_+}{2r_+}} \right) \right], \quad (2.2.21)$$

where $I_\nu(x)$ is the modified Bessel function of the first kind. The first solution is exponentially increasing in r_* at large r_* , so it does not correspond to modes that are localized in the zone region; it corresponds to decaying modes for signals sent from the far region. We thus focus on the second solution, which is exponentially damped at large r_* . This solution is approximated by a trigonometric function in the near horizon region

$$\varphi_{\ell m}(r_*) \propto \text{Im} \left[I_{2ir_+ \omega} \left(2\sqrt{\lambda_\ell} r_+ e^{\frac{r_* - r_+}{2r_+}} \right) \right] \xrightarrow[r_* < 0]{|r_*| \gg r_+} \sin \left[\omega r_* + \omega r_+ (\ln(\lambda_\ell r_+^2) - 1) \right]. \quad (2.2.22)$$

Given that the effective potential at the stretched horizon has the value $V_\ell(r_{*s}) = \lambda_\ell l_s^2 / 4r_+^2$, we find

$$\omega \gtrsim \frac{\sqrt{\lambda_\ell} l_s}{2r_+}. \quad (2.2.23)$$

We also find that imposing a boundary condition at the stretched horizon makes the spectrum discrete, with the gap between adjacent levels given by

$$\Delta\omega \equiv \omega_n - \omega_{n-1} \approx \frac{\pi}{2r_+ \left| \ln \frac{2}{\sqrt{\lambda_\ell} l_s} \right|}, \quad (2.2.24)$$

where we have imposed a simple Dirichlet boundary condition $\varphi_{\ell m}(r_{*s}) = 0$ for illustrative purposes. Note that for a given ℓ , each level has $(2\ell + 1)$ -fold degeneracy corresponding to $m = -\ell, -\ell + 1, \dots, \ell$. While the details of the spectrum do depend on the boundary condition, its basic structure—the discreteness and the scale characterizing the gaps—does not.

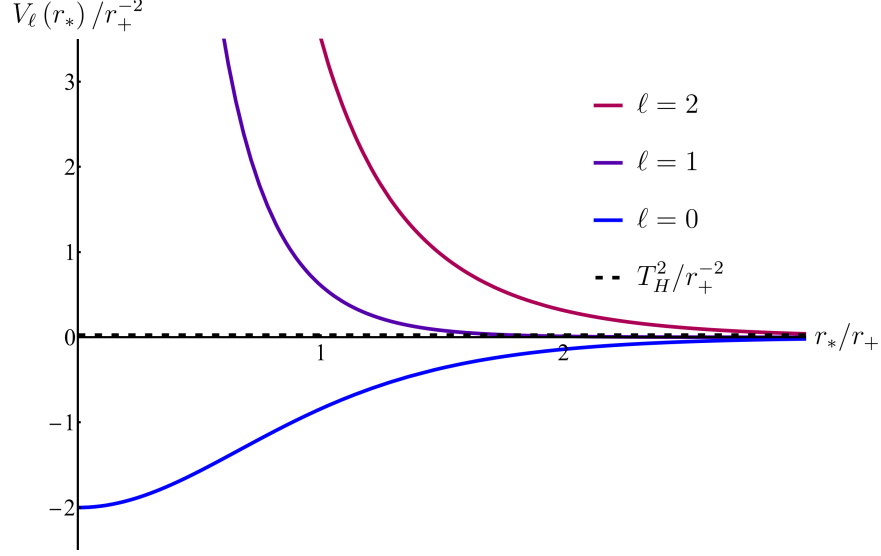


Figure 2.2: The effective potential V_ℓ of de Sitter spacetime for a minimally-coupled massless scalar field with $\ell = 0, 1, 2$ as a function of the tortoise coordinate r_* . The Hawking temperature $T_H = 1/2\pi r_+$ is indicated by the horizontal dashed line.

2.2.2 de Sitter spacetime

A similar analysis can be performed for de Sitter spacetime in static coordinates

$$f(r) = 1 - \frac{r^2}{r_+^2}, \quad (2.2.25)$$

leading to the same basic conclusion. Here, $r_+ = \alpha$ with α being the Hubble radius. The tortoise coordinate is given by

$$r_* = \frac{r_+}{2} \ln \frac{1 + \frac{r}{r_+}}{1 - \frac{r}{r_+}}, \quad (2.2.26)$$

which maps $r : (0, r_+)$ to $r_* : (0, \infty)$, and the effective potential is

$$V_\ell(r_*) = \left(1 - \frac{r^2}{r_+^2}\right) \left(-\frac{2}{r_+^2} + \frac{\ell(\ell+1)}{r^2} + m_\Phi^2\right), \quad (2.2.27)$$

which is plotted in Fig. 2.2 for $m_\Phi = 0$. The stretched horizon is located at

$$r_+ - r_s \approx \frac{l_s^2}{2r_+} \iff r_{*s} \approx r_+ \ln \frac{2r_+}{l_s}, \quad (2.2.28)$$

and the Hawking temperature is given by

$$T_H = \frac{1}{2\pi r_+}. \quad (2.2.29)$$

For $\ell \neq 0$, the effective potential blows up at small values of r_* . Therefore, if ω is small, we can restrict to $r_* > r_+$ which corresponds to the region near the cosmological horizon. For $\ell = 0$, the effective potential blows up at small r_* only when $m_\Phi > \sqrt{2}/r_+$; otherwise, the potential is confining, leading to a small number of bound states. Below, we focus on the case $m_\Phi > \sqrt{2}/r_+$ for $\ell = 0$.

With this restriction, the effective potential in the near horizon region is given by

$$V_\ell(r_*) \approx 4\lambda_\ell e^{-2\frac{r_*}{r_+}}, \quad \lambda_\ell = \frac{\ell^2 + \ell - 2}{r_+^2} + m_\Phi^2, \quad (2.2.30)$$

except for $\ell = 1$ with $m_\Phi = 0$, in which case

$$V_{\ell=1}(r_*) \approx \frac{32}{r_+^2} e^{-4\frac{r_*}{r_+}} \quad \text{for } m_\Phi = 0. \quad (2.2.31)$$

This approximation is valid only for $r_* \gtrsim r_+$. We can thus trust the solutions of Eq. (2.2.7) obtained using these $V_\ell(r_*)$'s only if $\omega^2 \ll V(r_* = r_+) \sim \lambda_\ell$ (or $\ll 1/r_+^2$ for $\ell = 1$, $m_\Phi = 0$).

The most general solution corresponding to the approximate potential in Eq. (2.2.30) is given by

$$\varphi_{\ell m}(r_*) = \text{Re} \left[A I_{ir_+\omega} \left(2\sqrt{\lambda_\ell} r_+ e^{-\frac{r_*}{r_+}} \right) \right], \quad (2.2.32)$$

where $A \in \mathbb{C}$ is an arbitrary constant. In order for the original field $\Phi(t, r_*, \theta, \phi)$ in Eq. (2.2.6) to be well-defined, the exact solution of $\varphi_{\ell m}(r_*)$ must vanish at least as fast as r_* when $r_* \approx r \rightarrow 0$. This condition fixes the phase of A in the approximate solution, leading to a single solution for $\varphi_{\ell m}(r_*)$ in the near horizon region:

$$\varphi_{\ell m}(r_*) \xrightarrow{r_* \gg r_+} \sin \left[\omega r_* - \frac{1}{2} \omega r_+ \ln(\lambda_\ell r_+^2) + \delta \right], \quad (2.2.33)$$

where δ is determined by the phase of A . For $\ell = 1$ and $m_\Phi = 0$, a similar analysis gives

$$\varphi_{\ell m}(r_*) \propto \text{Re} \left[A I_{\frac{i}{2}r_+\omega} \left(2\sqrt{2} e^{-\frac{2r_*}{r_+}} \right) \right] \xrightarrow{r_* \gg r_+} \sin \left[\omega r_* - \frac{\ln 2}{4} \omega r_+ + \delta \right]. \quad (2.2.34)$$

Given that the effective potential in Eq. (2.2.30) has the value $V_\ell(r_{*s}) = \lambda_\ell l_s^2/r_+^2$, we find

$$\omega \gtrsim \frac{\sqrt{\lambda_\ell} l_s}{r_+} \quad \text{for } (\ell, m_\Phi) \neq (1, 0), \quad (2.2.35)$$

and by imposing the boundary condition $\varphi_{\ell m}(r_{*s}) = 0$, we obtain a discrete spectrum with

$$\Delta\omega \approx \frac{\pi}{r_+ \left| \ln \frac{2}{\sqrt{\lambda_\ell} l_s} \right|} \quad \text{for } (\ell, m_\Phi) \neq (1, 0). \quad (2.2.36)$$

For $(\ell, m_\Phi) = (1, 0)$, the corresponding quantities are given instead by $V_\ell(r_{*s}) = 2l_s^4/r_+^6$ and

$$\omega \gtrsim \frac{\sqrt{2} l_s^2}{r_+^3}, \quad \Delta\omega \approx \frac{\pi}{r_+ \ln \frac{2^{3/4} r_+}{l_s}} \quad \text{for } (\ell, m_\Phi) = (1, 0). \quad (2.2.37)$$

Once again, the basic structure of the spectrum found here does not depend on the boundary condition at the stretched horizon.

2.2.3 General near horizon limit

We now see that the structure of the spectrum for

$$\omega \ll \sqrt{\frac{\ell^2}{r_+^2} + m_\Phi^2}, \quad \ell \gg 1 \quad (2.2.38)$$

in the near horizon region is universal for generic horizons, beyond the black hole and de Sitter spacetimes discussed so far. This reflects the universality of the near horizon limit, giving Rindler spacetime.

To see this, let us consider a spacetime with the metric in Eq. (2.2.1) which has a horizon at $r = r_+$. We assume that $f(r) > 1$ in the region $r > r_+$, which we call the allowed region. For a given horizon, this can always be arranged. Specifically, if $f(r) > 1$ in $r < r_+$, as in de Sitter spacetime, we redefine $r \rightarrow -r$ (and $r_+ \rightarrow -r_+$) to make the allowed region $r > r_+$. Note that this makes r and r_+ negative.

In the near horizon region, we then have $f(r) \approx f'(r_+)(r - r_+)$ with $f'(r_+) > 0$. Here, we have assumed that $f'(r_+)$ is not too suppressed compared with its natural size $f'(r_+) \sim 1/|r_+|$. (This excludes the horizon of a near extremal black hole from our consideration.) The tortoise coordinate in the near horizon region is then

$$r_* \approx \frac{1}{f'(r_+)} \ln \frac{r - r_+}{c|r_+|}, \quad (2.2.39)$$

where $c > 0$ is an unimportant $O(1)$ number defining the origin of r_* ,⁶ and the effective potential in this region is given by

$$V_\ell(r_*) \approx c|r_+|f'(r_+)\lambda_\ell e^{f'(r_+)r_*} \approx \lambda_\ell e^{f'(r_+)r_*}, \quad \lambda_\ell = \frac{f'(r_+)}{r_+} + \frac{\ell(\ell+1)}{r_+^2} + m_\Phi^2. \quad (2.2.40)$$

We want $\lambda_\ell > 0$ for this potential to trap modes in the near horizon region. (Here we exclude the nongeneric case of $\lambda_\ell = 0$ from consideration.) If $r_+ > 0$, this condition is satisfied for any values of ℓ and m_Φ . If $r_+ < 0$, we restrict our treatment to modes with large enough ℓ or m_Φ such that $\lambda_\ell > 0$.

For the above approximation to hold, we need $\omega^2 \ll V(r_* \sim -1/f'(r_+)) \sim \lambda_\ell$, which for $\ell \gg 1$ gives the condition in Eq. (2.2.38). Assuming that ω is in this range, we can obtain the general solution to Eq. (2.2.7) as

$$\varphi_{\ell m}(r_*) = \text{Re} \left[A I_{\frac{2i\omega}{f'(r_+)}} \left(2\sqrt{\frac{c|r_+|\lambda_\ell}{f'(r_+)}} e^{\frac{f'(r_+)r_*}{2}} \right) \right], \quad (2.2.41)$$

where $A \in \mathbb{C}$ is an arbitrary constant. The relevant solution is selected by a boundary condition at $r_* \approx O(r_+)$, which depends on spacetime under consideration. For Schwarzschild black hole and de Sitter spacetimes, these were given at $r \approx r_z$, and $r = 0$, respectively. This

⁶ In sections 2.2.1 and 2.2.2, c was taken to be $1/e$ and 2 , respectively.

fixes the phase of A and gives us a single solution for each ω . In the near horizon limit, it takes the form

$$\varphi_{\ell m}(r_*) \approx \sin\left(\omega r_* + \frac{\omega}{f'(r_+)} \ln \frac{c|r_+|\lambda_\ell}{f'(r_+)} + \delta\right) \quad \text{for } r_* < 0, |r_*| \gg |r_+|, \quad (2.2.42)$$

where δ is an $O(1)$ phase determined by the phase of A .

The existence of the stretched horizon, at $r_{*s} \approx -(1/f'(r_+)) \ln(4c|r_+|/l_s^2 f'(r_+))$, has two effects. First, since $V_\ell(r_{*s}) \approx \lambda_\ell l_s^2 / |r_+|^2$, frequency ω for given λ_ℓ is bounded from below:

$$\omega \gtrsim \frac{\sqrt{\lambda_\ell} l_s}{|r_+|}. \quad (2.2.43)$$

Second, imposing a boundary condition at $r_* = r_{*s}$ quantizes the spectrum:

$$\Delta\omega \sim \frac{\pi f'(r_+)}{2 \left| \ln \frac{2}{\sqrt{\lambda_\ell} l_s} \right|}. \quad (2.2.44)$$

We find that these are universal features of the spectrum in the regime of Eq. (2.2.38).

2.3 The Absence of Spacetime below the String Length

A key element of the analysis in the previous section is the hypothesis that spacetime, as we usually perceive, does not exist below the string length. In normal circumstances, this hardly affects low energy physics because of Wilsonian decoupling. In the presence of a horizon, however, large gravitational red/blue shift makes this fact relevant for low energy physics. Its significance can be best seen in the tortoise coordinate, in which the wavelength of a massless mode is preserved while propagating; in terms of this coordinate, the removal of the spacetime region within the string length from the horizon leads to excising a half line.

The proper distance, used in defining the stretched horizon in Eq. (2.2.10), is associated with the particular time foliation that leads to the static form of the metric, Eq. (2.2.1). In the case of an evaporating black hole, we apply our treatment to a sufficiently small time window, e.g. $\Delta t \lesssim r_+$, in which the system can be viewed as approximately static. The issue of time slicing will be discussed further in Section 2.4.

In this section, we demonstrate that the Bekenstein-Hawking entropy, S_{BH} , as well as Hawking temperature, T_{H} , follow only from two inputs from the ultraviolet (UV) physics (or two assumptions from the point of view of low energy theory):

- Spacetime does not exist below the string length, which introduces the “boundary” of space—the stretched horizon—at a proper distance l_s away from the horizon.

- Physics associated with the boundary is (maximally) quantum chaotic [67] across all low energy species [43], so the black hole and de Sitter vacuum microstates (or the microstates associated with any horizon) are typical in a suitable microcanonical ensemble.

In particular, these reproduce S_{BH} and T_{H} up to incalculable $O(1)$ coefficients. In the situation where the state evolves in time at the semiclassical level, for example when the black hole is evaporating or the system is perturbed by excitations, we assume that the typicality described above is reached quickly. In fact, it is believed that the dynamics associated with horizons is fast scrambling [68, 69].⁷

Next we discuss the UV sensitivities of the various modes, classifying them into modes that can be reliably described in the semiclassical theory and those that are intrinsically quantum gravitational. This provides a refinement of the concept of hard and soft modes introduced in Refs. [42, 43] to describe an analytic extension of spacetime at the microscopic level; the meaning of this analytic extension in quantum gravity will be discussed in more detail in Section 2.5. Finally, we will discuss how the vacuum and excited states of semiclassical theory are related to the microscopic description given here.

2.3.1 Entropy, temperature, and microstates for a semiclassical vacuum

As stated above, we assume that the unknown UV physics of quantum gravity appears in low energy physics as a lack of space below the proper length of order l_{s} . As we have seen in Section 2.2, this makes the spectrum discrete:

$$\omega_n^{(\ell)} = \omega_0^{(\ell)} + n\Delta\omega^{(\ell)} \quad n = 0, 1, 2, \dots, \quad (2.3.1)$$

where

$$\omega_0^{(\ell)} \approx \frac{\sqrt{\lambda_\ell} l_{\text{s}}}{|r_+|}, \quad \Delta\omega^{(\ell)} \sim \left\lfloor \frac{\pi f'(r_+)}{2 \ln \frac{2}{\sqrt{\lambda_\ell} l_{\text{s}}}} \right\rfloor. \quad (2.3.2)$$

Following Refs. [42, 43], we take the view that the energy of the system, which is usually attributed to the background, is carried by the quanta that fill these energy levels.

For concreteness, let us consider a Schwarzschild black hole. The case of de Sitter spacetime can be analyzed similarly, which we will discuss at the end of this subsection. Suppose that there are $N_n^{(\ell)}$ quanta of field Φ at the n -th level of angular momentum ℓ , which has $(2\ell + 1)$ -fold degeneracy. The total energy carried by Φ is then

$$E_\Phi = g_\Phi \sum_{\ell=0}^{\infty} \sum_{n=0}^{\infty} \omega_n^{(\ell)} N_n^{(\ell)}, \quad (2.3.3)$$

⁷ As discussed in Ref. [43], this assumption implies the absence of fundamental global symmetries (see, e.g., Refs. [94–97]); specifically, any linearly-realized global symmetry is explicitly broken by an $O(1)$ —or not exponentially suppressed—amount at the string scale.

where g_Φ is the number of degrees of freedom for field Φ . ($g_\Phi = 1$ for a real scalar field.) The picture of Refs. [42, 43] is that the sum of this energy for all low energy fields represents the total energy of the system (as measured at r satisfying $f(r) = 1$), which is determined self-consistently by the spacetime background used in calculating $\omega_n^{(\ell)}$. In the present case

$$E_{\text{bh}} = \sum_{\Phi} E_{\Phi} = M, \quad (2.3.4)$$

where the sum runs over all low energy fields that can be viewed as elementary in the effective field theory at a scale slightly below $1/l_s$.⁸

While the expressions for $\omega_0^{(\ell)}$ and $\Delta\omega^{(\ell)}$ are different for a field with spin, their values are of the same order as those of a scalar field. In particular, using Eq. (2.2.16) in footnote 5, one finds that λ_ℓ in Eq. (2.3.2) is simply replaced as

$$\lambda_\ell \rightarrow \lambda_{\ell,s} = \frac{\ell^2 + \ell + 1 - s^2}{r_+^2} \quad (\ell \geq |s|) \quad (2.3.5)$$

for $m_\Phi = 0$. This does not change the values of $\omega_0^{(\ell)}$ and $\Delta\omega^{(\ell)}$ much, as long as $|s| \approx O(1)$ which we assume here. In any case, the precise numbers for these quantities are not important, or trustable, as we will discuss in Section 2.3.2.

Because of the assumption of chaotic and fast scrambling dynamics at the stretched horizon, especially those across all low energy species, the distribution of energy among various species and levels is determined purely by the content of low energy fields. In particular, the distribution of quanta in each degree of freedom is given by maximizing the combinatorial numbers

$$C_\Phi(\{N_n^{(\ell)}\}) = \begin{cases} \prod_{\ell=0}^{\infty} \prod_{n=0}^{\infty} \frac{(N_n^{(\ell)} + 2\ell)!}{N_n^{(\ell)}!(2\ell)!} & (N_n^{(\ell)} \geq 0) & \text{for } \Phi: \text{ boson} \\ \prod_{\ell=0}^{\infty} \prod_{n=0}^{\infty} \frac{(2\ell+1)!}{N_n^{(\ell)}!(2\ell+1-N_n^{(\ell)})!} & (0 \leq N_n^{(\ell)} \leq 2\ell+1) & \text{for } \Phi: \text{ fermion} \end{cases} \quad (2.3.6)$$

under the constraint of Eq. (2.3.4).⁹ Following the standard analysis in statistical mechanics, we find

$$N_n^{(\ell)} = \frac{2\ell + 1}{e^{\beta\omega_n^{(\ell)}} \mp 1}, \quad (2.3.7)$$

where $\mp$ takes the minus and plus signs for bosonic and fermionic degrees of freedom, respectively, and β is determined by the condition

$$E_{\text{bh}} = \sum_{\Phi} g_{\Phi} \sum_{\ell=0}^{\infty} \sum_{n=0}^{\infty} \frac{(2\ell + 1)\omega_n^{(\ell)}}{e^{\beta\omega_n^{(\ell)}} \mp 1}. \quad (2.3.8)$$

⁸ We ignore the kinetic energy, which is not important for a black hole that is spherically symmetric at the classical level.

⁹ As in the standard statistical mechanics, this gives the most probable configuration of quanta. The probability of finding other configurations satisfying the constraint is not zero but exponentially suppressed.

Note that β in Eq. (2.3.8) does not depend on Φ , and there is no “chemical potential” for any Φ , since the dynamics at the stretched horizon is chaotic across all low energy species.

We can calculate the right-hand side of Eq. (2.3.8) by using Eq. (2.3.1) and replacing the sum over ℓ and n with the corresponding integrals. Assuming that

$$\ell \gg 1, \quad \frac{\ell^2}{r_+^2} \gg m_\Phi^2, \quad (2.3.9)$$

which is justified a posteriori as the integral is dominated by $\ell \sim r_+/l_s$, we may use

$$\omega_0^{(\ell)} \approx \frac{\ell l_s}{r_+^2}, \quad \Delta\omega^{(\ell)} \sim \frac{1}{\left| r_+ \ln \frac{r_+}{\ell l_s} \right|}. \quad (2.3.10)$$

This gives

$$E_{\text{bh}} \approx N_{\text{dof}} \frac{r_+^5}{\beta^4 l_s^2} \left(\left| \ln \frac{\beta}{r_+} \right| + O(1) \right), \quad (2.3.11)$$

where $N_{\text{dof}} = \sum_\Phi g_\Phi$ is the total number of degrees of freedom of low energy fields. Using

$$\frac{1}{l_P^2} \approx \frac{N_{\text{dof}}}{l_s^2} \quad (2.3.12)$$

(see, e.g., Ref. [98]) as well as $E_{\text{bh}} = M$ and $r_+ = 2Ml_P^2$, we find

$$\beta \approx Ml_P^2, \quad (2.3.13)$$

which is the inverse Hawking temperature. We also find that

$$S = \sum_\Phi g_\Phi \ln C_\Phi \approx M^2 l_P^2, \quad (2.3.14)$$

as indicated by the Bekenstein-Hawking entropy. Note that since the contribution to

$$\ln C_\Phi \approx \sum_{\ell=0}^{\infty} 2\ell \sum_{n=0}^{\infty} \beta \omega_n^{(\ell)} e^{-\beta \omega_n^{(\ell)}} \quad (2.3.15)$$

comes predominantly from the $\ell \approx O(r_+/l_s)$ modes, the entropy is given primarily by the number of different independent states that these large ℓ modes can take.

In the present method based on a low energy description of the system, the coefficients in Eqs. (2.3.13) and (2.3.14) cannot be obtained. This is because the quantities are dominated by the contributions from the $\ell \sim r_+/l_s$ modes that are localized near the stretched horizon, where the effect of unknown UV physics dominates. While this implies that the calculation is UV sensitive, its agreement with the results of Bekenstein and Hawking still gives us information about the UV physics; in particular, this unknown physics does not drastically increase the degrees of freedom compared to what is suggested by naively cutting off the spacetime at $r_* \approx r_{\text{HS}}$. Note that this calculation is equivalent to that in Refs. [42, 43], in which the mass and entropy of a black hole are obtained by integrating appropriate powers of the local Hawking temperature. The logic, however, is reversed here; once the geometry is given (within a time window of order r_+) by Eqs. (2.2.1) and (2.2.13), and correspondingly the

energy of the system by M , then a typical state represents a black hole vacuum microstate with the temperature and entropy of the black hole given by Eqs. (2.3.13) and (2.3.14). Microstates of the black hole vacuum correspond to different ways in which the energy levels in Eq. (2.3.1) are occupied under the energy constraint.

As in the standard thermal system, the microscopic state of a black hole changes generically in a timescale of order the inverse temperature $\beta \approx M l_{\text{P}}^2$. This implies that the energy of the system, i.e. the mass of the black hole, can be specified only up to the precision of $\Delta E \sim 1/\beta$; the state of a black hole comprises a superposition of energy eigenstates with the spread of eigenvalues of order $1/\beta$ or larger. In the following, we always assume that the state of the system of interest, e.g. a black hole or de Sitter spacetime, is specified with this maximal precision. A similar comment also applies to other quantities, such as the momentum of a black hole, whose minimal uncertainty is of order $\Delta p \sim \Delta(\sqrt{2ME}) \sim 1/M l_{\text{P}}^2$. The number of independent states consistent with this specification is given by the Bekenstein-Hawking entropy of Eq. (2.3.14). If the state involves superpositions of wider ranges of energy, momentum, and so on, e.g. as a result of backreaction of Hawking emission [99, 100], then our discussion below applies to each branch specified with the maximal precision for these quantities.

Note that a typical state in the space spanned by the independent microstates specified by $E = M$ and $p = 0$ within the minimal uncertainties of ΔE and Δp has angular momentum of order the uncertainty $\Delta J \sim \sqrt{N_{\text{dof}}} r_+/l_{\text{s}} \sim r_+/l_{\text{P}}$. This is consistent with the maximal precision with which angles can be specified consistently with the UV cutoff: $\Delta\theta \sim l_{\text{s}}/r_+$ for each species. We can thus regard these microstates as those of a non-rotating black hole in semiclassical theory.

We finally mention that we can discuss de Sitter spacetime in a similar manner. One difference is that we do not have a well-established notion of energy attributed to the spacetime in this case. However, by taking

$$E_{\text{dS}} = \frac{\alpha}{l_{\text{P}}^2}, \quad (2.3.16)$$

as implied by the Bekenstein-Hawking (or Gibbons-Hawking [57]) entropy $S_{\text{GH}} = \pi\alpha^2/l_{\text{P}}^2$ and Hawking temperature $T_{\text{H}} = 1/2\pi\alpha$, we reproduce all the properties associated with a static patch of the de Sitter spacetime [43]. (Physically, this energy can be specified only up to the precision of order T_{H} .) While the relationship of this energy to more conventionally defined energies is not clear, it represents some “energy” defined at $r = 0$, at which $f(r) = 1$, rather than at asymptotic infinity.¹⁰

¹⁰ It is interesting to note that if one considers the quantity $E = \int_{\Sigma} \sqrt{-g} \rho d^3x$, along the lines of Ref. [101], then one would get $E = \alpha/l_{\text{P}}^2$. Here, Σ is the $t = 0$ surface of *global* de Sitter spacetime comprising two static patches, $g = -r^4 \sin^2\theta$ is the determinant of the *spacetime* metric (in the static coordinates), $\rho = 3/8\pi l_{\text{P}}^2 \alpha^2$ is the energy density, and $d^3x = dr d\theta d\phi$.

2.3.2 UV (in)sensitivity: zone and horizon modes

Even though the density of states given by Eqs. (2.3.1) and (2.3.2) reproduces the entropy and temperature associated with the horizon, we do not expect that the precise spectrum is given by these expressions. This is because interactions near the stretched horizon, with the effective coupling given by $T_{\text{loc}}(r_s)^2 l_s^2 \sim O(1)$, deform the spectrum significantly from that of free theories (though the density of states will not change when coarse grained at a scale of order r_+ in the r_* coordinate).

To analyze this issue in more detail, let us consider the effective potential $V_\ell(r_*)$ for a fixed ℓ and negligible m_Φ . From Eq. (2.3.2), the average gap between adjacent energy levels is given by

$$\Delta\omega^{(\ell)} \sim \frac{1}{r_+ \ln \frac{r_+}{\ell l_s}}. \quad (2.3.17)$$

For a black hole, the height of the barrier is given by

$$\omega_{\text{barrier}}^{(\ell)} \sim \frac{\ell}{r_+}. \quad (2.3.18)$$

As we have seen, the temperature of the system is $\sim 1/r_+$, and the modes with

$$\omega \lesssim \frac{1}{r_+} \quad \text{and} \quad \ell \lesssim \frac{r_+}{l_s} \quad (2.3.19)$$

are significantly occupied. For each of these modes, the wavefunction is oscillatory between $r_* = r_{*s}$ and

$$r_* \approx -\frac{2}{f'(r_+)} \ln \ell \equiv r_{*b}^{(\ell)}, \quad (2.3.20)$$

outside of which it is exponentially damped. Thus the size of the region supporting these modes is given, in units of the wavelength of Hawking radiation, as

$$\frac{|r_{*b}^{(\ell)} - r_{*s}|}{r_+} \approx \left| \frac{1}{r_+ f'(r_+)} \ln \frac{r_+}{\ell^2 l_s^2 |f'(r_+)|} \right| \sim \left| \ln \frac{r_+}{\ell l_s} \right|, \quad (2.3.21)$$

where we have used $|f'(r_+)| \sim 1/r_+$ in the last expression.

For illustration, let us first consider the two extreme cases of $\ell \approx O(1)$ and $O(r_+/l_s)$. For $\ell \approx O(1)$, there are $1/\Delta\omega^{(\ell)} r_+ \sim \ln(r_+/l_s)$ independent modes that have $\omega \approx O(1/r_+)$ for each orbital and magnetic quantum numbers ℓ and m . We may take them to be wavepackets of width $\approx r_+$ in r_* distributed uniformly between r_{*s} and $r_{*b}^{(\ell)}$ without having a significant overlap with each other; see Fig. 2.3.

In order for these to form a basis, we need to prepare two sets of wavepackets, moving toward larger and smaller values of r_* . Among these (approximately orthogonal) wavepackets, the ones closest to the stretched horizon, i.e. those located within $\sim r_+$ in r_* from the stretched horizon, are special in that their dynamics cannot be described by a semiclassical theory. This is because the interaction strength of the unknown UV dynamics is strong there, which can also be seen from the fact that the unknown $O(1)$ coefficient in the definition of

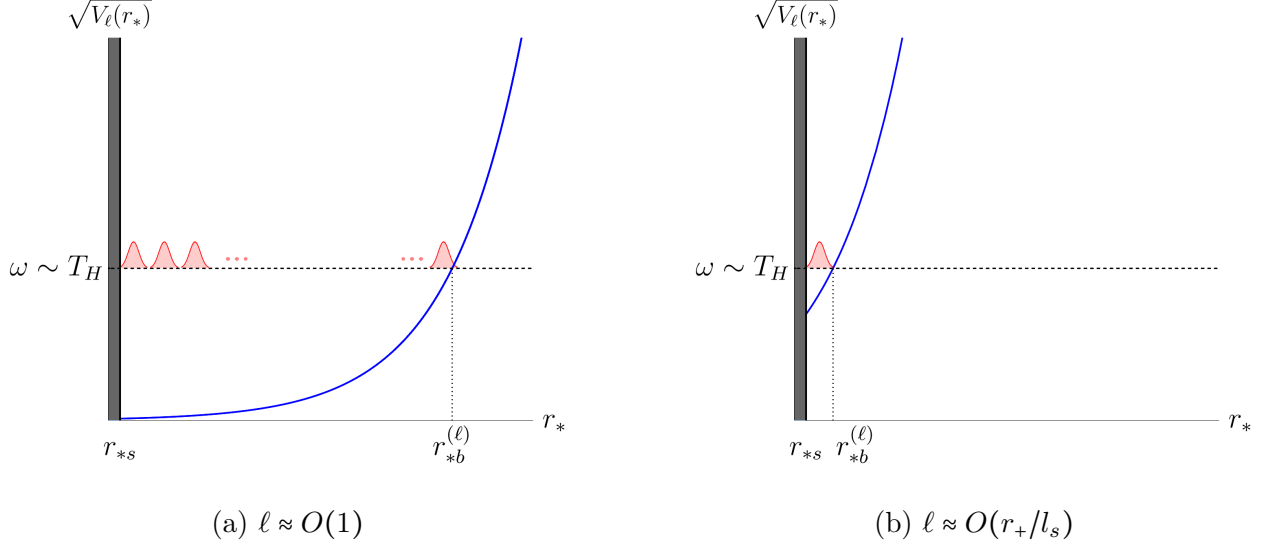


Figure 2.3: A schematic depiction of approximately orthogonal, independent modes having $\omega \approx O(T_H)$ and localized in the zone for a field of negligible mass. (a) For $\ell \approx O(1)$, there are $\sim \ln(r_+/l_s)$ independent ingoing and outgoing modes. Their approximate basis can be taken to be wavepackets of width $\sim r_+$ in r_* , distributed uniformly over the classically allowed region between r_{*s} and $r_{*b}^{(\ell)}$, where $r_{*b}^{(\ell)} - r_{*s} \sim r_+ \ln(r_+/l_s)$. (b) For modes with $\ell \approx O(r_+/l_s)$, the classically allowed region is near the stretched horizon with $r_{*b}^{(\ell)} - r_{*s} \sim r_+$, so all these modes must be regarded as horizon modes.

the stretched horizon in Eq. (2.2.10) translates into the ambiguity of the location of the stretched horizon of order r_+ in r_* . We call modes corresponding to these wavepackets, i.e. the wavepackets “next to” the stretched horizon, *horizon modes*. On the other hand, the dynamics of other $\sim \ln(r_+/l_s)$ wavepackets can be described by a semiclassical theory (at least) in the relevant timescale of order $1/T_H \sim r_+$. Modes associated with these semiclassically describable wavepackets are called *zone modes*.

For $\ell \approx O(r_+/l_s)$, the situation is different. In this case, there are only $1/\Delta\omega^{(\ell)}r_+ \approx O(1)$ independent modes having $\omega \approx O(1/r_+)$ for each ℓ and m , which, given Eq. (2.3.21), are all supported within $\sim r_+$ from the stretched horizon. Therefore, they are all horizon modes. As stated earlier, the entropy of a black hole (or de Sitter spacetime) is dominated by the number of independent states of these high ℓ modes

$$S \sim N_{\text{dof}} \sum_{\ell=0}^{O(r_+/l_s)} \sum_{m=-\ell}^{\ell} O(1) \sim \frac{r_+^2}{l_p^2}, \quad (2.3.22)$$

so the dynamics of the black hole (de Sitter) microstates cannot be described by a semiclassical theory. Indeed, the dynamics of these modes is expected to be nonlocal in the spatial directions along the horizon [68, 69].

In general, we call modes localized near the stretched horizon, i.e. within $\sim r_+$ in r_* , horizon modes while those away from it we call zone modes. We use the term “zone modes”

also in spacetimes other than the black hole spacetime, including de Sitter spacetime, even though there may not be real zones in these spacetimes. In the case where the region near the horizon is connected to another (ambient/bath) region across a barrier of the effective potential, we restrict the use of the term zone modes to those modes that are contained in the region near the horizon. For example, for a Schwarzschild black hole in asymptotically flat spacetime, zone modes only refer to modes within the zone $r \lesssim r_z$ of the black hole. Modes located outside the barrier, $r \gtrsim r_z$, are not directly involved in the near horizon dynamics, and we call them *far modes*.

It is important to realize that the terminologies introduced above are associated with the spatial position of modes at a given time, or more precisely a time interval of width $\Delta t \lesssim O(r_+)$ within which the system can be regarded as approximately static. This implies, for example, that a zone mode according to the classification at time t_1 can be a horizon, zone, or far mode (or a superposition of them) according to the classification at another time t_2 . In particular, if the mode is well localized in the zone at time t_1 and is propagating toward the stretched horizon, then it will be a horizon mode according to the classification at t_2 , a time after this mode has reached the stretched horizon. Generally, a horizon mode remains as a horizon mode for a long time, although it occasionally becomes an outgoing zone mode through interactions at the stretched horizon.

2.3.3 Vacuum microstates of semiclassical theories

Consider a state specified by the occupation numbers of all the levels in Eq. (2.3.1) (with the precise values of $\omega_n^{(\ell)}$ modified by interactions). This should be understood as a state of full quantum gravity. In particular, a typical pure state given in this way represents a microstate of a black hole or de Sitter vacuum. We now connect this description to the picture in semiclassical theory, wherein a vacuum state takes the form of a mixed, (approximately) thermal state.

For this purpose, we first consider a subset of zone modes which is relevant for describing a physical object (or objects) falling into the horizon. In the context of de Sitter spacetime, this means an object accelerating away from the observer toward the cosmological horizon. We can choose a subset of modes to accommodate *any* semiclassical object; the object is then described as an excitation of this set of modes over the semiclassical vacuum state.¹¹ We call these modes *hard modes*. They comprise only a tiny subset of the zone modes, since even the most entropic configuration of semiclassical matter has the entropy suppressed by powers of l_P/r_+ compared with the entropy of the Hawking cloud [7, 102], of which the entropy of the zone modes comprises an $O(1)$ fraction. The rest of the zone modes and the horizon modes are together called *soft modes*.

Similar to the zone and horizon modes, the hard and soft modes are defined at a specific

¹¹ This subset can be chosen to represent bound or meta-stable states (or an object made out of them). We may even choose it to represent a microscopic black hole, although such a construction would have to be made more precise.

time t . In particular, a mode defined as a hard (or soft) mode at time t may not be a hard (soft) mode at another time. This issue, however, is less relevant for these modes than for the zone and horizon modes, since the concept is used almost exclusively to construct an effective theory describing the region behind the horizon, which is erected at a given instantaneous time t as we will see below.

We label hard modes collectively by α , which includes all possible quantum numbers like species, level, and orbital and spin angular momenta. The state of these modes is specified by giving the occupation number n_α (≥ 0) for each α , which we denote by $|\{n_\alpha\}\rangle$ and normalize such that

$$\langle\{m_\alpha\}|\{n_\alpha\}\rangle = \delta_{\{m_\alpha\},\{n_\alpha\}} = \delta_{mn}, \quad (2.3.23)$$

where m and n are shorthand notations of $\{m_\alpha\}$ and $\{n_\alpha\}$. We assume that different hard-mode states are observationally distinguishable in that two different states do not have identical quantum numbers within the uncertainties ΔE , Δp , and so on.

Suppose that the system consists only of zone and horizon modes; examples include de Sitter spacetime.¹² In this class of systems, the state of the entire system is generally given as an entangled state of hard and soft modes. Because of the energy constraint, the state of soft modes that comes with the hard-mode state $|\{n_\alpha\}\rangle$ must have energy $E - E_n$, up to the uncertainty ΔE required by quantum mechanics, which is typically of the order of the Hawking temperature T_H . Here, E is the total energy of the system, e.g. E_{dS} in Eq. (2.3.16) for de Sitter spacetime, and

$$E_n = E_{\{n_\alpha\}} = \sum_\alpha n_\alpha \omega_\alpha \quad (2.3.24)$$

is the energy of the hard-mode state $|\{n_\alpha\}\rangle$.

The relevant Hilbert space of the system is then given by

$$\mathcal{H}(E) = \bigoplus_n \left(|\{n_\alpha\}\rangle \otimes \mathcal{H}_{\text{soft}}^{(n)} \right), \quad (2.3.25)$$

where $\mathcal{H}_{\text{soft}}^{(n)}$ is the Hilbert space spanned by the soft-mode states that carry energy $E - E_n$ within the uncertainty ΔE . Given that the number of hard modes is much smaller than the number of relevant soft modes, the effective dimension of the Hilbert space $\mathcal{H}_{\text{soft}}^{(n)}$ is given by

$$\ln \dim \mathcal{H}_{\text{soft}}^{(n)} = S(E - E_n), \quad (2.3.26)$$

where $S(E)$ is the entropy density of the system at energy E . For de Sitter spacetime, it is given by the Gibbons-Hawking entropy

$$S(E) = S_{\text{dS}}(E) = \pi E^2 l_{\text{P}}^2, \quad (2.3.27)$$

which gives the standard expression of $S_{\text{GH}} = \pi \alpha^2 / l_{\text{P}}^2$ for $E = E_{\text{dS}} = \alpha / l_{\text{P}}$.

¹² This is always the case if the effective potential increases monotonically as r_* moves away from the stretched horizon. Another system exhibiting a similar behavior is Rindler spacetime, although in this case the system has a planar rather than spherical symmetry.

A typical state described in Section 2.3.1 corresponds to a typical state in the Hilbert space $\mathcal{H}(E)$ of Eq. (2.3.25). For our purposes, we need not be very precise about what we mean by typical, but for concreteness one might imagine a state that is typical in $\mathcal{H}(E)$ under the Haar measure. Denoting a set of generic orthonormal basis states of $\mathcal{H}_{\text{soft}}^{(n)}$ by $|\psi_{i_n}^{(n)}\rangle$ ($i_n = 1, \dots, e^{S(E-E_n)}$), the state we are interested in can be written as

$$|\Psi(E)\rangle = \sum_n \sum_{i_n=1}^{e^{S(E-E_n)}} c_{ni_n} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle, \quad (2.3.28)$$

where the real and imaginary parts of complex coefficients c_{ni_n} 's can be viewed as taking random values following independently the Gaussian distributions with

$$\langle \text{Re } c_{ni_n} \rangle = \langle \text{Im } c_{ni_n} \rangle = 0, \quad \sqrt{\langle (\text{Re } c_{ni_n})^2 \rangle} = \sqrt{\langle (\text{Im } c_{ni_n})^2 \rangle} = \frac{1}{\sqrt{2e^{S_{\text{sys}}}}}. \quad (2.3.29)$$

Here, the brackets represent the ensemble average over (a sufficiently large portion of) the (n, i_n) space, and

$$e^{S_{\text{sys}}} = \sum_n e^{S(E-E_n)} \equiv z e^{S(E)} \quad (2.3.30)$$

is the number of independent states of the form of Eq. (2.3.28) with

$$z = \sum_n e^{-\frac{E_n}{T_H}}, \quad (2.3.31)$$

where we have used $\partial S(E)/\partial E = 1/T_H$ and $E_n \ll E$.¹³ In other words, we can say that c_{ni_n} is a complex Gaussian random variable, which implies that the phases of c_{ni_n} 's are distributed uniformly.

Note that by taking the basis of $\mathcal{H}_{\text{soft}}^{(n)}$ for each n , we find that soft-mode states are orthogonal

$$\langle \psi_{i_m}^{(m)} | \psi_{j_n}^{(n)} \rangle = \delta_{mn} \delta_{i_m j_n}. \quad (2.3.32)$$

This is because states $|\psi_{i_n}^{(n)}\rangle$ with different n can be observationally discriminated, which follows from the distinguishability of different hard-mode states $|\{n_\alpha\}\rangle$ as well as the fact that the state of combined hard and soft mode system is specified with the minimal uncertainty. In particular, even if some observables (e.g. E and p) may have to be intrinsically coarse grained (by ΔE and Δp), $|\psi_{i_n}^{(n)}\rangle$'s with different n can still be discriminated at the semiclassical level, and hence are orthogonal.¹⁴

We also note that $S_{\text{sys}} = S(E) + \ln z$ is equal to $S(E)$ at the leading order in l_P/r_+ and $1/E_n r_+$. This implies that the standard interpretation of the Gibbons-Hawking entropy as

¹³ We have used the equal sign for a relation that becomes exact in the thermodynamic limit.

¹⁴ We assume that such coarse grainings would be performed using smoothing functions which damp very rapidly outside the windows of order ΔE and Δp so that Eq. (2.3.23) is valid with sufficient accuracy. With this assumption, Eq. (2.3.32) is also valid at the same level of accuracy.

the entropy of de Sitter spacetime persists. A similar comment applies to a black hole, for which the density of soft-mode states is given by the Bekenstein-Hawking entropy

$$S(E) = S_{\text{bh}}(E) = 4\pi E^2 l_{\text{P}}^2, \quad (2.3.33)$$

although in this case the entire system also has degrees of freedom outside the zone, and $S_{\text{sys}} = S(E) + \ln z$ represents only the entropy of the black hole system (i.e. the zone and horizon modes) without including the contribution from the far modes.

We now take a complete set of orthonormal states of the form of Eq. (2.3.28), i.e. states having the energy E within ΔE , in a generic basis:

$$|\Psi_A(E)\rangle = \sum_n \sum_{i_n=1}^{e^{S(E-E_n)}} c_{ni_n}^A |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle \quad (A = 1, \dots, e^{S_{\text{sys}}}), \quad (2.3.34)$$

where

$$\langle \Psi_A(E) | \Psi_B(E) \rangle = \delta_{AB} \iff \sum_n \sum_{i_n=1}^{e^{S(E-E_n)}} c_{ni_n}^{A*} c_{ni_n}^B = \delta_{AB}. \quad (2.3.35)$$

In general, these states provide a basis for microstates of a semiclassical vacuum. For example, if the spectrum ω_α is given such that it represents hard modes inside the stretched de Sitter horizon, then the set $\{|\Psi_A(E)\rangle\}$ forms a basis for the microstates of the de Sitter vacuum.

We stress that in order for the states of Eq. (2.3.34) to be microstates of the spacetime, they need to be taken generically in the space of $\mathcal{H}(E)$. For example, if we took $e^{S_{\text{sys}}}$ orthonormal states that are, or approximately are, product states

$$|\Psi(E)\rangle \approx |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle, \quad (2.3.36)$$

then these states would still form a “basis” of the microstates in Eq. (2.3.34) in the sense that all the states of the form in Eq. (2.3.34) can be obtained by superposing them, although none of them is by itself a microstates of the spacetime under consideration—these special, exponentially rare, states (states whose entanglement structure is significantly different from generic states) are “firewall” [30] states which do not represent the spacetime under consideration. This is because spacetime is a manifestation of the entanglement structure of a holographic boundary state [13–15, 103–105], and entanglement cannot be represented as a linear operator—the concept of a linear vector space comprising the microstates of a spacetime is only an approximate one [58, 106]. This issue will be discussed further in the next section.

Since a semiclassical theory can describe only the dynamics of hard modes, it concerns only about the state of these modes. Thus, the vacuum state $\rho_{\text{vac}}(E)$ in a semiclassical theory is obtained by taking a typical vacuum microstate and then tracing out the soft modes. Specifically, it is given by

$$\begin{aligned} \rho_{\text{vac}}(E) &= \text{Tr}_{\text{soft}} |\Psi_A(E)\rangle \langle \Psi_A(E)| \\ &= \sum_n \left(\sum_{i_n=1}^{e^{S(E-E_n)}} |c_{ni_n}^A|^2 \right) |\{n_\alpha\}\rangle \langle \{n_\alpha\}| \end{aligned}$$

$$= \sum_n \frac{e^{-\frac{E_n}{T_H}}}{z} |\{n_\alpha\}\rangle \langle \{n_\alpha\}| + O\left(e^{-\frac{1}{2}S(E)}\right). \quad (2.3.37)$$

Here, in the last equation we have used

$$\varepsilon_n^{AB} \equiv z e^{\frac{E_n}{T_H}} e^{S(E-E_n)} \sum_{i_n=1} c_{ni_n}^{A*} c_{ni_n}^B - \delta_{AB} \approx O\left(e^{-\frac{1}{2}S(E-E_n)}\right), \quad (2.3.38)$$

which can be derived from Eq. (2.3.29). More precisely, when we vary A and B , ε_n^{AB} behave as complex Gaussian random variables with mean 0 and variance $e^{-S(E-E_n)}$, but obeying

$$(\varepsilon_n^{AB})^* = \varepsilon_n^{BA}, \quad \sum_n \frac{e^{-\frac{E_n}{T_H}}}{z} \varepsilon_n^{AB} = 0, \quad (2.3.39)$$

where the second relation follows from Eq. (2.3.35). In the case of de Sitter spacetime, the thermal state in Eq. (2.3.37) gives the vacuum state describing a static patch of the de Sitter spacetime with Hubble radius $\alpha = E l_p^2$.

The hard and soft modes described here provide a refinement of the hard and soft modes defined in Refs. [42–46]. In Refs. [42–46] a simple frequency space criterion was used to define the hard and soft modes, while in our new definition here, the hard modes are chosen to be a subset of the zone modes whose dynamics we intend to describe at the semiclassical level. This makes it possible, for example, to describe the dynamics of zone modes with $\omega \sim O(T_H)$ using a semiclassical theory, which was not possible with the previous definition. For many practical purposes, however, the two definitions are interchangeable. For a small object falling into the horizon, for example, the difference between the two definitions is not significant if we choose the frequency cutoff to be sufficiently larger than T_H . We can then employ the same construction for spacetime beyond the horizon, which we will discuss in Section 2.5.

Evaporating black hole

The situation is more complicated if the region near the horizon is coupled to an ambient/bath system. An important example of this is a black hole in asymptotically flat spacetime; other examples include a small black hole in asymptotically AdS spacetime and a large AdS black hole coupled to a separate bath system. In this case, the near horizon system, consisting of the zone and horizon, evolves in time, and this evolution modifies the vacuum states.

For concreteness, let us focus on a black hole in asymptotically flat spacetime. The state of the entire system then involves the hard, soft, and far (located outside the zone, $r > r_z$) modes. Thus, denoting orthonormal basis states of the far modes by $|\phi_a\rangle$, one would consider

the state of the system to be given by¹⁵

$$|\Psi_A(M)\rangle \stackrel{?}{=} \sum_n e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1}^{S_{\text{rad}}} \sum_{a=1}^{S_{\text{rad}}} c_{ni_na}^A |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \quad (2.3.40)$$

where we have assumed that the black hole system, comprising the hard and soft modes, is at rest and has energy M (up to the uncertainty of order T_{H}). Here, $e^{S_{\text{rad}}}$ is the number of independent far-mode states relevant here, i.e. those significantly entangled with the black hole system (typically Hawking radiation emitted earlier from the black hole), and A is the index for microstates running over

$$A = 1, \dots, e^{S_{\text{s+r}}}, \quad S_{\text{s+r}} = S_{\text{sys}} + S_{\text{rad}}, \quad (2.3.41)$$

with S_{sys} given by Eq. (2.3.30) with $S(E) = S_{\text{bh}}(E)$. The coefficients $c_{ni_na}^A$ satisfy the properties analogous to those in Eq. (2.3.29)

$$\langle \text{Re } c_{ni_na}^A \rangle = \langle \text{Im } c_{ni_na}^A \rangle = 0, \quad \sqrt{\langle (\text{Re } c_{ni_na}^A)^2 \rangle} = \sqrt{\langle (\text{Im } c_{ni_na}^A)^2 \rangle} = \frac{1}{\sqrt{2e^{S_{\text{s+r}}}}}, \quad (2.3.42)$$

with brackets representing the ensemble average over (a sufficiently large portion of) the (n, i_n, a) space. Note that $|\Psi_A(M)\rangle$ represent microstates of the system with the black hole put in the semiclassical vacuum, and a generic state in the Hilbert space of dimension $e^{S_{\text{s+r}}}$ has the black hole of mass M . Since black hole evaporation is a thermodynamically irreversible process [107, 108], most of these microstates do not become a state with a larger black hole in empty space when evolved backward in time—there is some junk radiation around it. This, however, does not change the fact that there are $e^{S_{\text{s+r}}}$ independent microstates relevant for the discussion here.

The fact that the height of the potential barrier is finite, however, implies that only modes with $\omega < \omega_{\text{barrier}}^{(\ell)}$ are thermalized in the zone. Given that the dynamics at the stretched horizon is strongly coupled, outgoing modes (i.e. modes moving toward larger r) with $\omega > \omega_{\text{barrier}}^{(\ell)}$ can still be viewed as obeying the thermal distribution, but this is not the case for ingoing modes with $\omega > \omega_{\text{barrier}}^{(\ell)}$. At the microscopic level, this implies that the microcanonical ensemble in Section 2.3.1 is taken with the extra constraint that the occupation numbers of ingoing

¹⁵ We assume that the standard issues for a factorization of Hilbert space in quantum field theory, such as those associated with short distance divergences and constraints from gauge invariance, are dealt with appropriately.

modes with $\omega > \omega_{\text{barrier}}^{(\ell)}$ are zero, leading to¹⁶

$$|\Psi_A(M)\rangle \stackrel{?}{\propto} \sum_n \prod_{\substack{\alpha' \in \text{ingoing} \\ \omega > \omega_{\text{barrier}}^{(\ell)}}} \delta_{n_{\alpha'}, 0} \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{S_{\text{rad}}} c_{ni_n a}^A |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \quad (2.3.43)$$

where the $\propto$ symbol represents “up to a normalization constant,” and the question mark above it indicates that this relation is tentative and will be updated momentarily. (The effect of the lack of ingoing soft modes with $\omega > \omega_{\text{barrier}}^{(\ell)}$ is negligible for our purpose.) This corresponds to taking the Hartle-Hawking [109] and Unruh [70] vacua for hard modes with $\omega < \omega_{\text{barrier}}^{(\ell)}$ and $> \omega_{\text{barrier}}^{(\ell)}$, respectively. The effect of the constraint on soft modes is negligible, since the vast majority of the relevant modes have $\omega \sim T_{\text{H}}$ and $\ell \gg 1$, and hence $\omega \ll \omega_{\text{barrier}}^{(\ell)}$.

The story, however, does not end here. Because of the coupling between the black hole system (zone + horizon) and the asymptotically flat spacetime around the region $r \sim r_z$, thermal quanta in the zone “leak” into the latter. This occurs mostly via s -wave modes that tunnel through the potential barrier. We emphasize that this process, occurring near the edge of the zone [110, 111] (for related discussions, see Refs. [112–114]), is governed by semiclassical physics—it does not involve strongly coupled, intrinsically quantum gravitational physics in any significant way. Given a black hole microstate in Eq. (2.3.43) (slightly modified due to backreaction; see below), the emission of quanta into the asymptotic region occurs unitarily following the dynamics of standard quantum field theory. The apparent violation of unitarity in Hawking’s analysis [17] occurs because we cannot calculate the configuration of zone mode quanta using semiclassical theory due to the strong dynamics *near the stretched horizon*. It is this incalculability that makes the semiclassical description of Hawking radiation, obtained after tracing out the soft modes, intrinsically thermal and hence leading to a mixed final state. Physics away from the stretched horizon can indeed be fully semiclassical, even within and at the edge of the zone.¹⁷

The emission of Hawking particles to the ambient space at $r \sim r_z$ gives a backreaction to the state of the black hole. In this region, quanta of the zone region is leaked into the ambient space through tunneling the potential barrier (and also via thermal hopping to some extent). This removes some of the quanta that would be reflected back to the zone by the potential, producing a deficit in ingoing zone mode quanta relative to those in the states of

¹⁶ Strictly speaking, there are small but nonzero amplitudes for outgoing hard modes with $\omega > \omega_{\text{barrier}}^{(\ell)}$ to be reflected back from the potential barrier. We mostly ignore this effect because it is not essential for our discussion. Including it, however, is straightforward; instead of taking the terms with $\exists \alpha', n_{\alpha'} \neq 0$ to be exactly absent, we keep these terms with small coefficients (compared to those of the terms with $\forall \alpha', n_{\alpha'} = 0$). Note that the size of these coefficients in general depends strongly on $\{n_{\alpha'}\}$, reflecting the $\omega_{\alpha'}$ dependence of the reflection amplitudes.

¹⁷ This implies that the process of black hole mining [115, 116] also occurs unitarily, which is governed by semiclassical physics if it is performed away from the stretched horizon. It also implies that the semiclassical calculation of the gray body factor, such as that in Ref. [117], is valid as long as r_+/l_s is sufficiently large; see below.

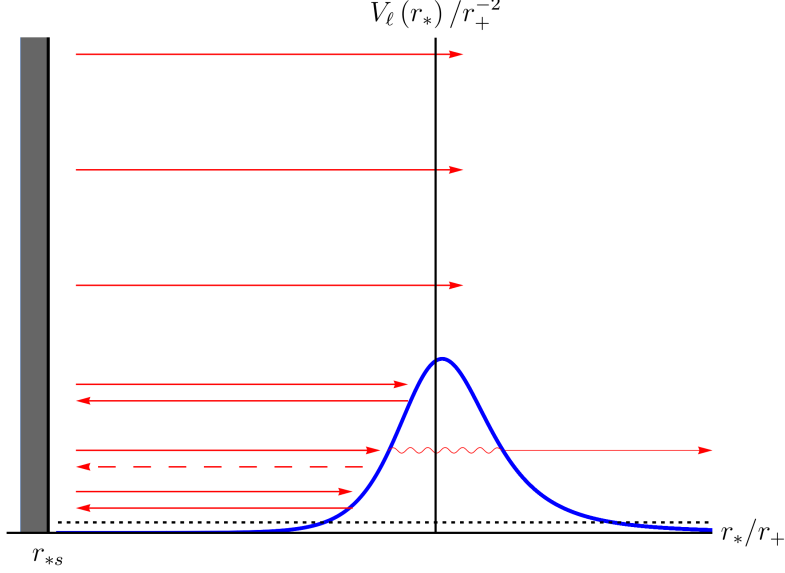


Figure 2.4: A schematic depiction of quanta occupying various modes. Solid arrows indicate that the occupation numbers of the modes are determined by the thermodynamic consideration as described in Section 2.3.1. For zone modes, this applies to the outgoing mode as well as the ingoing modes with $\omega \lesssim \omega_{\text{barrier}}^{(\ell)}$, except that backreaction of Hawking emission at $r \sim r_z$ leads to a deficit for some of these ingoing modes as indicated by the dashed arrow.

Eq. (2.3.43), i.e. with the Hartle-Hawking vacuum. Note that the process occurs only for low energy fields of mass $m_\Phi \lesssim T_H$. Given that an $O(1)$ number of quanta are emitted within each time interval of order $1/T_H \sim r_+$, there are $\sim \ln(r_+/l_s)$ quanta missing throughout the zone for each low energy field Φ of $m_\Phi \lesssim T_H$. Denoting the annihilation operators for hard mode quanta by $b_{\bar{\alpha}}$ (see below for more detail), we finally find that the microstates of an evaporating black hole are given by

$$|\Psi_A(M)\rangle \propto \prod_{\substack{\bar{\alpha} \in \text{ingoing} \\ m_\Phi \lesssim T_H, \omega \sim T_H}} b_{\bar{\alpha}} \sum_n \prod_{\substack{\alpha' \in \text{ingoing} \\ \omega > \omega_{\text{barrier}}^{(\ell)}}} \delta_{n_{\alpha'}, 0} \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{S_{\text{rad}}} c_{ni_n a}^A |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \quad (2.3.44)$$

where the number of annihilation operators $b_{\bar{\alpha}}$ in the product is at most of order $\ln(r_+/l_s)$ for each field; the precise number depends on the choice of the hard modes. A schematic depiction of the occupation of various modes is given in Fig. 2.4.

The deficit of ingoing modes described above implies that there is a negative energy flux carrying negative entropy for each field of $m_\Phi \lesssim T_H$. Note that these energy and entropy are measured with respect to the thermal, Hartle-Hawking vacuum state in Eq. (2.3.43) for modes with $\omega \sim T_H$. The flux has negative entropy because with the lack of some of the zone mode quanta, the number of independent states realizing the most probable configuration is smaller by $e^{O(1)\ln(r_+/l_s)}$ for each field. Thus, the microstate index A in Eq. (2.3.44) runs

effectively only for

$$A = 1, \dots, A_{\max}, \quad \ln A_{\max} - S_{\text{tot}} \sim -\ln \frac{r_+}{l_s}, \quad (2.3.45)$$

where S_{tot} is the coarse-grained entropy of the total system ignoring the backreaction

$$e^{S_{\text{tot}}} = \sum_n \prod_{\substack{\alpha' \in \text{ingoing} \\ \omega > \omega_{\text{barrier}}}} \delta_{n_{\alpha'}, 0} e^{S_{\text{bh}}(M - E_n)} e^{S_{\text{rad}}} \equiv z' e^{S_{\text{bh}}(M) + S_{\text{rad}}}, \quad (2.3.46)$$

and we have taken the number of low energy degrees of freedom relevant for the emission process to be of $O(1)$.¹⁸ As discussed in Refs. [110, 111], this negative entropy is essential for the unitarity of the black hole evolution, in particular for the black hole to keep relaxing into a lower mass black hole while absorbing the negative energy flux. Again, we emphasize that semiclassical physics is sufficient to understand the unitary emission process at $r \sim r_z$ and the resulting emergence of a negative energy and entropy flux. The unknown UV physics enters only in the process occurring at the stretched horizon, in which the ingoing negative energy-entropy flux is absorbed into horizon modes and the black hole relaxes into its semiclassical vacuum state.

So far, we have assumed that our black hole is large, in particular $\ln(r_+/l_s) \gg 1$. In this limit, the difference of energies between adjacent discrete levels for modes relevant for Hawking emission, $\sim r_+/\ln(r_+/l_s)$, is much smaller than the uncertainty of energy of Hawking quanta, $\sim r_+$, so that the effect of the discreteness of energy levels is negligible in calculating the spectrum of Hawking radiation. In particular, the semiclassical calculation of the spectrum, including the gray body factor, persists with high precision. If the value of $\ln(r_+/l_s)$ is reduced, however, the effect of the discreteness of levels may become important. In particular, if the size of the classically allowed region $r_{*b}^{(\ell)} - r_{*s} \approx 2r_+ \ln(r_+/l_s)$, becomes smaller than a half wavelength of a Hawking quantum $\lambda_H/2 \approx \pi/T_H \approx 4\pi^2 r_+$, i.e.

$$r_+ \lesssim e^{2\pi^2} l_s, \quad (2.3.47)$$

then the effect may become non-negligible. (This condition can also be obtained by requiring $\Delta\omega$ in Eq. (2.2.24) to be larger than T_H .) A naive guess is that in this regime, the energy of each Hawking quantum is larger than that obtained by the semiclassical calculation, since the frequency of the lowest energy level is expected to become larger than T_H . An interesting point is that the black hole enters the regime of Eq. (2.3.47) before its mass is reduced to the Planck mass, since this condition can be written as

$$M \lesssim \frac{e^{2\pi^2} l_s}{2l_p^2}. \quad (2.3.48)$$

For $l_s/l_p \approx 20$, as we might expect in our universe, the right-hand side is 4×10^9 times larger than the Planck mass (corresponding to the black hole of $T_H \approx 1 \times 10^8$ GeV). We leave further discussion of this issue, including its possible phenomenological implications, for the future.

¹⁸ This implies that the range of i_n in Eq. (2.3.44) is, strictly speaking, smaller than that in Eqs. (2.3.40) or (2.3.43). For the lack of a better notation, we interpret the sum in Eq. (2.3.44), and analogous expressions later, to include this minor effect.

2.3.4 Excited states

Before concluding this section, let us discuss semiclassical excitations in the zone. The states we have considered so far are typical states in a suitably defined microcanonical ensemble. For a black hole, for example, the relevant ensemble consists of the states that contain a fixed energy E in a spatial region $r \lesssim r_z$ within an uncertainty ΔE . These states are all vacuum states from the point of view of semiclassical theory as indicated by the fact that the Bekenstein-Hawking entropy is associated with the background (vacuum) spacetime in semiclassical theory.

We can, however, consider atypical states obtained by acting annihilation and/or creation operators of hard modes

$$b_\gamma = \sum_n \sqrt{n_\gamma} |\{n_\alpha - \delta_{\alpha\gamma}\}\rangle \langle \{n_\alpha\}|, \quad (2.3.49)$$

$$b_\gamma^\dagger = \sum_n \sqrt{n_\gamma + 1} |\{n_\alpha + \delta_{\alpha\gamma}\}\rangle \langle \{n_\alpha\}| \quad (2.3.50)$$

A state obtained in this way is not typical as it does not have the most probable configuration among the states in the ensemble. On these vacuum states, where γ specifies the mode which is annihilated/created by the operator. (States obtained by acting annihilation operators can become relevant when one considers backreaction of the Hawking emission or black hole mining process. In particular, some of the operators $b_{\tilde{\alpha}}^{(z)}$ in Eq. (2.3.44) can be superpositions of the b_γ operators, if we choose the hard modes to include the relevant zone modes.)

We are interested in semiclassical excitations whose backreaction to the geometry is negligible, or regarded as being small, for example a baseball falling into an astronomical black hole. This implies that the number of creation/annihilation operators that can be acted on a vacuum state is limited, so that their algebra defined on a vacuum state does not close in a strict mathematical sense. This type of structure is in fact common in erecting a semiclassical theory in quantum gravity, and here we simply treat the space of states obtained in this way as a Hilbert space; for a mathematically more rigorous treatment, see, e.g., Ref. [118]. In holography, one can think of this space as a code subspace embedded in a physical Hilbert space [11, 119, 120], although we do not discuss the error correcting nature of operators in this paper.

Strictly speaking, the space of semiclassical excitations generated by b_γ and $b_\gamma^\dagger$ acted on each vacuum state is not orthogonal to the space of vacuum microstates; namely, the Hilbert space cannot be strictly written as $\mathcal{H}_{\text{exc}} \otimes \mathcal{H}_{\text{vac}}$ [46]. One can see this by calculating inner products between states obtained by exciting black hole microstates in Eq. (2.3.43):

$$\langle \Psi_A(M) | b_\beta^\dagger b_\gamma | \Psi_B(M) \rangle = \delta_{\beta\gamma} \sum_n \prod_{\alpha' \in \substack{\text{ingoing} \\ \omega > \omega_{\text{barrier}}}} \delta_{n_{\alpha'}, 0} \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{e^{S_{\text{rad}}}} n_\gamma c_{ni_n a}^{A*} c_{ni_n a}^B, \quad (2.3.51)$$

$$\langle \Psi_A(M) | b_\beta b_\gamma^\dagger | \Psi_B(M) \rangle = \delta_{\beta\gamma} \sum_n \prod_{\alpha' \in \substack{\text{ingoing} \\ \omega > \omega_{\text{barrier}}}} \delta_{n_{\alpha'}, 0} \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{e^{S_{\text{rad}}}} (n_\gamma + 1) c_{ni_n a}^{A*} c_{ni_n a}^B. \quad (2.3.52)$$

These are not proportional to δ_{AB} in general, so that excited states built on different vacuum microstates are not necessarily orthogonal. However, for $A \neq B$, the right-hand sides of the above equations are exponentially suppressed by a factor of $e^{-S_{\text{tot}}/2}$, where we have taken

$$\langle \Psi_A(M) | \Psi_B(M) \rangle = \delta_{AB} \iff \sum_n \prod_{\substack{\alpha' \in \text{ingoing} \\ \omega > \omega_{\text{barrier}}}} \delta_{n_{\alpha'}, 0} \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{e^{S_{\text{rad}}}} c_{ni_n a}^{A*} c_{ni_n a}^B = \delta_{AB}. \quad (2.3.53)$$

Therefore, the deviation from the product space structure is exponentially small.

Incidentally, the annihilation and creation operators in Eqs. (2.3.49) and (2.3.50) satisfy the standard commutation relations

$$[b_\beta, b_\gamma^\dagger] = \delta_{\beta\gamma} \sum_n |\{n_\alpha\}\rangle \langle \{n_\alpha\}|, \quad [b_\beta, b_\gamma] = [b_\beta^\dagger, b_\gamma^\dagger] = 0 \quad (2.3.54)$$

as operators, without having an exponentially small correction. In erecting a semiclassical theory, one regards all the microstates $|\Psi_A(M)\rangle$ as representing the same geometry $\mathcal{M}$. This allows us to define field operators

$$\Phi_\Gamma(x) = \sum_{\gamma'} (b_\gamma u_{\gamma'}(x) + b_\gamma^\dagger v_{\gamma'}(x)), \quad (2.3.55)$$

where we have split the index γ into the index for species Γ and that for other quantum numbers γ' , e.g. a component of spin, whose structure may depend on Γ : $\gamma = (\Gamma, \gamma')$. Here, $u_{\gamma'}(x)$ and $v_{\gamma'}(x)$ are mode functions defined in the allowed region of $\mathcal{M}$. Because of Eq. (2.3.54), these field operators and their conjugate momenta obey the standard equal-time commutation relations, so that the resulting semiclassical theory respects causality exactly. For black hole and de Sitter spacetimes, this theory can be used to describe physics in the regions $r > r_s$ and $< r_s$, respectively.

2.4 Holographic Description

So far, we have assumed that the fundamental description of a system represents spacetime in the allowed region (with the horizon degrees of freedom included). For example, this corresponds to taking a distant view for a black hole and a static patch view for de Sitter spacetime. Why is this the case?

In this section, we discuss this issue from the perspective of holography. We assume that these spacetimes arise in setups in which the holographic description has only a single boundary; in particular, the black hole is formed by a gravitational collapse, and de Sitter spacetime arises in a cosmological context, for example as a universe created by bubble nucleation [121] which is filled with a positive cosmological constant. These setups are, arguably, more “realistic” than those with multiple boundaries, which will be discussed in Section 2.6. We conclude the section with a discussion on how the (effective) boundary Hilbert spaces describing the spacetimes considered here are related to the infinite-dimensional “fundamental” Hilbert space.

2.4.1 Black hole spacetime

We begin by considering a collapse formed black hole in an asymptotically AdS spacetime. We would like to know what spacetime picture can be obtained from the boundary CFT by reconstructing the bulk in a simple manner. By simple reconstruction, we mean reconstruction of the bulk using only low-complexity, causally-propagating operators and sources in the CFT [37]. In particular, we are interested in obtaining a “gauge-fixed” bulk description where the spacetime is foliated by equal-time hypersurfaces, using which the canonical formulation of quantum mechanics can be employed.

One way to obtain such a description is to “pull” the boundary into the bulk by coarse graining boundary degrees of freedom [35, 36, 122]. Based on intuition from tensor networks [119, 120, 123], one can consider a series of states [124] defined on successfully “renormalized” boundaries obtained by moving the original boundary inward to the bulk. The coarse-grained degrees of freedom are distributed locally on a renormalized boundary, although the dynamics they obey are not necessarily local.

This procedure is expected to work beyond the AdS/CFT context, which we assume to be the case, and henceforth we will not necessarily assume that the spacetime is asymptotically AdS. A particular way of performing this renormalization is to “continuously” coarse grain boundary degrees of freedom uniformly throughout the renormalized boundary space. This leads to a specific spacelike or null hypersurface swept by a series of renormalized boundaries called a *holographic slice* [35], which we will discuss in more detail below. This surface plays the role of an equal-time hypersurface in the bulk, providing a gauge fixing necessary for the canonical formulation.

In Fig. 2.5(a), we depict holographic slices in a black hole spacetime obtained without including quantum effects in the bulk.

We find that the slices do not enter the black hole interior, hence providing a distant view of the black hole. This is because the black hole horizon, sufficiently after the black hole has stabilized, plays a role of a barrier [125] for extremal surfaces, which are used to move renormalized boundaries. In fact, the slices approach Schwarzschild time slices near the black hole; they then stay near the horizon for long time and are eventually capped off at $r = 0$.

In order to describe an evaporating black hole, quantum effects in the bulk must be included. The procedure of forming a holographic slice can be extended to incorporate these effects [36]. In this case, the holographic slice is smoothly capped off at $r = 0$ initially, but at some (boundary) time it becomes asymptoting to a quantum extremal surface located (approximately) at the horizon; and as time progresses further, it again becomes a surface without a hole. This is depicted in Fig. 2.5(b). This behavior of holographic slices, in fact, can be used to define what we mean by the “formation” and “evaporation” of a black hole. The holographic slices approximate Schwarzschild time slices in this case as well, giving a distant description of the black hole. Note that the region swept by the holographic slices is essentially the simple wedge [37], which can be reconstructed in a simple manner using the Hamilton-Kabat-Lifschytz-Lowe (HKLL) procedure [33, 34, 126] together with boundary time evolutions with sources.

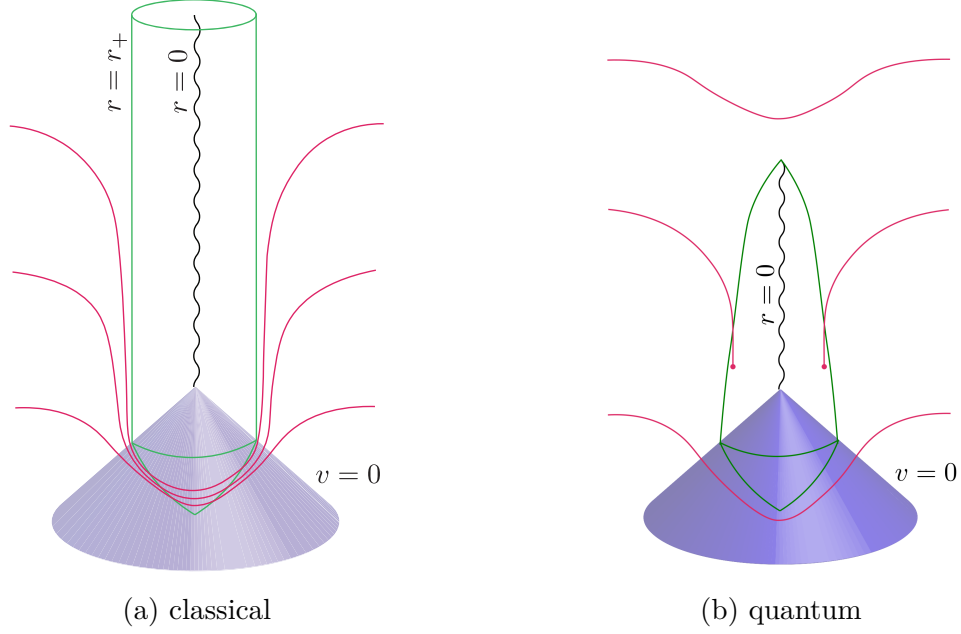


Figure 2.5: Holographic slices in a black hole spacetime in ingoing Eddington-Finkelstein coordinates (v, r) obtained (a) without and (b) with quantum effects in the bulk. These slices correspond to equal-time hypersurfaces in the bulk, which approach Schwarzschild time slices near the black hole. The figures are taken from Refs. [35] and [36].

In any event, the description of a black hole that naturally results from a holographic theory with a single boundary is an exterior/distant one. In this description, the unitarity of black hole evolution is not an issue, since the stretched horizon behaves as a regular material surface from the point of view of quantum information flow. The question, rather, is in what sense the near empty interior region exists for an infalling observer as predicted by general relativity. We will come back to this issue in Section 2.5.

2.4.2 de Sitter (or cosmological) spacetime

To discuss holographic descriptions beyond AdS/CFT, we need to introduce a “boundary” in spacetime on which a holographic state can be defined. One way to do this is to consider a renormalized boundary deep in the bulk—called a renormalized leaf [35, 36] in this context—and “unrenormalize” it by successively integrating in relevant degrees of freedom, making the holographic slice grow toward the “non-renormalized” boundary. This procedure depends on the background spacetime, but we can perform it in each branch of a state representing a well-defined spacetime in the semiclassical limit.

General picture

To be specific, let us adopt the scheme in Refs. [35, 36], in which successively renormalized boundaries $\sigma(\lambda)$, parameterized by “renormalization scale” λ , are obtained by the flow equation

$$\frac{dx^\mu}{d\lambda} = s^\mu, \quad (2.4.1)$$

where x^μ are the embedding coordinates of the codimension-2 surface $\sigma(\lambda)$, on which the holographic states are defined, and s^μ is the evolution vector given by

$$s^\mu = \frac{1}{2}(\Theta_k l^\mu + \Theta_l k^\mu). \quad (2.4.2)$$

Here, $\{k^\mu, l^\mu\}$ are the future-directed null vectors orthogonal to $\sigma(\lambda)$, normalized such that $k_\mu l^\mu = -2$,¹⁹ and $\Theta_{k,l}$ represent quantum expansions in the corresponding directions [127], which reduce to the classical expansions $\theta_{k,l}$ in the limit that bulk quantum effects are ignored.²⁰ We assume that $\sigma(\lambda)$ is neither quantum trapped nor antitrapped, so that the quantum expansion Θ_s associated with the evolution vector s^μ satisfies

$$\Theta_s = \Theta_k \Theta_l \leq 0. \quad (2.4.3)$$

This condition is needed for the consistency of the interpretation that the flow (toward larger λ) corresponds to coarse graining of the boundary degrees of freedom.

Since a holographic slice is nothing but a codimension-1 surface swept by $\sigma(\lambda)$, we can extend it “outward” using the flow equation. For a generic spacetime, we can perform this extension up to the point where one of the $\Theta_{k,l}$ becomes zero, i.e. until $\sigma(\lambda)$ becomes a quantum marginally trapped or antitrapped surface. Suppose that $\Theta_k = 0$ throughout this surface, which we call $\sigma(0)$. Then, the dimension of the Hilbert space associated with states on the holographic slice is bounded [6, 127] by the generalized entropy of $\sigma(0)$, so that the $\sigma(0)$ can be identified as an equal-time surface of a non-renormalized boundary—or a leaf—on which a non-renormalized boundary state is given.²¹ A codimension-1 surface foliated by such leaves is called a holographic screen [128], or Q -screen in the quantum context [129], on which a holographic theory for general spacetimes is supposed to live [58–60]. This framework is indeed consistent with the hypothetical relationship [124] between bulk spacetime and the entanglement structure of a boundary state. A sketch of a holographic screen, leaves, and a holographic slice is given in Fig. 2.6.

¹⁹ This condition does not fix the normalizations of k^μ and l^μ separately, but it is sufficient to ensure the validity of the following treatment. In particular, s^μ is invariant under rescalings of k^μ and l^μ satisfying $k_\mu l^\mu = -2$. Below, we fix this freedom conveniently in each setup when we give explicit expressions for k^μ , l^μ , $\theta_{k,l}$, and $\Theta_{k,l}$.

²⁰ More precisely, we must use a modified version of the quantum expansion which includes a bulk entropy contribution from an “exterior” region, as described in Ref. [36].

²¹ If we follow the flow further beyond $\lambda = 0$, then we lose this property; in this sense, $\sigma(0)$ is the “maximally unrenormalized” leaf.

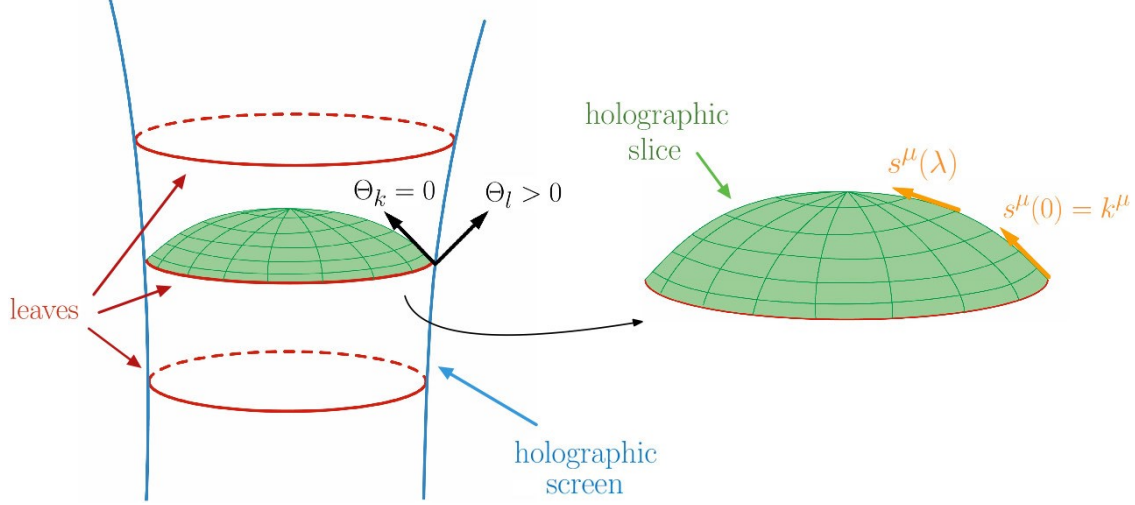


Figure 2.6: A holographic theory resides on a holographic screen, which is a hypersurface foliated by quantum marginally antitrapped (or trapped) surfaces called leaves. At a given boundary time, a boundary state lives on a leaf, from which one can reconstruct a bulk equal-time hypersurface using the flow equation in Eq. (2.4.1).

In some cases, non-renormalized leaves cannot be reached with a finite evolution in λ . In particular, this occurs in asymptotically AdS and flat spacetimes. In an asymptotically AdS spacetime, the metric in the asymptotic region can be expanded in a Fefferman-Graham series [130]

$$ds^2 = \frac{L^2}{z^2} \{g_{ab}(x^a, z)dx^a dx^b + dz^2\}, \quad (2.4.4)$$

where L is the AdS length scale, $a, b = 0, \dots, d-1$, and $g_{ab}(x^a, z) = g_{ab}^{(0)}(x^a) + z^2 g_{ab}^{(2)}(x^a) + \dots$ with $g_{ab}^{(0)}$ being the conformal boundary metric. Suppose that $\sigma(\lambda)$ is taken to be a constant $t = x^0$ surface at a constant $z = \epsilon$. The null normals are then given by $k_\mu = (L/\epsilon)(-dt + dz)$ and $l_\mu = (L/\epsilon)(-dt - dz)$, yielding $\Theta_k = -(d-1)/L$ and $\Theta_l = (d-1)/L$ up to higher order corrections in ϵ . Here, we have used the fact that the quantum expansions $\Theta_{k,l}$ approach classical expansions $\theta_{k,l}$ in the asymptotic region because of the lack of matter there. This implies that a leaf of any regularized boundary located at $z > 0$ is, in fact, a renormalized one, and that the non-renormalized holographic screen can lie formally only at spacelike infinity,²² where a nonregularized holographic CFT lives in AdS/CFT. It also gives the evolution vector

$$s^\mu = \frac{(d-1)\epsilon}{L^2} \frac{\partial}{\partial z} + O(\epsilon^2), \quad (2.4.5)$$

showing that the holographic slice extending inward from a regularized boundary evolves initially in the z direction, up to corrections suppressed by ϵ ; see Fig. 2.7(a). This behavior

²² We adopt a definition of the non-renormalized holographic screen such that it is a hypersurface foliated by leaves, on which *at least one* of the $\Theta_{k,l}$ vanishes.

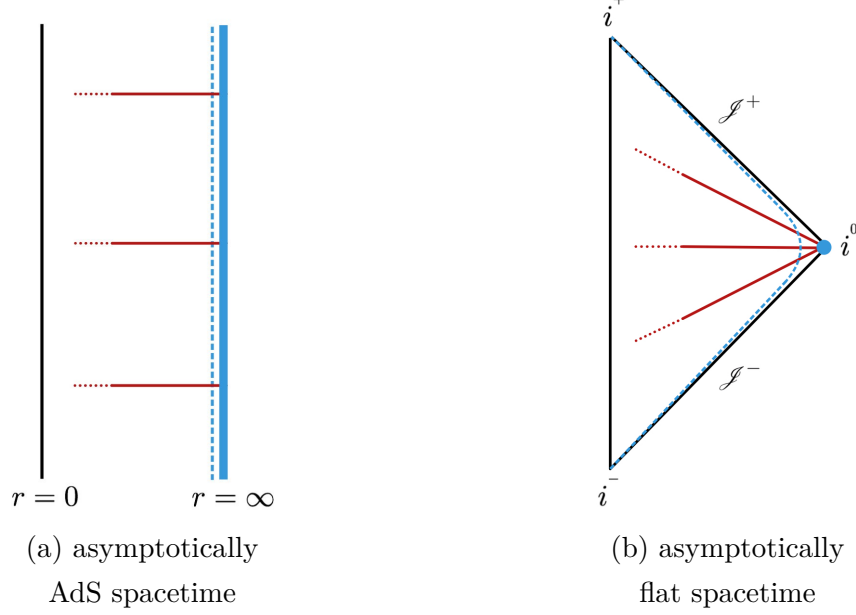


Figure 2.7: Renormalized holographic screens, on which regularized holographic theories reside, are timelike hypersurfaces in an asymptotically AdS and flat spacetimes, which are depicted by the dotted blue lines in the Penrose diagrams in (a) and (b), respectively. In both cases, holographic slices agree with the time slices given by timelike Killing vectors in the asymptotic regions. In the non-renormalized limit, physics occurring in a finite time is described by holographic theories located at the conformal boundary ($r = \infty$) and spatial infinity (i^0) in the cases of asymptotically AdS and flat spacetimes, respectively.

is the same as that of the conventional holographic renormalization group flow [131–135] in the AdS/CFT correspondence. We also find that the flow freezes, $|dz/d\lambda| \rightarrow 0$, as $z \rightarrow 0$.

For an asymptotically flat spacetime, the metric in the asymptotic region can be expanded in the Bondi-Sachs form [136, 137] as

$$ds^2 = -\frac{V}{r}e^{2\beta}du^2 - 2e^{2\beta}dudr + r^2h_{AB}(dx^A - U^A du)(dx^B - U^B du), \quad (2.4.6)$$

where $A, B = 1, \dots, d-1$, and each function admits a large r expansion of the form $V = r + O(1)$, $\beta = O(r^{-2})$, $U^A = O(r^{-2})$, and $h_{AB} = O(1)$. As in the case of an asymptotically AdS spacetime, let us consider $\sigma(\lambda)$ which is a constant time slice of the surface $r = R$. The null normals are then given by $k_\mu = du$ and $l_\mu = (V/r)du + 2dr$, leading to $\Theta_k = -2/R$ and $\Theta_l = 2/R$ at the leading order in $1/R$. We thus find that a leaf of a regularized holographic screen is a renormalized one, and the flow vector is

$$s^\mu = \frac{2}{R} \left(-\frac{\partial}{\partial u} + \frac{\partial}{\partial r} \right) + O\left(\frac{1}{R^2}\right), \quad (2.4.7)$$

showing that the generated flow follows a hypersurface of equal Minkowski time in the inward radial direction, up to corrections suppressed by $1/R$, and that it freezes as $R \rightarrow \infty$. This is

illustrated in Fig. 2.7(b), from which we see that any process occurring in the bulk within a finite time interval can be described by a boundary theory located at spatial infinity i^0 . By taking the time interval to infinity, one recovers an S -matrix description in the bulk. This view is consistent with discussions of flat space holography advanced, e.g., in Refs. [138–142].

de Sitter spacetime

Let us now discuss de Sitter spacetime. As we will see below, care is needed to consider holography of de Sitter spacetime in a static patch. Our view is that in this context, the concept of exact de Sitter spacetime arises only as a result of idealization, very much analogous to an eternal single-sided black hole in an asymptotically flat spacetime (which does not exist because of Hawking radiation).²³

For now, we bypass this issue by considering de Sitter spacetime in a cosmological setup. In particular, we consider an empty bubble universe in which there is a positive vacuum energy density ρ_Λ . The interior of the bubble is an open Friedmann-Lemaître-Robertson-Walker (FLRW) universe described by the metric

$$ds^2 = -d\tau^2 + a(\tau)^2 \left[d\chi^2 + \left\{ \frac{1}{\sqrt{-\kappa}} \sinh(\sqrt{-\kappa}\chi) \right\}^2 d\Omega^2 \right]. \quad (2.4.8)$$

Here,

$$a(\tau) = \sqrt{\frac{-\kappa}{\tilde{\Lambda}}} \sinh(\sqrt{\tilde{\Lambda}}\tau) \quad (2.4.9)$$

is the scale factor, where $\tilde{\Lambda}$ is related to the vacuum energy density by $\tilde{\Lambda} = 8\pi\rho_\Lambda l_p^2/3$, and $\kappa < 0$ is a curvature parameter related to the physical curvature radius by $r_{\text{curv}}(\tau) = a(\tau)/\sqrt{-\kappa}$. The holographic screen is located at the apparent horizon, where $\theta_k = 0$ and $\theta_l > 0$:

$$\chi = \chi_{\text{sc}}(\tau) = \frac{1}{\sqrt{-\kappa}} \ln \left[\coth \frac{\sqrt{\tilde{\Lambda}}\tau}{2} \right], \quad (2.4.10)$$

where we have taken k^μ and l^μ to be the future-directed ingoing and outgoing null vector orthogonal to a leaf, respectively, and we have indexed leaves by FLRW times at their locations. Here, we have ignored the difference between classical and quantum expansions, which is valid because we consider the semiclassical vacuum state, implying that Gibbons-Hawking radiation is not extracted as semiclassical radiation.

At late times, the universe behaves like a de Sitter spacetime:

$$a(\tau) \approx \frac{\sqrt{-\kappa}}{2\sqrt{\tilde{\Lambda}}} e^{\sqrt{\tilde{\Lambda}}\tau} \quad \text{for } \tau \gg \frac{1}{\sqrt{\tilde{\Lambda}}}. \quad (2.4.11)$$

The metric in Eq. (2.4.8) then becomes flat slicing of the de Sitter spacetime; see Fig. 2.8. On the other hand, the physical distance to the holographic screen approaches the Hubble

²³ This is consistent with the expectation from string theory that there is no absolutely stable de Sitter vacuum [143, 144].

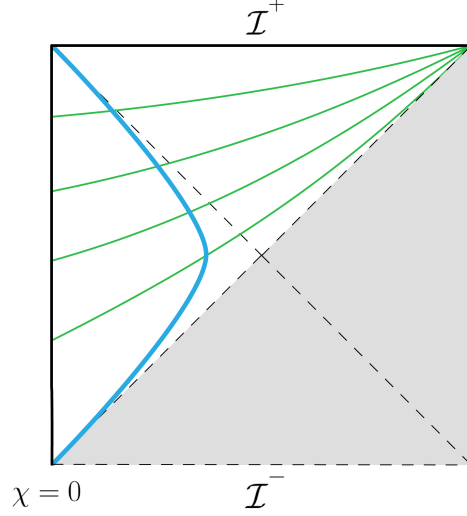


Figure 2.8: Penrose diagram of de Sitter spacetime. A de Sitter bubble corresponds to the upper-left half of this diagram, where FLRW equal-time hypersurfaces are drawn by green lines. The holographic screen is located along the trajectory of the apparent horizon, which is depicted by the solid blue curve.

radius $1/\sqrt{\tilde{\Lambda}}$:

$$a(\tau)\chi_{\text{sc}}(\tau) = \frac{1}{\sqrt{\tilde{\Lambda}}} \sinh(\sqrt{\tilde{\Lambda}}\tau) \ln \left[\coth \frac{\sqrt{\tilde{\Lambda}}\tau}{2} \right] \xrightarrow{\tau \gg \frac{1}{\sqrt{\tilde{\Lambda}}}} \frac{1}{\sqrt{\tilde{\Lambda}}} \left(1 - \frac{2}{3} e^{-2\sqrt{\tilde{\Lambda}}\tau} \right). \quad (2.4.12)$$

The holographic theory thus describes a static patch of the late-time de Sitter spacetime, i.e. the patch inside the holographic screen.

The holographic slices approach static de Sitter time slices at late times. To see this, we can calculate k^μ and l^μ for arbitrary (τ, χ)

$$k^\mu = \begin{pmatrix} 1 \\ -\sqrt{\frac{\tilde{\Lambda}}{-\kappa}} \frac{1}{\sinh(\sqrt{\tilde{\Lambda}}\tau)} \end{pmatrix}, \quad l^\mu = \begin{pmatrix} 1 \\ \sqrt{\frac{\tilde{\Lambda}}{-\kappa}} \frac{1}{\sinh(\sqrt{\tilde{\Lambda}}\tau)} \end{pmatrix} \quad (2.4.13)$$

and their associated quantum expansions

$$\begin{cases} \Theta_k \approx \theta_k = 2\sqrt{\tilde{\Lambda}} \left[\coth(\sqrt{\tilde{\Lambda}}\tau) - \frac{\coth(\sqrt{-\kappa}\chi)}{\sinh(\sqrt{\tilde{\Lambda}}\tau)} \right], \\ \Theta_l \approx \theta_l = 2\sqrt{\tilde{\Lambda}} \left[\coth(\sqrt{\tilde{\Lambda}}\tau) + \frac{\coth(\sqrt{-\kappa}\chi)}{\sinh(\sqrt{\tilde{\Lambda}}\tau)} \right]. \end{cases} \quad (2.4.14)$$

The evolution vector is then

$$\begin{pmatrix} s^\tau \\ s^\chi \end{pmatrix} = 2\sqrt{\tilde{\Lambda}} \begin{pmatrix} \coth(\sqrt{\tilde{\Lambda}}\tau) \\ -\sqrt{\frac{\tilde{\Lambda}}{-\kappa}} \frac{\coth(\sqrt{-\kappa}\chi)}{\sinh^2(\sqrt{\tilde{\Lambda}}\tau)} \end{pmatrix}. \quad (2.4.15)$$

We are interested in the spacetime region well inside the holographic screen

$$\chi \ll \chi_{\text{sc}}(\tau) \xrightarrow{e^{\sqrt{\Lambda}\tau} \gg 1} \sqrt{-\kappa} \chi \ll 2e^{-\sqrt{\Lambda}\tau} \quad (2.4.16)$$

at late times

$$\alpha = \frac{1}{\sqrt{\Lambda}} \ll r_{\text{curv}}(\tau) \longrightarrow e^{\sqrt{\Lambda}\tau} \gg 1, \quad (2.4.17)$$

where α is the Hubble radius of the late time universe. In this region, the relation between the FLRW coordinates (τ, χ) and the de Sitter static coordinates (t, r) is given by

$$\begin{cases} \alpha e^{\frac{\tau}{\alpha}} = \sqrt{\alpha^2 - r^2} e^{\frac{t}{\alpha}}, \\ \frac{\alpha}{2} \sqrt{-\kappa} \chi e^{\frac{\tau}{\alpha}} = r, \end{cases} \quad (2.4.18)$$

leading to the s^μ vector in the static coordinates

$$\begin{pmatrix} s^t \\ s^r \end{pmatrix} \approx \begin{pmatrix} -\frac{8}{3\alpha^3} \frac{r^2}{(1-r^2/\alpha^2)^2} e^{-\frac{2t}{\alpha}} \\ -\frac{2}{r} \left(1 - \frac{r^2}{\alpha^2}\right) \end{pmatrix} \approx \begin{pmatrix} -\frac{8r^2}{3\alpha^3} e^{-\frac{2t}{\alpha}} \\ -\frac{2}{r} \end{pmatrix}. \quad (2.4.19)$$

We thus find that both s^t and s^r are negative, and $|s^t| \ll |s^r|$ in the relevant region.

From the above analysis, we conclude that in the holographic description of de Sitter spacetime, the region swept by holographic slices is the interior of the static patch. Thus, what is analogous to the exterior of a single-sided black hole is the interior of a static patch in de Sitter spacetime. This indicates that the interior of the static patch is the region in which semiclassical field operators can be reconstructed in a simple manner. We note that the same conclusion can also be obtained by regularizing de Sitter spacetime in different ways, for example by considering a big-bang universe filled with two fluid components with the equation-of-state parameters $w = -1$ and $w > -1$, or with a single fluid component of $w = -1 + \epsilon$, where $\epsilon (> 0)$ is taken sufficiently small that the system can be viewed as in a de Sitter vacuum [58, 59]. This provides a justification for the description of de Sitter spacetime adopted in Sections 2.2 and 2.3, focusing on a single static patch.²⁴

It is important that the holographic description based on a single static patch, which we may call “single-sided” de Sitter spacetime, assumes an appropriate physical regularization. Had we started with exact de Sitter spacetime, then the location of a leaf would be on the bifurcation surface or the future horizon. In this case, a subregion of the leaf would have degenerate extremal surfaces, all of which are located on the future horizon and have areas equal to the volume of the subregion. This would imply that holographic slices sweep only a codimension-1 surface in the bulk, i.e. the future horizon, failing to reconstruct the codimension-0 spacetime [58].²⁵ For $\Theta_k = 0$ and $\Theta_l \neq 0$, this can also be seen from the fact

²⁴ It is also comforting that there is a perturbative positive energy theorem in a static patch of de Sitter spacetime [145].

²⁵ More precisely, if we use global information of the boundary state, the interior of the static patch may be reconstructed. However, the entanglement wedge of a boundary subregion in this case contains either the entirety or none of the interior, depending on the size of the subregion. This indicates a “singular” nature of the limit of de Sitter spacetime.

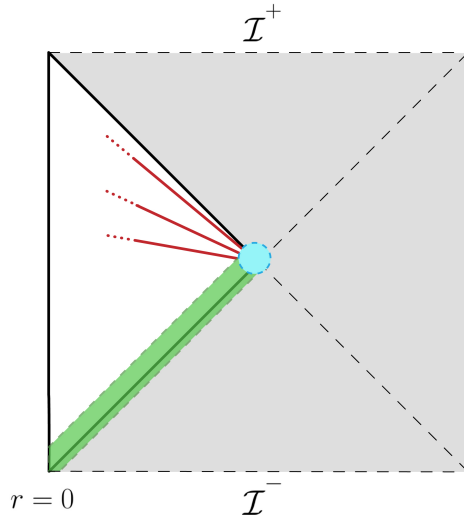


Figure 2.9: A holographic description of “single-sided” de Sitter spacetime. Holographic slices agree (approximately) with equal-time hypersurfaces in static coordinates. The smaller the amount of UV renormalization becomes, the closer renormalized leaves are to the bifurcation surface. The geometry in the vicinity of the past horizon (indicated by light green) and the bifurcation surface (light blue) is regularization dependent.

that $s^\mu \propto k^\mu$.

The picture of a single-sided de Sitter spacetime is shown in Fig. 2.9, in which the region near the bifurcation surface and the past horizon should be viewed as regularization dependent. This picture can also be obtained if we begin with a renormalized holographic screen foliated by renormalized leaves located deep inside the static patch and then push the screen outward by unrenormalizing it using the flow equation. Renormalized leaves then approach the bifurcation surface (or the future horizon if the assumed deviation from exact de Sitter spacetime is significant), but they never get there. Given the non-decoupling of bulk gravity on these leaves, we expect that the holographic theory on the screen is gravitational, but we do not make any further speculation about this theory here.

de Sitter entropy

The fact that holographic states in a single-sided de Sitter spacetime are analogous to those in a single-sided black hole suggests that we can interpret the de Sitter entropy in an analogous manner to the black hole case [146]. At the classical level, de Sitter spacetime is parameterized by one continuous number: the Hubble radius α . At the quantum level, this freedom leads to the corresponding independent quantum states, which are discretized. The number of independent states corresponding to the value of the Hubble radius between α and $\alpha + \delta\alpha$ is

$$\mathcal{N} \sim e^{S_{\text{GH}}} \frac{\delta\alpha}{\alpha}, \quad (2.4.20)$$

where

$$S_{\text{GH}} = \frac{\pi\alpha^2}{l_{\text{P}}^2} \quad (2.4.21)$$

is the Gibbons-Hawking entropy. As in standard statistical mechanics, this result does not depend on the detailed choice of $\delta\alpha$ (unless $\delta\alpha$ is taken to be exponentially small in S_{GH}).

Now, suppose that the time scale for the evolution of a de Sitter microstate is given by Δt . Using the energy of Eq. (2.3.16), we find that the required uncertainty of α , determined by $\Delta E \Delta t \sim 1$, is

$$\Delta\alpha \sim \frac{l_{\text{P}}^2}{\Delta t}. \quad (2.4.22)$$

We expect that $\Delta t \sim 1/T_{\text{H}} \sim \alpha$ because of large redshift between the location of the stretched horizon, where everything is controlled by the string scale, and the location at which the time t is measured, $r = 0$. This gives

$$\Delta\alpha \sim \frac{l_{\text{P}}^2}{\alpha}. \quad (2.4.23)$$

In the semiclassical regime $\alpha \gg l_{\text{P}}$, this uncertainty is smaller than the Planck length, $\Delta\alpha \ll l_{\text{P}}$, which is consistent with the fact that the Hubble radius can be precisely specified (treated classically) in a semiclassical theory.

Note that the situation described above is analogous to the black hole case. For a black hole, $E = M$ and $r_+ \sim M l_{\text{P}}^2$ give $\Delta r_+ \sim l_{\text{P}}^2/\Delta t$. Assuming $\Delta t \sim 1/T_{\text{H}} \sim M l_{\text{P}}^2$, this leads to

$$\Delta r_+ \sim \frac{1}{M}. \quad (2.4.24)$$

In the semiclassical regime $M \gg 1/l_{\text{P}}$, we find $\Delta r_+ \ll l_{\text{P}}$.

There is one apparent difference between the de Sitter and black hole cases. Since the de Sitter spacetime is an “inside-out” version of the black hole spacetime, simple operators can cause excitations inside the stretched horizon, $r < r_{\text{s}}$, and with many such excitations the geometry will be backreacted. In particular, we can form a black hole inside the de Sitter horizon, leading to a different semiclassical geometry. We will now try to understand microscopic entropies of such geometries and their relations to the de Sitter entropy.²⁶

We naturally expect that the original theory on the holographic screen can accommodate all such solutions. This interpretation is consistent if we regard a holographic theory of de Sitter spacetime to be associated with a fixed vacuum energy ρ_{Λ} (within uncertainty), and not a fixed horizon radius. In this case, the solution with a black hole has an entropy smaller than the solution with no black hole [57], so the dimension of the Hilbert space of the holographic theory is

$$\dim \mathcal{H} \sim \sum_{\Delta S \leq \Delta S_{\text{max}}} e^{S_{\text{GH}} - \Delta S} \Rightarrow \ln \dim \mathcal{H} \sim S_{\text{GH}} \quad (2.4.25)$$

even including all the spacetimes with varying sizes of black holes. Here, ΔS is the entropy deficit of a spacetime with a black hole(s), S_{GH} is the Gibbons-Hawking entropy expressed

²⁶ A similar issue was also discussed in Ref. [63].

in terms of the vacuum energy ρ_Λ or the Hubble radius α_0 of the de Sitter spacetime without a black hole

$$S_{\text{GH}} = \frac{\pi \alpha_0^2}{l_{\text{P}}^2} = \frac{3}{8 \rho_\Lambda l_{\text{P}}^4}, \quad (2.4.26)$$

and ΔS_{max} is the entropy deficit when the cosmic and black hole horizons have the same radius, which occurs with $r_{+, \text{bh}} = r_{+, \text{dS}} = \alpha_0 / \sqrt{3}$:

$$\Delta S_{\text{max}} = S_{\text{GH}} - 2 \times \frac{\pi \alpha_0^2}{3 l_{\text{P}}^2} = \frac{\pi \alpha_0^2}{3 l_{\text{P}}^2} = \frac{1}{8 \rho_\Lambda l_{\text{P}}^4}, \quad (2.4.27)$$

where the factor of 2 comes from the fact that we have both cosmological and black hole horizons. The fact that ΔS_{max} is positive serves as a consistency check for our interpretation.²⁷

In fact, because of the inside-out nature, forming black holes inside the de Sitter horizon is analogous to forming (small) black holes outside the black hole horizon, specifically in the zone, while keeping the total energy fixed (though the latter necessarily breaks the spherical symmetry). Such excited states can be included in the Hilbert space associated with the central black hole without changing its dimensions, $\ln \dim \mathcal{H} = S_{\text{bh}}$, at the leading order.

2.4.3 Holographic Hilbert spaces

Before concluding this section, we discuss the structure of holographic boundary Hilbert spaces in a more general setting. The picture presented here builds on the analyses performed in Refs. [35, 58] and is suggested by relations between bulk geometries and boundary entanglement entropies.

Effective boundary Hilbert spaces $\mathcal{H}_{\text{eff}}(\mathcal{A})$

Consider a set of states associated with a leaf σ characterized by its volume $\mathcal{A}$ on the boundary. ($\mathcal{A}$ is an area from the point of view of bulk spacetime.) Holography implies that

²⁷ In $d + 1$ dimensions, this occurs when the black hole mass becomes

$$M = \frac{d-1}{d-2} \left(\frac{d-2}{d} \right)^{d/2} \frac{\text{vol}(\Omega_{d-1})}{8\pi l_{\text{P}}^{d-1}} \alpha_0^{d-2} \quad (2.4.28)$$

with

$$r_{+, \text{bh}} = r_{+, \text{dS}} = \sqrt{\frac{d-2}{d}} \alpha_0. \quad (2.4.29)$$

Here, $\text{vol}(\Omega_{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$ is the volume (area) of the $(d-1)$ -dimensional unit sphere. The maximal entropy deficit is thus

$$\Delta S_{\text{max}} = \frac{\text{vol}(\Omega_{d-1}) \alpha_0^{d-1}}{4 l_{\text{P}}^{d-1}} \left[1 - 2 \times \left(\frac{d-2}{d} \right)^{\frac{d-1}{2}} \right], \quad (2.4.30)$$

which is indeed positive for all $d \geq 2$.

the number of degrees of freedom on the leaf is given by

$$\mathcal{N} = \frac{\mathcal{A}}{4G_{\text{N}}}, \quad (2.4.31)$$

which we will assume to be uniformly distributed over the leaf. Here, we have used G_{N} instead of l_{P} , since discussion in this subsection does not depend on the number of spacetime dimensions. We denote the Hilbert space comprising states of these $\mathcal{N}$ degrees of freedom by $\mathcal{H}_{\text{eff}}(\mathcal{A})$, so that

$$\ln \dim \mathcal{H}_{\text{eff}}(\mathcal{A}) = \frac{\mathcal{A}}{4G_{\text{N}}}. \quad (2.4.32)$$

The reason for the subscript “eff” will become clear later.

We expect that states of these $\mathcal{N}$ degrees of freedom can represent various different spacetimes, or more precisely the domain of dependence D_σ of a spacelike hypersurface bounded by σ in these spacetimes. In addition, for a given bulk spacetime, there may be many independent states in $\mathcal{H}_{\text{eff}}(\mathcal{A})$ that span the space of microstates for the spacetime, as in the case of black hole and de Sitter spacetimes. How can these happen?

Let us assume that the entanglement entropy of subregions of a boundary state $|\psi\rangle$ dual to a semiclassical geometry can be calculated via the Hubeny-Rangamani-Ryu-Takayanagi (HRRT) prescription [13, 103, 104] (or its quantum extension [14, 15, 105]). In particular, we assume that the boundary and bulk Hilbert spaces can be appropriately factorized, which may involve a gauge choice or the introduction of edge modes [147–152]. Here we consider the “classical limit,” meaning that all the subregions we consider contain $O(\mathcal{N})$ degrees of freedom. Given a bulk spacetime, one can then find the corresponding entanglement entropies for all subregions of the boundary. The collection of all boundary subregions and their corresponding entanglement entropies will be referred to as the entanglement structure of the state, which we denote by $\mathcal{S}(|\psi\rangle)$.

First, we note that for a given entanglement structure $\mathcal{S}_0$, we can always find a basis of the Hilbert space $\mathcal{H}_{\text{eff}}(\mathcal{A})$ in which all basis states have the specified entanglement structure. This is because by applying local unitaries to a state, one can generate $e^{\mathcal{N}}$ orthogonal states while preserving the entanglement structure of the original state. This fact, however, does not mean that these $e^{\mathcal{N}}$ states span an $e^{\mathcal{N}}$ -dimensional space of microstates for bulk spacetime corresponding to the entanglement structure $\mathcal{S}_0$. Indeed, by generically superposing $e^{O(\mathcal{N})}$ of these states, one would obtain a state whose entanglement structure is drastically different from $\mathcal{S}_0$, so that the resulting state is dual to a completely different spacetime, if it represents bulk spacetime at all.

Of course, given an entanglement structure, there exists a subspace of dimension $e^{O(\mathcal{N}^p)}$ with $p < 1$ in which generic states have this same entanglement structure up to $O(\mathcal{N}^p)$ corrections. This is because we generally have

$$\mathcal{S}\left(\sum_{i=1}^{e^M} c_i |\psi_i\rangle\right) = \mathcal{S}_0 + O(M), \quad (2.4.33)$$

where $\mathcal{S}(|\psi_i\rangle) = \mathcal{S}_0$ for all i , so that for $M = O(\mathcal{N}^p)$ the corrections are suppressed by powers of $\mathcal{N}$ compared to $\mathcal{S}_0$, which is of order $\mathcal{N}$. The subspace obtained in this way, however,

comprises only an exponentially small subset of $\mathcal{H}_{\text{eff}}(\mathcal{A})$; in particular, it is a measure zero subset of $\mathcal{H}_{\text{eff}}(\mathcal{A})$ in the classical limit.

A nontrivial thing is that for $\mathcal{N} \gg 1$, there exist subspaces of dimension $e^{O(\mathcal{N})}$ in which generic states have the same entanglement structure $\mathcal{S}_0$ up to small corrections. Specifically, for such a subspace spanned by basis states $|\psi_i\rangle$ ($i = 1, \dots, e^{Q\mathcal{N}}$), we have

$$\mathcal{S}\left(\sum_{i=1}^{e^{Q\mathcal{N}}} c_i |\psi_i\rangle\right) = \mathcal{S}_0 + O(\mathcal{N}^p; p < 1), \quad (2.4.34)$$

where $Q \leq 1$ does not scale with $\mathcal{N}$. The existence of these subspaces with entanglement structures invariant under superpositions is expected from canonical typicality (also referred to as the general canonical principle) [153, 154], which states that generic states in such a subspace have the same reduced density matrix for small subsystems (up to small corrections). This is the case despite the fact that the size of the subspace is large enough that one would naively think that superpositions would ruin the entanglement structure at $O(\mathcal{N})$. The proof of this statement is purely kinematical and hence applies generally. In fact, according to canonical typicality the correction term in Eq. (2.4.34) is exponentially small, $O(e^{-Q\mathcal{N}/2})$.

In Fig. 2.10, we show a sketch of the collection $\mathcal{C}$ of states in $\mathcal{H}_{\text{eff}}(\mathcal{A})$ which have the same entanglement structure $\mathcal{S}_0$ up to corrections higher order in $1/\mathcal{N}$

$$\mathcal{C} = \{|\psi\rangle \mid \mathcal{S}(|\psi\rangle) = \mathcal{S}_0 + O(\mathcal{N}^p; p < 1)\} \quad (2.4.35)$$

in the case that they form a subspace of dimension $e^{Q\mathcal{N}}$ with $Q < 1$. In addition to this subspace represented by the pink plane, $\mathcal{C}$ contains $e^{\mathcal{N}} - e^{Q\mathcal{N}}$ states (slightly “thickened” in $\mathcal{H}_{\text{eff}}(\mathcal{A})$) orthogonal to it, which are schematically represented by the red arrow. Furthermore, in the $e^{Q\mathcal{N}}$ -dimensional subspace, there are exponentially rare states that do not have the entanglement structure $\mathcal{S}_0$. These states can be obtained by fine-tuning coefficients when we expand $|\psi\rangle$ in terms of the basis states $|\psi_i\rangle$ of the subspace and are represented by white arrows.

In Ref. [58], it was argued that it is this $e^{Q\mathcal{N}}$ -dimensional subspace that comprises the space of microstates for a spacetime. In particular, for $Q < 1$ the corresponding entanglement structure $\mathcal{S}_0$ can be non-maximal,²⁸ and generic states in this subspace can have dual bulk spacetimes which are simply reconstructable. In this case, even if one considers an exponentially large superposition of microstates, geometric operators are effectively linear so long as the state is generic within the subspace [106]. On the other hand, if $Q = 1$, the “subspace” is the whole Hilbert space $\mathcal{H}_{\text{eff}}(\mathcal{A})$, so applying Page’s analysis [155], we see that the only entanglement structure consistent with Eq. (2.4.34) is that of maximal entropy. In this case, the resulting spacetime is not in a simple wedge, and reconstruction of the bulk requires some level of nonlinearity, or state dependence.²⁹

²⁸ By the entanglement structure being maximal, we mean that for any subregion A its entanglement entropy S_A is maximal, i.e. $S_A = \|A\|/4G_N$, at the leading order in $1/\|A\|$. Here $\|A\|$ is the volume of A on the boundary.

²⁹ This is indeed the case for the interior of a black hole; see Section 2.5.1.

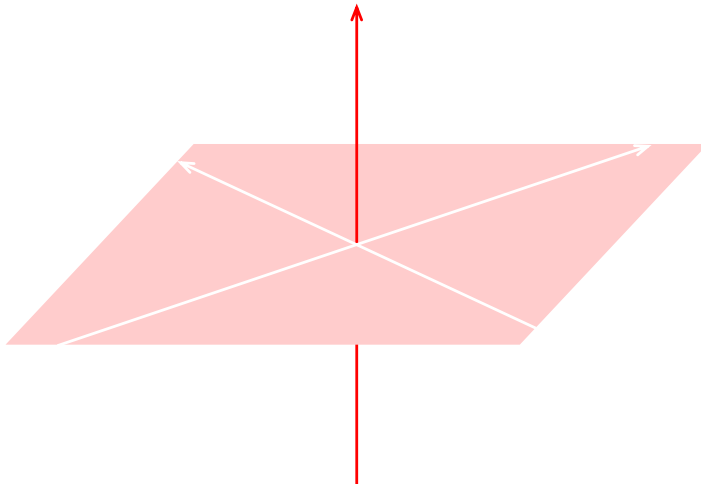


Figure 2.10: A sketch of the collection of states in holographic Hilbert space $\mathcal{H}_{\text{eff}}(\mathcal{A})$ which have the same entanglement structure at leading order in $1/\mathcal{N}$. It forms an $e^{Q\mathcal{N}}$ -dimensional subspace of $\mathcal{H}_{\text{eff}}(\mathcal{A})$ (represented by the pink plane) except that exponentially rare nongeneric states are excluded (white arrows), and that $e^{\mathcal{N}} - e^{Q\mathcal{N}}$ isolated states orthogonal to it are added (red arrow).

The structure discussed above allows for a single holographic Hilbert space $\mathcal{H}_{\text{eff}}(\mathcal{A})$ to harbor effective subspaces dual to different geometries. In fact, similarly to Eq. (2.4.25), one can show that $\mathcal{H}_{\text{eff}}(\mathcal{A})$ satisfying Eq. (2.4.32) can support a number of $e^{Q\mathcal{N}}$ -dimensional subspaces with $Q < 1$. Geometric operators are approximately linear in each of these subspaces, which gives the effective linear space of microstates for a fixed semiclassical geometry.

Fundamental boundary Hilbert space $\mathcal{H}_{\text{UV}}$

So far, we have considered boundary Hilbert space with fixed volume $\mathcal{A}$: $\mathcal{H}_{\text{eff}}(\mathcal{A})$. However, general spacetime involves boundary evolution in which the volume of a leaf (a boundary equal-time surface) changes [60]. Such an evolution can occur in the “fundamental” Hilbert space $\mathcal{H}_{\text{UV}}$ which contains, at least effectively, $\mathcal{H}_{\text{eff}}(\mathcal{A})$ ’s with different $\mathcal{A}$ ’s:³⁰

$$\mathcal{H}_{\text{UV}} \supset \{\mathcal{H}_{\text{eff}}(\mathcal{A}) \mid \dim \mathcal{H}_{\text{eff}}(\mathcal{A}) \in \mathbb{N}\}. \quad (2.4.36)$$

A naive possibility is to literally have $\mathcal{H}_{\text{UV}} \supset \bigoplus_{\mathcal{A}} \mathcal{H}_{\text{eff}}(\mathcal{A})$, but this need not be the case. In fact, motivated by the relation between geometric objects in the bulk and quantum information theoretic quantities on the boundary, which has been learned in AdS/CFT and is expected to apply beyond, one can imagine that $\mathcal{H}_{\text{eff}}(\mathcal{A})$ ’s are contained in $\mathcal{H}_{\text{UV}}$ in a more intricate manner.

³⁰ A semiclassical description is valid only when $\dim \mathcal{H}_{\text{eff}}(\mathcal{A}) \gg 1$, but we will be sloppy about it in writing Eq. (2.4.36).

In Refs. [35, 124], it was envisioned that $\mathcal{H}_{\text{eff}}(\mathcal{A})$ is embedded in $\mathcal{H}_{\text{UV}}$ as an effective subspace (in the sense discussed before; see, e.g., Fig. 2.10) in which a generic state has the property

$$\sum_i S_i = \frac{\mathcal{A}}{4G_{\text{N}}}, \quad (2.4.37)$$

where S_i represents the entanglement entropy of the state in a sufficiently small subregion, A_i , of the holographic space Ω , on which $\mathcal{H}_{\text{UV}}$ is defined; the sum runs over all of these small subregions such that $\Omega = \cup_i A_i$ and $A_i \cap A_j = \emptyset$ ($i \neq j$). This allows us to consider holographic states of all bulk spacetimes, and also their dynamics, in a single Hilbert space $\mathcal{H}_{\text{UV}}$.

The Hilbert space $\mathcal{H}_{\text{UV}}$ will be defined by introducing a short distance cutoff δ in Ω and then sending δ to zero, so that

$$\dim \mathcal{H}_{\text{UV}} \rightarrow \infty. \quad (2.4.38)$$

For $(d+1)$ -dimensional asymptotically AdS and flat bulk spacetimes, S_i will behave as

$$S_i \sim \frac{\|\partial A_i\|}{\delta^{d-2}} \quad \text{and} \quad S_i = f(A_i) \frac{\|A_i\|}{\delta^{d-1}}, \quad (2.4.39)$$

respectively. Here, $\|x\|$ represents the volume of the object x , ∂A_i is the boundary of A_i , and $f(A_i)$ is a function of $O(1)$.³¹

According to this picture, states representing a cosmological spacetime with finite leaf area comprise a tiny effective subspace of $\mathcal{H}_{\text{UV}}$, obtained from a generic state in $\mathcal{H}_{\text{UV}}$ by an infinite number (in the limit $\delta \rightarrow 0$) of fine-tunings.³² It will be interesting to study the dynamics in $\mathcal{H}_{\text{UV}}$ in the thermodynamic limit, using information from semiclassical theory.

2.5 Analytic Extension of Spacetime in Quantum Gravity

In a single-sided system, the spacetime region obtained from analytic extension in general relativity—e.g. the interior of a black hole and the region outside a static patch in de Sitter spacetime—emerges as a collective phenomenon involving horizon (and possibly other) modes. This can occur because a huge gravitational red/blueshift at the stretched horizon makes the string dynamics relevant in a static description, which makes the state take a generic, universal form across all low energy species. In this section, we review this construction [42–46] and refine it to include the effect of black hole evolution analyzed in Section 2.3.3.

³¹ In asymptotically flat spacetime, for example, the ratio of $f(A_i)$ for A_i being a half of Ω , $A_{1/2}$, to that for A_i being a cutoff size region, A_δ , is $f(A_{1/2})/f(A_\delta) = (1/\sqrt{\pi})\{\Gamma[d/2]/\Gamma[(d+1)/2]\}$.

³² A similar picture was discussed for de Sitter spacetime in Ref. [156].

2.5.1 The interior of a black hole

Consider a state of a black hole of mass M at some time $t = t_*$. We assume that it is in the semiclassical vacuum state, which is achieved typically more than one scrambling time t_{scr} after the last perturbation, where³³

$$t_{\text{scr}} = 2r_+ \left[\ln \frac{r_+}{l_P} + O(1) \right]. \quad (2.5.3)$$

As discussed in Section 2.3.1, we assume that the mass of the black hole (as well as other quantities such as the momentum) is specified with the maximal precision allowed by the uncertainty principle, $\Delta M \sim T_H$. If the state of the system involves a superposition of a black hole of a wider range of masses, then our consideration below applies to each of the branches containing a black hole with minimal uncertainties.³⁴

Ignoring evolution effects for now, a black hole vacuum state is given as in Eq. (2.3.40). This state picks out a set of special states of the combined system of soft and far modes: those multiplied by the hard-mode states $|\{n_\alpha\}\rangle$ in Eq. (2.3.40). We denote these states using the double-ket symbol

$$\|\{n_\alpha\}_A\gg = \varsigma_n^A \sum_{i_n=1}^{e^{S_{\text{bh}}(M-E_n)}} \sum_{a=1}^{e^{S_{\text{rad}}}} c_{ni_na}^A |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle. \quad (2.5.4)$$

Here, A is the index for the microstate, specified by the coefficients $c_{ni_na}^A$ in Eq. (2.3.40), and

³³ This expression is obtained from the fact that the scrambling time represented in ingoing Eddington-Finkelstein time $v = t + r_*$ is [18]

$$v_{\text{scr}} = \frac{1}{2\pi T_H} \ln S_{\text{bh}} \approx 4r_+ \ln \frac{r_+}{l_P}. \quad (2.5.1)$$

The scrambling time in Schwarzschild time is then related to this time by [45]

$$t_{\text{scr}} = v_{\text{scr}} + r_{*s} \approx 2r_+ \ln \frac{r_+}{l_P}, \quad (2.5.2)$$

where the r_{*s} term in the middle expression comes from the fact that the scrambling time in Eq. (2.5.1) is defined as the minimal time needed to recover information about an object falling into the stretched horizon (at $r_* = r_{*s}$) at a location sufficiently far from the horizon, $r_* \approx O(r_+)$. In the last equation, we have used $r_{*s} \approx -2r_+ \ln(r_+/l_P)$, obtained from Eq. (2.2.17) by identifying l_s with l_P .

³⁴ In practice, different branches decohere with environment, so focusing on a single branch is phenomenologically forced on us when we discuss the physics of the black hole itself, such as its interior, using the pure state language. Of course, when discussing more global aspects, such as the full unitarity of a black hole formation and evaporation process, we need to take into account all these branches [99, 100]. However, the effect from such a superposition, i.e. a superposition of “macroscopically distinguishable” black holes, is subdominant in entropic consideration, compared with the Bekenstein-Hawking entropy associated with a black hole with minimal uncertainties.

ς_n^A is the normalization constant

$$\varsigma_n^A = \frac{1}{\sqrt{\sum_{i_n=1} e^{S_{\text{bh}}(M-E_n)} \sum_{a=1} e^{S_{\text{rad}}} c_{ni_na}^{A*} c_{ni_na}^A}} = \sqrt{z} e^{\frac{E_n}{2T_H}} \left(1 - \frac{1}{2} \tilde{\varepsilon}_n^{AA}\right), \quad (2.5.5)$$

where z is given by Eq. (2.3.31). The last expression is obtained by using statistical properties of $c_{ni_na}^A$ in Eq. (2.3.42), and $\tilde{\varepsilon}_n^{AB}$ is the quantity analogous to ε_n^{AB} in Eq. (2.3.38):

$$\tilde{\varepsilon}_n^{AB} \equiv z e^{\frac{E_n}{T_H}} \sum_{i_n=1} e^{S_{\text{bh}}(M-E_n)} \sum_{a=1} e^{S_{\text{rad}}} c_{ni_na}^{A*} c_{ni_na}^B - \delta_{AB} \approx O\left(e^{-\frac{1}{2}\{S_{\text{bh}}(M-E_n)+S_{\text{rad}}\}}\right), \quad (2.5.6)$$

which satisfies

$$(\tilde{\varepsilon}_n^{AB})^* = \tilde{\varepsilon}_n^{BA}, \quad \sum_n \frac{e^{-\frac{E_n}{T_H}}}{z} \tilde{\varepsilon}_n^{AB} = 0. \quad (2.5.7)$$

Substituting Eq. (2.5.4) into Eq. (2.3.40), we see that the state can be written in the thermofield double form

$$|\Psi_A(M)\rangle = \frac{1}{\sqrt{z}} \sum_n e^{-\frac{E_n}{2T_H}} |\{n_\alpha\}\rangle \|\{n_\alpha\}_A\rangle, \quad (2.5.8)$$

up to exponentially suppressed corrections of order $e^{-\{S_{\text{bh}}(M)+S_{\text{rad}}\}/2}$. We can also check that the states $\|\{n_\alpha\}_A\rangle$ are orthonormal in $\{n_\alpha\}$ as well as the microstate index A , up to exponentially small corrections [46]:

$$\langle\langle \{m_\alpha\}_A | \{n_\alpha\}_B \rangle\rangle = \delta_{mn} \eta_n^{AB}, \quad (2.5.9)$$

where

$$\eta_n^{AB} \equiv \begin{cases} 1 & \text{for } A = B \\ \tilde{\varepsilon}_n^{AB} \approx O\left(e^{-\frac{1}{2}\{S_{\text{bh}}(M-E_n)+S_{\text{rad}}\}}\right) & \text{for } A \neq B. \end{cases} \quad (2.5.10)$$

Since the spectrum of the states $|\{n_\alpha\}\rangle$ represents semiclassical physics in the zone of the single-sided black hole, the corresponding states $\|\{n_\alpha\}_A\rangle$ of the soft and far modes can be identified as the states in the second exterior of an effective two-sided black hole, given by Eq. (2.5.8).

With this identification, one can construct annihilation and creation operators for modes in the second exterior

$$\tilde{b}_\gamma^A = \sum_n \sqrt{n_\gamma} \|\{n_\alpha - \delta_{\alpha\gamma}\}_A\rangle \langle\langle \{n_\alpha\}_A | \quad (2.5.11)$$

$$= \sum_n \sqrt{n_\gamma} \varsigma_{n_\gamma}^A \varsigma_n^{A*} \sum_{i_{n_\gamma}=1} e^{S_{\text{bh}}(M-E_{n_\gamma})} \sum_{j_n=1} e^{S_{\text{bh}}(M-E_n)} \sum_{a,b=1} e^{S_{\text{rad}}} c_{n_\gamma i_{n_\gamma} a}^A c_{nj_n b}^{A*} |\psi_{i_{n_\gamma}}^{(n_\gamma)}\rangle |\phi_a\rangle \langle\psi_{j_n}^{(n)}| \langle\phi_b|, \quad (2.5.12)$$

$$\tilde{b}_\gamma^{A\dagger} = \sum_n \sqrt{n_\gamma + 1} \|\{n_\alpha + \delta_{\alpha\gamma}\}_A\rangle \langle\langle \{n_\alpha\}_A | \quad (2.5.13)$$

$$= \sum_n \sqrt{n_\gamma + 1} \varsigma_{n_\gamma}^A \varsigma_n^{A*} \sum_{i_{n_\gamma}=1} e^{S_{\text{bh}}(M-E_{n_\gamma})} \sum_{j_n=1} e^{S_{\text{bh}}(M-E_n)} \sum_{a,b=1} e^{S_{\text{rad}}} c_{n_\gamma i_{n_\gamma} a}^A c_{nj_n b}^{A*} |\psi_{i_{n_\gamma}}^{(n_\gamma)}\rangle |\phi_a\rangle \langle\psi_{j_n}^{(n)}| \langle\phi_b|, \quad (2.5.14)$$

in addition to those in the first exterior, Eqs. (2.3.49) and (2.3.50). Here, $n_{\pm\gamma} \equiv \{n_\alpha \pm \delta_{\alpha\gamma}\}$, and we have used the same symbol γ to specify both the first and second exterior modes. From these operators, one can then construct annihilation and creation operators for interior modes, as well as the infalling time evolution operator, through an appropriate Bogoliubov transformation [42–46].³⁵ We will see this construction more explicitly below for an evaporating black hole.

In this picture, the second exterior of the black hole emerges effectively as a collective phenomenon associated with the soft and far modes. Note that while the energy of each soft or early Hawking mode may be tiny, $\lesssim T_H$, their collective excitation can have a much larger energy, $\gg T_H$. In particular, a “quasi particle” created by $\tilde{b}_\gamma^{A\dagger}$ in the second exterior has energy E_γ , which is negative and $|E_\gamma| > T_H$. The fact that this energy is negative implies that the Hamiltonian H of the original single-sided black hole is mapped to the generator of the timelike isometry in the effective two-sided picture

$$H \mapsto H_R - H_L, \quad (2.5.15)$$

where H_R and H_L are the Hamiltonian operators acting on the first and second exteriors, respectively. As we will see below, excitation in the black hole interior is a superposition of quasi-particle excitations built by $\tilde{b}_\gamma^A$ ’s and $\tilde{b}_\gamma^{A\dagger}$ ’s and original excitations in the zone.

It is worth mentioning that we have no freedom in choosing the basis in the space of $\|\{n_\alpha\}_A\>\>$ ’s. In other words, the interpretation that $\|\{n_\alpha\}_A\>\>$ is a state in which n_α of the mode corresponding to α in the first exterior is excited in the second exterior, is not invariant under unitary transformation

$$\|\{n_\alpha\}_A\>\> \rightarrow \|\{n_\alpha\}_A\>\> = \sum_{n'} U_{nn'} \|\{n'_\alpha\}_A\>\>, \quad (2.5.16)$$

so that a state obtained by acting $\tilde{b}_\gamma^A$ or $\tilde{b}_\gamma^{A\dagger}$ on $|\Psi_A(M)\>\>$ in Eq. (2.3.40) cannot be regarded as the vacuum state at the semiclassical level. This is because the minimal uncertainty condition imposed on $|\Psi_A(M)\>\>$ allows us to physically distinguish between different $\|\{n_\alpha\}_A\>\>$ ’s, and hence

³⁵ This construction differs from a similar construction in Refs. [47, 52, 53], in which the degrees of freedom that are identified as those in the first exterior of the effective two-sided black hole increase as the black hole evaporates; in particular, Hawking radiation emitted earlier composes degrees of freedom in the first exterior. In the construction here, the number of degrees of freedom composing the first exterior (zone modes) decreases as the evaporation progresses; in particular, early Hawking radiation is identified as a part of the degrees of freedom in the second exterior.

The idea that the construction of interior operators involves early Hawking radiation was promoted in Ref. [50], but its specific realization is different here. In contrast to the picture laid out in Ref. [50], the second exterior of the effective two-sided geometry arises primarily from degrees of freedom directly associated with the black hole (soft modes), and the involvement of Hawking radiation is indirect (although it is significant for an old black hole, i.e. a black hole that is nearly maximally entangled with the rest of the system). In particular, the structure of entanglement is not bipartite between the first exterior and early Hawking radiation degrees of freedom as envisioned in Ref. [50].

there is no ambiguity in defining the corresponding $\|\{n_\alpha\}_A\>\>$'s. Our framework, therefore, does not suffer from the “frozen vacuum” problem posed in Ref. [157].

We also comment that $|\phi_a\rangle$ in Eq. (2.5.4) may not consist of only the early Hawking radiation; if a major portion of the radiation interacts with other degrees of freedom, such degrees of freedom must also be included in $|\phi_a\rangle$. This is because $|\phi_a\rangle$'s represent states of all the degrees of freedom that are entangled significantly with the hard and soft modes. It follows that the degrees of freedom that were once the second-exterior degrees of freedom keep playing that role even if they no longer take the form of Hawking radiation. This situation, however, does not last forever. The role as a second exterior emerges only relationally with respect to the hard modes. Thus, after the black hole is evaporated completely, the Hawking radiation as well as any other degrees of freedom that have interacted with it become regular matter that does not have any spacetime interpretation.

Finally, we emphasize that the construction described here is performed in a theory with gravity in the bulk, implying that the precise spectrum of the horizon and zone modes (represented by the boundary conditions on the stretched horizon in the low energy field theory) are not arbitrary; rather, they are determined by a consistent UV theory. In other words, the construction here should be viewed as being performed in a holographic boundary theory (see, e.g., footnote 42 below). The physics in a nongravitational bulk can be reproduced if we send $l_P \rightarrow 0$ (with fixed $r_+ = 2Ml_P^2$), which implies $l_s \rightarrow 0$ because of Eq. (2.3.12). The change from the distant description to the infalling description considered here is then reduced to the change of reference frames in quantum field theory (from the Rindler to Minkowski frame for $r_+ \rightarrow \infty$).

Effective theory of the interior of an evaporating black hole

As we have seen in Section 2.3.3, for an evaporating black hole some of the ingoing zone modes are not excited in the black hole vacuum; see, e.g., Eq. (2.3.44) and Fig. 2.4. Strictly speaking, however, the coefficients of the terms involving zone-mode states in which these modes are excited are not exactly zero (though they are exponentially suppressed); see footnote 16. We can therefore define the double-ket states as in Eq. (2.5.4) for all $\{n_\alpha\}$ and construct annihilation and creation operators, Eqs. (2.5.11)–(2.5.14), for the second exterior mode corresponding to any hard mode γ in the first exterior.

Nevertheless, it is true that in the black hole vacuum state, the hard-mode states in which these ingoing modes are excited do not have sizable coefficients. We thus have to use Eq. (2.5.4) in Eq. (2.3.44) instead of Eq. (2.3.40), and we obtain

$$|\Psi_A(M)\rangle \propto \prod_{\substack{\bar{\alpha} \in \\ m_\Phi \lesssim T_H, \omega \sim T_H}} b_{\bar{\alpha}} \sum_n \prod_{\substack{\alpha' \in \\ \omega > \omega_{\text{barrier}}}} \delta_{n_{\alpha'}, 0} e^{-\frac{E_n}{2T_H}} |\{n_\alpha\}\rangle \|\{n_\alpha\}_A\>, \quad (2.5.17)$$

up to exponentially small corrections. We thus find that the ingoing modes α' with the frequency larger than the barrier height are virtually not entangled with the corresponding second exterior modes in the vacuum microstate.

This implies that the spacetime is not smooth across the future horizon of the second

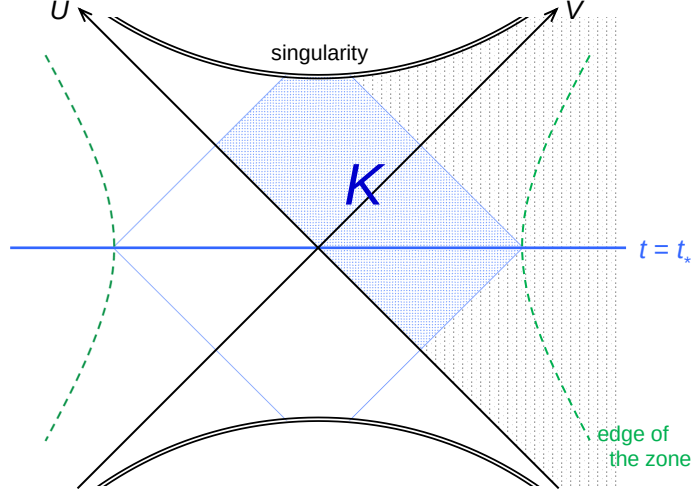


Figure 2.11: The spacetime region K described by the effective theory of the interior erected at a boundary (Schwarzschild) time $t = t_*$.

exterior, or the past horizon of the first exterior, since the missing entanglement is essential for the connectedness of the spacetime there. The spacetime region which the effective theory built on the state in Eq. (2.5.17) describes is thus the shaded region in Fig. 2.11, which we refer to as region K . Here, U and V are the Kruskal-Szekeres coordinates erected at $t = t_*$:

$$ds^2 = -dUdV + r^2 d\Omega_{d-1}^2, \quad (2.5.18)$$

which are given in the near horizon region by

$$\begin{cases} U = -R e^{-\omega} \\ V = R e^{\omega}, \end{cases} \quad \begin{cases} U = R e^{-\omega} \\ V = R e^{\omega}, \end{cases} \quad \begin{cases} U = R e^{-\omega} \\ V = -R e^{\omega}, \end{cases} \quad \begin{cases} U = -R e^{-\omega} \\ V = -R e^{\omega} \end{cases} \quad (2.5.19)$$

for Region I ($U < 0, V > 0$), Region II ($U, V > 0$), Region III ($U > 0, V < 0$), and Region IV ($U, V < 0$), respectively, with

$$R = 2\sqrt{r_+|r - r_+|}, \quad \omega = \frac{1}{2r_+}(t - t_*). \quad (2.5.20)$$

With the state given by Eq. (2.5.17), the interior of the black hole is in the semiclassical vacuum in the infalling frame.

We now study in more detail how the interior region is described in this effective theory. Let us first take the Schrödinger picture (as we have implicitly been doing). We want to understand what an object located in the zone and falling toward the black hole will experience as it crosses the horizon. An excited state with N particles in the zone can be obtained by applying appropriate superpositions of creation operators $b_\gamma^\dagger$ on a black hole vacuum state:

$$|\Psi(t = t_*)\rangle \equiv \prod_{i=1}^N \left(\sum_{\gamma} f_{\gamma}^{(i)} b_{\gamma}^\dagger \right) |\Psi_A(M)\rangle, \quad (2.5.21)$$

where $|\Psi_A(M)\rangle$ is the black hole vacuum microstate in Eq. (2.3.44), and $f^{(i)}$ is the wave-function of the i -th particle represented in the γ space. This state can be straightforwardly mapped to that in the effective theory—we simply have to take $|\Psi_A(M)\rangle$ to be that in Eq. (2.5.17) instead of Eq. (2.3.44). We denote this state in the effective theory by $|\Psi(\tau = 0)\rangle$:

$$|\Psi(t = t_*)\rangle \mapsto |\Psi(\tau = 0)\rangle, \quad (2.5.22)$$

by choosing the origin of the infalling time τ to match the boundary time $t = t_*$.

The operators used to interpret a state in the effective theory is given by b_γ and $b_\gamma^\dagger$ in Eqs. (2.3.49) and (2.3.50) and $\tilde{b}_\gamma^A$ and $\tilde{b}_\gamma^{A\dagger}$ in Eqs. (2.5.11) and (2.5.13). In particular, we can form infalling mode operators out of superpositions of these operators

$$a_\xi^A = \sum_\gamma (\alpha_{\xi\gamma} b_\gamma + \beta_{\xi\gamma} b_\gamma^\dagger + \zeta_{\xi\gamma} \tilde{b}_\gamma^A + \eta_{\xi\gamma} \tilde{b}_\gamma^{A\dagger}), \quad (2.5.23)$$

$$a_\xi^{A\dagger} = \sum_\gamma (\beta_{\xi\gamma}^* b_\gamma + \alpha_{\xi\gamma}^* b_\gamma^\dagger + \eta_{\xi\gamma}^* \tilde{b}_\gamma^A + \zeta_{\xi\gamma}^* \tilde{b}_\gamma^{A\dagger}), \quad (2.5.24)$$

where ξ is the label in which the frequency ω with respect to boundary time t is traded with the frequency Ω associated with infalling time τ , and $\alpha_{\xi\gamma}$, $\beta_{\xi\gamma}$, $\zeta_{\xi\gamma}$, and $\eta_{\xi\gamma}$ are the Bogoliubov coefficients calculable using the standard field theory method [70, 71].³⁶

To obtain the generator of infalling time evolution $\tilde{H}$, we start from the generator of boundary time evolution

$$H = H_{\text{soft}} + \sum_\gamma \omega_\gamma b_\gamma^\dagger b_\gamma + H_{\text{int}}(\{b_\gamma\}, \{b_\gamma^\dagger\}), \quad (2.5.26)$$

where H_{soft} gives time evolution of the soft modes as well as interactions between the hard and soft modes. The infalling time evolution generator is then given by

$$\tilde{H} = \sum_\xi \Omega_\xi a_\xi^{A\dagger} a_\xi^A + \tilde{H}_{\text{int}}(\{a_\xi^A\}, \{a_\xi^{A\dagger}\}), \quad (2.5.27)$$

where $\tilde{H}_{\text{int}}(\{a_\xi^A\}, \{a_\xi^{A\dagger}\})$ is determined by matching it with $H_{\text{int}}(\{b_\gamma\}, \{b_\gamma^\dagger\})$. The state in the effective theory evolves as

$$|\Psi(\tau)\rangle = e^{-i\tilde{H}\tau} |\Psi(\tau = 0)\rangle. \quad (2.5.28)$$

Note that the generator in Eq. (2.5.27) contains a term that involves both b_γ and $\tilde{b}_\gamma^A$ operators, e.g. $b_\gamma^\dagger \tilde{b}_\gamma^A$, even in the free part. Thus, it acts on the hard and soft (and far) modes simultaneously, as can be seen in Eqs. (2.5.12) and (2.5.14).

³⁶ For a massless scalar field, for example, Eq. (2.5.23) takes the form

$$a_\xi^A = \pm \frac{i}{2\pi\sqrt{\Omega_\xi T_H}} \int_0^\infty d\omega_\gamma \left[\frac{\Xi}{\sqrt{1 - e^{-\frac{\omega_\gamma}{T_H}}}} b_\gamma + \frac{\Xi^*}{\sqrt{e^{\frac{\omega_\gamma}{T_H}} - 1}} b_\gamma^\dagger - \frac{\Xi^*}{\sqrt{1 - e^{-\frac{\omega_\gamma}{T_H}}}} \tilde{b}_\gamma^A - \frac{\Xi}{\sqrt{e^{\frac{\omega_\gamma}{T_H}} - 1}} \tilde{b}_\gamma^{A\dagger} \right] \quad (2.5.25)$$

in the near horizon limit. Here, we have adopted the continuum notation for the sum over the frequency, and $\Xi = (\Omega_\xi/2\pi T_H)^{\pm \frac{i\omega_\gamma}{2\pi T_H}} \Gamma\left(1 \pm \frac{\omega_\gamma}{2\pi i T_H}\right) / \left|\Gamma\left(1 \pm \frac{\omega_\gamma}{2\pi i T_H}\right)\right|$ is a pure phase. The $\pm$ symbol in these equations takes $+$ and $-$ for ingoing and outgoing modes, respectively.

The effective theory is defined with the initial state on Σ , the zone at $t = t_*$, and its mirror in the second exterior $\tilde{\Sigma}$, and the spacetime is not smoothly connected across the U axis (or formally has an infinite energy shock wave along the U axis). We can, therefore, trust results obtained using the theory only in the region

$$K = D(\Sigma \cup \tilde{\Sigma}) \cap \{(U, V) | V > 0\}, \quad (2.5.29)$$

shown in Fig. 2.11, where $D(X)$ represents the domain of dependence of X . This is, however, all we need to describe the fate of the fallen object. Note that here we consider the fate of an object falling into a black hole in its semiclassical vacuum. Also, if an object leaves the region $D(\Sigma \cup \tilde{\Sigma})$ before hitting the singularity, we can choose to erect an effective theory at a different time such that the entire trajectory of the object (until it hits the singularity) is contained in the region described by the effective theory; such a choice is always possible [42].

The description of the interior which is more along the lines of a conventional quantum field theory treatment can be obtained by adopting the Heisenberg picture [46]. With the operators in Eqs. (2.5.23) and (2.5.24), quantum field operators at $\tau = 0$ are given by

$$\tilde{\Phi}_\Gamma(\mathbf{x}, 0) = \sum_{s, \Omega, \mathbf{L}} \left(f_s(\Omega, \mathbf{L}) \varphi_{\Omega, \mathbf{L}}(\mathbf{x}) a_\xi^A + g_s(\Omega, \mathbf{L}) \varphi_{\Omega, \mathbf{L}}^*(\mathbf{x}) a_{\xi^c}^{A\dagger} \right), \quad (2.5.30)$$

where we have decomposed index ξ into Γ , s , Ω , and $\mathbf{L}$ which represent species, spin, frequency, and orbital angular momentum quantum numbers, respectively, with ξ^c representing the CPT conjugate of ξ . Here, $f_s(\Omega, \mathbf{L})$ and $g_s(\Omega, \mathbf{L})$ are the standard factors providing Lorentz representation of the field (Dirac spinors, polarization vectors, etc), and $\varphi_{\Omega, \mathbf{L}}(\mathbf{x})$ are the spatial wavefunctions, determined by matching $\tilde{\Phi}_\Gamma(\mathbf{x}, 0)$ with quantum field operators of the original theory at $t = t_*$.³⁷

The Heisenberg picture field operators can then be defined as

$$\tilde{\Phi}_\Gamma(\mathbf{x}, \tau) = e^{i\tilde{H}\tau} \tilde{\Phi}_\Gamma(\mathbf{x}, 0) e^{-i\tilde{H}\tau}, \quad (2.5.31)$$

where $\tilde{H}$ is given in Eq. (2.5.27). The quantities we are interested in are the correlators

$$\langle \tilde{\Phi}_{\Gamma_1}(x_1) \tilde{\Phi}_{\Gamma_2}(x_2) \cdots \tilde{\Phi}_{\Gamma_n}(x_n) \rangle = \text{Tr} [\tilde{\rho}(0) \tilde{\Phi}_{\Gamma_1}(x_1) \tilde{\Phi}_{\Gamma_2}(x_2) \cdots \tilde{\Phi}_{\Gamma_n}(x_n)], \quad (2.5.32)$$

where $x_i = \{\mathbf{x}_i, \tau_i\}$, and $\tilde{\rho}(0) = |\Psi(\tau = 0)\rangle \langle \Psi(\tau = 0)|$. Since these are expectation values in the state at a fixed time $\tau = 0$, we must adopt the in-in formalism rather than the more conventional in-out formalism to calculate them. This ultimately comes from the fact that the S -matrix cannot be defined at the semiclassical level for an object falling into a black hole.

Using the Schwinger-Keldysh method, Eq. (2.5.32) can be written as a path integral over an appropriate closed time contour with the boundary condition given by $\tilde{\rho}(0)$. We can also calculate it using perturbation theory in the canonical in-in formalism. Note that fields in Eq. (2.5.32) need not be in the interior of the black hole; they only need to be in the region

³⁷ This match requires information of the coefficients in Eqs. (2.5.23) and (2.5.24).

K . We can thus compute correlators between fields inside and outside the horizon using Eq. (2.5.32).

In the construction of the effective theory described so far, we have used an input from semiclassical theory to determine the coefficients in Eqs. (2.5.23) and (2.5.24). However, we expect that this is ultimately not needed. In particular, we expect that these coefficients are determined (though not uniquely) by the requirement that the generator $\tilde{H}$ can be written in the local form in terms of the original quantum fields [43]; namely, when a_ξ^A 's in $\tilde{H}$ are represented by b_γ 's and $\tilde{b}_\gamma^A$'s which in turn are represented by quantum fields and their canonical conjugates, $\tilde{H}$ takes a local form in the first and emergent second exteriors. To find the coefficients directly from the boundary theory, one possibility is to use a physical probe to construct the infalling Hamiltonian, along the lines of Refs. [158, 159]. We leave a detailed study of these issues for the future.

State dependence and intrinsic ambiguity

The operators in the effective theory constructed so far, $\tilde{b}_\gamma^A$, $\tilde{b}_\gamma^{A\dagger}$, a_ξ^A , and $a_\xi^{A\dagger}$, depended on the vacuum microstate, indexed by A , on which the excited states are built. This dependence, however, can be relaxed in such a way that a single set of operators can describe an object in the zone even if it is entangled arbitrarily with the vacuum microstates [46, 51] (see also Ref. [53]).

The Hilbert space spanned by all the independent vacuum microstates of a black hole of mass M is given by

$$\mathcal{M} = \left\{ \sum_{A=1}^{e^{S_{\text{tot}}}} a_A |\Psi_A(M)\rangle \left| a_A \in \mathbb{C}, \sum_{A=1}^{e^{S_{\text{tot}}}} |a_A|^2 = 1 \right. \right\}, \quad (2.5.33)$$

where S_{tot} is given by Eq. (2.3.46). We consider a subspace of $\mathcal{M}$ spanned by $e^{S_{\text{eff}}}$ independent microstates

$$\hat{\mathcal{M}} = \left\{ \sum_{A'=1}^{e^{S_{\text{eff}}}} a_{A'} |\Psi_{A'}(M)\rangle \left| a_{A'} \in \mathbb{C}, \sum_{A'=1}^{e^{S_{\text{eff}}}} |a_{A'}|^2 = 1 \right. \right\}, \quad (2.5.34)$$

where $S_{\text{eff}} < S_{\text{tot}}$. By choosing the bases of $\mathcal{M}$ and $\hat{\mathcal{M}}$ appropriately, we can take $\{|\Psi_{A'}(M)\rangle\}$ to be a subset of $\{|\Psi_A(M)\rangle\}$, so $\langle \Psi_{A'}(M) | \Psi_{B'}(M) \rangle = \delta_{A'B'}$.

We now define the following operators associated with the Hilbert subspace $\hat{\mathcal{M}}$

$$\tilde{\mathcal{B}}_\gamma = \sum_{A'=1}^{e^{S_{\text{eff}}}} \tilde{b}_\gamma^{A'}, \quad \tilde{\mathcal{B}}_\gamma^\dagger = \sum_{A'=1}^{e^{S_{\text{eff}}}} \tilde{b}_\gamma^{A'\dagger}, \quad (2.5.35)$$

where $\tilde{b}_\gamma^{A'}$ and $\tilde{b}_\gamma^{A'\dagger}$ are given by Eqs. (2.5.12) and (2.5.14). These operators will work as desired if the excited states obtained by acting $\tilde{b}_\gamma^{A'}$'s and $\tilde{b}_\gamma^{A'\dagger}$'s on $|\Psi_{A'}(M)\rangle$ are effectively orthogonal for different A' 's. In fact, we can show that the algebra of these operators in the black hole Hilbert space built on $\hat{\mathcal{M}}$, i.e. the space spanned by the vacuum microstates of $\hat{\mathcal{M}}$ and the states in which these microstates have been excited, is the same as that of the mode

operators in the second exterior of the corresponding two-sided black hole, up to correction of order [46]

$$\epsilon = \max \left\{ \frac{e^{\frac{E_{\max}}{2T_H}}}{e^{\frac{1}{2}\{S_{\text{bh}}(M)+S_{\text{rad}}\}}}, \frac{e^{\frac{E_{\max}}{2T_H}+S_{\text{eff}}}}{e^{S_{\text{bh}}(M)+S_{\text{rad}}}} \right\}, \quad (2.5.36)$$

where $E_{\max}$ is the maximum energy which an excitation can carry in the semiclassical theory. Given that $E_{\max} \ll M$, we find that the error is exponentially small for

$$S_{\text{eff}} < S_{\text{bh}}(M) + S_{\text{rad}} \approx S_{\text{tot}}, \quad (2.5.37)$$

where the symbol $<$ here means that S_{eff} is smaller than $S_{\text{bh}}(M) + S_{\text{rad}}$ and that the fractional difference between S_{eff} and $S_{\text{bh}}(M) + S_{\text{rad}}$ is not exponentially small, specifically

$$S_{\text{bh}}(M) + S_{\text{rad}} - S_{\text{eff}} \gg \frac{E_{\max}}{2T_H}. \quad (2.5.38)$$

Below, we use the symbols $<$ and $>$ to mean similar relations.

Similarly, we can define infalling mode operators

$$\mathcal{A}_\xi = \sum_\gamma (\alpha_{\xi\gamma} b_\gamma + \beta_{\xi\gamma} b_\gamma^\dagger + \zeta_{\xi\gamma} \tilde{\mathcal{B}}_\gamma + \eta_{\xi\gamma} \tilde{\mathcal{B}}_\gamma^\dagger), \quad (2.5.39)$$

$$\mathcal{A}_\xi^\dagger = \sum_\gamma (\beta_{\xi\gamma}^* b_\gamma + \alpha_{\xi\gamma}^* b_\gamma^\dagger + \eta_{\xi\gamma}^* \tilde{\mathcal{B}}_\gamma + \zeta_{\xi\gamma}^* \tilde{\mathcal{B}}_\gamma^\dagger), \quad (2.5.40)$$

which act linearly in the black hole Hilbert space built on $\hat{\mathcal{M}}$. Here, the coefficients $\alpha_{\xi\gamma}$, $\beta_{\xi\gamma}$, $\zeta_{\xi\gamma}$, and $\eta_{\xi\gamma}$ are the same as those in Eqs. (2.5.23) and (2.5.24). The matrix elements of products of $\mathcal{A}_\xi$ and $\mathcal{A}_\xi^\dagger$ in the black hole Hilbert space built on $\hat{\mathcal{M}}$, then, are the same as the corresponding field theory values on the two-sided black hole background, up to corrections suppressed by ϵ in Eq. (2.5.36).³⁸

The existence of operators $\mathcal{A}_\xi$ and $\mathcal{A}_\xi^\dagger$ allows us to erect the effective theory such that the dependence of operators on states is invisible in the effective theory. Suppose that the state at $t = t_*$ is an entangled state between semiclassical excitations and black hole vacuum microstates

$$|\Psi(t_*)\rangle = \sum_{A=1}^{S_{\text{tot}}} \sum_I d_{AI}(t_*) |\Psi_{A,I}(M)\rangle, \quad (2.5.41)$$

where $|\Psi_{A,I}(M)\rangle$ represents the state in which the semiclassical excitation I exists on the black hole vacuum state $|\Psi_A(M)\rangle$. Even in this case, given that the logarithm of the dimension of the excitation Hilbert space, S_{exc} , is much smaller than S_{tot} , we can write the state using the Schmidt decomposition as

$$|\Psi(t_*)\rangle = \sum_{I=1}^{\mathcal{K}} g_I |\Psi_{A(I),I}(M)\rangle, \quad (2.5.42)$$

³⁸ Precisely speaking, the operators $\mathcal{A}_\xi$ and $\mathcal{A}_\xi^\dagger$ (and $\tilde{\mathcal{B}}_\gamma$ and $\tilde{\mathcal{B}}_\gamma^\dagger$) can be used for states built on a larger vacuum microstate space. Specifically, the algebra of these operators is the same as the corresponding semiclassical algebra in the space of black hole states built on a typical state in subspace $\hat{\mathcal{M}}'$ of $\mathcal{M}$ as long as $\dim(\hat{\mathcal{M}}' \cap \hat{\mathcal{M}}) > \dim \hat{\mathcal{M}}'/2$ [46].

where $\sum_{I=1}^{\mathcal{K}} |g_I|^2 = 1$, $g_I > 0$, and $\mathcal{K}$ is the Schmidt number. The point is that since $\mathcal{K}$ satisfies

$$\mathcal{K} \leq S_{\text{exc}} \ll S_{\text{bh}}(M) + S_{\text{rad}}, \quad (2.5.43)$$

$\hat{\mathcal{M}}$ can always be taken to contain the space of vacuum microstates spanned by $\{|\Psi_{A(I)}(M)\rangle | I = 1, \dots, \mathcal{K}\}$. This guarantees that the effective theory respects the standard tenet of quantum mechanics that physical observables are given by linear operators acting on the Hilbert space of the theory.

Since S_{eff} only needs to satisfy Eq. (2.5.37), one might think that a single, fixed set of $\mathcal{A}_\xi$, $\mathcal{A}_\xi^\dagger$ operators can cover the states built on most of the black hole vacuum microstates in $\mathcal{M}$ of Eq. (2.5.33) by taking

$$S_{\text{eff}} = c \{S_{\text{bh}}(M) + S_{\text{rad}}\} \quad (2.5.44)$$

with c close to (but not exponentially close to) 1. This is, however, not the case. The dimension of the space $\hat{\mathcal{M}}_\perp$ of vacuum microstates that are orthogonal to the states in $\hat{\mathcal{M}}$ is

$$\dim \hat{\mathcal{M}}_\perp = e^{S_{\text{tot}}} - e^{S_{\text{eff}}}, \quad (2.5.45)$$

which is much larger than $\dim \hat{\mathcal{M}} = e^{S_{\text{eff}}}$ even for c close to 1. In fact, for $c > 1/2$, there is a simple relation between the fraction of $\mathcal{M}$ which a fixed set of operators can cover and the size of error for using these operators:

$$\frac{\dim \hat{\mathcal{M}}}{\dim \mathcal{M}} \approx \epsilon. \quad (2.5.46)$$

This makes it clear that we cannot use a fixed set of operators to cover a significant fraction of states in $\mathcal{M}$ while keeping the error ϵ small. In fact, to cover all states in $\mathcal{M}$ by fixed sets of operators, we need double exponentially large number, $O(e^{e^{S_{\text{tot}} - S_{\text{eff}}}})$, of sets.³⁹

Finally, we note that the construction of the effective theory described here has an intrinsic ambiguity coming from the fact that the actions of infalling mode operators are not strictly orthogonal to $\hat{\mathcal{M}}$ in the space of black hole microstates. Specifically, we find that the inner product between states obtained by operating mode operators on two vacuum microstates $|\Psi_A(M)\rangle$ and $|\Psi_B(M)\rangle$ reproduces the field theory value multiplied by δ_{AB} , but with corrections of order ϵ in Eq. (2.5.36) which are not proportional to δ_{AB} . The existence of these exponentially suppressed corrections means that an excited state cannot have an exact and unambiguous association with a unique vacuum microstate.

The fact that the corrections are only of order ϵ , however, implies that up to these exponentially suppressed corrections, the mode operators b_γ , $b_\gamma^\dagger$, $\tilde{\mathcal{B}}_\gamma$, $\tilde{\mathcal{B}}_\gamma^\dagger$, $\mathcal{A}_\xi$, and $\mathcal{A}_\xi^\dagger$ act only on the excitation index I , and not on the vacuum index A' . In other words, ignoring these corrections, the Hilbert space can be viewed as

$$\mathcal{H} \approx \mathcal{H}_{\text{exc}} \otimes (\mathcal{H}_{\text{vac}} \cong \hat{\mathcal{M}}), \quad (2.5.47)$$

³⁹ This is related to the well-known fact that in a Hilbert space of dimension $e^S \gg 1$, there are $O(e^{e^S})$ approximately orthogonal states with exponentially small overlaps of $O(e^{-S/2})$.

where these mode operators act only on $\mathcal{H}_{\text{exc}}$, and an excited state can be associated “uniquely” with a vacuum microstate. The exponentially suppressed corrections discussed here constitute an intrinsic ambiguity of semiclassical physics, resulting from the fact that the black hole system (consisting of zone and horizon modes as well as the relevant degrees of freedom of far modes) is finite dimensional.

Consistency with the semiclassical expectation

Suppose that a falling object hits the stretched horizon at $t = t_*$ when viewed from the exterior. The object then disappears from the zone, but this does not mean that the state immediately becomes the quasi-equilibrium form of Eq. (2.3.44) with Eq. (2.3.42) at $t = t_*$; rather, it stays in a state with excited horizon modes for a while. Now, consider that we erect an effective theory of the interior shortly after t_* : $t = t_* + \delta t$. In this case, the effective theory can have semiclassical excitations in the interior reflecting the fact that there is an object that has fallen into the horizon at t_* . These include everything that the object does to the spacetime region described by the effective theory; in particular, the excitations existing in the effective theory need not be the object itself.⁴⁰

On the other hand, if we erect an effective theory more than one scrambling time after the last disturbance to the stretched horizon, $t = t_* + \Delta t$ with $\Delta t > t_{\text{scr}}$, then we expect that the effective theory finds the semiclassical vacuum in the interior, since the black hole state has already equilibrated by then. (There can be zone mode excitations corresponding to semiclassical objects in the zone.) Is this consistent with the semiclassical expectation?

In Fig. 2.12, we show by the central blue diamond the spacetime region described by the effective theory erected at $t = t_*$, at which the object reaches the stretched horizon. The coordinates U and V are given by Eqs. (2.5.19) and (2.5.20). Now, consider the Kruskal-Szekeres coordinates $\tilde{U}$ and $\tilde{V}$ adapted to the effective theory erected at $t = t_* + \Delta t$, which are related to U and V by

$$\tilde{U} = U e^{\frac{1}{2r_+} \Delta t}, \quad \tilde{V} = V e^{-\frac{1}{2r_+} \Delta t} \quad (2.5.48)$$

(so that $\tilde{U} = U$ and $\tilde{V} = V$ for $\Delta t = 0$). The spacetime region that can be described by this theory, depicted by the red shaded region in Fig. 2.12 for $\Delta t > 0$, must thus satisfy

$$\tilde{U} \lesssim r_z = O(r_+) \quad \Rightarrow \quad U \lesssim r_+ e^{-\frac{1}{2r_+} \Delta t}. \quad (2.5.49)$$

⁴⁰ Suppose, for example, that the object hits the stretched horizon at $t = t_*$ with almost the speed of light. Then, the object will not appear in the effective theory erected at $t = t_* + \delta t$, since it is squeezed into the U axis so strongly that it cannot exist in the effective theory, a theory that cannot describe physics below the Planck length. Even in this case, however, the object may emit a high energy quantum in the outward direction shortly after it crosses the horizon. The quantum would then have to appear in the effective theory as a signal emerging from somewhere on the U axis to the positive V direction, if δt is sufficiently small. (This modification of the boundary condition along the U axis must occur through a map of horizon excitations in the microscopic theory to the effective theory, which requires a UV physics.)

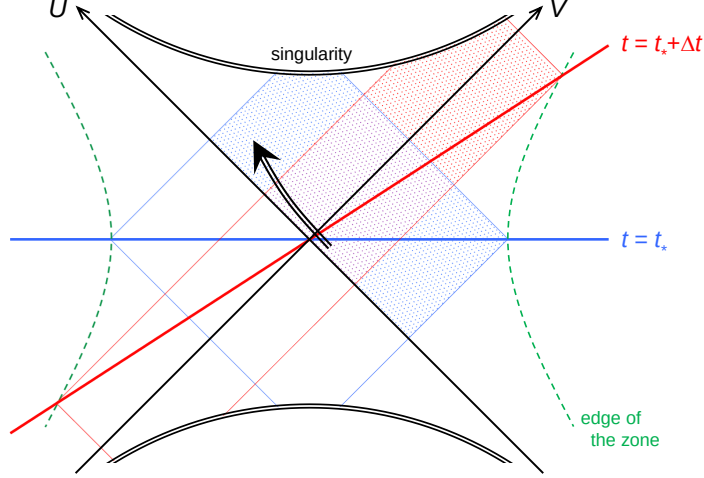


Figure 2.12: The spacetime regions described by effective theories erected at $t = t_*$ (blue) and $t = t_* + \Delta t$ with $\Delta t > 0$ (red). The black arrow indicates a falling object that reaches the stretched horizon at $t = t_*$.

It is then clear that the object that fell into the horizon at $t = t_*$ spends only tiny proper time in this region, given by

$$\Delta\tau \lesssim r_+ e^{-\frac{1}{2r_+} \Delta t}. \quad (2.5.50)$$

Since holography limits the maximum amount of information an object can handle to be of $O(1)$ per Planck time, this implies that the object cannot cause any physical effect in an effective theory erected after $t = t_* + \Delta t_{\max}$, with

$$\Delta t_{\max} = 2r_+ \ln \frac{r_+}{l_P}, \quad (2.5.51)$$

where the expression is reliable up to fractional corrections of order $1/\ln(r_+/l_P)$. Comparing this with the scrambling time in Eq. (2.5.3), we find that the picture is indeed consistent with what we expect from semiclassical theory.

Young black hole and the role of the Page time

The construction of the effective theory described so far applies to a black hole of any age. However, if the black hole is young, i.e. if it is not maximally entangled with other systems, then we can have an alternative construction in which the $\tilde{b}_\gamma^A$ and $\tilde{b}_\gamma^{A\dagger}$ operators act only on soft-mode states, without involving far mode degrees of freedom [44, 46].

This is done by projecting operators in Eqs. (2.5.11) and (2.5.13) on the space of soft-mode states using the Petz map [160, 161]. Specifically, we can take

$$\tilde{b}_\gamma^A = \sigma^{-1/2} \left[\sum_a \langle \phi_a | \left(\sum_n \sqrt{n_\gamma} \|\{n_\alpha - \delta_{\alpha\gamma}\}_A\| \|\{n_\alpha\}_A\| \right) | \phi_a \rangle \right] \sigma^{-1/2}, \quad (2.5.52)$$

$$\tilde{b}_\gamma^{A\dagger} = \sigma^{-1/2} \left[\sum_a \langle \phi_a | \left(\sum_n \sqrt{n_\gamma + 1} \|\{n_\alpha + \delta_{\alpha\gamma}\}_A\| \|\{n_\alpha\}_A\| \right) | \phi_a \rangle \right] \sigma^{-1/2}. \quad (2.5.53)$$

Here,

$$\sigma = \sum_a \langle \phi_a | \left(\sum_n \|\{n_\alpha\}_A\| \right) \|\{n_\alpha\}_A\| | \phi_a \rangle, \quad (2.5.54)$$

and $\|\{n_\alpha\}_A\|$'s are given by Eq. (2.5.4). Note that $\tilde{b}_\gamma^A$, $\tilde{b}_\gamma^{A\dagger}$, and σ are operators acting only on the space of soft-mode states. Infalling mode operators acting only on the hard and soft modes can then be constructed by substituting these $\tilde{b}_\gamma^A$ and $\tilde{b}_\gamma^{A\dagger}$ in Eqs. (2.5.23) and (2.5.24).

The algebra of the above mode operators in the black hole Hilbert space, i.e. the space obtained by acting mode operators on the vacuum state in Eq. (2.3.44), follows that of semiclassical theory up to errors of order

$$\epsilon_{\text{young}} = \max \left\{ \frac{1}{e^{\frac{1}{2}S_{\text{bh}}(M)}}, \frac{e^{S_{\text{rad}}}}{e^{S_{\text{bh}}(M)}} \right\}. \quad (2.5.55)$$

Therefore, if the black hole is young, i.e. $S_{\text{rad}} < S_{\text{bh}}(M)$, then the operators $\tilde{b}_\gamma^A$ and $\tilde{b}_\gamma^{A\dagger}$ in Eqs. (2.5.52) and (2.5.53) as well as a_ξ^A and $a_\xi^{A\dagger}$ constructed using them can be used to describe the interior of the black hole (up to corrections suppressed exponentially in a macroscopic entropy). This elucidates the role of the Page time in constructing the effective theory: if a black hole is young, then interior operators *can be* represented purely using the black hole degrees of freedom (i.e. zone and horizon modes), while if it is old, then the operators *must* involve the early Hawking radiation (i.e. far modes).

The promotion of $\tilde{b}_\gamma^A$, $\tilde{b}_\gamma^{A\dagger}$, a_ξ^A , and $a_\xi^{A\dagger}$ operators to act linearly in $\hat{\mathcal{M}}$ can be made similarly as before. In particular, errors of the promoted operators are of order

$$\hat{\epsilon}_{\text{young}} = \max \left\{ \frac{1}{e^{\frac{1}{2}S_{\text{bh}}(M)}}, \frac{e^{S_{\text{rad}}+S_{\text{eff}}}}{e^{S_{\text{bh}}(M)}} \right\}, \quad (2.5.56)$$

so that these operators work correctly as long as

$$S_{\text{eff}} < S_{\text{bh}}(M) - S_{\text{rad}}. \quad (2.5.57)$$

This agrees with the result of the general analysis in Ref. [51].

Incidentally, in the way of constructing an effective theory of the interior described here, operators $\tilde{b}_\gamma^A$ and $\tilde{b}_\gamma^{A\dagger}$ cannot be represented using only far modes, even if the black hole is old. Namely, unlike Eqs. (2.5.52) and (2.5.53), projecting operators in Eqs. (2.5.11) and (2.5.13) on the space of far-mode states does not work. Technically, this is because of the energy constraint imposed on the black hole, i.e. the combined system of zone and horizon modes. The relation of this statement to entanglement wedge reconstruction, in which the interior of an old black hole is reconstructed only using the early Hawking radiation, will be discussed below.

Relation to entanglement wedge reconstruction

As we have seen, operators in our effective theory describing the interior must involve horizon degrees of freedom, regardless of the age of the black hole. On the other hand, the analysis [18–20, 54, 55] based on holographic entanglement wedge reconstruction [12, 162–166]

says that after the Page time operators acting on early radiation are sufficient to reconstruct a portion of the black hole interior. What is the relation between these two statements?

A key ingredient to understanding this is the boundary time evolution [45]. In general, entanglement wedge reconstruction assumes that we know the time evolution operator of the boundary theory; in models discussed in Refs. [18–20, 54, 55], for example, the Hamiltonian of a system consists of boundary conformal field theory as well as an auxiliary theory coupling to it. With this knowledge, one can reconstruct the state of some of the horizon modes at $t = t_h$ from the state of the radiation at $t = t_R$ if

$$t_R > t_h + t_{\text{scr}}. \quad (2.5.58)$$

Note that as discussed in Section 2.3.2, the notion of zone, horizon, and far modes is associated with a specific time t , so that a component of a horizon mode at some time can become a far mode at a later time.

A detailed way in which the reconstruction described above works was discussed in Ref. [45]. It is essentially the Hayden-Preskill protocol [68] applied to the horizon and zone modes. Recall that zone modes at $t = t_h - t_{\text{sig}}$ either become horizon modes (for ingoing modes) or far modes (for outgoing modes) by $t = t_h$. Here, $t_{\text{sig}} \approx |r_{*s}|$ is the signal propagation time: the time it takes for a massless quantum to propagate from the stretched horizon to the edge of the zone. Hence, radiation at t_R satisfying Eq. (2.5.58) can reconstruct some of the zone modes. In fact, we can arbitrarily choose which zone modes to reconstruct, and we can take them to be hard modes. This implies that radiation at $t = t_R$ can reconstruct an object in the interior of the effective theory erected at $t = t_*$ if

$$t_* < t_R - t_{\text{scr}} - t_{\text{sig}} \quad (2.5.59)$$

and $t_* + t_{\text{scr}} > t_{\text{Page}}$. In other words, an object in the black hole interior can be reconstructed from radiation at $t = t_R$ if it exists in the inner wedge of the stretched horizon at

$$t = t_R - t_{\text{scr}}, \quad (2.5.60)$$

assuming $t_R > t_{\text{Page}}$. This is sketched in Fig. 2.13(a). The entanglement island, $\mathcal{I}_{\text{rad}}$, of the radiation, $R(t_R)$, represents the spacetime region in which some information about the region is reconstructed. Thus, its edge, i.e. the minimal quantum extremal surface for $R(t_R)$, is located near the stretched horizon at $t = t_R - t_{\text{scr}}$. Given that the discussion for reconstruction here is intrinsically semiclassical, and hence does not have a precision of resolving the Planck scale, this is consistent with what was found in Refs. [18, 54, 55].

We stress that while the entanglement island $\mathcal{I}_{\text{rad}}$ represents the interior region which one can reconstruct from the radiation $R(t_R)$, the amount of information that can be reconstructed, or the size of code subspace [11, 119] one can construct, depends on the location in this region. Suppose one wants to reconstruct an interior object that has a significant amount of information, say having (coarse-grained) entropy S_{obj} . Then, the radiation must contain the corresponding amount of information. This modifies Eqs. (2.5.59) and (2.5.60) to

$$t_* < t_R - t_{\text{scr}} - t_{\text{sig}} - \frac{S_{\text{obj}}}{T_H} \quad (2.5.61)$$

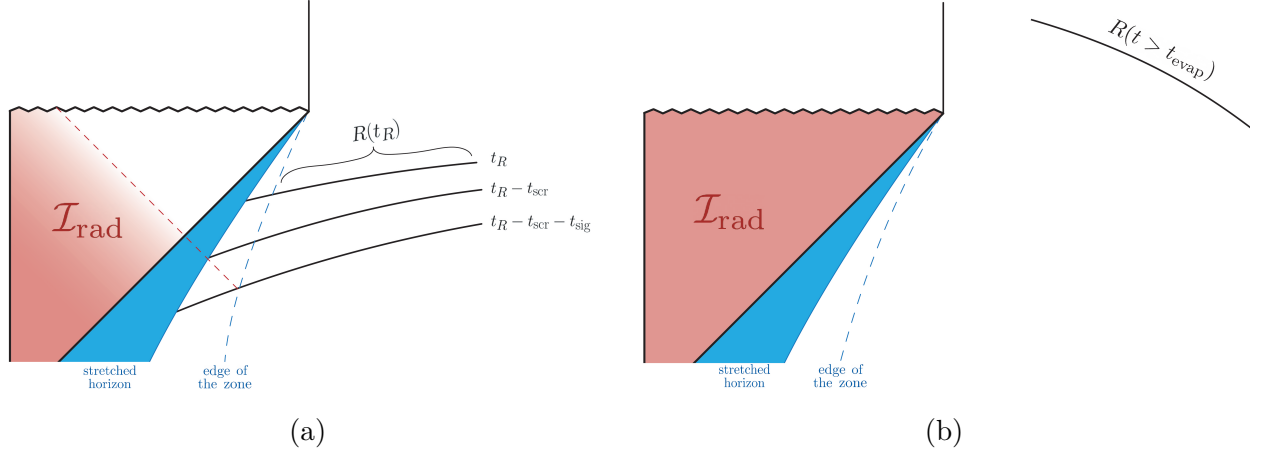


Figure 2.13: (a) Radiation at $t = t_R$, $R(t_R)$, contains information about some of the horizon and zone modes at $t < t_R - t_{\text{scr}} - t_{\text{sig}}$. This allows us to reconstruct a portion $\mathcal{I}_{\text{rad}}$ of the interior spacetime corresponding to the inner wedge of the stretched horizon at $t = t_R - t_{\text{scr}}$. The amount of information that can be reconstructed, however, depends on the location of an object in ingoing time, which is indicated by the gradation of the red shade. (b) Radiation after the black hole is fully evaporated, $R(t > t_{\text{evap}})$, allows for reconstruction of the full interior region.

and

$$t = t_R - t_{\text{scr}} - \frac{S_{\text{obj}}}{T_{\text{H}}}, \quad (2.5.62)$$

where t_R must satisfy $t_R > t_{\text{Page}} + S_{\text{obj}}/T_{\text{H}}$. This feature inherits from the use of the Hayden-Preskill protocol, and more fundamentally boundary time evolution, in the reconstruction. Because of the scrambling and quantum error correcting nature of black hole dynamics, one can arbitrarily choose *which* information to reconstruct [68], but the *amount* of information to be reconstructed is bounded from above by

$$S_{\text{obj}} \leq T_{\text{H}}(t_R - t_{\text{scr}} - v_{\text{obj}}), \quad (2.5.63)$$

where v_{obj} is the ingoing Eddington-Finkelstein time v at the location of the object, chosen so that $v = t$ at the stretched horizon. This is represented by the gradation of the red shade in Fig. 2.13(a).

The discussion above makes it clear that entanglement wedge reconstruction is nothing but the statement of unitarity as viewed from the exterior. Indeed, the entanglement island of radiation after the black hole is fully evaporated can be viewed as the entire interior; see

Fig. 2.13(b).⁴¹

The understanding of entanglement wedge reconstruction described above also elucidates why an operator acting only on radiation $R(t_R)$, in particular a unitary operator representing a physical manipulation of $R(t_R)$, can affect the interior of a black hole $\mathcal{I}_{\text{rad}}$ in a seemingly acausal manner. In the manifestly unitary picture adopted so far in this paper, based on an external view of the black hole, an excitation X in the interior $\mathcal{I}_{\text{rad}}$ is, in fact, “located” in the zone and/or stretched horizon at $t < t_R$ (and possibly in the early radiation as well for an old black hole), which is timelike separated from $R(t_R)$. There is, therefore, no a priori reason why the operator generating X commutes with operators acting on the radiation at $t = t_R$. In particular, an operator $\mathcal{O}_R$ acting on $R(t_R)$ can be a “precursor” [167, 168] of X ; i.e., when evolved backward in time, $\mathcal{O}_R$ induces a change of the state of an object falling into the black hole at an earlier time, which changes the interior of the black hole.

On the other hand, in the global spacetime picture of general relativity, it is still true that radiation at $t = t_R$ is spacelike separated from the interior region $\mathcal{I}_{\text{rad}}$. This raises the question of why there can be any unitary operator acting on $R(t_R)$ which affects $\mathcal{I}_{\text{rad}}$, without completely jeopardizing the intuition coming from the semiclassical, global spacetime picture. This issue was studied in Ref. [169], in which it was argued that if the dynamics of a black hole is sufficiently complex (e.g. leading to a pseudorandom state), then the information about the interior $\mathcal{I}_{\text{rad}}$ cannot be accessed by any simple operator acting on $R(t_R)$ which does not have exponential computational complexity. This reproduces the causality of the semiclassical theory, assuming that it represents only a feature of simple operations performed on the system in each observer frame.

In the present context, the required complexity arises from the dynamics of the stretched horizon as viewed from the exterior of the black hole. Note that the time t_* at which the effective theory describing $\mathcal{I}_{\text{rad}}$ is erected must satisfy Eq. (2.5.61). This implies that ingoing zone modes (as well as most of the horizon modes) at $t = t_*$ must go through the complex horizon dynamics before they become radiation modes at $t = t_R$. In particular, this is true for the ingoing hard modes for any choice of hard modes describing the object carrying the entropy S_{obj} . The semiclassical picture emerging from our framework, therefore, respects causality, at least in the sense of Ref. [169].

In entanglement wedge reconstruction, reconstructing the interior on a Hawking radiation state involves backward time evolution, so that the resulting boundary operators acting on the radiation degrees of freedom are highly fragile; i.e., a small deformation of the operators destroys the success of the reconstruction. In the case of the effective theory, on the other hand, interior operators are given simply by Eqs. (2.5.12), (2.5.14), (2.5.23), and (2.5.24) in terms of hard, soft, and far mode operators at time t_* , which can be obtained easily in the boundary theory from the structure of entanglement of the whole state at $t = t_*$, if

⁴¹ Strictly speaking, the existence of an island in this case is not rigorously established, since its edge cannot be obtained as a surface that is quantum extremized in the regime in which a semiclassical description is valid. However, from physical consideration as well as continuity of $\mathcal{I}_{\text{rad}}$ in t , we expect that $R(t)$ ($t > t_{\text{evap}}$) has the entire interior region as the “island.”

such a state is given.⁴² In this “equal-time” conversion between the distant and infalling descriptions, which use modes at a given time t_* directly, there is no upper limit on the amount of information about the infalling object described, such as that in Eq. (2.5.63); the only restriction on the size of the object is the one coming from the validity of the semiclassical picture.

2.5.2 The outside of a de Sitter static patch

A construction analogous to the black hole interior can be applied to a state describing de Sitter spacetime [43]. As we have argued in Section 2.4.2, a microstate representing empty de Sitter spacetime consists of the state of the “zone modes” inside a static patch as well as that of the horizon modes. Thus, by choosing a subset of the zone modes as hard modes, it can be written in the form of Eq. (2.3.34) with $S(E - E_n) \rightarrow S_{\text{ds}}(E - E_n)$. Here, $S_{\text{ds}}(E)$ is given by Eq. (2.3.27), and E is the “energy” given in terms of the Hubble radius α by $E = \alpha/l_{\text{P}}^2$; see Eq. (2.3.16). As in the case of a black hole, we can define the “double-ket” states using soft modes

$$\|\{n_\alpha\}_A\gg = \varsigma_n^A \sum_{i_n=1}^{e^{S_{\text{ds}}(E-E_n)}} c_{ni_n}^A |\psi_{i_n}^{(n)}\rangle, \quad (2.5.64)$$

where $A = 1, \dots, e^{S_{\text{sys}}}$, and ς_n^A is the normalization constant

$$\varsigma_n^A = \frac{1}{\sqrt{\sum_{i_n=1}^{e^{S_{\text{ds}}(E-E_n)}} c_{ni_n}^{A*} c_{ni_n}^A}} = \sqrt{z} e^{\frac{E_n}{2T_{\text{H}}}} \left[1 + O\left(e^{-\frac{1}{2}S_{\text{ds}}(E-E_n)}\right) \right], \quad (2.5.65)$$

where $T_{\text{H}} = 1/2\pi\alpha$, and z is given by Eq. (2.3.31). In terms of these states, the de Sitter vacuum microstate can be rewritten as

$$|\Psi_A(E)\rangle = \frac{1}{\sqrt{z}} \sum_n e^{-\frac{E_n}{2T_{\text{H}}}} |\{n_\alpha\}\rangle \|\{n_\alpha\}_A\gg. \quad (2.5.66)$$

Since the state $|\{n_\alpha\}\rangle$ is specified by the occupation numbers n_α of modes inside the stretched horizon of a static patch, which we refer to as the polar region here, the thermofield double state in Eq. (2.5.66) represents a vacuum state at $\tau = 0$ of global de Sitter spacetime of which the original static patch is a portion. Here, τ is global de Sitter time, or time associated with closed slicing. This implies that $\|\{n_\alpha\}_A\gg$ should be identified as the state of the other hemisphere of the static patch at $\tau = 0$. States in which modes in the polar region are excited or deexcited are those obtained by acting corresponding creation operators Eq. (2.3.50) or

⁴² A specific construction can go as follow. We can first construct zone mode operators b_γ and $b_\gamma^\dagger$ using the HKLL procedure [33, 34, 126] or its extension [37]. We can then expand the full state $|\Psi(t_*)\rangle$ at $t = t_*$ in terms of the eigenstates of number operators $b_\alpha^\dagger b_\alpha$ ’s and identify the $\|\{n_\alpha\}\gg$ states associated with $|\Psi(t_*)\rangle$. This allows us to define $\tilde{b}_\gamma$ and $\tilde{b}_\gamma^\dagger$ operators acting on $\|\{n_\alpha\}\gg$ ’s, and hence infalling operators a_ξ and $a_\xi^\dagger$ through Eqs. (2.5.23) and (2.5.24). The effective theory of the interior (erected at $t = t_*$) can then be obtained using the generator $\tilde{H}$ of time evolution, given in Eq. (2.5.27).

	Evaporating black hole	Cosmological de Sitter spacetime
microscopic level	zone region	polar region
	stretched horizon	stretched horizon
	far region	—
effective theory	two-sided black hole	global de Sitter spacetime
	the second exterior	the other hemisphere

Table 2.1: Correspondence between an evaporating black hole and cosmological de Sitter spacetime.

annihilation operators Eq. (2.3.49), respectively, to a de Sitter vacuum microstate.⁴³ These states are naturally mapped to excited states in the effective theory built on Eq. (2.5.66).

As in the case of a black hole, we can introduce annihilation and creation operators for modes in the other hemisphere in the effective theory as

$$\tilde{b}_\gamma^A = \sum_n \sqrt{n_\gamma} \varsigma_{n_\gamma}^A \varsigma_n^{A*} \sum_{i_{n_\gamma}=1}^{e^{S_{\text{dS}}(E-E_{n_\gamma})}} \sum_{j_n=1}^{e^{S_{\text{dS}}(E-E_n)}} c_{n_\gamma i_{n_\gamma}}^A c_{n j_n}^{A*} |\psi_{i_{n_\gamma}}^{(n_\gamma)}\rangle \langle \psi_{j_n}^{(n)}|, \quad (2.5.67)$$

$$\tilde{b}_\gamma^{A\dagger} = \sum_n \sqrt{n_\gamma + 1} \varsigma_{n_\gamma}^A \varsigma_n^{A*} \sum_{i_{n_\gamma}=1}^{e^{S_{\text{dS}}(E-E_{n_\gamma})}} \sum_{j_n=1}^{e^{S_{\text{dS}}(E-E_n)}} c_{n_\gamma i_{n_\gamma}}^A c_{n j_n}^{A*} |\psi_{i_{n_\gamma}}^{(n_\gamma)}\rangle \langle \psi_{j_n}^{(n)}|, \quad (2.5.68)$$

which can be used to form annihilation and creation operators for global time slicing

$$a_\xi^A = \sum_\gamma (\alpha_{\xi\gamma} b_\gamma + \beta_{\xi\gamma} b_\gamma^\dagger + \zeta_{\xi\gamma} \tilde{b}_\gamma^A + \eta_{\xi\gamma} \tilde{b}_\gamma^{A\dagger}), \quad (2.5.69)$$

$$a_\xi^{A\dagger} = \sum_\gamma (\beta_{\xi\gamma}^* b_\gamma + \alpha_{\xi\gamma}^* b_\gamma^\dagger + \eta_{\xi\gamma}^* \tilde{b}_\gamma^A + \zeta_{\xi\gamma}^* \tilde{b}_\gamma^{A\dagger}), \quad (2.5.70)$$

where ξ is the label in which the frequency ω with respect to static time t is traded with the frequency Ω associated with global time τ , and $\alpha_{\xi\gamma}$, $\beta_{\xi\gamma}$, $\zeta_{\xi\gamma}$, and $\eta_{\xi\gamma}$ are the Bogoliubov coefficients (which, of course, differ from the black hole case). The correspondence between the black hole and de Sitter cases is summarized in Table 2.1. Note that for de Sitter spacetime, there is no region corresponding to the region outside the zone of an evaporating black hole.

Physical quantities in global de Sitter spacetime can be calculated using the global time evolution operator

$$U(\tau) = e^{-i\tilde{H}\tau}, \quad (2.5.71)$$

where

$$\tilde{H} = \sum_\xi \Omega_\xi a_\xi^{A\dagger} a_\xi^A + \tilde{H}_{\text{int}}(\{a_\xi^A\}, \{a_\xi^{A\dagger}\}). \quad (2.5.72)$$

⁴³ Deexcited states become relevant if Gibbons-Hawking radiation is extracted by a material.

In the Heisenberg picture, this can be done by evolving quantum fields $\tilde{\Phi}_a(\mathbf{x}, 0)$ formed from a_ξ^A and $a_\xi^{A\dagger}$ at $\tau = 0$ as

$$\tilde{\Phi}_a(\mathbf{x}, \tau) = U(\tau)^\dagger \tilde{\Phi}_a(\mathbf{x}, 0) U(\tau) \quad (2.5.73)$$

and sandwiching their products by the state $|\Psi(\tau = 0)\rangle$ of the effective theory at $\tau = 0$ obtained by matching with the state of the microscopic theory at $t = 0$. ($|\Psi(\tau = 0)\rangle = |\Psi_A(E)\rangle$ if the system is in the semiclassical vacuum.) In the Schrödinger picture, the state must be evolved with $U(\tau)$ while $\tilde{\Phi}_a(\mathbf{x}, 0)$ are inserted at intermediate stages of the evolution.

The promotion of operators to a less state-dependent form can be made in a similar way to the back hole case, Eqs. (2.5.35), (2.5.39) and (2.5.40). The analysis of errors of operator algebras can be performed similarly. In particular, the global description in the effective theory has an intrinsic error of order

$$\epsilon = \max \left\{ \frac{e^{\frac{E_{\max}}{2T_H}}}{e^{\frac{1}{2}S_{\text{GH}}}}, \frac{e^{\frac{E_{\max}}{2T_H} + S_{\text{eff}}}}{e^{S_{\text{GH}}}} \right\}, \quad (2.5.74)$$

where S_{GH} is the Gibbons-Hawking entropy, $E_{\max}$ is the maximum energy that an excitation can carry in the effective theory, and S_{eff} represents the size of the microscopic Hilbert space covered by the promoted operators, as defined in Eq. (2.5.34).

Like the case of a black hole, the effective theory of global de Sitter spacetime is intrinsically semiclassical in that the algebra of operators in the theory is defined only up to an uncertainty of $e^{-S_{\text{GH}}/2}$. This is consistent with the observation in Ref. [75] that symmetries of classical de Sitter spacetime cannot be implemented exactly in a finite-dimensional Hilbert space.

Consistency with the semiclassical expectation

Suppose an object hits the stretched horizon in the static patch description at some time $t = t_{\text{obj}}$. The state shortly after it will not take the form of Eq. (2.3.34) with random $c_{n_{in}}^A$ which leads to the vacuum state of Eq. (2.5.66). Instead, it stays in an excited state for a while, reflecting the existence of an object just outside the horizon. In the effective theory, this is represented as the existence of an excitation in the other hemisphere at $\tau = 0$. Such an excitation has a physical significance, since we can retrieve (a part of) information about it if the system leaves the de Sitter phase after $t = t_{\text{obj}}$, e.g. by tunneling into a Minkowski vacuum; see Fig. 2.14.

According to the static patch description, the system is expected to relax into an equilibrium state of the form of Eq. (2.3.34) when one scrambling time⁴⁴

$$t_{\text{scr}} = \alpha \left[\ln \frac{\alpha}{l_{\text{P}}} + O(1) \right] \quad (2.5.77)$$

⁴⁴ To obtain this expression, we can consider an analogue of ingoing Eddington-Finkelstein time in de Sitter spacetime, $v = t - r_*$, and use the general expression for scrambling time to obtain

$$v_{\text{scr}} = \frac{1}{2\pi T_H} \ln S_{\text{dS}} \approx 2\alpha \ln \frac{\alpha}{l_{\text{P}}}. \quad (2.5.75)$$

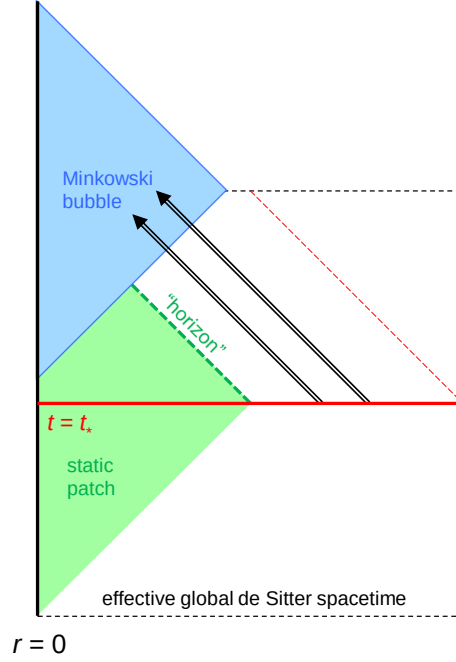


Figure 2.14: Information outside the static patch in the effective theory can be retrieved if the system leaves the de Sitter phase at a later time, for example by tunneling into a Minkowski bubble universe.

has passed after the object hits the stretched horizon. This implies that if we erect the effective theory at $t_* > t_{\text{obj}} + t_{\text{scr}}$, then the effective theory sees the vacuum in the other hemisphere. Is this consistent with the semiclassical expectation?

In Fig. 2.15, we depict the Penrose diagram of the situation in which an object hits the stretched horizon at t_{obj} , and a Minkowski bubble is nucleated at the location of the observer, $r = 0$, at $t = t_{\text{obj}} + \Delta t$. We expect that if $\Delta t > t_{\text{scr}}$, the observer cannot receive any information from the object, since then the effective theory erected at $t = t_{\text{obj}} + \Delta t$ will not have an excitation in the other hemisphere of the effective global de Sitter spacetime, and so there is nothing that can send a signal to the observer. From the viewpoint of the object, this implies that it cannot send any signal to the observer after it crosses the $t = t_{\text{obj}} + \Delta t$ hypersurface.

This is indeed consistent with what semiclassical theory implies. We can show that in order for the signal to reach the Minkowski bubble, and hence the observer, the object must

The scrambling time in static time is then

$$t_{\text{scr}} = v_{\text{scr}} - r_{*s} \approx \alpha \ln \frac{r_+}{l_P}, \quad (2.5.76)$$

where we have used $r_{*s} \approx \alpha \ln(r_+/l_P)$, obtained from Eq. (2.2.28) by identifying l_s with l_P .

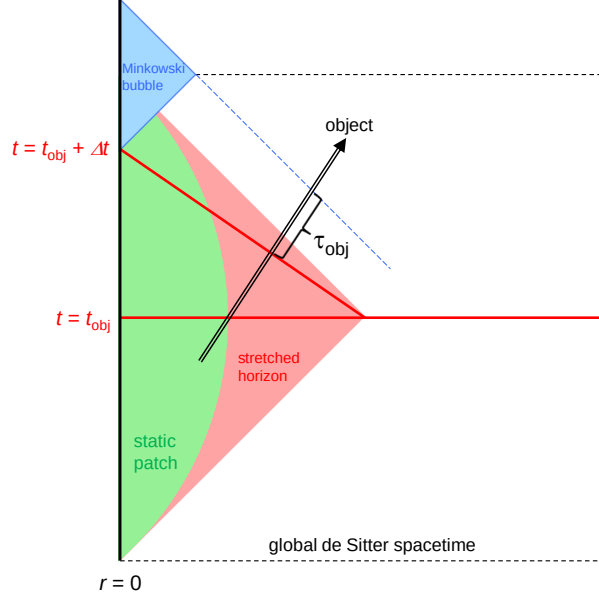


Figure 2.15: An object falling into the stretched horizon at $t = t_{\text{obj}}$ can try to send a physical signal to the observer at $r = 0$ who enters a Minkowski bubble nucleated at $r = 0$ at $t = t_{\text{obj}} + \Delta t$. The holographic static-patch description of de Sitter spacetime suggests that this is not possible if Δt is larger than the scrambling time t_{scr} . This expectation is consistent with the semiclassical picture, since for $\Delta t > t_{\text{scr}}$ the object would have to send the signal within proper time $\tau_{\text{obj}} \lesssim l_{\text{P}}$, which is not possible.

send it within the proper time

$$\tau_{\text{obj}} \lesssim l_{\text{P}} \quad (2.5.78)$$

after it crosses the $t = t_{\text{obj}} + \Delta t$ hypersurface, but this is not possible because the holographic principle does not allow for the object to send physical information within l_{P} . Incidentally, this conclusion does not change even if the object is accelerated toward $r = 0$ after it crosses the stretched horizon, as long as the invariant acceleration is smaller than $O(1/l_{\text{s}})$.⁴⁵

Schwarzschild de Sitter spacetime

Our analysis applies straightforwardly to a Schwarzschild de Sitter spacetime which has a black hole of radius r_{bh} (centered at $r = 0$) in de Sitter spacetime of radius r_{dS} , as long as r_{bh} is sufficiently smaller than r_{dS} . (The case in which r_{bh} and r_{dS} are comparable is analogous

⁴⁵ Note that here we have considered the case in which the system was in a meta-stable de Sitter phase before entering a Minkowski phase. We expect that the situation is different if the earlier, de Sitter phase is only approximate in that there is noticeable time dependence of background spacetime, as in the case of slow-roll inflation. A different analysis will be needed for such cases.

to the case of a near extremal black hole, which we do not consider in this paper.)

Specifically, if $r_{\text{bh}} \ll r_{\text{dS}}$, the black hole can be viewed as a small excitation for the purpose of erecting the effective theory outside the de Sitter horizon, so that the construction described so far goes through without significant modifications. The situation is similar for the effective theory of the black hole interior. A notable thing is that if the black hole is old enough that it is maximally entangled and that the majority of emitted Hawking radiation has reached the stretched de Sitter horizon, then operators describing the black hole interior must involve degrees of freedom associated with the de Sitter horizon.

2.6 Intrinsically Extended Spacetimes

So far, we have been considering “realistic,” single-sided systems and seen how the effective two-sided pictures emerge as collective phenomena associated with the stretched horizon degrees of freedom. However, there is nothing theoretically wrong in considering “intrinsically two-sided” systems which involve two copies of holographic theories discussed so far. In fact, such systems have been considered in various contexts in black hole [25, 170–180] and de Sitter [61–66, 181–186] physics.

In this section, we discuss how such intrinsically two-sided systems can be understood in the framework described so far. In particular, we will see how these and single-sided systems lead to similar semiclassical physics, despite the fact that states in the two cases have rather different structures at the microscopic level. While we focus here on the two-sided case for a black hole, we expect it is relatively straightforward to extend it to black holes with more than two exterior regions [187].

2.6.1 Two-sided eternal black hole

Let us consider a static, two-sided eternal black hole. Strictly speaking, for a finite black hole mass, this system exists only as a large AdS black hole, but we can imagine that a two-sided black hole in asymptotically flat spacetime also behaves in approximately the same manner if it is sufficiently large compared with the scale that we are interested in.

We specifically consider a thermofield double state which is prepared by the Euclidean path integral over a half of the time compactified on a circle of length β [170]. In the context of the AdS/CFT correspondence, this is a thermally entangled state of two CFTs. Similar to the case of a single-sided black hole, we divide modes in each side of the black hole into zone and horizon modes at the state preparation time, which we refer to as $t = 0$.⁴⁶ We then have four classes of modes, i.e. zone and horizon modes in both (right and left) sides of the black hole.

⁴⁶ For a large AdS black hole, we define zone modes to be all the modes other than horizon modes defined as in Section 2.3.2.

Assuming that the two sides of the black hole have the same microscopic structure (which corresponds to the case that the two CFTs are the same), the black hole vacuum state at $t = 0$ as viewed from the exterior is given by

$$|\Psi(M)\rangle = \frac{1}{\sqrt{\sum_{\mu'} e^{-\beta E_{\mu'}}}} \sum_{\mu} e^{-\frac{\beta}{2} E_{\mu}} |\mu\rangle_{\text{R}} |\mu\rangle_{\text{L}}, \quad (2.6.1)$$

where $|\mu\rangle_{\text{R}}$ represent energy eigenstates of the holographic theory describing the right side with energy E_{μ} and similarly for $|\mu\rangle_{\text{L}}$. The mass M of the black hole is related to β by a smoothness condition for spacetime [188]:

$$M = \frac{r_+}{2l_{\text{p}}^2} \left(1 + \frac{r_+^2}{L^2}\right), \quad \beta = \frac{4\pi r_+ L^2}{3r_+^2 + L^2}, \quad (2.6.2)$$

where L is the AdS length. In the context of holography, this relation can be viewed as arising from the requirement that the boundary spacetime can be smoothly extended to the bulk spacetime [189].

We note that the state $|\Psi(M)\rangle$ has exactly zero energy under the two-sided (modular) Hamiltonian $H = H_{\text{R}} - H_{\text{L}}$, which is possible because the system is static under time evolution generated by H .⁴⁷ This time evolution corresponds to Schwarzschild time evolution in a single-sided black hole; see Eq. (2.5.15).

In general, the energy eigenstates $|\mu\rangle_{\text{R}}$ do not agree with the number eigenstates of the zone and horizon modes for the right side (and similarly for the left side). However, to describe the dynamics of a zone mode that has energy $E \gtrsim 1/\beta = T_{\text{H}}$, this issue can be ignored. Consider, for example, a zone mode of energy $\sim E$ localized at $\bar{r}_*$ in the zone and having a Gaussian tale of the form $\sim e^{-O(1)E^2(r_* - \bar{r}_*)^2}$. The correction arising from the existence of the stretched horizon to the quantum theory of matter on classical black hole background is then suppressed by the exponential of $E^2(\bar{r}_* - r_{*,\text{s}})^2$. This implies that the effect of the deviation of zone-mode states from the thermofield double form coming from their interactions with the stretched horizon is negligible compared with the factor of $e^{-\beta E/2}$ if the mode is sufficiently hard ($E \gg 1/\beta$) or away from the stretched horizon ($\bar{r}_* - r_{*,\text{s}} \gg \beta$).

We now take a subset of the zone modes on the right side as right-side hard modes, whose state we denote by $|n\rangle_{\text{R}}$. Here, $n = \{n_{\alpha}\}$ represents the set of occupation numbers of these modes. The corresponding zone modes on the left side are then left-side hard modes, whose state is $|n\rangle_{\text{L}}$. The rest of the modes on the right and left sides are right- and left-side soft modes, respectively, whose states we denote by $|\psi_i\rangle_{\text{R}}$ and $|\psi_i\rangle_{\text{L}}$. Here, the index i runs over all soft-mode states, not just those in a specific energy window, since we no longer impose the energy constraint for modes on one side.

The state in Eq. (2.6.1) can then be approximated, for the purpose of describing the

⁴⁷ We can extend the state in Eq. (2.6.1) to the case in which two CFTs are different. In this case, the state must have some, though exponentially small, energy uncertainty, reflecting the fact that the spectra of the two CFTs are not identical.

dynamics of hard modes, as

$$|\Psi(M)\rangle \approx \frac{1}{\sqrt{Z}} \sum_n \sum_i e^{-\frac{\beta}{2}(E_n+E_i)} |n\rangle_{\text{R}} |\psi_i\rangle_{\text{R}} |n\rangle_{\text{L}} |\psi_i\rangle_{\text{L}}, \quad (2.6.3)$$

where E_n and E_i are the energies carried by $|n\rangle_{\text{R}}$ and $|\psi_i\rangle_{\text{R}}$ defined with respect to H (or H_{R}), and

$$Z = \sum_n \sum_i e^{-\beta(E_n+E_i)}. \quad (2.6.4)$$

We can then build an effective theory of the interior on the vacuum microstate of Eq. (2.6.3).⁴⁸

It is important to realize that the microscopic entanglement structure of $|\Psi(M)\rangle$ in Eq. (2.6.3) is different from that of a microstate of a single-sided black hole, which takes the form of Eq. (2.3.28) (ignoring the entanglement with far modes). In particular, in Eq. (2.6.3), the hard modes on the right side, which we identify as “our” side, are entangled directly with those on the left side, while in Eq. (2.3.28) they are entangled with the soft modes on the same side.

The way in which the thermal nature is introduced is also different in two setups. In Eq. (2.6.3), the Boltzmann factors are introduced already at the microscopic level, while in Eq. (2.3.28) the coefficients take random values, and the Boltzmann factors for the hard modes arise only statistically after tracing out the soft modes; see Eq. (2.3.37). In the case of the two-sided black hole, tracing out the left side gives

$$\rho_{\text{R}} = \frac{1}{Z} \sum_n \sum_i e^{-\beta(E_n+E_i)} |n\rangle_{\text{R}} |\psi_i\rangle_{\text{R}} \langle n|_{\text{R}} \langle \psi_i|. \quad (2.6.5)$$

Further tracing out the soft modes lead to

$$\rho_{\text{R,hard}} = \frac{1}{z} \sum_n e^{-\beta E_n} |n\rangle_{\text{R}} \langle n|, \quad (2.6.6)$$

where $z = \sum_n e^{-\beta E_n}$. At the microscopic level, the two-sided black hole is a model of a single-sided black hole only in the sense that Eq. (2.6.6) takes the same form as Eq. (2.3.37).

Excited states in which there are objects in the right exterior (our side) of the black hole are obtained by acting with the annihilation/creation operators

$$b_{\text{R}\gamma} = \sum_n \sqrt{n_\gamma} |\{n_\alpha - \delta_{\alpha\gamma}\}_{\text{R}} \langle \{n_\alpha\}|, \quad (2.6.7)$$

$$b_{\text{R}\gamma}^\dagger = \sum_n \sqrt{n_\gamma + 1} |\{n_\alpha + \delta_{\alpha\gamma}\}_{\text{R}} \langle \{n_\alpha\}| \quad (2.6.8)$$

on the vacuum state in Eq. (2.6.3). As in the case of a single-sided black hole, describing the interior requires time evolution operator other than that generated by $H = H_{\text{R}} - H_{\text{L}}$. The

⁴⁸ In this particular case, we can take the hard modes to be the entire zone modes, though this is not the case in general.

appropriate generator $\tilde{H}$ can be constructed using annihilation/creation operators acting on left-side hard modes

$$b_{L\gamma} = \sum_n \sqrt{n_\gamma} |\{n_\alpha - \delta_{\alpha\gamma}\}\rangle_L |\{n_\alpha\}\rangle, \quad (2.6.9)$$

$$b_{L\gamma}^\dagger = \sum_n \sqrt{n_\gamma + 1} |\{n_\alpha + \delta_{\alpha\gamma}\}\rangle_L |\{n_\alpha\}\rangle \quad (2.6.10)$$

as

$$\tilde{H} = \sum_\xi \Omega_\xi a_\xi^\dagger a_\xi + \tilde{H}_{\text{int}}(\{a_\xi\}, \{a_\xi^\dagger\}), \quad (2.6.11)$$

where

$$a_\xi = \sum_\gamma (\alpha_{\xi\gamma} b_{R\gamma} + \beta_{\xi\gamma} b_{R\gamma}^\dagger + \zeta_{\xi\gamma} b_{L\gamma} + \eta_{\xi\gamma} b_{L\gamma}^\dagger), \quad (2.6.12)$$

$$a_\xi^\dagger = \sum_\gamma (\beta_{\xi\gamma}^* b_{R\gamma} + \alpha_{\xi\gamma}^* b_{R\gamma}^\dagger + \eta_{\xi\gamma}^* b_{L\gamma} + \zeta_{\xi\gamma}^* b_{L\gamma}^\dagger) \quad (2.6.13)$$

are infalling mode operators with the coefficients $\alpha_{\xi\gamma}$, $\beta_{\xi\gamma}$, $\zeta_{\xi\gamma}$, and $\eta_{\xi\gamma}$ taking the same values as those in Eqs. (2.5.23) and (2.5.24) for near horizon modes. The reason why the construction here need not involve soft modes is that the black hole vacuum state given in Eq. (2.6.3), in fact, factors into hard-mode and soft-mode parts:

$$|\Psi(M)\rangle \approx \left(\frac{1}{\sqrt{z}} \sum_n e^{-\frac{\beta}{2} E_n} |n\rangle_R |n\rangle_L \right) \left(\frac{1}{\sqrt{\sum_j e^{-\beta E_j}}} \sum_i e^{-\frac{\beta}{2} E_i} |\psi_i\rangle_R |\psi_i\rangle_L \right), \quad (2.6.14)$$

so that the soft-mode piece can simply be ignored. Note that unlike the construction in Refs. [178, 179], we do not claim that the generator $\tilde{H}$ of infalling time evolution can be represented purely on the right operators.⁴⁹ The evolution with $\tilde{H}$ in Eq. (2.6.11) allows us to describe the semiclassical physics in the domain of dependence of the union of zones on both sides at $t = 0$.

This erection of the effective theory without involving soft modes is essentially nothing other than the construction of the semiclassical theory from the beginning. The unknown UV physics, including the effect of ignoring horizon modes, is reflected in the choice of the vacuum in the effective theory. For example, for a large AdS black hole in thermal equilibrium with the ambient AdS spacetime, the correct choice for typical soft-mode states is the Hartle-Hawking vacuum. This choice, however, cannot be derived from the low energy consideration alone.

We note that our starting point of Eq. (2.6.1) involves a choice of time. Suppose we

⁴⁹ In the large N limit construction of Refs. [178, 179], operators $b_{R\gamma}$, $b_{R\gamma}^\dagger$, $b_{L\gamma}$, and $b_{L\gamma}^\dagger$ generate a type III₁ von Neumann algebra (which becomes type II_∞ at the next order in $1/N$ [180]), so that Hilbert spaces for $|n\rangle_R$, $|n\rangle_L$, $|\psi_i\rangle_R$, and $|\psi_i\rangle_L$ are not defined. This is not the case for finite N as envisioned here.

evolve the state $|\Psi(M)\rangle$ under the full, microscopic Hamiltonian $H_R + H_L$:⁵⁰

$$|\Psi(M)\rangle \rightarrow e^{-i(H_R+H_L)t}|\Psi(M)\rangle = \frac{1}{\sqrt{\sum_{\mu'} e^{-\beta E_{\mu'}}}} \sum_{\mu} e^{-\left(\frac{\beta}{2}+2it\right)E_{\mu}} |\mu\rangle_R |\mu\rangle_L. \quad (2.6.15)$$

In general, this state is not equivalent to $|\Psi(M)\rangle$; the phases on the right-hand side have a physical meaning. For example, as t increases from 0, the maximal interior volume of a spatial surface anchored to the boundaries at time t grows, although for $t \gtrsim e^{S_{\text{BH}}}$ the mixing between the hard and soft modes becomes so important that the naive semiclassical picture of interior volume growth ceases to apply [190].

In the two-sided state in Eq. (2.6.3), the von Neumann entropy of the right side is given by

$$\begin{aligned} S_R &= -\text{Tr}[\rho_R \ln \rho_R] \\ &= \sum_n \sum_i \frac{e^{-\beta(E_n+E_i)}}{Z} [\beta(E_n + E_i) + \ln Z], \end{aligned} \quad (2.6.16)$$

where ρ_R is given by Eq. (2.6.5). Approximating that the density of states of the soft modes does not depend on E_i and denoting it by $S_{\text{soft}}(M)$, we have up to $O(1)$ coefficients

$$\sum_i \rightarrow \beta \int dE_i e^{S_{\text{soft}}(M)}, \quad (2.6.17)$$

where β is used to match the dimension. This gives

$$Z \sim \sum_n e^{S_{\text{soft}}(M)-\beta E_n} = z e^{S_{\text{soft}}(M)}, \quad (2.6.18)$$

and

$$S_R = \ln[z e^{S_{\text{soft}}(M)}] + \beta \langle E_n \rangle_R + O(1). \quad (2.6.19)$$

Here,

$$\langle E_n \rangle_R = \frac{1}{z} \sum_n E_n e^{-\beta E_n} \quad (2.6.20)$$

is the thermal energy of the hard modes in the right exterior, as measured by H . Thus, up to the thermal contribution of the hard modes (and an unimportant $O(1)$ term), the entropy S_R agrees with S_{sys} given by Eqs. (2.3.30) and (2.3.31), which is the entropy of a single-sided black hole. Namely, for a two-sided black hole in the state of Eq. (2.6.3), the entropy of the black hole is given at the leading order by the entanglement entropy between the two sides (i.e. two CFTs in the AdS/CFT context).

⁵⁰ It is important that this evolution is performed using the full, microscopic Hamiltonian (or equivalently performed in the boundary theory), since the analogous operator $H_R + H_L$ in the bulk field theory is singular at the bifurcation surface.

Evaporation and the destruction of the wormhole

Let us now consider coupling the two-sided black hole discussed so far to a reservoir so that the black hole can radiate into it. We do this on the right side (i.e. our side) of the black hole and see its effect on an object falling into the black hole from the same side.

Before coupling the two systems, the state is given by the product of the black hole state in Eq. (2.6.3) and the ground state of the reservoir system $|\phi_0\rangle$:

$$|\Psi_0\rangle = \frac{1}{\sqrt{Z}} \sum_m \sum_j e^{-\frac{\beta}{2}(E_m + E_j)} |m\rangle_{\text{R}} |\psi_j\rangle_{\text{R}} |m\rangle_{\text{L}} |\psi_j\rangle_{\text{L}} |\phi_0\rangle, \quad (2.6.21)$$

where Z is given by Eq. (2.6.4). After the coupling of the two systems, this state is no longer an energy eigenstate, so it starts evolving. Specifically, the coupling injects positive energy shock waves into the two systems, after which the Hawking emission process begins. This process backreacts and produces a superposition of black holes having different masses and momenta (at different locations) in the right-side space. As discussed in Section 2.3.1, we focus on a branch in which the black hole has a well-defined mass and (vanishing) momentum, within the minimum uncertainty required by quantum mechanics. We can then say that after time t is passed, the state for each (m, j) changes as⁵¹

$$|m\rangle_{\text{R}} |\psi_j\rangle_{\text{R}} |\phi_0\rangle \xrightarrow[\text{+ projection}]{\text{time evolution}} \sum_n e^{S_{\text{bh}}(M_{\text{R}} - E_n)} \sum_{i_n=1} \sum_a c_{ni_na}^{mj} |n\rangle_{\text{R}} |\psi_{i_n}^{(n)}\rangle_{\text{R}} |\phi_a^{(m,j)}\rangle, \quad (2.6.22)$$

where

$$\sum_n e^{S_{\text{bh}}(M_{\text{R}} - E_n)} \sum_{i_n=1} \sum_a |c_{ni_na}^{mj}|^2 = 1, \quad (2.6.23)$$

M_{R} is the mass of the black hole as viewed from the right side, and $|\phi_a^{(m,j)}\rangle$ represents the state of the reservoir. Here, $|n\rangle_{\text{R}} |\psi_{i_n}^{(n)}\rangle_{\text{R}}$ and $|\phi_a^{(m,j)}\rangle$ carry energies M_{R} and $E_m + E_j - M_{\text{R}}$, respectively (within the uncertainty), and we have assumed that t is not too large that the structure of the black hole is dramatically altered (for example, that it is fully evaporated). The index of $|\psi_{i_n}^{(n)}\rangle_{\text{R}}$ now carries subindex n because the energy of $|\psi_{i_n}^{(n)}\rangle_{\text{R}}$ is correlated with that of $|n\rangle_{\text{R}}$.

Substituting Eq. (2.6.22) into Eq. (2.6.21), we obtain the state at time t

$$|\Psi_t\rangle = \frac{1}{\sqrt{Z}} \sum_n \sum_m e^{S_{\text{bh}}(M_{\text{R}} - E_n)} \sum_{i_n=1} \sum_j \sum_a e^{-\frac{\beta}{2}(E_m + E_j)} c_{ni_na}^{mj} |n\rangle_{\text{R}} |\psi_{i_n}^{(n)}\rangle_{\text{R}} |m\rangle_{\text{L}} |\psi_j\rangle_{\text{L}} |\phi_a^{(m,j)}\rangle. \quad (2.6.24)$$

We assume that for $t > t_{\text{scr}}$, the coefficients $c_{ni_na}^{mj}$ take Gaussian random values across (n, i_n, a) for each (m, j) , as in Eq. (2.3.42), although this assumption is less justified than that for a

⁵¹ We regard the left states $|m\rangle_{\text{L}} |\psi_j\rangle_{\text{L}}$ as (approximate) eigenstates of the left boundary Hamiltonian so that they do not evolve.

single-sided black hole. Now, consider the (non-normalized) states multiplying $|n\rangle_{\text{R}}$ in the sum over n in Eq. (2.6.24).

$$\|n\rangle_{\text{R},nn} = \frac{1}{\sqrt{Z}} \sum_m \sum_{i_n=1}^{e^{S_{\text{bh}}(M_{\text{R}}-E_n)}} \sum_j \sum_a e^{-\frac{\beta}{2}(E_m+E_j)} c_{ni_na}^{mj} |\psi_{i_n}^{(n)}\rangle_{\text{R}} |m\rangle_{\text{L}} |\psi_j\rangle_{\text{L}} |\phi_a^{(m,j)}\rangle. \quad (2.6.25)$$

With the assumption stated above, the norms of these states are given by

$$\begin{aligned} {}_{\text{R},nn}\langle n \| n \rangle_{\text{R},nn} &= \frac{1}{Z} \sum_m \sum_{i_n=1}^{e^{S_{\text{bh}}(M_{\text{R}}-E_n)}} \sum_j \sum_a e^{-\beta(E_m+E_j)} |c_{ni_na}^{mj}|^2 \\ &= \frac{1}{(\sum_n e^{-\beta_{\text{R}} E_n})} e^{-\beta_{\text{R}} E_n}, \end{aligned} \quad (2.6.26)$$

up to corrections exponentially suppressed in $S_{\text{bh}}(M_{\text{R}})$. Here, in the last line we have used

$$\langle |c_{ni_na}^{mj}|^2 \rangle = \frac{1}{\sum_n \sum_a e^{S_{\text{bh}}(M_{\text{R}}-E_n)}} \quad (2.6.27)$$

$$\implies \sum_{i_n=1}^{e^{S_{\text{bh}}(M_{\text{R}}-E_n)}} \sum_a |c_{ni_na}^{mj}|^2 = \frac{e^{S_{\text{bh}}(M_{\text{R}}-E_n)}}{(\sum_n e^{S_{\text{bh}}(M_{\text{R}}-E_n)})} = \frac{1}{(\sum_n e^{-\beta_{\text{R}} E_n})} e^{-\beta_{\text{R}} E_n}, \quad (2.6.28)$$

and

$$\beta_{\text{R}} = \left. \frac{\partial S_{\text{bh}}(E)}{\partial E} \right|_{E=M_{\text{R}}} \quad (2.6.29)$$

is the temperature of the black hole as viewed from the right side, which is different from β . The state $\|n\rangle_{\text{R},nn}$ is thus related to the corresponding normalized state $\|n\rangle_{\text{R}}$ by

$$\|n\rangle_{\text{R},nn} = \frac{1}{\sqrt{\sum_n e^{-\beta_{\text{R}} E_n}}} e^{-\frac{\beta_{\text{R}}}{2} E_n} \|n\rangle_{\text{R}}. \quad (2.6.30)$$

From Eqs (2.6.25) and (2.6.30), we find that the state in Eq. (2.6.24) can be written in the thermofield double form

$$|\Psi_t\rangle = \frac{1}{\sqrt{\sum_n e^{-\beta_{\text{R}} E_n}}} \sum_n e^{-\frac{\beta_{\text{R}}}{2} E_n} |n\rangle_{\text{R}} \|n\rangle_{\text{R}}. \quad (2.6.31)$$

This allows us to erect the effective theory of the interior, following the construction described in Section 2.5. In particular, it implies that an object located in the zone of the right side at time t will smoothly pass through the horizon (of the black hole of mass M_{R}) from the right side.⁵²

⁵² We can also erect an effective theory as viewed from the left exterior, i.e. the opposite to the side from which the black hole evaporates. The construction is analogous to that described here—we identify $\|m\rangle_{\text{L}}$ as the state that comes with $|m\rangle_{\text{L}}$, leading to the state in the effective theory

$$|\Psi'_t\rangle = \frac{1}{\sqrt{\sum_n e^{-\beta E_n}}} \sum_n e^{-\frac{\beta}{2} E_n} |n\rangle_{\text{L}} \|n\rangle_{\text{L}}, \quad (2.6.32)$$

where $\beta = \partial S_{\text{bh}}(E)/\partial E|_{E=M}$. The analysis presented below also applies to this case with similar conclusions.

It is interesting to consider what happens if the initial state has an excitation in the second exterior, i.e. the exterior on the left side. Suppose that the state at the time of the coupling, $t = 0$, has an excitation in the zone, e.g.,

$$|\Psi_{\text{init}}\rangle = \prod_{i=1}^N \left(\sum_{\gamma} f_{\gamma}^{(i)} b_{L\gamma}^{\dagger} \right) |\Psi_0\rangle, \quad (2.6.33)$$

where $|\Psi_0\rangle$ is given by Eq. (2.6.21). In this case, the state after $t > t_{\text{scr}}$ is given by Eq. (2.6.24) with the replacement

$$e^{-\frac{\beta}{2}E_m} c_{ni_na}^{mj} |\phi_a^{(m,j)}\rangle \rightarrow \sum_p e^{-\frac{\beta}{2}E_p} U_p^m c_{ni_na}^{pj} |\phi_a^{(p,j)}\rangle, \quad (2.6.34)$$

where U_p^m is a unitary matrix which has the indices $p = \{p_{\alpha}\}$ and $m = \{m_{\alpha}\}$ and depends on $f_{\gamma}^{(i)}$. Similarly, if there is an excitation on the left side that fell into the stretched horizon at an earlier time, then the state of the left-side soft modes at $t = 0$ deviates from that in Eq. (2.6.21), causing the change of the state at time t

$$e^{-\frac{\beta}{2}E_j} c_{ni_na}^{mj} |\phi_a^{(m,j)}\rangle \rightarrow \sum_k e^{-\frac{\beta}{2}E_k} V_k^j c_{ni_na}^{mk} |\phi_a^{(m,k)}\rangle, \quad (2.6.35)$$

where V_k^j is a unitary matrix acting on the space of the left-side soft states. With the changes in Eqs. (2.6.34) and (2.6.35), we can define $\|n\|_{\text{R},nn}$ analogously to Eq. (2.6.25) and calculate its norm

$$\begin{aligned} \text{R},nn \langle\langle n \| n \rangle\rangle_{\text{R},nn} &= \frac{1}{Z} \sum_m e^{S_{\text{bh}}(M_{\text{R}} - E_n)} \sum_{i_n=1} \sum_j \sum_{p,p'} \sum_{k,k'} \sum_{a,a'} \left\{ e^{-\frac{\beta}{2}(E_p + E_{p'} + E_k + E_{k'})} U_p^{m*} U_{p'}^m V_k^{j*} V_{k'}^j \right. \\ &\quad \left. \times c_{ni_na}^{pk*} c_{ni_na'}^{p'k'} \langle \phi_a^{(p,k)} | \phi_{a'}^{(p',k')} \rangle \right\} \end{aligned} \quad (2.6.36)$$

$$= \frac{1}{(\sum_n e^{-\beta_{\text{R}} E_n})} e^{-\beta_{\text{R}} E_n}, \quad (2.6.37)$$

which we find is the same as Eq. (2.6.26) up to exponentially suppressed corrections. Here, we have assumed that there is no intricate (and unexpected) cancellation between $V_k^{j*} V_{k'}^j$ and $c_{ni_na}^{pk*} c_{ni_na'}^{p'k'}$ in the sum over (k, k') , which would jeopardize the scaling in the last line.⁵³

With this assumption, we thus find that the state of the effective theory is still given by Eq. (2.6.31). This implies that the existence of an excitation on the left exterior at $t = 0$ cannot affect the effective theory erected at $t > t_{\text{scr}}$ after evaporation began on the right side. In other words, the Einstein-Rosen bridge between the two sides is broken by Hawking radiation for $t > t_{\text{scr}}$, although an object falling into the black hole from either side sees smooth spacetime when it crosses the horizon.

At the technical level, this occurs because the energy constraint in Eq. (2.6.22) (i.e. the condition that the mass of the black hole as viewed from the right side is M_{R}) breaks

⁵³ A similar assumption is not needed for $U_p^{m*} U_{p'}^m$, since the number of degrees of freedom of semiclassical excitations is too small to jeopardize this scaling anyway.

entanglement between the right and left side modes necessary to have a bridge [42]. To see this, we can trace out the soft and far modes in the state of Eq. (2.6.24)

$$\text{Tr}_{\text{soft+far}} |\Psi_t\rangle\langle\Psi_t| = \frac{1}{(\sum_{m'} e^{-\beta_R E_{m'}})(\sum_{n'} e^{-\beta E_{n'}})} \sum_m \sum_n e^{-\beta_R E_m} e^{-\beta E_n} |m\rangle_{\text{R}} \langle n|_{\text{L}} \langle m|_{\text{L}} \langle n| \quad (2.6.38)$$

and find that it takes a different form than

$$\text{Tr}_{\text{soft+far}} |\Psi_0\rangle\langle\Psi_0| = \frac{1}{\sum_{n'} e^{-\beta E_{n'}}} \sum_m \sum_n e^{-\frac{\beta}{2}(E_m + E_n)} |m\rangle_{\text{R}} \langle m|_{\text{L}} \langle n|_{\text{L}} \langle n|, \quad (2.6.39)$$

obtained from the thermofield double state which has a connected Einstein-Rosen bridge. In particular, we see that the hard modes on the right side are mostly entangled with the right-side soft modes and far modes in $|\Psi_t\rangle$, and not with modes on the left side as in $|\Psi_0\rangle$.

Physically, this breaking of the Einstein-Rosen bridge occurs due to decoherence caused by Hawking emission from the right side. After the Hawking emission, black holes of different masses (as viewed from the right side) can be semiclassically discriminated, so that they should be viewed as living in different branches of the wavefunction. Thus, unless a falling object somehow preserves coherence among these different branches,⁵⁴ it cannot see any signal sent from the left exterior inside the black hole. In general, the Einstein-Rosen bridge—or wormhole—prepared by the thermofield double state in Eq. (2.6.3) is fragile under a realistic physical process occurring to it.⁵⁵

Entangled black holes

For a reason similar to the case above, entangled black holes of the form

$$|\Psi(M_1, M_2)\rangle = \sum_{A=1}^{e^{S_{\text{sys}}(M_1)}} \sum_{B=1}^{e^{S_{\text{sys}}(M_2)}} \eta_{AB} |\Psi_A(M_1)\rangle |\Psi_B(M_2)\rangle \quad (2.6.40)$$

generically do not have an Einstein-Rosen bridge connecting them; an object falling into one of the black holes sees a smooth horizon but cannot receive any signal sent in from the other black hole [42]. The same comment applies to more than two entangled black holes.

Interior holography?

We now discuss the possibility of formulating a “holographic theory of the interior.” In the discussion so far, the fundamental degrees of freedom are assumed to live on boundaries (e.g. holographic screens or AdS boundaries) outside the black hole horizon. However, discussion

⁵⁴ It is not clear to what extent this is possible because the falling object necessarily feels gravity generated by the black hole in question.

⁵⁵ If the decoherence is imperfect, the two sides can be “weakly connected,” e.g. connected in the interior but with a physical domain wall between the two horizons.

66], which we will address later.

2.6.2 Global de Sitter spacetime

We now discuss a possible description of global de Sitter spacetime analogous to a static, two-sided black hole. As in the case of a two-sided black hole in asymptotically flat spacetime, we consider it to be an approximate description of a sufficiently long-lived meta-stable de Sitter spacetime.

Specifically, we assume that global de Sitter spacetime at $t = 0$ can be described by a state of a thermofield double-like form between two holographic theories each of which describes a static patch. Following the case of a two-sided black hole, we may divide modes in each theory into “zone” and horizon modes. We then take a subset of the zone modes to be hard modes while leaving all the other modes as soft modes. In the case of de Sitter spacetime, however, we expect that gravity is not decoupling in the holographic theories (because the boundary is not in an asymptotic region; see Section 2.4.2), so that the two theories are interacting through it.⁵⁶ We thus denote by $|n\rangle_N$, $|\psi_i\rangle_N$, $|n\rangle_S$, and $|\psi_i\rangle_S$ the states of the hard and soft modes in the limit that gravity is turned off in these holographic theories. Here, the subscripts N and S specify the theory under consideration (referring to the north and south hemispheres, respectively).

Assuming that the two theories have the same microscopic structure, the relevant thermofield double state can be written as

$$|\Psi'(E)\rangle \approx \frac{1}{\sqrt{Z}} \sum_n \sum_i e^{-\frac{\beta}{2}(E_n + E_i)} |n\rangle_N |\psi_i\rangle_N |n\rangle_S |\psi_i\rangle_S, \quad Z = \sum_n \sum_i e^{-\beta(E_n + E_i)}, \quad (2.6.42)$$

in the limit that gravity is turned off in holographic theories. Here, $E = \alpha/l_P^2$, and E_n and E_i are the energies carried by the hard and soft modes in the north hemisphere defined with respect to the Hamiltonian $H = H_N - H_S$. We expect that the state in Eq. (2.6.42) can be prepared by the Euclidean path integral over a half of the time compactified on a circle of length β (in the limit that gravity is turned off); see Fig. 2.17(a). While the location of the boundary is now reversed compared with the case of a black hole, depicted in Fig. 2.17(b), we assume that the periodicity β of the time direction is still related to the temperature of de Sitter spacetime as $\beta = 1/T_H = 2\pi\alpha$, as in Ref. [188].

When gravity is turned on in holographic theories, the Euclidean path integral preparing the state is expected to receive extra contributions suppressed by the gravitational coupling which are not diagonal in the space spanned by $|n\rangle|\psi_i\rangle$. This gives the correction to the state of the form

$$|\Psi'(E)\rangle \longrightarrow |\Psi(E)\rangle = |\Psi'(E)\rangle + \sum_{m,n} \sum_{j,i} \varepsilon_{m,n;j,i} |m\rangle_N |\psi_j\rangle_N |n\rangle_S |\psi_i\rangle_S, \quad (2.6.43)$$

⁵⁶ By gravity we mean the full dynamics associated with gravity at short and long distances. Note that in 2+1 dimensions, there is no massless propagating graviton.

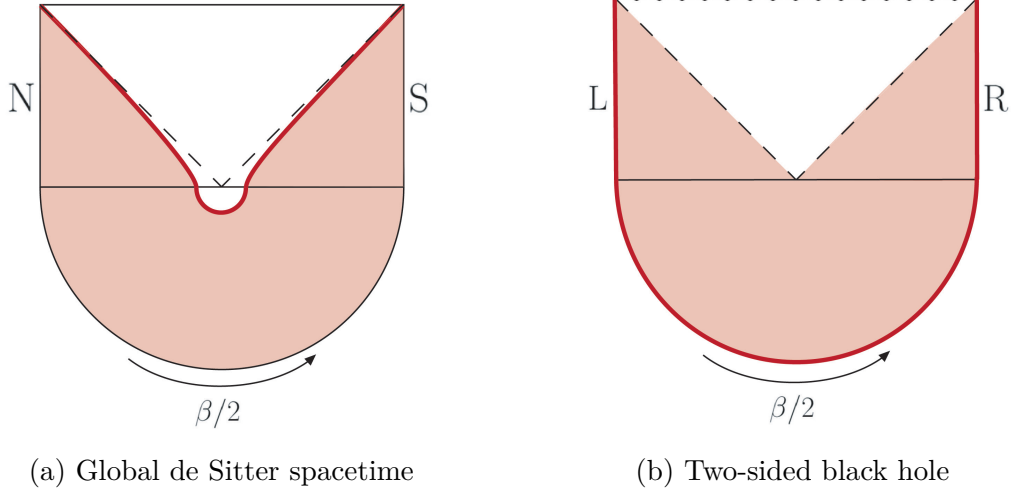


Figure 2.17: Preparation of a holographic state of the thermofield double form by Euclidean path integral. In each panel, the thick red line represents the location at which the holographic theory resides, while the shaded region corresponds to the bulk region that can be simply reconstructed, e.g., by the procedure of continuously pulling in the boundary.

where $\varepsilon_{m,n;j,i}$ are coefficients of order l_p^2/α^2 . We suspect that these off-diagonal parts are related to the fact that, unlike the case of a black hole, a positive energy shockwave in de Sitter spacetime gives a traversable “wormhole” between the two hemispheres [182, 185].

Excited states in which there are objects in the north and south hemispheres of the global de Sitter spacetime can be obtained by acting annihilation/creation operators, given by Eqs. (2.6.7)–(2.6.10) with $R \rightarrow N$ and $L \rightarrow S$, on the vacuum state in Eq. (2.6.43). Describing the region outside the horizons of the two static patches requires time evolution operator other than that generated by $H = H_N - H_S$. The appropriate generator $\tilde{H}$ can be constructed as

$$\tilde{H} = \sum_{\xi} \Omega_{\xi} a_{\xi}^{\dagger} a_{\xi} + \tilde{H}_{\text{int}}(\{a_{\xi}\}, \{a_{\xi}^{\dagger}\}), \quad (2.6.44)$$

where

$$a_{\xi} = \sum_{\gamma} (\alpha_{\xi\gamma} b_{R\gamma} + \beta_{\xi\gamma} b_{R\gamma}^{\dagger} + \zeta_{\xi\gamma} b_{L\gamma} + \eta_{\xi\gamma} b_{L\gamma}^{\dagger}), \quad (2.6.45)$$

$$a_{\xi}^{\dagger} = \sum_{\gamma} (\beta_{\xi\gamma}^* b_{R\gamma} + \alpha_{\xi\gamma}^* b_{R\gamma}^{\dagger} + \eta_{\xi\gamma}^* b_{L\gamma} + \zeta_{\xi\gamma}^* b_{L\gamma}^{\dagger}) \quad (2.6.46)$$

are mode operators with the coefficients $\alpha_{\xi\gamma}$, $\beta_{\xi\gamma}$, $\zeta_{\xi\gamma}$, and $\eta_{\xi\gamma}$ determined by semiclassical calculation.

Since the $t = 0$ hypersurface is a Cauchy surface, this effective theory describes the entire global de Sitter spacetime. The theory, however, is intrinsically semiclassical. The choice of the vacuum cannot be derived from the first principle, although we expect that the correct choice corresponding to typical soft-mode states is the Bunch–Davies vacuum [191]. The description of field fluctuations on the de Sitter background is also only statistical. To go

beyond this, e.g. to describe the details of a state depending also on the horizon microstate, we must resort to the microscopic theory.

In the $|\Psi(E)\rangle$ state, the von Neumann entropy of the north hemisphere is

$$S_N = \sum_n \sum_i \frac{e^{-\beta(E_n + E_i)}}{Z} [\beta(E_n + E_i) + \ln Z] + O\left(\frac{l_P^2}{\alpha^2}\right). \quad (2.6.47)$$

Assuming that the density of states of the soft modes does not depend on E_i and denoting it by $S_{\text{soft}}(E)$, we have

$$S_N = \ln [z e^{S_{\text{soft}}(E)}] + \beta \langle E_n \rangle_N + O(1), \quad (2.6.48)$$

where $z = \sum_n e^{-\beta E_n}$, and $\langle E_n \rangle_N = \sum_n E_n e^{-\beta E_n} / z$ is the thermal energy of the hard modes in the north hemisphere, as measured by H . Thus, the entropy S_N agrees with S_{sys} given by Eqs. (2.3.30) and (2.3.31) at the leading order. Namely, in the global state of Eq. (2.6.43), the entropy of de Sitter spacetime can be interpreted as the entanglement entropy between the two hemispheres, at the leading order.

Relation to the DS/dS correspondence

The theory we are considering consists of two holographic systems located on the stretched horizons of two static patches, each of which covers one half of the spatial section of global de Sitter spacetime at $t = 0$, the global time at which the spatial volume becomes minimal. These two systems, each of which is expected to be strongly coupled, are weakly interacting through gravity. This structure is reminiscent of that in the DS/dS correspondence [61, 62], a proposed holographic description of de Sitter spacetime.

While the two theories have similar structures at $t = 0$, they can be different at $t \neq 0$. In particular, if we evolve our theory with $H_N + H_S$, the places where the holographic systems are located move toward the future along the stretched horizons of the two static patches, as indicated by the red drawings in Fig. 2.18. This makes simply reconstructed regions stay within the static patches. On the other hand, in the DS/dS correspondence, each of the holographic systems is completed itself into de Sitter spacetime of one lower dimensions, leading to the simply reconstructed region called a DS/dS patch, which is depicted by the blue drawings in the figure.

One might speculate that the two theories represent the same system evolved differently in holographic space. It will be interesting to study this possible relation, but we leave it for the future.

Relation to the Shaghoulian-Susskind proposal

At the end of Section 2.6.1, we discussed the possibility of having a holographic theory of the black hole interior by placing the relevant degrees of freedom on the apparent horizon. We can consider an analogous situation for de Sitter spacetime, in which case the degrees of freedom relevant for describing the region outside the horizons of the two static patches are placed on the horizon; see Fig. 2.19.

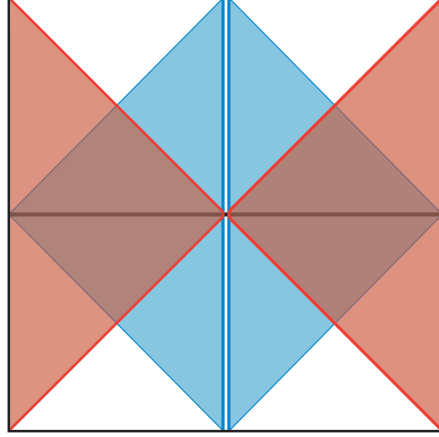


Figure 2.18: The holographic theory based on static patches consists of two weakly interacting systems located on the stretched horizons of two static patches, with the simply reconstructed region staying within the two patches (red). In the DS/dS correspondence, each of the holographic systems is located by itself on de Sitter spacetime with one lower dimensions, leading to the simply reconstructed region called the DS/dS patch (blue). The structures of the two theories are similar at $t = 0$.

As in the case of the black hole, we can test the consistency of this picture by computing the entanglement entropy between the two patches, which can be done using the maximin procedure. By minimizing a surface homologous to one of the horizons on a spacelike surface bounded by the two horizons at a fixed time, we find the resulting surface is nothing other than one of the horizons. The subsequent maximization over spacelike surfaces thus gives us the same surface, so the de Sitter horizon is the maximin surface. This gives the entanglement entropy

$$S = \frac{\mathcal{A}_{\text{dS}}}{4G_{\text{N}}}, \quad (2.6.49)$$

where $\mathcal{A}_{\text{dS}}$ is the area of the de Sitter horizon, which is consistent with the picture that the evolution of the system is unitary. This seems to be what is proposed in Refs. [63, 64] as a holographic theory of de Sitter spacetime (referred to as the monolayer proposal in Ref. [66]), although our interpretation here says that the theory describes only the shaded region in Fig. 2.19; in particular, it does not describe the regions inside the two static patches.

In our picture, the degrees of freedom used for the theory described above are a subset of the degrees of freedom of holographic theories describing the static patches. Therefore, the former are not really independent of the latter at the microscopic level. However, we may treat them independent for the purpose of describing semiclassical physics in the bulk. If this is the case, we can regard the holographic theory as consisting of two layers of degrees of freedom on the horizons, one describing the outside and the other describing the interiors of the horizons. This seems to be the proposal of Ref. [65], called the bilayer proposal in Ref. [66].

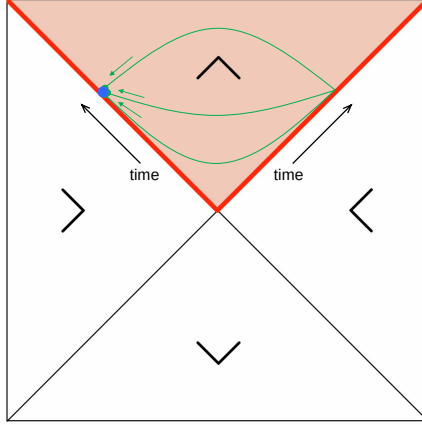


Figure 2.19: Holographic theory placed on the de Sitter horizons of two static patches can describe the region outside the horizons (as viewed by polar observers). Entanglement entropy between the two sides is computed by finding the maximin surface, which is nothing other than one of the horizons (blue dot). The entanglement entropy, therefore, is given by the Gibbons-Hawking entropy.

2.7 Gravitational Path Integral

The approach we have described so far is based on the canonical formalism of quantum mechanics. In particular, we have assumed the existence of a Hilbert space factor representing states of the horizon degrees of freedom, with which we could construct an effective theory describing a spacetime region behind the horizon—the interior in the case of a black hole—using only the low energy input that the dynamics of these degrees of freedom is maximally chaotic (and fast scrambling).

Quantum mechanics, however, can also be formulated using path integrals, and we expect that the same physical conclusions would be obtained from this formalism. In this section, we discuss what the picture looks like in this case. We will see that the corresponding picture is that of Refs. [19, 20], in which aspects of unitary evolution can be reproduced by gravitational path integral that fixes the “boundary condition” based on the quantity of interest and then integrates over all possible semiclassical geometries consistent with it, including those with nontrivial topologies (in particular, replica wormholes). This allows us to relate the framework described so far to the treatment based on the quantum extremal surface prescription [18, 54, 55]. While we focus on the case of a black hole here, we suspect that a similar story can be developed for de Sitter spacetime as well.

The picture presented here was outlined in Refs. [45, 56]. The ensemble nature of the semiclassical description arising from an ensemble of microstates was also discussed in Refs. [76–80]. The understanding of the Page curve presented here is based on the picture developed in Refs. [81–87].

2.7.1 Ensemble from coarse graining

The starting point for the path integral formalism is very different from that of the canonical formalism. Specifically, in the present context it should start from a collection of classical field configurations on classical geometries, which will then be integrated over. We are interested in a “low energy” framework in which the detailed microscopic knowledge is not necessary to understand the physics. This implies that a black hole must be treated as a (semi)classical object in which the detailed microscopic structure cannot be discriminated.

This treatment, in fact, is required by quantum mechanics. As we have argued in Sections 2.2 and 2.3, microstates of a black hole can be regarded as independent quantum states arising from superposing energy eigenstates in a small energy window, e.g. of order T_H , around M . Thus, to discriminate these microstates, one would need an exponentially long measurement time $\sim e^{S_{\text{bh}}}/T_H$. The black hole, however, would already have evaporated by the time such a measurement would be completed (or the state of the black hole will be altered significantly for a large AdS black hole thermalized with the environment). One therefore cannot operationally discriminate different microstates by performing a measurement on the black hole.⁵⁷

We thus have to regard a black hole spacetime appearing in gravitational path integral as representing a maximally mixed ensemble of the microstates consistent with the classical specification of the black hole [45, 56]. Note that this picture is associated with a single-sided (or a single-sided description) of a black hole. To see this, one can imagine preparing a black hole state using Euclidean path integral in the approximation that the black hole is static. As represented in Fig. 2.20, the “ ψ_+ - ψ_- component” of the density matrix of the exterior region can then be obtained by the path integral in which the spatial field configurations are fixed to be ψ_+ and ψ_- above and below the cut corresponding to the exterior region of the black hole.

At the semiclassical level, this simply gives the thermal density matrix which we regard as the semiclassical black hole vacuum state. We can, however, imagine following the same procedure at the microscopic level to get the corresponding density matrix

$$\rho_{\text{micro}} \propto \sum_n e^{-\frac{\beta}{2}\Delta H} |n\rangle \langle n| e^{-\frac{\beta}{2}\Delta H}, \quad (2.7.1)$$

where $\{|n\rangle\}$ is a complete set of black hole vacuum microstates consistent with the background used in the semiclassical path integral, $\beta = 1/T_H$ is the periodicity of Euclidean time in the angular direction, and ΔH is the microscopic exterior (boost) Hamiltonian. Here, the zero of ΔH is chosen to be a typical energy associated with the space $\mathcal{H}_{\text{vac}}$ spanned by $|n\rangle$'s.⁵⁸

Since the matrix elements of the $\beta\Delta H/2$ operator in $\mathcal{H}_{\text{vac}}$ are of order or smaller than 1, we find that many microstates contribute to ρ_{micro} . In particular, there are exponentially

⁵⁷ One can, of course, *infer* the microstate of the black hole if we prepare a specific state of initial collapsing matter and then use the microscopic theory to simulate its time evolution.

⁵⁸ The index n here corresponds to the index A in the notation in Sections 2.3 and 2.5.

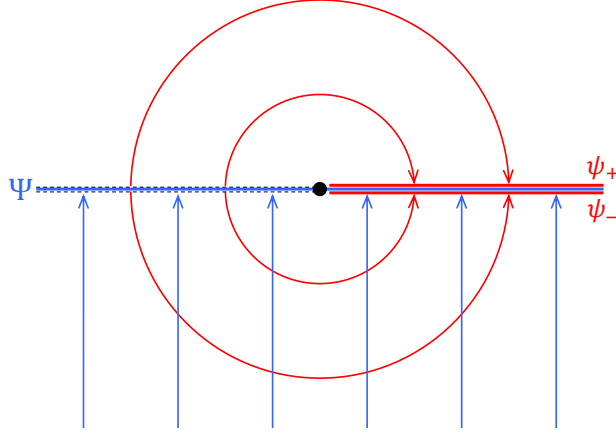


Figure 2.20: The density matrix of the exterior region, whose element is specified by the field configurations ψ_+ and ψ_- , can be computed by Euclidean path integral, which can be interpreted to receive equal contributions from all the black hole microstates. Euclidean path integral can also be used to prepare a two-sided state Ψ , which is pure and unique.

many microstates that contribute dominantly and almost equally; these are the microstates among which the matrix elements $\beta\Delta H/2$ is much smaller than 1. The fact that semiclassical theory cannot resolve the detailed structure of microstates implies that it perceives ρ_{micro} to be the maximally mixed state⁵⁹

$$\rho_{\text{sc}} \propto \sum_{n=1}^{e^{S_{\text{bh}}}} |n\rangle\langle n|, \quad (2.7.2)$$

where the sum runs over a sufficiently large set of vacuum microstates $|n\rangle$, whose precise specification is not important as discussed in Section 2.3 (or as in standard statistical mechanics). The number of elements of this set is of order $e^{S_{\text{bh}}}$, which we have already indicated in Eq. (2.7.2).

This interpretation indeed reproduces many features which are attributed to the ensemble nature of holographic *theories* in lower dimensional quantum gravity [19] in terms of an ensemble of microscopic *states* [45] (see also Refs. [76–80]). Incidentally, this should be contrasted with the situation in which a global (two-sided) state is prepared by Euclidean path integral. In this case, the generated state is a pure state

$$|\Psi_{\text{micro}}\rangle \propto \lim_{\tau \rightarrow \infty} e^{-\Delta\tilde{H}\tau} |i\rangle \quad (2.7.3)$$

obtained by evolving some initial state $|i\rangle$ at $\tau = -\infty$ to $\tau = 0$. Here, $|i\rangle$ is chosen in the space of black hole (not necessarily vacuum) microstates $\mathcal{H}_{\text{bh}}$, and $\Delta\tilde{H}$ is the microscopic

⁵⁹ The contributions to ρ_{micro} from states with $|\beta\Delta H| \gg 1$ are negligible either because of the Boltzmann suppression (for $\Delta H > 0$) or because the number of states is too small (for $\Delta H < 0$). The chaotic nature of the black hole dynamics then implies that the remaining contribution can be well approximated to come from the maximally mixed ensemble of microstates with $|\beta\Delta H| \lesssim 1$.

“infalling” (inertial) Hamiltonian with its zero chosen to be a typical energy of vacuum microstates. This leads to the unique, lowest energy eigenstate, determined by the choice of $\mathcal{H}_{\text{bh}}$; see also discussion around Eq. (2.6.1).

2.7.2 Replica method, entanglement island, and the Page curve

With the understanding that the black hole spacetime appearing in gravitational path integral represents the (maximally mixed) ensemble of black hole microstates, the results of Refs. [19, 20] can be understood in a simple manner.

Euclidean gravitational path integral

Let us first consider a setup similar to Ref. [20] in which a black hole living in a gravitational region is coupled to a nongravitational region.⁶⁰ The black hole then radiates into the nongravitational region. Our interest is to compute the von Neumann entropy of the emitted radiation.

Since the radiation lives in a nongravitational region, we must be able to describe its state semiclassically. In particular, assuming that the system can be viewed as quasi-static at each instance of time, we expect that the density matrix of the radiation in a region R can be calculated using a Euclidean path integral by specifying its element by spatial configurations of the radiation field above and below the cut on R ; see Fig. 2.21(a). The obtained density matrix is not normalized, which we denote by $\hat{\rho}_R$ with the hat over ρ_R indicating that the density matrix is not normalized.⁶¹

The path integral performed, however, includes the gravitational region in which a black hole resides. Thus, the obtained density matrix, in fact, involves an ensemble average over the black hole microstates in the sense of Eq. (2.7.2)

$$\overline{\hat{\rho}_R} = \frac{1}{e^{S_{\text{bh}}}} \sum_{n=1}^{e^{S_{\text{bh}}}} \hat{\rho}_{R,n}, \quad (2.7.4)$$

where $\hat{\rho}_{R,n} = \text{Tr}_{R^c} |n\rangle\langle n|$ represents the (nonnormalized) density matrix of R for the microstate

⁶⁰ The nongravitational region is a proxy for a weakly gravitational region. The distinction between gravitational and nongravitational regions is necessary to correctly identify semiclassical degrees of freedom in the latter region, in particular Hawking radiation quanta, of which the von Neumann entropy is calculated. Without this separation, the Page curve of our interest cannot be computed using the method adopted here [192], although we expect that there is a way to reproduce the same result without involving an artificial separation of regions, at least in principle.

⁶¹ The density matrix ρ_R here refers to the fine-grained (reduced) density matrix of fields in R . In the language of Ref. [174], the radiation here is the “radiation in boldface.”

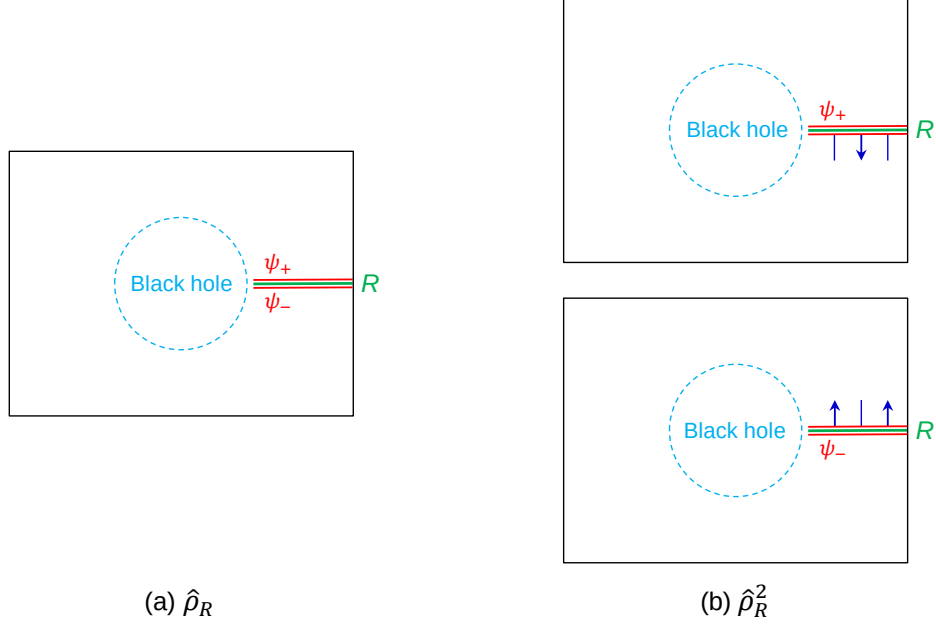


Figure 2.21: (a) The (nonnormalized) density matrix $\hat{\rho}_R$ of radiation in nongravitational region R can be calculated by performing path integral with the boundary condition fixing the spatial field configurations above and below the cut along R . (b) The square of the density matrix, $\hat{\rho}_R^2$ is calculated using the replica method, which replicates the spacetime into two and imposes the boundary condition such that one side of the cuts along R on the two sheets are appropriately sewn together (as represented by blue arrows).

$|n\rangle$ with R^c being the complement of R . Defining the normalized density matrix by

$$\overline{\rho}_R = \frac{\overline{\hat{\rho}_R}}{\text{Tr} \overline{\hat{\rho}_R}}, \quad (2.7.5)$$

we can compute the von Neumann entropy of R as

$$S_R^{(\text{sc})} = -\text{Tr} [\overline{\rho}_R \ln \overline{\rho}_R]. \quad (2.7.6)$$

Since this involves coarse-graining, i.e. the ensemble average over microstates, the quantity $S_R^{(\text{sc})}$ is well approximated by the thermal entropy of the radiation, which increases monotonically in time (until the black hole is fully evaporated, if it is not eternal). This is Hawking's result [17] showing the apparent violation of unitarity in the black hole evaporation process.

However, what we really want to understand is the behavior of

$$S_R^{(\text{micro})} = -\text{Tr} [\rho_{R,n} \ln \rho_{R,n}] \quad (2.7.7)$$

(for each microstate $|n\rangle$), which must show the Page behavior [41] if the evolution of the system is unitary. In particular, $S_R^{(\text{micro})}$ must go down to zero when the black hole is fully

evaporated and all the emitted radiation is included in R . A question is how (or if) we can see this behavior in a semiclassical analysis.

The idea is that while the semiclassical description necessarily involves the ensemble average over black hole microstates, it still allows for calculating the ensemble average of many different quantities. For example, we can calculate the ensemble average of the square of the density matrix for the radiation using the replica method [104, 193, 194], i.e. by replicating the spacetime into two copies and imposing the boundary condition for the path integral such that (one side of) the cuts along R on two sheets are appropriately sewn together; see Fig. 2.21(b). Performing path integral with this boundary condition gives

$$\overline{\hat{\rho}_R^2} = \frac{1}{e^{S_{\text{bh}}}} \sum_{n=1}^{e^{S_{\text{bh}}}} \hat{\rho}_{R,n}^2. \quad (2.7.8)$$

An important point is that the performed path integral must involve all possible geometries in the gravitational region, including those having a topology that geometrically connects the two sheets (the replica wormhole) [19, 20]. This is because the gravitational path integral should not predetermine the geometry on which quantum fields are integrated over.⁶² In general, the inclusion of geometries with nontrivial topologies makes Eq. (2.7.8) different from the square of Eq. (2.7.4): $\overline{\hat{\rho}_R^2} \neq (\overline{\hat{\rho}_R})^2$.

In Fig. 2.22, we illustrate the calculation of the trace of the square of the nonnormalized density matrix $\hat{\rho}_R$, which we denote by $Z_2 = \text{Tr } \hat{\rho}_R^2$. As found in Refs. [19, 20], there are two contributions to this. The first is the one in which the gravitational region is filled separately for two sheets; Fig. 2.22(a). Restricting the radiation configuration to those resembling the emitted Hawking radiation at a coarse-grained level, the integral of semiclassical fields provides a factor of $e^{S_{\text{rad}}}$, since the relevant four spatial configurations of the fields (both sides of the two cuts) are all related by Euclidean evolution as indicated by the red arrows in the figure. On the other hand, the gravitational path integral gives $e^{S_{\text{bh}}}$ for each sheet, so that this contribution is given by

$$Z_2^{(\text{disconnected})} \sim e^{2S_{\text{bh}} + S_{\text{rad}}}. \quad (2.7.9)$$

The other contribution is the one coming from a replica wormhole; Fig. 2.22(b). In this case, the radiation contribution is $e^{2S_{\text{rad}}}$, since we now have two independent cycles for the evolution, so that the four relevant configurations are related only pairwise (the red and blue arrows). On the other hand, the gravitational contribution is now $e^{S_{\text{bh}}}$ because it comes from a single, connected component. This therefore gives

$$Z_2^{(\text{wormhole})} \sim e^{S_{\text{bh}} + 2S_{\text{rad}}}. \quad (2.7.10)$$

⁶² As elucidated, e.g., in Ref. [86], this prescription calculates $\overline{\hat{\rho}_R^2}$ with exponential accuracy, taking into account the detailed microscopic structure of radiation (radiation in boldface in the language of Ref. [174]). If we instead calculate $\overline{\hat{\rho}_R^2}$ including only trivial topology, i.e. the disconnected contribution, then we would fail to capture the effect from the exponentially complex microscopic structure, yielding $\overline{\hat{\rho}_R^2}$ as computed in the semiclassical theory (i.e. the squared density matrix of non-boldface radiation). This latter prescription would lead to Hawking's result [17] when we compute the von Neumann entropy using $S_R = -\lim_{n \rightarrow 1} \partial(\text{Tr } \overline{\rho_R^n})/\partial n$, where $\overline{\rho_R^n} = \overline{\hat{\rho}_R^n}/(\text{Tr } \overline{\hat{\rho}_R})^n$.

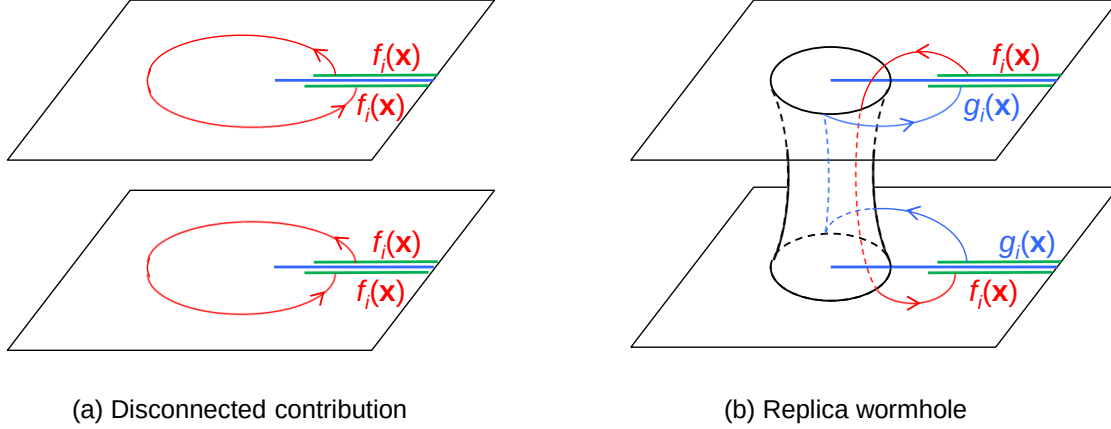


Figure 2.22: Two contributions to replica method calculation of $\overline{\text{Tr } \hat{\rho}_R^2}$. (a) The normal contribution in which the gravitational region is filled separately on each sheet. Integration of semiclassical matter gives $e^{S_{\text{rad}}}$ coming from Euclidean evolution relating all 4 field configurations above and below the cuts on two sheets (red arrows), while gravitational path integral gives the $e^{S_{\text{bh}}}$ factor for each sheet. (b) For the replica wormhole configuration, the gravitational regions on two sheets are geometrically connected. The contribution from gravitation path integral is thus only $e^{S_{\text{bh}}}$, while the contribution from matter integration gives $e^{2S_{\text{rad}}}$ since there are two independent cycles for evolution (red and blue arrows).

Adding the two contributions together, we find that $Z_2 \sim \max\{e^{2S_{\text{bh}}+S_{\text{rad}}}, e^{S_{\text{bh}}+2S_{\text{rad}}}\}$.

A similar analysis can be performed for Z_n for $n \in \mathbb{N}$. Assuming that the replica symmetric wormhole connecting all n sheets dominates for $S_{\text{rad}} > S_{\text{bh}}$, we find

$$Z_n \sim \begin{cases} e^{nS_{\text{bh}}+S_{\text{rad}}} & \text{for } S_{\text{bh}} > S_{\text{rad}} \\ e^{S_{\text{bh}}+nS_{\text{rad}}} & \text{for } S_{\text{rad}} > S_{\text{bh}}. \end{cases} \quad (2.7.11)$$

This is illustrated in Fig. 2.23. Since the density matrix ρ_R is given by $\hat{\rho}_R / \text{Tr } \hat{\rho}_R$, we thus find

$$\overline{\text{Tr } \rho_R^n} = \overline{\left(\frac{\text{Tr } \hat{\rho}_R^n}{(\text{Tr } \hat{\rho}_R)^n} \right)} \approx \frac{\overline{\text{Tr } \hat{\rho}_R^n}}{(\overline{\text{Tr } \hat{\rho}_R})^n} = \frac{Z_n}{Z_1^n} \sim \begin{cases} e^{(1-n)S_{\text{rad}}} & \text{for } S_{\text{bh}} > S_{\text{rad}} \\ e^{(1-n)S_{\text{bh}}} & \text{for } S_{\text{rad}} > S_{\text{bh}}, \end{cases} \quad (2.7.12)$$

where in the second equation we have used the fact that the standard deviation of $\text{Tr } \hat{\rho}_R^n$ is smaller than its typical size in the ensemble, which we can easily be convinced. By analytically continuing this result in n , we can now obtain the ensemble average of the von Neumann entropy of the radiation

$$\overline{S_R} = -\lim_{n \rightarrow 1} \frac{\partial}{\partial n} \overline{\text{Tr } \rho_R^n} \sim \begin{cases} S_{\text{rad}} & \text{for } S_{\text{bh}} > S_{\text{rad}} \\ S_{\text{bh}} & \text{for } S_{\text{rad}} > S_{\text{bh}}, \end{cases} \quad (2.7.13)$$

which reproduces the Page curve. This is because we have calculated the ensemble average of the von Neumann entropy (which obeys the Page curve for all members of the ensemble), and not the von Neumann entropy of the averaged density matrix as in Eq. (2.7.6).

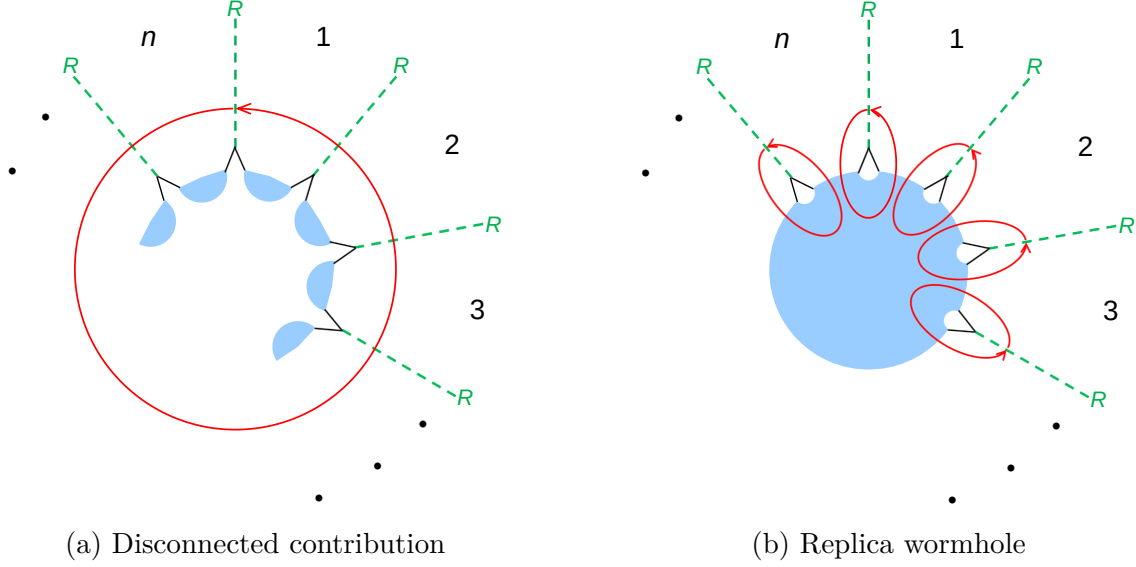


Figure 2.23: (a) The contribution to Z_n with the gravitational region filled separately (light blue) on each sheet. All semiclassical matter configurations above and below the cuts R (green dashed) are related by Euclidean evolution (red arrow), giving the factor of $e^{S_{\text{rad}}}$. (b) For the replica symmetric wormhole configuration, the gravitational regions on n sheets are geometrically connected. This allows for n independent cycles for evolution (red arrows), giving the contribution from matter integration of order $e^{nS_{\text{rad}}}$.

Note that this observation is similar to that in Ref. [81], but here we do not consider any ensemble of holographic theories. Instead, the relevant ensemble arises from the coarse graining of black hole microstates, which is *forced* on us if we adopt any formalism involving the semiclassical picture of the black hole, such as gravitational path integral. See also Refs. [86, 87] for related discussion.

Entanglement island and the Lorentzian picture

It was shown in Ref. [20], building on the technique developed in Refs. [104, 105, 195], that the prescription using the replica method in gravitational path integral is equivalent to the quantum extremal surface prescription [15] in the original unreplicated spacetime.⁶³ Going to the Lorentzian signature, this therefore leads to the following picture [18, 54, 55].

In order to calculate the von Neumann entropy of radiation S_{rad} in some region R , we

⁶³ Strictly speaking, this was shown only in the limit that the contribution from the graviton to the semiclassical field integration is negligible compared with those from other quantum fields. Below we assume that the equivalence of the two prescriptions persists beyond this limit.

need to find an entanglement island I which extremizes the following quantity⁶⁴

$$S_{\text{gen}}(R, I) = \frac{\mathcal{A}(\partial I)}{4G_N} - \text{Tr}[\rho_{\text{sc}}(R \cup I) \ln \rho_{\text{sc}}(R \cup I)], \quad (2.7.14)$$

where $\mathcal{A}(\partial I)$ is the area of the boundary of I , and $\rho_{\text{sc}}(X)$ is the reduced density matrix of the region X in the semiclassical theory. Note that I can be a null region. In general, there can be multiple such I 's, and the entropy S_{rad} is given by the minimum of $S_{\text{gen}}(R, I)$'s associated with all such I 's:

$$S_{\text{rad}} = \min_I \text{ext}_I S_{\text{gen}}(R, I). \quad (2.7.15)$$

This entropy is the same as that calculated by the replica method in path integral

$$S_{\text{rad}} = \overline{S_R}, \quad (2.7.16)$$

so that it obeys the Page curve.

It is important that this extremization procedure is performed on a global spacetime of general relativity. In particular, for a black hole spacetime, it must be performed on the whole spacetime including the interior of the black hole. This, therefore, gives a complementary picture to that described in Sections 2.2 – 2.6. In the picture described here, the existence of the black hole interior is obvious—in fact, the framework assumes it—while to understand the unitary nature of black hole evolution, one needs to resort to a method that appropriately incorporates nonperturbative effects of quantum gravity, such as replica wormholes. On the other hand, in the framework of Sections 2.2 – 2.6, the unitarity of the evolution is an assumption and, as a consequence, the existence of the interior is not manifest—the interior emerges only effectively as a collective phenomenon involving horizon degrees of freedom, which are subject to universally chaotic and fast-scrambling dynamics.

Manifestly unitary vs global spacetime descriptions

We can, in fact, understand the existence of the two frameworks discussed above—based on the manifestly unitary and global spacetime pictures, respectively—from the viewpoint of gauge symmetries of the underlying theory [45, 56]. As emphasized in Ref. [196], a theory of quantum gravity has nonperturbative gauge redundancies much larger than the standard diffeomorphism. A particular manifestation of this is the apparent violation of the Bekenstein-Hawking entropy bound in the semiclassical description of a black hole [45, 56, 197]. This violation arises from huge spatial volume inside the black hole [198, 199] (including the so-called bags-of-gold configurations), which leads to the number of independent quantum states exceeding the Bekenstein-Hawking bound. Many of these semiclassically independent states, however, are equivalent under the nonperturbative gauge symmetries, making the number of truly independent states satisfy the Bekenstein-Hawking bound.

The two frameworks describe this phenomenon in very different, though equivalent, ways. In the framework discussed in Sections 2.2 – 2.6, the nonperturbative gauge redundancies (as well as a part of the standard diffeomorphism) are explicitly fixed by employing a

⁶⁴ The regions R and I can be viewed either as spatial regions or causal regions associated with them.

Schwarzschild-like time foliation, which is motivated by holography. In this framework, making unitarity manifest, the number of independent black hole states is simply $e^{S_{\text{BH}}}$, and the apparently much larger number of independent interior states in semiclassical theory comes from the fact that the Hilbert space spanned by the $e^{S_{\text{BH}}}$ states admit $e^{e^{S_{\text{BH}}}}$ approximately orthogonal states, between which inner products are of order $e^{-S_{\text{BH}}/2}$ or smaller. Since the semiclassical theory cannot detect such small inner products, it appears to accommodate independent quantum states larger than the Bekenstein-Hawking bound in the interior of a black hole.

On the other hand, in the framework making the interior manifest, the argument runs in the other direction. In this case, the starting point is the global spacetime picture, which is highly redundant. In the canonical formulation of quantum mechanics, this implies the existence of over-entropic semiclassical interior states $|\psi_i\rangle$. After including nonperturbative effects of gravity, however, these semiclassically orthogonal states develop small overlaps, $\langle\psi_i|\psi_j\rangle \sim e^{-S_{\text{BH}}/2}$, in a such way that the rank of the matrix $M_{ij} \equiv \langle\psi_i|\psi_j\rangle$ is reduced drastically to $e^{S_{\text{BH}}}$. Such a large number of null states are a manifestation of the large nonperturbative gauge redundancies [196, 200], which relate even spaces with different topologies [196, 201]. In the path integral formulation, this reduction of the semiclassical Hilbert space to the physical one is achieved by including additional contributions to the path integral (e.g. replica wormholes) which projects states onto those invariant under the relevant gauge symmetries as we have seen in this section.

2.8 Conclusion

In quantum gravity, there has been a difficulty in reconciling fundamental principles of physics in the presence of a black hole (or other) horizon, particularly the unitarity of quantum mechanics and the equivalence principle of general relativity [17]. This difficulty is, in fact, an avatar of the conventional difficulty of describing UV physics at the Planck scale, albeit in a redshifted form. Because of a huge gravitational redshift between the horizon and asymptotic regions, the degrees of freedom represented by the Bekenstein-Hawking entropy [1, 31], which obey an intrinsically “stringy” dynamics, appear to have exponentially degenerate states that cannot be discriminated operationally by performing a measurement on the black hole. A low energy description of quantum gravity treats these states in a thermal way, so that unitarity appears to be lost.

While a semiclassical theory, as a low energy description of quantum gravity, cannot describe all the microscopic dynamics of the fundamental theory, it can still be used to obtain a coarse-grained understanding of how unitarity and the equivalence principle can coexist in a black hole system, with only a few inputs from the UV theory. This is what we have explored in this paper.

One way of doing this is simply to postulate that the evolution of a black hole is unitary when viewed from the exterior [38–41]. This is a view motivated by holography [10]. In this case, the existence of the black hole interior appears to be in jeopardy [30]. We have seen, however, that with the assumption that the dynamics of horizon degrees of freedom

is chaotic and fast scrambling across all low energy species, the interior emerges at a semiclassical level as a collective phenomenon involving the horizon, and possibly other, degrees of freedom [42–46]. An important point is that for a black hole with minimal uncertainties (i.e. a specific black hole at the semiclassical level), the assumption of Gaussian randomness of the coefficients for microstates allows us to learn many features associated with the construction of the effective theory of the interior. We do not need any further details about the microscopic theory to figure these things out. A similar construction applies to de Sitter spacetime, which we have also elaborated in this paper.

The compatibility between unitarity and the equivalence principle can alternatively be seen by starting from the global spacetime of general relativity. In this case, the existence of the interior is manifest, but at the apparent expense of unitarity as Hawking’s original calculation indicates [17]. However, while a semiclassical picture necessarily involves an average over microstates, we can directly calculate the average of the quantity we are interested in, e.g. the von Neumann entropy of emitted Hawking radiation, rather than the quantity in the averaged state. This is what is done in Refs. [18–20, 54, 55] by employing the replica method in gravitational path integral, or equivalently the quantum extremal surface prescription. In this way, one can see the unitarity of the underlying theory without knowing its detailed dynamics.

Despite the fact that the two frameworks described above appear very different, they give the same physical conclusions. In particular, a black hole evolves unitarily and has a smooth horizon. Viewed in this way, one can conclude that historical confusions about black hole physics come from the fact that only one of these features can be made manifest in a given low energy description; the other one appears in a highly nontrivial manner. We have seen that the description making unitarity (quantum mechanics) manifest comes more naturally with the canonical/Hamiltonian formulation of quantum mechanics, while the one making the interior (general relativity) manifest is associated with the path integral/Lagrangian formulation. It remains to be seen if there is a microscopic formulation of quantum gravity in which both these features are manifest at the same time.

Chapter 3

Complementarity for a Dynamical Black Hole

3.1 Introduction

Studying the quantum mechanics of a black hole has revealed surprising facts about the nature of spacetime [1, 3, 31]. Of particular interest is that the black hole interior and Hawking radiation are not independent quantum degrees of freedom. This idea, called complementarity [40], has been tested in various simplified setups [18–20, 54, 55], using methods [13–15, 103] developed by analyzing the holographic nature of quantum gravity [7, 8], in particular the AdS/CFT correspondence [10].

This lack of independence allows us to construct operators representing the interior which act solely on the Hawking radiation, as is indeed implied by entanglement wedge reconstruction [12, 16, 162–166, 202] with entanglement islands [18, 54, 55]. Since the radiation and the interior are spacelike separated, an exterior observer’s ability to affect interior degrees of freedom leads to an apparent violation of causality. This issue was analyzed by Kim, Tang, and Preskill [169], who studied a partially evaporating black hole and proposed that the spacelike nature of the interior and radiation in the semiclassical description is protected because the operators representing the interior effectively commute with any operators that do not have computational complexity exponential in the entropy of the black hole. They, however, did not address how to interpret operators acting on the radiation after the black hole is fully evaporated, which would still be able to affect the interior acausally. Bousso and Penington studied a process involving such operators and concluded that it leads to a breakdown of the semiclassical spacetime [203].

In this paper, we study how complementarity is manifested without the issue of causality described above. We claim that to fully understand this issue, we need to properly analyze how a dynamical black hole behaves under time evolution. For this purpose, we adopt an exterior description [42–46, 204] in which an effective theory describing a portion of the black hole interior is constructed from the horizon (and Hawking radiation) degrees

of freedom *at each exterior time*. This portion of the interior is obtained by evolving the modes on the corresponding exterior time slice by the *infalling* Hamiltonian, not by exterior time evolution, and therefore must be viewed as associated with the exterior time at which the effective theory is erected. A sketch of this construction is given in Fig. 3.3. In this description, the black hole interior is a transient phenomenon obtained by reshuffling the relevant degrees of freedom while an infalling object is being scrambled into the black hole.

Using this construction, we show the following:

- The black hole interior must be viewed as existing in the causal past of the corresponding Hawking radiation despite the fact that they are spacelike separated in the semiclassical picture. In the fundamental description, the relationship between operators corresponding to the interior and their reconstruction in the Hawking radiation is timelike.
- The relationship between these operators becomes effectively spacelike in the semiclassical description because of limitations on an infalling observers’ ability to access complex information. With this limitation, the chaotic dynamics on the horizon, as viewed from the exterior, convert any effect caused by an interior observer to information in the Hawking radiation which cannot be accessed at the semiclassical level.
- An operation on the Hawking radiation cannot affect the experience of an observer in the interior no matter how complex it is, providing a solution to the causality issue different from that in [169]. Executing such an operation can only hope to “rewrite the record” to make it appear that the past was different. We show that the same applies also to black hole mining [115, 116], so that the semiclassical spacetime is robust under information extraction.

Note that this provides a simple solution to the problem of possible information cloning in the interior and the radiation. The no-cloning theorem [205] is not violated simply because an infalling object and its information in Hawking radiation are the same timelike-related degrees of freedom.

We will formulate a unitary exterior description in which the flow of information is strictly causal throughout the black hole’s formation and evaporation. This is not trivial, especially during the black hole formation era. We will describe the challenge and propose a solution which applies bulk holography to the stretched *event* horizon. In this description, the stretched event horizon provides the inner “edge” of the spacetime, on which the information about the interior is holographically encoded.

The picture presented here implies that a description based on the global spacetime is necessarily coarse-grained. We will discuss the relation between the exterior description and calculations employing the global spacetime description, such as replica wormhole calculation [19, 20], entanglement islands [18, 54, 55], and α -states arising from baby universes [206–208] created by the evaporation of a black hole [209, 210].

The organization of this paper is as follows. After laying out the exterior description in Section 3.2, we discuss in Section 3.3 how the interior of the black hole and the global spacetime of general relativity emerge from the exterior description. In Section 3.4, we argue

that the global spacetime description cannot be used to define unitary quantum mechanics consistent with interior reconstruction, and show how the framework in Section 3.3 addresses this issue. In particular, we see that the black hole interior must be viewed as existing in the past of the Hawking radiation containing its information, despite the fact that they are spacelike separated in the semiclassical description. We discuss how the timelike relation can be regarded as spacelike at the semiclassical level because of the limitation of an infaller’s ability and the chaotic dynamics of the horizon degrees of freedom.

In Section 3.5, we extend the framework of Section 3.3 to the entire history of black hole evolution, including its formation era. This gives a description in which the flow of information is causally trivial throughout the history. Our construction employs bulk holography associated with the stretched event horizon. In Section 3.6, we discuss the relation of our picture with other works that use the global spacetime description. Section 3.7 gives conclusions.

Throughout the paper, we adopt natural units $\hbar = c = k = 1$.

3.2 Exterior Description

Black holes form and then evaporate, and a distant observer expects the evolution to be unitary in agreement with quantum mechanics [38–41]. Unitarity is further supported by the AdS/CFT correspondence[10]. To preserve unitarity, we must follow *all* the relevant degrees of freedom, including the “intrinsically quantum gravitational” (or “Planckian”) degrees of freedom associated with the stretched horizon [40]. For a static black hole, this horizon is located at a proper distance of order the cutoff/string length l_s away from the mathematical horizon. The spacetime region relevant for the description is the exterior of the stretched horizon.

A simple way to implement the exterior description is to adopt the canonical formalism of quantum mechanics. The relevant time slices are Schwarzschild-like time slices, i.e. slices that foliate the spacetime region outside the horizon; see Fig. 3.1.

At times when the black hole exists, these time slices cover the region outside the black hole. In the context of holography, these slices can be obtained, for example, by continuously pulling the boundary into the bulk using the method in [35, 36].

A feature of the exterior description is that if the state before the formation of the black hole is pure, then the states with the black hole present and those after it evaporates are also pure. The relevant degrees of freedom in each period are

- **Before the black hole formation:** low energy degrees of freedom whose dynamics can be described by a semiclassical theory. In particular, they include matter that collapses gravitationally to form the black hole.
- **While the black hole is present:** the degrees of freedom outside the horizon. At each time, we can divide these degrees of freedom into three classes (see [204] for more detail):

- *Horizon modes*: modes on the stretched horizon, which may be viewed as “located” in the region $r \leq r_s$, where r_s represents the location of the stretched horizon in the Schwarzschild radial coordinate r (although the classical spacetime picture is not valid in this region).¹
- *Zone modes*: modes localized in the region $r_s < r \leq r_z$, where r_z is the location of the outer edge of the near black-hole region called the zone (i.e. the region inside the barrier of the effective gravitational potential generated by the black hole).
- *Far modes*: modes located in the asymptotic region, $r > r_z$.

This decomposition can be performed at the gauge-fixed level with positive semi-definite inner products between the states, up to the intrinsic ambiguity of separation between horizon and zone modes (we assume that the usual issue of factorizing modes located in different regions in quantum field theory is appropriately taken care of). The dynamics of the horizon modes cannot be described by a low energy semiclassical theory, while those of the zone and far modes can.

- **After the black hole evaporation:** low energy degrees of freedom whose dynamics can be described by the semiclassical theory.² In particular, they include Hawking radiation emitted from the black hole.

As has long been argued [38–41], the exterior description should be unitary, since the stretched horizon behaves as a regular material surface as far as the flow of quantum information is concerned. On the other hand, general relativity implies the existence of an interior. While there are proposals otherwise, such as fuzzballs [211–213] and firewalls [30, 214, 215], we take the view that nothing special happens at the horizon for an infalling observer. The question then is how we can understand the emergence of a smooth interior for the infaller in the exterior description.

3.3 Emergence of Global Spacetime

In [42–46, 204], it was shown that an effective theory of a smooth interior can be constructed from the exterior degrees of freedom. In particular, the smooth interior emerges as a secondary concept associated with the modes in the zone describing an object falling into the black hole. Coherent excitations of the rest of the zone and horizon modes (as well as far modes for an old black hole) play the role of the modes in the second exterior of a two-sided

¹ Without an explicit ultraviolet theory, it is scheme dependent if we view these degrees of freedom as additional degrees of freedom beyond the low energy quantum fields or their cutoff-scale excitations near the stretched horizon. In any event, the microscopic dynamics of these degrees of freedom cannot be described by low energy semiclassical theory.

² This does not mean that the state of these degrees of freedom is fully specified by operators within the low energy semiclassical theory; see Section 3.4.

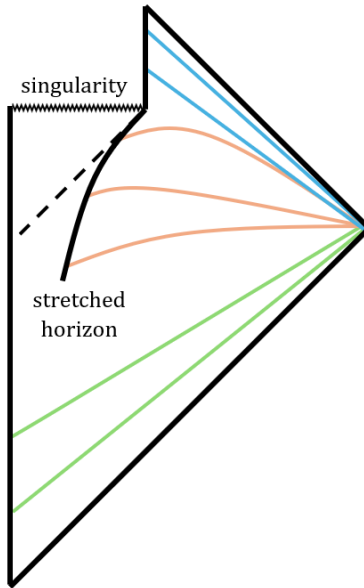


Figure 3.1: In the exterior description, equal-time slices foliate the spacetime region outside the “inner boundary” of space, which for a (quasi-)static black hole corresponds to the stretched horizon.

geometry, which is emergent in the effective theory. As emphasized in [42–46, 204], it is important to apply this construction to a state *at a fixed time* t in the exterior description. This provides *a portion of* the black hole interior, which can be used to describe the experience of the infalling object at the semiclassical level. To cover the entire black hole interior, we need multiple effective theories erected at different times.

In Section 3.3.1, we review this construction. We will focus on the main idea, referring the readers to [204] for more details. In Section 3.3.2, we discuss properties of the construction relevant for our later discussions, especially focusing on how the picture of global spacetime may emerge.

3.3.1 Effective theory of the interior

For concreteness, we focus on a collapse-formed black hole with negligible spin and charge in 4-dimensional asymptotically flat spacetime. We assume that the black hole has a mass M at some time t , where t is the time in the exterior description.³ The horizon, zone, and far modes are defined based on their locations *at this time*, according to the criterion given

³ More precisely, we focus on a branch of the state at t in which the black hole has mass M and is located at a fixed position, up to irreducible quantum uncertainties of order the Hawking temperature T_H and the proper length l_s , respectively [42]. Our construction of the effective theory of the interior applies to each such branch separately.

in Section 3.2. A *vacuum* microstate of the semiclassical black hole at time t , i.e. a state in which excitations described by the semiclassical theory are absent, can then be written as follows.

We first select a subset of the zone modes that will be relevant for describing a falling object which is assumed to lie in the zone at time t . These are called *hard modes*, which we collectively label as α . The semiclassical vacuum microstate, together with Hawking radiation emitted earlier, can then be written as⁴

$$|\Psi(M)\rangle = \sum_n e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} \sum_a c_{ni_n a} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \quad (3.3.1)$$

where $|\{n_\alpha\}\rangle$ represent states of the hard modes, specified by the occupation number n_α for each mode α , while $|\psi_{i_n}^{(n)}\rangle$ represent orthonormal states of the horizon and zone modes that are not included in the hard modes, which are called *soft modes* (any basis in the i_n space can be taken for each n). Finally, $|\phi_a\rangle$ are the states of the far modes, which generally include Hawking radiation emitted earlier. Note that since the total energy of the black hole system, consisting of the hard and soft modes, is M , the number of basis states $|\psi_{i_n}^{(n)}\rangle$ appearing with $|\{n_\alpha\}\rangle$ in (3.3.1) is given by $S_{\text{bh}}(M-E_n)$. Here, E_n is the energy carried by $|\{n_\alpha\}\rangle$, $S_{\text{bh}}(E) = 4\pi E^2 l_{\text{P}}^2$ is the Bekenstein-Hawking entropy density at energy E , and l_{P} is the Planck length, which we take to be not too different from l_s , assuming that the number of low energy species is not enormous.

Because of fast-scrambling [68, 69] and chaotic [67] dynamics at the horizon, (the relevant portion of) the coefficients $c_{ni_n a}$ in (3.3.1) take random values across all low energy species.⁵ (The specific values of $c_{ni_n a}$ correspond to a specific black hole microstate.) This implies that the state of the system can be written in the thermofield double form

$$|\Psi(M)\rangle = \frac{1}{\sqrt{\sum_m e^{-\frac{E_m}{T_{\text{H}}}}}} \sum_n e^{-\frac{E_n}{2T_{\text{H}}}} |\{n_\alpha\}\rangle \|\{n_\alpha\}\rangle, \quad (3.3.2)$$

up to exponential accuracy, where $T_{\text{H}} = 1/(\partial S_{\text{bh}}(E)/\partial E|_{E=M}) = 1/8\pi M l_{\text{P}}^2$ is the Hawking temperature. The state with the double-ket symbol is defined as the state of soft and far modes which couples to the state $|\{n_\alpha\}\rangle$ in (3.3.1)

$$\|\{n_\alpha\}\rangle = \sqrt{\sum_m e^{-\frac{E_m}{T_{\text{H}}}} e^{\frac{E_n}{2T_{\text{H}}}}} e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} \sum_a c_{ni_n a} |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \quad (3.3.3)$$

where the overall coefficient is determined by the normalization condition, $\langle\langle \{n_\alpha\} | \{n_\alpha\} \rangle\rangle = 1$. We see that the double-ket states play the role of the states in the second exterior of the effective two-sided black hole, as depicted in Fig. 3.2. This allows us to define, along the

⁴ Here we ignore the effect of backreaction of Hawking emission on the occupation numbers of hard modes, which has some implication for the region that can be directly reconstructed from the state $|\Psi(M)\rangle$ without resorting to exterior time evolution; for details, see [204].

⁵ More precisely, for each a , the coefficients $c_{ni_n a}$ are expected to behave as independent and identically distributed complex Gaussian random variables with mean 0.

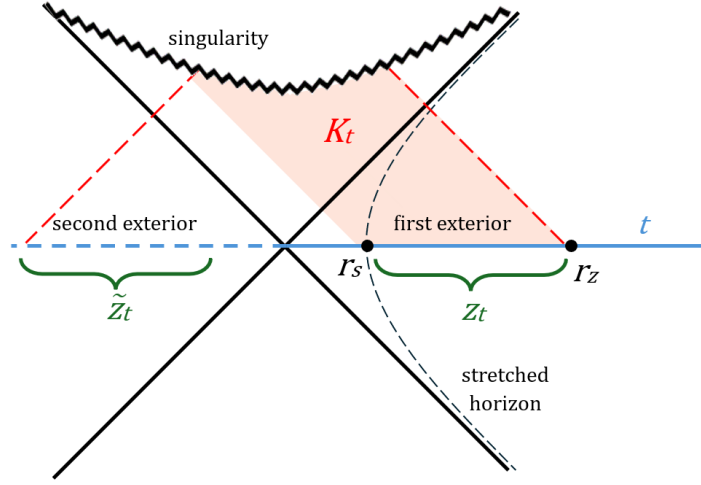


Figure 3.2: The effective theory of the interior erected at time t gives an emergent two-sided black hole geometry. The degrees of freedom associated with the zone $\tilde{Z}_t$ of the second exterior arise from soft and far mode states coupled to the states in the zone Z_t of the first exterior in the vacuum microstate. The spacetime region described by this effective theory is the region indicated as K_t .

lines of [47–50, 52], annihilation and creation operators

$$b_\gamma = \sum_n \sqrt{n_\gamma} |\{n_\alpha - \delta_{\alpha\gamma}\}\rangle \langle \{n_\alpha\}|, \quad b_\gamma^\dagger = \sum_n \sqrt{n_\gamma + 1} |\{n_\alpha + \delta_{\alpha\gamma}\}\rangle \langle \{n_\alpha\}|, \quad (3.3.4)$$

$$\tilde{b}_\gamma = \sum_n \sqrt{n_\gamma} \|\{n_\alpha - \delta_{\alpha\gamma}\}\rangle\langle\langle \{n_\alpha\}\|, \quad \tilde{b}_\gamma^\dagger = \sum_n \sqrt{n_\gamma + 1} \|\{n_\alpha + \delta_{\alpha\gamma}\}\rangle\langle\langle \{n_\alpha\}\|, \quad (3.3.5)$$

from which we can construct operators describing the interior of the black hole, using only the degrees of freedom on the exterior region, i.e. the hard, soft, and far modes defined on the time slice t .⁶

The operators describing semiclassical excitations in the interior region, obtained from Eqs. (3.3.4) and (3.3.5), act on all of the hard, soft, and far modes. However, if the black hole is young, i.e. if the combined system of the hard and soft modes is not maximally entangled with far modes at time t , then the interior operators can be constructed without involving far modes [44, 46]. This can be done by projecting the operators acting on all modes onto the Hilbert space of hard and soft modes using the Petz map. Schematically, when the entanglement between the black hole and far modes is negligible, the vacuum microstate can be written as

$$|\Psi(M)\rangle \approx \sum_n e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} c_{ni_n} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle. \quad (3.3.6)$$

⁶ While the operators in (3.3.5) depend sensitively on the microstate of the black hole, i.e. the precise values of $c_{ni_n a}$ in (3.3.1), we can promote them to apply to a wide range of microstates to the extent that it is sufficient to ensure the validity of the semiclassical description of the interior [46, 51, 53].

The double-ket states can then be defined by

$$\|\{n_\alpha\}\rangle\rangle \approx \sqrt{\sum_m e^{-\frac{E_m}{T_H}} e^{\frac{E_n}{2T_H}}} e^{S_{\text{bh}}(M-E_n)} \sum_{i_n=1} c_{ni_n} |\psi_{i_n}^{(n)}\rangle, \quad (3.3.7)$$

instead of (3.3.3), giving interior operators acting only on hard and soft modes. For a young black hole, these operators reproduce the same algebra as that obtained using the “full” double-ket states in (3.3.3), up to exponentially suppressed corrections.

We stress that in the case of either a young or an old black hole, interior operators constructed as above *must* involve soft modes, i.e. modes near the black hole region. This implies, in particular, that the construction does not work when a black hole does not exist at time t . This is different from entanglement wedge reconstruction [12, 16, 162–166, 202], which allows for reconstructing interior operators only using far modes (early radiation) [18, 54, 55]. The difference comes from the fact that the construction described here does not use exterior time evolution with respect to t in converting the fundamental modes to infalling modes, while entanglement wedge reconstruction (implicitly) uses it by assuming the knowledge of microscopic time evolution [45, 204].

For an evaporating black hole, the spacetime region K_t that can be obtained by the construction above is a portion of the interior which corresponds to the intersection of the domain of dependence of the union of Z_t and $\tilde{Z}_t$ with the causal future of Z_t [204]:

$$K_t = D(Z_t \cup \tilde{Z}_t) \cap J^+(Z_t), \quad (3.3.8)$$

where Z_t represents the zone at time t , i.e. the partial Cauchy surface $r_s < r \leq r_z$ at t , and $\tilde{Z}_t$ is its second exterior mirror in the effective two-sided geometry arising in the effective theory; see Fig. 3.2. Any operator of the semiclassical theory contained in this region can be constructed using the time evolution operator $\tilde{U}(\tau) = e^{-i\tilde{H}\tau}$ associated with an infalling frame. Here,

$$\tilde{H} = \sum_{\xi} \Omega_{\xi} a_{\xi}^{\dagger} a_{\xi} + \tilde{H}_{\text{int}}(\{a_{\xi}\}, \{a_{\xi}^{\dagger}\}) \quad (3.3.9)$$

is the generator of the infalling time evolution, and τ is the infalling time; a_{ξ} and $a_{\xi}^{\dagger}$ are the annihilation and creation operators for infalling mode ξ , obtained from b_{γ} , $b_{\gamma}^{\dagger}$, $\tilde{b}_{\gamma}$, and $\tilde{b}_{\gamma}^{\dagger}$ by Bogoliubov transformations, and Ω_{ξ} is the frequency of the mode ξ . Note that this time evolution is *different* from the original time evolution, which acts on all the microscopic degrees of freedom. In fact, the effective theory of the interior obtained in this way is intrinsically semiclassical, possessing intrinsic ambiguities of order $e^{-S_{\text{tot}}/2}$, where S_{tot} is the coarse-grained entropy of the combined black hole and radiation system [46].

The picture presented here implies that a generic state in the Hilbert space of microstates for a collapse-formed, single-sided black hole does not have a firewall. In [215], it was argued that since on a black hole background “arbitrarily many” configurations of matter can be excited without costing energy as viewed from a distance (by putting them near the horizon), a generic state of the black hole Hilbert space has a firewall. These configurations, however, correspond precisely to black hole microstates; specific values of c_{ni_n} in (3.3.6), or $c_{ni_n a}$ in (3.3.1), correspond simply to a specific *vacuum* microstate. For generic values of these

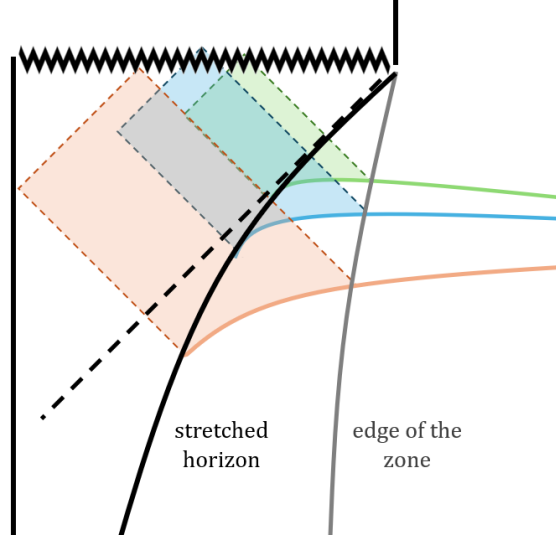


Figure 3.3: The effective theory of the interior erected at an exterior time t describes the region of the black hole interior corresponding to the interior of the effective two-sided black hole composed of the zone at t and its second exterior mirror. To obtain the whole black hole interior, we need to patch the pictures obtained by effective theories erected at different exterior times.

coefficients, the present construction works, giving an effective description in which a falling object goes through the horizon smoothly. In particular, the fate of a semiclassical object in the zone, represented by a set of $b_j^\dagger$ operators *acted on* the vacuum microstate, can be calculated by evolving the system by $\hat{U}(\tau)$ using the in-in formalism, reproducing the result consistent with general relativity. It is only for an exponentially atypical state (e.g. a state in which one of the coefficients is exponentially larger than the others) that this construction does not work, leading to a firewall at the horizon.⁷

3.3.2 Global spacetime from a patchwork

From the exterior description, the global spacetime picture is obtained by erecting effective theories of the interior described above at different times t and “patching” them [42], as sketched in Fig. 3.3. We first highlight two important aspects in this procedure:

- (i) There are many different ways to represent the same interior operator in fundamental (hard, soft, and far) degrees of freedom. To see this, consider the state at time $t' >$

⁷ This can be phrased such that a generic single-sided state is mapped onto a special (thermofield double) two-sided state in the effective description. This occurs because of the state-dependent [46–53] definition of infalling mode operators. For refutations against other arguments [29, 30, 157, 214] for firewalls, see [42–46, 204] and references therein.

$t + O(1/T_H)$ of the system that was described by (3.3.1) at time t . This state can be written as

$$|\Psi(M')\rangle \approx \sum_{n'} e^{S_{\text{bh}}(M' - E_{n'})} \sum_{i'_{n'}=1} \sum_{a'} c_{n'i'_{n'}a'} |\{n'_{\alpha'}\}\rangle |\psi_{i'_{n'}}^{(n')}\rangle |\phi_{a'}\rangle, \quad (3.3.10)$$

where $M' < M - O(T_H)$ is the black hole mass at t' , and $|\{n'_{\alpha'}\}\rangle$, $|\psi_{i'_{n'}}^{(n')}\rangle$, and $|\phi_{a'}\rangle$ are the states of the hard, soft, and far modes *defined at that time*. If $t' - t$ is smaller than the signal propagating time between the edge of the zone and the stretched horizon, then the interior portions of K_t and $K_{t'}$ constructed from the two states in Eqs. (3.3.1) and (3.3.10) overlap. This allows us to represent the same interior operator in $K_t \cap K_{t'}$ differently—as O and O' —using the microscopic degrees of freedom appearing in two vacuum microstates Eqs. (3.3.1) and (3.3.10).

For $M - M' > O(T_H)$, the hard mode states appearing in Eqs. (3.3.1) and (3.3.10) are orthogonal (possibly up to exponentially small overlaps), since the black holes at times t and t' are macroscopically distinguishable; the same is true for the soft and far mode states. It follows that O and O' are not related simply by the original exterior time evolution $U(t', t)$:

$$O' \neq U(t', t) O U^\dagger(t', t). \quad (3.3.11)$$

This can be seen, for example, from the fact that some of the hard modes at time t evolve into far modes at t' , and that for a young black hole, interior operators can be constructed without involving far mode degrees of freedom. The argument here demonstrates that a version of quantum error correction discussed in [11] also applies here in constructing interior operators from the bulk degrees of freedom, i.e. the hard, soft, and far modes.

- (ii) Different operators in the interior region, in particular operators that are independent at the semiclassical level, involve the same microscopic degrees of freedom. This is, in a sense, the converse of the phenomenon described in (i). This violation of independence comes from the fact that soft modes at some time t are not necessarily independent from hard modes at some other time t' , as they can be related by time evolution $U(t, t')$. However, since the number of independent soft mode and radiation states is exponentially larger than that of hard mode states, the degree of violation arising in this way is suppressed exponentially by the factor $e^{-S_{\text{tot}}/2}$.

These two features reinforce the conclusion that the picture of the black hole interior is intrinsically semiclassical: it is defined only up to the precision of order $e^{-S_{\text{tot}}/2}$.

We now discuss the third aspect, which is particularly relevant for our discussion:

- (iii) If the state at time t is given by (3.3.1), then the operator $\Phi_\Gamma(x)$ that acts *on this state* and represents a local Heisenberg-picture operator at $x \in K_t$ in (3.3.8) is given by

$$\Phi_\Gamma(x) = \sum_{s, \Omega, \mathbf{L}} \left(f_s(\Omega, \mathbf{L}) \varphi_{\Omega, \mathbf{L}}(x) a_\xi + g_s(\Omega, \mathbf{L}) \varphi_{\Omega, \mathbf{L}}^*(x) a_{\xi^c}^\dagger \right). \quad (3.3.12)$$

Here, we have decomposed index ξ of the infalling mode operators a_ξ and $a_\xi^\dagger$ into Γ , s , Ω , and $\mathbf{L} = \{\ell, m\}$ which represent species, spin, frequency, and orbital angular

momentum quantum numbers, respectively, with ξ^c representing the CPT conjugate of ξ ; $f_s(\Omega, \mathbf{L})$ and $g_s(\Omega, \mathbf{L})$ are the standard factors providing Lorentz representation of the field (Dirac spinors, polarization vectors, etc), and $\varphi_{\Omega, \mathbf{L}}(x)$ are the mode functions. Note that since a_ξ and $a_\xi^\dagger$ are given by Bogoliubov transformations of $b_\gamma, b_\gamma^\dagger, \tilde{b}_\gamma$, and $\tilde{b}_\gamma^\dagger$ in Eqs. (3.3.4) and (3.3.5), these operators act on the hard, soft, and far mode states defined at t .

As depicted in Fig. 3.3, in order to cover the entire black hole interior I , we need to erect the effective theory of the interior in at least every time interval of order $1/2\pi T_H$. In other words, in order to obtain the complete set of interior operators $\{\Phi_\Gamma(x) | x \in I\}$, we need to use a_ξ and $a_\xi^\dagger$ defined at different times t_x in constructing operators at different locations x , since not all $x \in I$ belong to the same K_t for some t . While operators obtained in this way act generally on the hard, soft, and far mode states at different times t_x , we can time evolve them to obtain the set of operators that all act on the radiation state at time t_r after the black hole is fully evaporated, i.e. $t_r > t_{\text{evap}}$:

$$\mathcal{I}_{t_r} = \{\mathcal{O}_{\Gamma, x}(t_r) \equiv U(t_r, t_x)\Phi_\Gamma(x)U^\dagger(t_r, t_x) | x \in I\}. \quad (3.3.13)$$

As discussed in [45, 204], this is how entanglement wedge reconstruction [12, 16, 162–166, 202] represents all the interior operators in the final state Hawking radiation. We stress that from this perspective, entanglement wedge reconstruction of interior operators from Hawking radiation involves evolution *backward in time*, despite the fact that the interior region and the region in which the radiation exists are spacelike separated in the global spacetime picture.

The operators $\mathcal{O}_{\Gamma, x}(t_r)$ acting on final-state Hawking radiation represent local operators in the interior of the black hole, a region *spacelike separated* from the region where the radiation resides. This seems to imply that an operation on the Hawking radiation corresponding to the application of $\mathcal{O}_{\Gamma, x}(t_r)$ can affect the experience of an observer falling into the black hole, potentially violating the causality of general relativity.⁸ This issue was studied in [169] (see also [21, 23]) for radiation before complete evaporation (the case in footnote 8), wherein it was shown that the operation of acting $\mathcal{O}_{\Gamma, x}(t_r)$ is so highly complex that locality in the semiclassical description (i.e. the spacelike separated nature between the interior and radiation regions) is preserved for all practical purposes.

However, as we will see in the next section, this leaves a major puzzle associated with causality in the global spacetime picture. We will see that the construction given in [42–46, 204] and described here addresses this puzzle. In particular, we will see that in our picture, unlike what was envisioned in earlier works [169, 203, 216], one *cannot* affect the experience of an observer who has fallen into the black hole at an earlier time by performing an (even complex) operation on radiation emitted from the black hole. This is because, at

⁸ This argument can be extended to radiation emitted from an old black hole, i.e. to the case $t_{\text{Page}} < t_r < t_{\text{evap}}$, where t_{Page} is the Page time [41]. In this case, the amount of information in the relevant interior region, $(\cup_{t < t_r - t_{\text{scr}}} K_t) \cap I$, accessible in the radiation is limited and depends on the location in this region, where t_{scr} is the scrambling time [68, 69]. See Fig. 13 of [204] and discussion associated with it for more details.

the microscopic level, the region inside the black hole should be regarded as being *timelike separated* from the region in which the radiation resides; more specifically, the former is in the past of the latter. The fact that operators like $\mathcal{O}_{\Gamma,x}(t_r)$ exist simply reflects the standard statement that there are operators at a time t corresponding to changing the *record* of what happened at times earlier than t . Together with the fact that the $\mathcal{O}_{\Gamma,x}(t_r)$ effectively commute with a class of simple operators acting on radiation—the feature arising from the involvement of complex horizon dynamics—this allows us to treat the past event represented by $\mathcal{O}_{\Gamma,x}(t_r)$ as occurring in a spacelike separated region in semiclassical spacetime.

3.4 A Puzzle in the Global Spacetime Picture and Its Resolution

In this section, we discuss a problem with global spacetime when it is used to formulate standard quantum mechanics at the fundamental level. We will then show that by using the description of the black hole interior presented in the previous section, this problem is avoided. In this description, the Hawking radiation containing information about the interior always lies in the causal future of the corresponding interior region. It is only in the semiclassical description that the interior appears causally disconnected from the radiation region.

We will discuss in detail how this apparent change of the description arises from the complex dynamics at the horizon, as viewed from the exterior. By the semiclassical description, we mean the description including only finite products of local operators in the limit that the Newton constant is sent to zero, $G_N \rightarrow 0$. For operators acting on the final-state Hawking radiation, this corresponds to operators that act on a number of degrees of freedom which does not scale with the entropy of the radiation.⁹ We will show that these operators all effectively commute with operators describing the interior, so that they can be viewed as spacelike separated at the semiclassical level.

3.4.1 A problem with the microscopic description of global spacetime

We argue that the following three postulates cannot be satisfied at the same time:

- (i) Diffeomorphism invariance persists throughout inextendible (global) spacetime, i.e. the region in which the classical description of spacetime is available in general relativity.
- (ii) The canonical formalism of quantum mechanics can be built using Cauchy surfaces in global spacetime, at least when they satisfy certain “niceness” properties [217–219].

⁹ This implies that the semiclassical theory does not describe *everything* about the low energy degrees of freedom.

- (iii) An observer in the exterior region can reconstruct local operators in the interior using only Hawking radiation degrees of freedom.

We will show this by demonstrating that the above three postulates, taken together, lead to a violation of the no-cloning theorem of quantum mechanics.

To see this, consider a setup in which a detector staying outside the horizon throughout the history of the black hole evolution extracts information about the interior at some time t_r after the Page time t_{Page} .¹⁰ For simplicity, we focus on the case in which t_r is after the full evaporation of the black hole, $t_r > t_{\text{evap}}$, though the extension to the case with $t_{\text{Page}} < t_r < t_{\text{evap}}$ is straightforward (see footnote 8).

While the operation performed by the detector to achieve the information extraction needs to be highly nonlocal, acting on most of the radiation degrees of freedom, we can fine-tune the initial configuration of the detector so that it reads off information corresponding to a semiclassical excitation in the interior.¹¹ As a concrete example, we can set up the initial configuration of the detector such that it interacts with the degree of freedom in the Hawking radiation which is represented by a product of operators $\mathcal{O}_{\Gamma,x}(t_r)$ for some spacetime point $x \in I$. Note that this requires an exponential fine-tuning of the configuration of the detector, which must be determined by the microscopic theory using the knowledge of the initial state of collapsing matter forming the black hole. This is difficult in practice but possible in principle, because we can have as much time as we need to prepare the detector before actually forming the black hole.

With this setup, we find it problematic to consider the global spacetime as a platform for a regular quantum mechanical system. In particular, we cannot employ the canonical formalism of quantum mechanics using a series of Cauchy surfaces, such as nice slices [217–219], in which the information extracted by the detector is removed from the future of x in the interior and appears in the future of r in the exterior region. This can be seen by considering a foliation of spacetime accommodating a time slice that goes through both the past of x and the future of the region in which the detector extracts the information; see Fig. 3.4. On this slice, the information extracted by the detector is cloned, since the destruction of the same information in the interior occurs at x and hence has not yet happened on this time slice. This contradicts the no-cloning theorem, hence violating the superposition principle of quantum mechanics. We stress that this problem occurs even when we only use slices that are nice in the semiclassical description.

One might think that this problem is avoided by claiming that the microscopic description of the global spacetime forces us to use a special foliation in which the time slice going through

¹⁰ Strictly speaking, t_r is not exactly a single time but has a “width,” since it will take some time for the detector to extract the information from Hawking radiation through interactions with it; but we will loosely refer to the period of time in which the detector extracts the information as t_r .

¹¹ The corresponding operator acting on the radiation comprises a product of S_{rad} local operators, where S_{rad} is the coarse-grained entropy of the radiation, so that the required fine-tuning is of order $e^{-O(S_{\text{rad}})}$. We do not consider such an operator to be a part of the semiclassical theory. Thus, the problem discussed here is not that of the semiclassical description but of the fundamental description applied to the global spacetime.

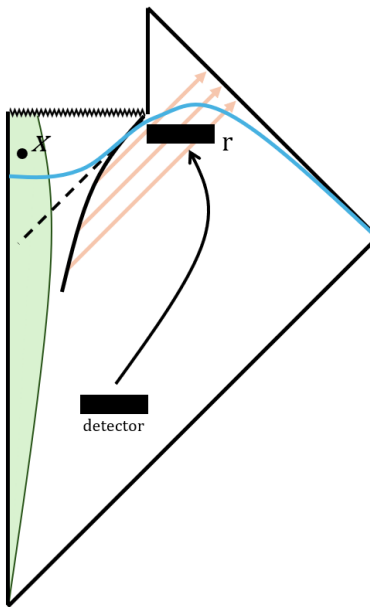


Figure 3.4: A detector can be set up such that it extracts from Hawking radiation the information that is located at a spacetime point x in the interior of the black hole. Interpreting this as being caused by nonlocal interactions between x and the spacetime region r in which the radiation resides, we encounter the problem of the existence of a time slice (blue line) that goes through both the future of r and the past of x . On such a slice, the information in question is cloned, leading to a contradiction with the no-cloning theorem of quantum mechanics.

t_r must also go through x . We find it difficult, however, to make sense of such a claim given that the arbitrariness of choosing time foliation, i.e. diffeomorphism invariance, plays an important role in ensuring the consistency of theories with dynamical gravity, even at the perturbative level.¹²

3.4.2 A solution: the interior as an effective description of dissipation on the horizon

The above argument implies that, if we are to keep (iii), we must relinquish (i) or (ii). The understanding of the black hole interior described in Section 3.3 addresses the puzzle by abandoning postulate (ii). In this framework, the interior spacetime—in fact only a portion of it—emerges as an effective description associated with a given exterior time. The effective

¹² This is a difficulty which we encounter in addition to the fact that the time t_r has a width, so that the relevant time slice would generally be ill-defined.

time evolution foliating the interior region is different from the more fundamental, exterior time evolution, which foliates only the spacetime region outside the horizon. Consequently, any information in the effective interior is contained in modes *outside the horizon*, and hence exists in the causal past of the radiation. By viewing the interior this way, we need not demand that the interior is foliable by global time slices, and hence do not take postulate (ii). The property in postulate (i) emerges only effectively after patching the pictures obtained by effective theories erected at different times. This results in postulate (iii) not implying that an operation on Hawking radiation can affect the physics in the interior, avoiding the problem described in Section 3.4.1.

Recall that in the framework of Section 3.3, physics occurring at the interior point x is described by an effective theory erected at a time t_x ($< t_{\text{evap}}$) satisfying

$$x \in K_{t_x}, \quad (3.4.1)$$

where K_t is given by (3.3.8).¹³ As can be seen in Fig. 3.5, the spacetime region K_{t_x} described by the effective theory contains (a portion of) the zone, $r_s < r < r_z$, only up to the time

$$T_{t_x} = t_x + 2r_+ \ln \frac{r_+}{l_P}, \quad (3.4.2)$$

where r_+ is the horizon radius of the black hole *at time* t_x , and we have ignored the term of order r_+ which is not enhanced by the logarithm $\ln(r_+/l_P)$. (We will also do the same below.) This implies that in the exterior description, the physics occurring in the interior portion of K_{t_x} should be viewed as occurring in the time frame

$$t_x < t < T_{t_x}, \quad (3.4.3)$$

since the effective theory simply reorganizes correlations between the degrees of freedom associated with the time slices in (3.4.3) such that they can be interpreted as arising from the infalling time evolution generated by $\tilde{H}$ in (3.3.9), *without* involving a further exterior time evolution. The earliest effective theory to which x can be associated is the one erected at time $t_{\min} = \min\{t_x : x \in K_{t_x}\}$. The latest effective theory which x can be associated with is similarly the one erected at $t_{\max} = \max\{t_x : x \in K_{t_x}\}$. Consequently, x must be associated with exterior modes between $t_{\min}$ and $T_{t_{\max}}$. In other words, the physics occurring in the interior spacetime point x must be viewed as occurring in the time frame

$$t_{\min} < t < T_{t_{\max}} \quad (3.4.4)$$

in the exterior description.

¹³ It is important that K_{t_x} only contains the future of the region *outside* the stretched horizon; it does not contain the region which is in the future of $r < r_s$ at t_x but not in the future of $r > r_s$ at t_x , as illustrated in Fig. 3.5. This ultraviolet cutoff is a manifestation of the fact that the effective theory of the interior is intrinsically semiclassical, without which all the radial null geodesics entering the stretched horizon before t_x would pile up along the horizon of the second exterior in the effective theory causing large backreaction on the spacetime [42].

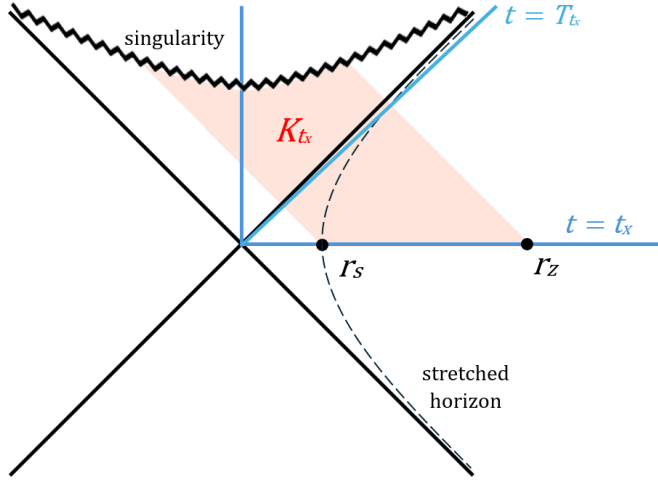


Figure 3.5: The physics occurring in the region described by the effective theory erected at time t_x cannot be obtained by any effective theory erected after $T_{t_x} = t_x + 2r_+ \ln(r_+/l_P)$, implying that it must be viewed as having occurred before T_{t_x} in the exterior description.

Importantly, for any point x in the interior of the black hole

$$T_{t_{\max}} < t_{\text{evap}}. \quad (3.4.5)$$

This implies that in the exterior description, the physics which an interior observer experiences at the spacetime point x must have *already happened* before the black hole evaporates. Thus, whatever the detector does to the final state Hawking radiation, it *cannot affect* their experience at x nor events occurring in its future. Exerting operators $\mathcal{O}_{\Gamma,x}(t_r)$ on the radiation at t_r can *reconstruct* the event at x , or even *rewrite the record* of what happened at x and in its future; however, it cannot change the actual experience which the interior observer has had inside the black hole.

In this respect, the situation is not so different from the standard situation in which a complicated operation can be applied to the system “after the fact” so that it can change the record of what has happened before the operation.¹⁴ In our setup, if we reconstruct from Hawking radiation what happened in the interior of the black hole before applying the complex operation, i.e. at $t < t_r$, then we would conclude that nothing extraordinary happened at x in the interior. On the other hand, if we do such a reconstruction after the detector exerts the complex operation, i.e. at $t > t_r$, then we would conclude that an event occurred at x “out of the blue” which did affect the future of x in the interior. (The

¹⁴ Throughout the paper, we assume that things happen when they actually occur through time evolution. In other words, we consider that a measurement occurs when the relevant information corresponding to the measurement outcome (which is not full quantum information) is amplified into various subsystems [220–222]. In particular, we do not take an extreme (and what we think an exotic) standpoint [223] that a measurement occurs only when the relevant information reaches to a boundary or horizon region.

past of x is unaffected because the boundary condition for the interior is determined by the exterior state at t_x , fixing the state in the past light cone of x .) This “nonlocal influence” by the operation, however, is an illusion—what a physical observer *actually experiences* in the interior is not affected by the operation of the detector.

In the exterior description, the event x must thus be viewed as having occurred *in the past of r* , despite the fact that they are spacelike separated in the semiclassical description. This picture becomes evident because we have employed the construction in [42–46, 204], which makes manifest the dynamical nature of interior reconstruction, i.e., how the exterior time evolution is involved in the reconstruction. In fact, this feature is not visible in a model without appropriately taking into account time evolution, for example in the case of a tensor network corresponding to a single global time slice.

An alternative way of viewing all of this is that to correctly reconstruct the interior, we need to include the agent, e.g. a detector, which has exerted operators $\mathcal{O}_{\Gamma,x}(t_r)$ on the radiation. We should then evolve the *combined* system of the agent and radiation backward in time and use the construction of [42–46, 204] at time t_x . By doing this, rather than simply adopting the entanglement wedge reconstruction protocol to the radiation after $\mathcal{O}_{\Gamma,x}(t_r)$ ’s are operated, we will indeed reconstruct the interior in which nothing extraordinary happens at x . The inclusion of the agent affects the interior significantly, despite the fact that it comprises a much smaller number of degrees of freedom than those of the radiation. This is because the standard process of black hole formation and evaporation is one in which the coarse-grained entropy increases [107, 108], and hence the reverse time evolution employed during interior reconstruction is a highly fine-tuned, fragile process.

The analysis here can be extended to the case in which we act operators $\mathcal{O}_{\Gamma,x}(t_r)$ before the black hole is fully evaporated, which is possible if $t_r > t_{\text{Page}}$. To see this, consider an infalling observer who reaches the stretched horizon at some exterior time t_f . While the experience of the observer shortly after crossing the horizon is described by the effective theory of the interior erected at t_f , this does not necessarily mean that all the experiences of this observer are described in this effective theory. For example, the observer may accelerate toward the horizon right after they cross it, eventually leaving K_{t_f} ; see Fig. 3.6. Even in this case, however, the experience of the infallen observer should be viewed as fully completed by

$$t_{\text{comp}} = t_f + t_{\text{scr}} \quad (3.4.6)$$

in the exterior description, where

$$t_{\text{scr}} = 4r_+ \ln \frac{r_+}{l_{\text{P}}} \quad (3.4.7)$$

is the scrambling time. One way to see this is to notice that the late trajectory of the observer is fully contained in the effective theory erected at $t'_f = t_f + 2r_+ \ln(r_+/l_{\text{P}})$, even in the worst case that the observer starts moving outward at the speed of light one cutoff time after crossing the mathematical horizon. This is indicated by the arrow in Fig. 3.6. The argument that led to (3.4.3) then says that any experience of the observer in the interior must have occurred by

$$T'_{t'_f} = t_f + t_{\text{scr}}. \quad (3.4.8)$$

On the other hand, the general argument in [68] (or entanglement wedge reconstruction [18]) says that the information about an object that fell into the black hole at t_f emerges in

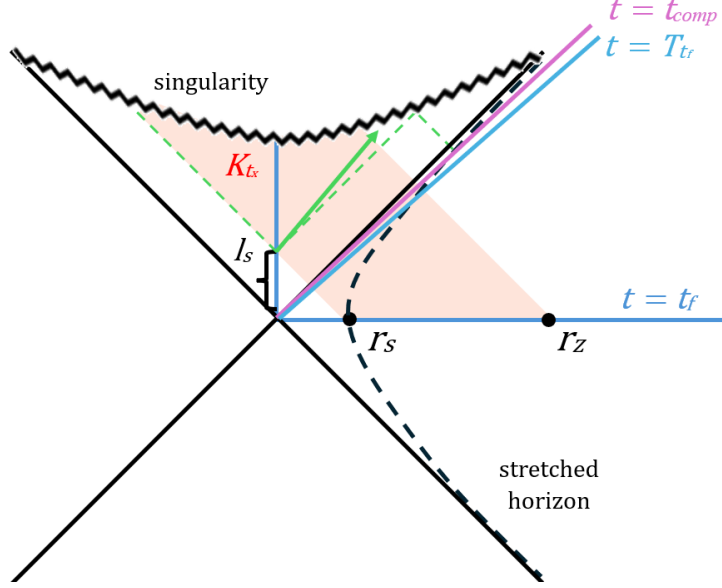


Figure 3.6: The physics associated with an object that fell into the horizon at time t_f cannot be described by any effective theory erected after $t_{\text{comp}} = t_f + t_{\text{scr}}$. This implies that it must be viewed as having occurred before t_{comp} in the exterior description.

Hawking radiation only after $t_f + t_{\text{scr}}$. Thus, again any operation on Hawking radiation—even if it is performed before t_{evap} , and no matter how complex it is—cannot affect the experience of an observer in the interior. This conclusion holds even if we mine [115, 116] radiation from the Hawking cloud in the zone around the black hole.¹⁵

Finally, our analysis yields the following interpretation [42] of the black hole’s singularity. From an exterior perspective, an object does not immediately thermalize when it encounters the stretched horizon. Instead, it manifests as a distinct excitation on the horizon for some duration. Over the scrambling time, t_{scr} , this excitation will gradually dissipate into the horizon. During this period, the dissipation process has an effective description as an object traversing the near-empty interior of the black hole. Here, coherent “quasi-particle” horizon modes—and beyond the Page time, early radiation modes as well—denoted by $\tilde{b}_\gamma$, serve as modes in the effective second exterior of the black hole. The singularity of the semiclassical picture signifies the limit of the applicability of this effective description. After the object’s excitation is fully integrated into the horizon, the effective theory describing the dissipation process is no longer available.

¹⁵ To reconstruct the interior of the black hole at an earlier time t from radiation at t_r ($> t, < t_{\text{evap}}$), we need to include all semiclassically distinguishable branches (with different black hole masses and positions) that have resulted from the backreaction [99, 100] of Hawking emission occurring after t .

3.4.3 Converting relations from timelike to spacelike using complexity

It is highly nontrivial to regard two sets of degrees of freedom that are not independent with each other as spacelike separated at the semiclassical level. If interior reconstruction indeed works, Hawking radiation for an old black hole is not independent of the interior degrees of freedom. To regard these as spacelike separated, while preserving causality at the semiclassical level, an operator acting on one set must commute with that acting on the other, possibly up to corrections suppressed exponentially in a macroscopic entropy. Here we argue that our framework indeed implies that a process involving only low energy degrees of freedom cannot affect the semiclassical physics in a region that can be viewed as spacelike separated.

In [169], Kim, Tang, and Preskill (KTP) analyzed this issue of causality, assuming that the black hole's dynamics can be represented by pseudorandom unitaries [24, 169, 175]. Since it is expected that chaotic dynamics generally produce pseudorandom states sufficiently after the scrambling time, we assume that the same applies to the dynamics of our horizon modes as viewed from the exterior. KTP studied a partially evaporated black hole and concluded that if an operator $\mathcal{O}_{\text{rad}}$ acting on the radiation can be modeled by a quantum circuit with size polynomial in the entropy of the black hole, i.e. $\mathcal{C}(\mathcal{O}_{\text{rad}}) \sim \text{poly}(S_{\text{BH}})$, then $\mathcal{O}_{\text{rad}}$ commutes with operators $\mathcal{O}(x)$ describing the interior up to corrections suppressed exponentially in S_{BH}

$$\|[\mathcal{O}_{\text{rad}}, \mathcal{O}(x)]\| \sim O(e^{-S_{\text{BH}}}). \quad (3.4.9)$$

Applying the same argument to the final-state Hawking radiation, we could conclude that an operator $\mathcal{O}'_{\text{rad}}$ acting only on a portion S'_{rad} ($< S_{\text{rad}}$) of the radiation whose complexity goes as

$$\mathcal{C}(\mathcal{O}'_{\text{rad}}) \sim \text{poly}(S_{\text{rad}} - S'_{\text{rad}}) \quad (3.4.10)$$

would effectively commute with operators describing the interior of the black hole:

$$\|[\mathcal{O}'_{\text{rad}}, \mathcal{O}(x)]\| \sim O(e^{-(S_{\text{rad}} - S'_{\text{rad}})}). \quad (3.4.11)$$

However, this by itself does not solve the issue, since one would still be able to act on an operator on the entire final-state radiation ($S'_{\text{rad}} \approx S_{\text{rad}}$) and affect the interior; indeed in that case the operator can be simple.¹⁶ This would lead to the problem discussed in Section 3.4.1. As we have seen, however, the Hawking radiation containing information about the interior is, in fact, *timelike* related with the interior degrees of freedom. The former is in the future

¹⁶ Here we have assumed, following [169], that the state of final-state Hawking radiation is efficiently generated. It seems possible to us that the generation of final-state Hawking radiation could be hard, which can happen if, for example, the computational basis of Hawking radiation (the basis determined by spatial locality) and the natural basis of the horizon dynamics are related by a generic unitary. If this is the case, then the operator affecting the interior must be complex, with the complexity $\mathcal{C}$ scaling exponentially in S_{rad} .

of the latter. Therefore, one cannot change the physics that has already occurred in the interior by acting with an operator on the Hawking radiation, *no matter how complex it is*.

This leaves the question about the effect of an interior operator (past) on Hawking radiation (future). In order to preserve causality, we want the semiclassical physics associated with the radiation not to be affected by an operation in the interior. By semiclassical physics, we mean the collection of processes represented by operators which act on degrees of freedom whose numbers do not scale with S_{rad} , which is proportional to the initial entropy of the black hole and hence to $1/G_N$. To study this question, we can employ a variant of the KTP analysis; the interior operator in KTP is replaced with a semiclassical operator acting on the radiation, and the combined system of the black hole and radiation in KTP is replaced with the black hole degrees of freedom at the time when the effective theory of the interior is erected.¹⁷ The number of degrees of freedom representing semiclassical physics in the interior is much smaller than that representing the black hole, S_{BH} . The KTP argument then implies that an interior operator $\mathcal{O}_{\text{int}}$ corresponding to a circuit of size

$$\mathcal{C}(\mathcal{O}_{\text{int}}) \sim \text{poly}(S_{\text{BH}}) \quad (3.4.12)$$

effectively commutes with semiclassical operators $\tilde{\mathcal{O}}_{\text{rad}}$ acting on the radiation, up to corrections suppressed exponentially in S_{BH} .

$$\|[\mathcal{O}_{\text{int}}, \tilde{\mathcal{O}}_{\text{rad}}]\| \sim O(e^{-S_{\text{BH}}}). \quad (3.4.13)$$

In other words, in order to affect the semiclassical physics of the radiation, an operation performed in the interior must have complexity exponential in S_{BH} . This is, however, impossible because any physical “experimenter” in the interior hits the singularity in polynomial time before any such operation can be performed. Note that we have not imposed any constraints on the computational complexity of $\tilde{\mathcal{O}}_{\text{rad}}$; we have just imposed the restriction on the number of degrees of freedom on which it acts.

Any physical operation in the black hole interior, therefore, cannot affect semiclassical physics occurring in the radiation at later times, except for corrections suppressed exponentially in the entropy of the black hole. The only operators acting on the radiation which do not commute with interior operators are those acting on a large (macroscopic) number of degrees of freedom scaling as S_{rad} . Of course, these noncommutativities are what allow us to decode the information about the interior from the Hawking radiation.

Finally, we note that the analogous statement cannot be made for reverse time evolution. We *can* perform a physical operation on the (now initial-state) radiation which affects semiclassical physics in the interior (of the white hole). In this sense, we may say that the Penrose diagram of Fig. 3.1 faithfully represents physics for the standard black hole evolution, a process in which the coarse-grained entropy increases, but not for the reverse evolution, a process with decreasing coarse-grained entropy caused by a finely-tuned initial condition.

¹⁷ The relation between the degrees of freedom representing the interior and Hawking radiation always involves the horizon dynamics, which we assume to be pseudorandom. The analysis of KTP can thus be adopted in the present context.

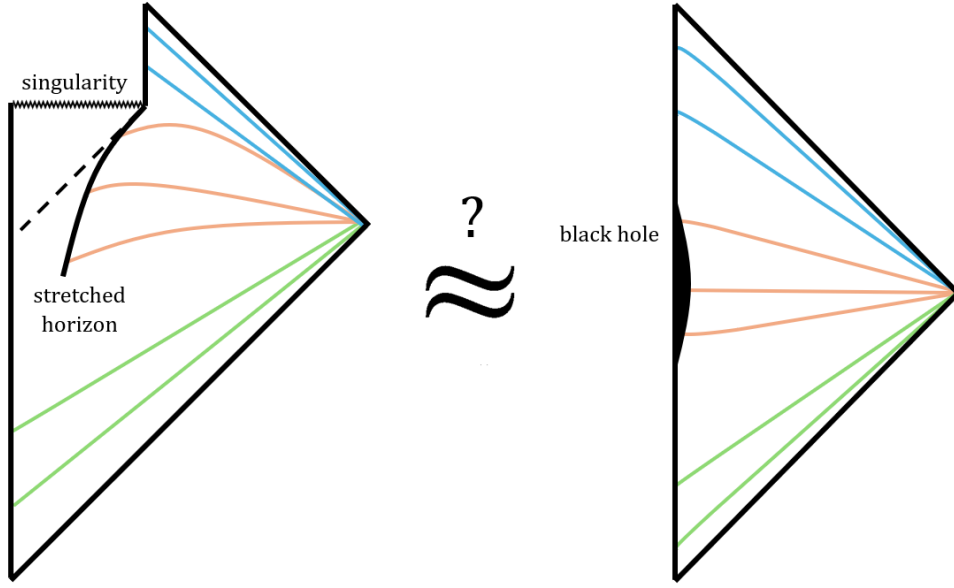


Figure 3.7: In the exterior description, the causal flow of information is obvious after the black hole is formed. The question is if there exists an analogous picture throughout the history of the black hole evolution, including the times during its formation.

3.5 Causal Structure of Black Hole Formation and Evaporation

One virtue of the exterior picture is that the causal flow of information is clear. When an object falls into a black hole, its information stays on the stretched horizon, whose space-time trajectory is timelike, and later comes back to ambient space in the form of Hawking radiation. One may thus wonder if the causal structure of the entire spacetime can in fact be trivial in the exterior description (or at least in some version of the exterior description). We depict how such a description would look in the right panel of Fig. 3.7.

This would indeed be the case if the “stretched horizon,” at which the spacetime terminates, follows a timelike trajectory *throughout* the history of the black hole evolution, including during formation. Taking this possibility seriously, however, raises an interesting question. To see this, imagine that a diary has existed at $r = 0$ for an extended period of time in near-empty spacetime. Then, we fire a null shell of particles toward $r = 0$, creating a black hole of mass M ; see Fig. 3.8. Before the null ray, the spacetime is Minkowski. After the null ray, the spacetime is Schwarzschild. Given the unitarity of the S -matrix, the information about the diary must eventually be emitted from the stretched horizon in the form of Hawking radiation. In order to avoid a mysterious spacelike motion of the information from the interior to the exterior, the timelike trajectory of the stretched horizon must extend to the spacetime region before the null ray, as depicted by a curved dashed line in Fig. 3.8.

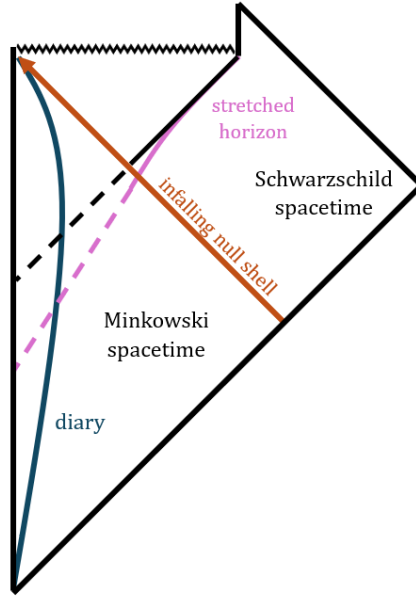


Figure 3.8: Consider a diary in the region near $r = 0$ far in the past. We then fire a null shell of infalling particles toward $r = 0$, creating a black hole of mass M . The unitarity of the S -matrix implies that the information about the diary must eventually be emitted from the stretched horizon in the form of Hawking radiation. An exterior description with trivial causality, the right panel of Fig. 3.7, is obtained by extending the stretched horizon toward the past using the event horizon.

However, the spacetime in this region is Minkowski, and there is no obvious locally defined surface which can play the role of the stretched horizon.

3.5.1 Exterior description with trivial causal structure

We argue that the description corresponding to the right panel of Fig. 3.7, however, is still possible. In particular, we will provide below a surface which plays the role of the stretched horizon before the null shell, which in the case of Fig. 3.8 is the appropriately stretched horizon of a spherical Rindler observer. Before the black hole’s formation, the information about the diary stays in modes near this horizon, and it will eventually come out in a causal way in the form of Hawking radiation. This is possible because in quantum gravity the spatial location of information is not uniquely determined.

In quantum gravity, the physics associated with some spacetime region can be encoded in the degrees of freedom on the boundary surrounding it, if certain conditions are met [7–9]. In particular, for a closed codimension-2 surface Σ , the physics occurring in the domain of dependence of a partial Cauchy surface bounded by Σ can fully be described using the intrinsically quantum gravitational degrees of freedom on Σ and the degrees of freedom

outside of it.

We apply this to each Cauchy surface before the formation of the black hole, with Σ identified as the (stretched) *event horizon*. Specifically, in the setup of Fig. 3.8, we may adopt the ingoing Vaidya metric

$$ds^2 = -\left(1 - \frac{2M(v)}{r}\right)dv^2 + 2dvdr + r^2d\Omega^2, \quad (3.5.1)$$

where

$$M(v) = \begin{cases} 0 & \text{for } v < v_0 \\ M & \text{for } v > v_0 \end{cases}, \quad (3.5.2)$$

with v_0 the ingoing time at which the null shell is shot. For $v > v_0$, this corresponds to time slices of the exterior description, as can be seen from the fact that the coordinate transformation

$$v = t + r + 2Ml_{\text{P}}^2 \ln \frac{r - 2Ml_{\text{P}}^2}{2Ml_{\text{P}}^2} \quad (3.5.3)$$

gives the Schwarzschild metric. For $v < v_0$, on the other hand, the metric also covers a region inside the event horizon: by the coordinate transformation

$$v = t + r, \quad (3.5.4)$$

we obtain the flat space metric

$$ds^2 = -dt^2 + dr^2 + r^2d\Omega^2. \quad (3.5.5)$$

The trajectory of the event horizon in this metric is given by

$$r = t + 4Ml_{\text{P}}^2 - v_0 \quad (3.5.6)$$

for $v_0 - 4Ml_{\text{P}}^2 \leq t \leq v_0 - 2Ml_{\text{P}}^2$.

We may convert the metric (3.5.5) into that corresponding to the “exterior picture,” by adopting the spherical Rindler coordinates [224] tailored to the present setup

$$\tau = \operatorname{arctanh} \frac{t - t_*}{r}, \quad \rho = \sqrt{r^2 - (t - t_*)^2}, \quad (3.5.7)$$

where

$$t_* = v_0 - 4Ml_{\text{P}}^2 \quad (3.5.8)$$

is the time at which the event horizon forms. This gives the metric

$$ds^2 = -\rho^2 d\tau^2 + d\rho^2 + \rho^2 \cosh^2 \tau d\Omega^2, \quad (3.5.9)$$

applicable in the region $v < v_0$, i.e. $\rho < 4Ml_{\text{P}}^2 e^{-\tau}$, and $\tau \geq 0$. The Rindler horizon agrees with the event horizon (3.5.6), which connects smoothly with the black hole horizon at $v = v_0$. While the system is not in the global thermal equilibrium as the metric (3.5.9) is not static, observers at constant ρ follow trajectories of constant acceleration and see Unruh

radiation [70, 225, 226]. The stretched horizon can thus be defined to be located where the local (Tolman) Unruh temperature $1/2\pi\rho$ becomes the string scale $1/2\pi l_s$ [45, 227]:

$$\rho = \rho_s \approx l_s. \quad (3.5.10)$$

This surface connects smoothly with the stretched horizon of the black hole at $v = v_0$. At the time of the formation, the area of the stretched horizon is of the cutoff-scale size, $\mathcal{A}(0) \approx 4\pi l_s^2$. The nonstretched and stretched horizons are depicted by dashed lines in Fig. 3.8.

We claim that when the theory is quantized canonically, there is a description which keeps only the portion of the spacetime outside the stretched horizon, $\rho > \rho_s$, and in which the stretched horizon carries the coarse-grained entropy (the number of effective degrees of freedom)

$$S_h(\tau) = \frac{\mathcal{A}(\tau)}{4l_p^2}, \quad (3.5.11)$$

where

$$\mathcal{A}(\tau) = 4\pi\rho_s^2 \cosh^2 \tau \quad (3.5.12)$$

is the area of the stretched horizon at time τ (≥ 0).¹⁸ When the stretched horizon intersects the null shell at

$$\tau = \ln \frac{v_0 - t_*}{\rho_s} = \ln \frac{4Ml_p^2}{\rho_s}, \quad (3.5.13)$$

$S_h(\tau)$ and $\mathcal{A}(\tau)$ agree with the entropy and the area of the stretched horizon of the black hole of mass M , respectively. We postulate that, as suggested by the holographic principle, the degrees of freedom associated with the stretched horizon, represented by (3.5.11), encode semiclassical physics in the domain of dependence of a partial Cauchy surface bounded by the stretched horizon at time τ

$$D(\Xi), \quad \Xi: r < \rho_s \cosh \tau, t = t_* + \rho_s \sinh \tau. \quad (3.5.14)$$

Thus, when the diary hits this past-extended stretched horizon, its information is stored in these degrees of freedom. This information will then be transferred to the black hole horizon degrees of freedom when the stretched horizon is hit by the incoming null shell.

To develop intuition about our picture, we can employ a tensor network model [119, 120, 123]. In a tensor network, a semiclassical excitation, in this case the diary, can be regarded as a perturbation on the network. We can insert this perturbation on a specific location in the network, but the same information can then be “pushed” through the network and be viewed as existing in external legs of the network, which typically are identified as the boundary degrees of freedom in holography. The location of this information, however, need not be either the original location of the perturbation or the external legs of the network. In particular, we can cut the network into two at a closed surface Σ which encloses the original location of the perturbation; see Fig. 3.9. The information about the perturbation can then be viewed as existing on legs that go through Σ . In fact, if all the information about the

¹⁸ The calculation of [224] gives the von Neumann entropy, which is zero in the present setup.

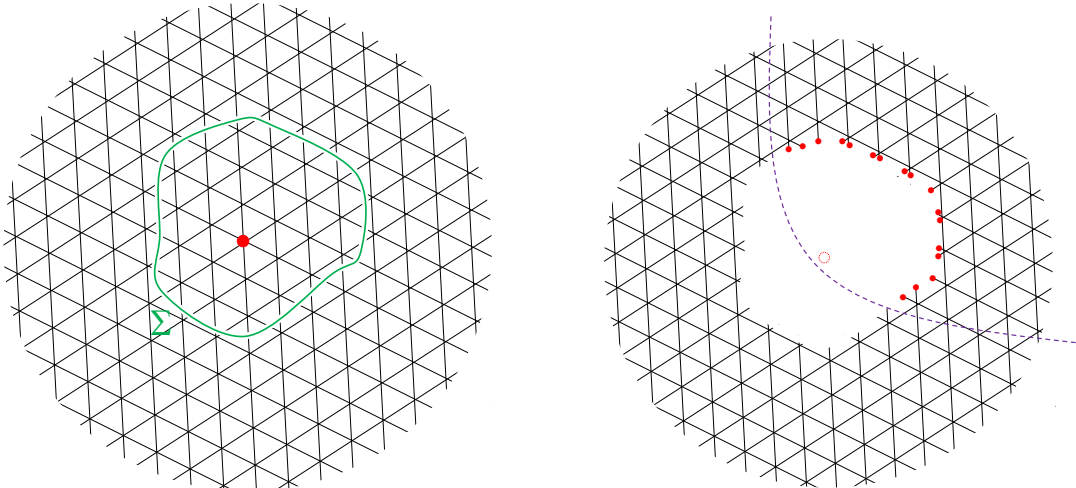


Figure 3.9: An excitation in a semiclassical theory can be viewed as a perturbation on the network of entanglement representing a spatial slice (left). The network can be cut at a closed surface Σ surrounding the perturbation and be replaced with an equivalent network in which the interior of Σ is removed (right). An operator corresponding to the original excitation can be represented by a nonlocal operator acting on the internal legs of the new network, obtained by pushing the original operator through the network.

inner portion of the network can be represented in this way on the cut legs of the outer portion, then we need not even keep the inner portion to describe the system.

Our description conforms with the recent proposal of Bousso and Penington [203, 216] which extends the concept of entanglement wedge reconstruction to bulk subregions. According to their criterion, the max-entanglement wedge $e_{\max}$ [203] of the partial Cauchy surface outside the extended stretched horizon at any time is the entire spacetime. This is consistent with our claim, given that entanglement wedge reconstruction assumes the full knowledge about the (exterior) time evolution, so that it can reconstruct the region $D(\Xi)$ (3.5.14) for all t .

Our description may also be obtained by “pulling in” the boundary of the (putative) holographic theory [35, 36, 124]. Imagine that there is a “screen” in the asymptotic region $r \rightarrow \infty$ on which the holographic theory lives. (In asymptotically AdS spacetime, this can be the conformal boundary.) Each time slice of the screen is called a leaf. We can then renormalize the leaf at each time, which moves the leaf inward in the bulk. As discussed in [36], when doing this at the quantum level, we need to keep the information outside the renormalized leaf $\sigma(\lambda)$, where λ is the flow parameter. Specifically, the flow is given by the equation

$$\frac{dx^\mu}{d\lambda} = \frac{1}{2}(\Theta_k l^\mu + \Theta_l k^\mu), \quad (3.5.15)$$

where x^μ are the embedding coordinates of the renormalized leaf σ , and $\{k^\mu, l^\mu\}$ are the future-directed null vectors orthogonal to σ . Here, $\Theta_{k,l}$ are the modified versions [36] of the quantum expansions [127] in the corresponding directions, whose computation requires

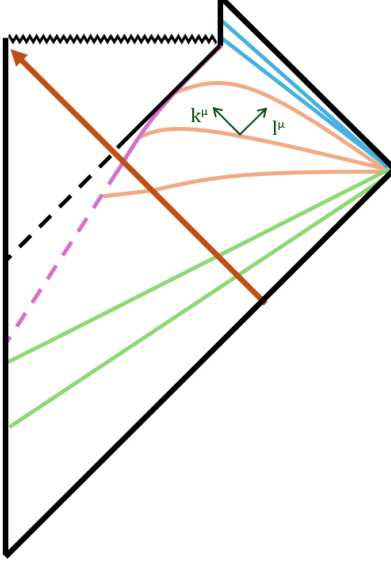


Figure 3.10: An equal-time hypersurface of the holographic boundary, called a leaf, can be pulled inward by a renormalization flow. The trajectory of each leaf, depicted in the figure, then provides an equal-time hypersurface in the bulk called a holographic slice in [35, 36]. We can terminate the flow when the renormalized leaf reaches the stretched event horizon, giving the manifestly causal exterior description.

knowledge about the state outside σ . In the Schwarzschild region the flow terminates near the horizon, but in the Minkowski region it can go all the way toward $r = 0$. Our description can, thus, be viewed as the holographic theory in which each leaf is renormalized down to the extended stretched horizon $r = r_s(t)$ in the Minkowski region. The region $r < r_s(t)$ is then obviously encoded in the renormalized theory defined in the region $r \geq r_s(t)$. For a sketch of the trajectory of the renormalized leaves, see Fig. 3.10.

We expect that the exterior description presented here can be generalized to the case of arbitrary (not necessarily spherically symmetric) collapsing matter. The basic idea is to stretch the event horizon in such a way that the system of the stretched and nonstretched horizons appears locally Rindler. One possible procedure to extend the stretch horizon to earlier times is as follows: consider a congruence of past-directed timelike curves $x_p^\mu(\tau)$, each of which starts from a point p on the stretched horizon at a late time t_0 in the direction orthogonal to the equal-time hypersurface $t = t_0$. Here, τ is the proper time along each curve, with $\tau = 0$ on the stretched horizon. We require that each curve maintains proper acceleration a^μ of

$$\sqrt{a^\mu a_\mu} \approx \frac{1}{l_s} \quad (3.5.16)$$

in the direction orthogonal to the hypersurface given by the spacetime trajectory of the congruence. An equal-time hypersurface in the exterior spacetime region can be chosen in such a way that it agrees with a constant τ surface on the congruence. If caustics develop in

the congruence, we terminate the corresponding curves there. We expect that a procedure like this works, at least, for simple spacetimes, but we leave a detailed study of the stretching of the event horizon for future work.

3.5.2 Significance of the event horizon

The discerning reader may note that the argument presented in Section 3.5.1 relies primarily on the holographic principle, and may be applied to any timelike surface outside the event horizon, not just the stretched event horizon. For example, one might consider in the setup of Fig. 3.8 an extension of the stretched horizon defined by some time-dependent radius $r_t(t)$, whose spacetime trajectory is timelike and which agrees with the black hole horizon at the infalling null shell, such as

$$\frac{dr_t(t)}{dt} > 0, \quad r_t(t_{\text{null}}) = r_s. \quad (3.5.17)$$

Here, t_{null} is the time at which the trajectory $r_t(t)$ of the surface intersects the null shell, and r_s is the radius of the stretched horizon of the black hole of mass M . We call this trajectory the “tube.” One can imagine that the tube is capped off in the past at some time t_0 : $r_t(t_0) = 0$.

Such a description may indeed exist, but the dynamics on the tube would then be very different from the dynamics of the horizon. To see this, consider a trajectory of a particle which enters and then leaves the tube in the global spacetime picture. In the exterior picture, this should be described as a process in which a particle hits the surface of the tube at some time and then some time later the same particle is spit out from a different point on the surface. This implies that, unlike the case of the stretched horizon, the dynamics of the surface are not scrambling and chaotic. We do, however, expect that these properties are associated with the (stretched) null surface, since a particle inside the surface becomes causally disconnected from the exterior.

What is the significance, then, of the event horizon compared with other null surfaces? For these other surfaces, such as the horizons of the standard and spherical Rindler spaces, the area of the horizon is either infinite or increases without bound. We thus expect that information absorbed into the corresponding stretched horizon will not be sent back to the ambient space, since the horizon degrees of freedom will not be fully scrambled. On the other hand, for an event horizon, its stretched surface becomes the stretched horizon of the black hole, whose area will eventually shrink to zero due to Hawking radiation. Any information contained within the stretched event horizon, therefore, will be transferred to ambient space at late times, either in the form of Hawking radiation or through mining.

The expectation given above is consistent with the fact that neither standard nor spherical Rindler space contains a singularity. In these spacetimes, an “exterior” observer cannot retrieve information about an object which crossed the horizon at an earlier time. Thus, there is no problem of cloning even if the observer jumps into the horizon and catches up with the original object, which is generally possible if there is no singularity. On the other hand, for an event horizon, an exterior observer at late times *can* retrieve information behind

the horizon. Therefore, an observer must not be able to jump into the horizon and obtain a second copy of the information they have retrieved. This requires that every causal curve in the interior region ends at a singularity, as is indeed the case for a Schwarzschild black hole (and presumably for other unstable black holes as well, after appropriate quantum effects are taken into account).

3.5.3 Topology change and quantum global hyperbolicity

In the exterior description given here, the flow of information is causally trivial. This description, however, does not include the region inside the black hole. This requires that the spatial topology changes as time progresses. For example, if the black hole is formed in $(d+1)$ -dimensional asymptotically flat spacetime, the spatial topologies in the periods before the formation of the stretched event horizon, during the presence of the stretched (event and black hole) horizon, and after the black hole evaporation are $\mathbb{R}^d$, $\mathbb{R}^d \setminus B^d$, and $\mathbb{R}^d$, respectively. These spacetimes are not globally hyperbolic classically. This is due to the singularity at the endpoint of black hole evaporation, and the fact that equal-time hypersurfaces do not enter into the interior of the stretched horizon.

The unitarity of time evolution in the exterior description, however, implies that the spacetime is “globally hyperbolic” at the fully quantum level, i.e. if we include intrinsically quantum gravitational degrees of freedom on the horizon and use the microscopic Hamiltonian for time evolution. In this sense, equal-time hypersurfaces in the exterior description play the role of Cauchy surfaces in quantum gravity, and may be called quantum Cauchy surfaces.

3.6 Global Spacetime as a Coarse-Grained Description

We have argued that the global spacetime is only emergent at the level of semiclassical gravity and cannot be used to formulate unitary quantum mechanics at the fundamental level.¹⁹ On the other hand, the concept of global spacetime itself is well defined in the sense that the physics occurring there is robust under physical processes, including nonperturbative information extraction processes performed on Hawking radiation. In this section, we discuss how the description based on the global description can emerge as a coarse-grained description, and explain why global methods nevertheless yield important results, including hints of unitarity.

The global spacetime description can be obtained at the semiclassical level by evolving the

¹⁹ In the context of AdS/CFT, this implies that any putative nonperturbative bulk theory should not be formulated on a global slice (or should not include a smooth interior), unless an intricate nonlocal mechanism reducing the redundancies on an equal-time hypersurface is involved.

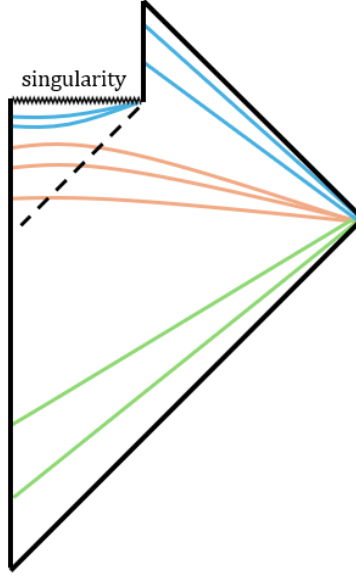


Figure 3.11: The global spacetime description can be obtained by evolving the initial equal-time hypersurface in such a way that at each time, the equal-time hypersurface is a Cauchy slice. This description, however, is consistent only if we disallow a process involving an exponentially fine-tuned detector configuration accessing information about the interior; see Section 3.4.

initial equal-time hypersurface in such a way that at each time, the equal-time hypersurface is a Cauchy slice; see Fig. 3.11.²⁰ The famous calculation of Hawking [17] suggests that this corresponds to one which is coarse-grained over black hole microstates. This results in the final state of Hawking radiation being thermal. In other words, the global spacetime description regards states after the black hole formation as the (maximal) mixture over black hole microstates [45]. We will show that this interpretation is consistent with results obtained using replica wormholes [19, 20], entanglement islands [18, 54, 55], and baby universes [209, 210], which see some traits of unitarity despite the fact that they employ the global spacetime description.

3.6.1 The Page curve from replica wormholes and entanglement islands

While the global spacetime description represents an ensemble average over microstates, it can still allow us to calculate the ensemble averages of many different quantities by changing

²⁰ Strictly speaking, equal-time hypersurfaces after black hole evaporation are not Cauchy surfaces at the classical level because of the singularity at the endpoint of the evaporation; see the discussion in Section 3.5.3.

the boundary conditions.²¹ In particular, by adopting the replica boundary conditions [19, 20], gravitational path integrals in the global spacetime can calculate the ensemble averages of the traces of powers of the density matrix of Hawking radiation $\overline{\text{Tr} \rho_R^n}$ for different n . From this, the ensemble average of the von Neumann entropy of the radiation can be obtained as

$$\overline{S_R} = -\lim_{n \rightarrow 1} \frac{\partial}{\partial n} \overline{\text{Tr} \rho_R^n} \sim \min\{S_{\text{rad}}, S_{\text{bh}}\}, \quad (3.6.1)$$

where S_{rad} and S_{bh} are the coarse-grained entropies of the radiation and the black hole. Since this quantity is the ensemble average of microscopic von Neumann entropies, which obey the Page curve [41] for all members of the ensemble, it also obeys the Page curve [45, 81–87]. This is how the replica wormhole calculation in [19, 20] yields the Page curve, despite the fact that it employs the global spacetime description, which is not unitary.²²

The replica prescription for the gravitational path integral is virtually equivalent to the entanglement island prescription [18, 54, 55] for calculating entropies. Thus, the success of the calculation in [18, 54, 55], which adopts the global spacetime description, can also be interpreted in the same way. Note that because of the ensemble averaging, Hawking radiation in a single copy of spacetime does not contain information about the microstate by itself. This is why the semiclassical entropy must be used for Hawking radiation when applying the island formula (in a single copy of the global spacetime).

3.6.2 Baby universe α -states as fine-grained information in Hawking radiation

The transition from connected to disconnected Cauchy surfaces depicted in Fig. 3.11 can be viewed as the creation of a baby universe [206–208, 210], which will be separate from the ambient space at late times. This picture was studied by Marolf and Maxfield in [209]. Assuming that the singularity at the end of the black hole evaporation is mild, we may envision the process as in Fig. 3.12, in which a closed universe corresponding to the black hole interior is separated from the ambient space describing the exterior of the black hole.

Marolf’s and Maxfield’s interpretation is that after the black hole evaporates, its microscopic information is encoded in the α -states of the baby universes arising in the system of replicated black holes. For a single black hole, different α sectors comprise superselection sectors, and hence its information cannot be accessed in a single experiment involving a single black hole. Our analysis implies that this is an artifact of the global spacetime description, in which the Hawking radiation involves the ensemble average over the microstates. In the full exterior description, the microscopic information *can* be accessed even in a single experiment

²¹ In lower dimensional gravity, the relevant ensemble can consist of different microscopic boundary theories [228, 229].

²² To recover the full unitary content of the theory, one would calculate the ensemble averages of all of the moments, $\overline{\text{Tr} \rho_R^n}$ for $1 \leq n \leq e^{S_{\text{rad}}}$. Such a calculation, however, would not be able to be performed within the framework of semiclassical gravity.

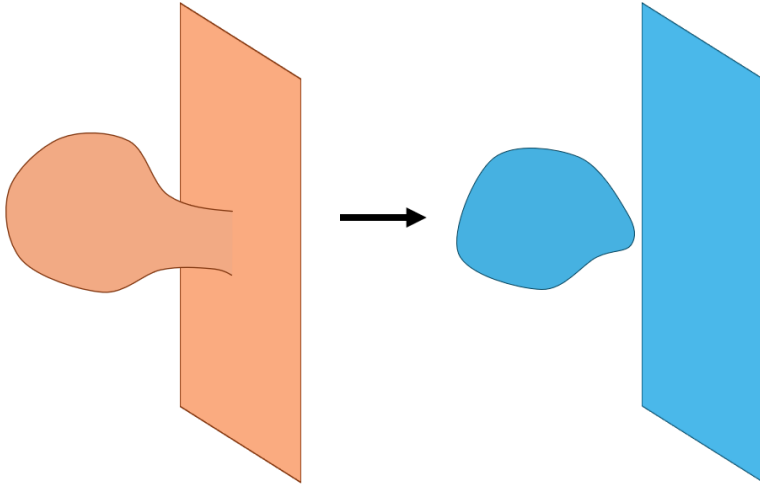


Figure 3.12: Assuming that the singularity at the endpoint of black hole evaporation is sufficiently weak, the evaporation of a black hole can be viewed as the creation of a baby universe in (but only in) the effective global spacetime description.

by an appropriately fine-tuned detector, as depicted in Fig. 3.4. In other words, α sectors in the global spacetime description are not superselection sectors in the exterior description, which is unitary and hence, unlike [209], gives a pure final state.

This interpretation is consistent with the conjecture [196, 200] that in a complete theory of quantum gravity, the Hilbert space of baby universes is trivial

$$\dim \mathcal{H}_{\text{BU}} = 1. \quad (3.6.2)$$

A nontrivial Hilbert space for baby universe states emerges only in an effective description in which certain microscopic information in the fundamental description is coarse-grained.

We expect that this conclusion applies more generally. For example, we may imagine that states representing a semiclassical closed universe can be embedded, possibly like bound states [156, 230], in a theory accommodating an asymptotic region (such as null infinities of flat space and the conformal boundary of AdS space). Different α sectors associated with the semiclassical closed universe would then correspond to different microstates in the putative boundary theory in which this universe is embedded.

3.7 Conclusion

In this paper we have described a consistent view of black hole complementarity. In this picture, the fundamental description of the spacetime ends at the stretched horizon. A semiclassical picture of spacetime including the interior can be obtained only effectively, by the mechanism described in Section 3.3. As discussed there, the interior emerges as a way of reorganizing the degrees of freedom participating in the transient phenomenon of an

infalling object dissipating into the stretched horizon. We argued using this framework that no operation performed on Hawking radiation, however complex, can affect the experience of an infalling observer. Rather, reconstructions of interior operators from Hawking radiation should be interpreted as acting on a historical record of the effective interior, which has causally propagated from the vicinity of the stretched horizon (and early Hawking radiation). While in the effective global picture the interior operators appear spacelike separated from the exterior operators used for reconstruction, in the fundamental picture they are timelike related. We explained this phenomenon as a consequence of the chaotic dynamics of the horizon modes.

It is natural to ask whether this construction can be extended to describe a black hole formed from collapse. To do so, it would be necessary to extend the stretched horizon before the formation of the black hole. In Section 3.5, we discussed a scheme for doing this, in which the extended stretched horizon is defined as a fixed radius surface in spherical Rindler coordinates, where the Unruh temperature reaches the string scale. We speculate, by analogy to the black hole case, that degrees of freedom in the interior might be constructed using the Unruh modes near this extended stretched horizon. We leave exploration of this and extension to spacetimes without spherical symmetry for future work.

Part II

Entanglement and Complexity for Gravitating Subregions

Chapter 4

Hollow-grams: Generalized Entanglement Wedges from the Gravitational Path Integral

4.1 Introduction

In recent years, the gravitational path integral has emerged as a central tool in our understanding of quantum gravity, providing a framework for addressing profound questions about the nature of spacetime and quantum mechanics. Although it has not been derived as a controlled approximation to an underlying theory,¹ it has nevertheless been successful in computing various coarse-grained features expected of quantum gravitational systems such as the spectral form factor at late times [232].

One of its most striking applications has been the derivation of the Page curve [19, 20], which captures the unitary evolution of black hole evaporation. Although these calculations were performed using an auxiliary non-gravitational bath to collect Hawking radiation for mathematical convenience, the physically relevant setting is one where all regions, including the radiation, are themselves gravitating. This raises fundamental questions about how entanglement and entropy should be defined in gravitational systems and has been critiqued in Refs. [233, 234].

The challenge of extending the notion of entanglement entropy to fully gravitational systems was addressed by Bousso and Penington (BP) [203, 216], who made a proposal

¹ An exception is two spacetime dimensions where Jackiw-Teitelboim (JT) gravity and its variants have been defined using the gravitational path integral, and correspond to ensemble averages over boundary theories [88, 228, 229, 231].

to identify entanglement wedges for gravitating regions.² In a time-reflection symmetric setting, the BP proposal states that the generalized entanglement wedge $E(a)$ of an input bulk region a is the smallest generalized entropy region that includes a and has the same conformal boundary as a .

The BP proposal provides a natural generalization of the Ryu-Takayanagi formula [13] to the case of bulk regions and has many non-trivial properties expected of an entanglement wedge [216]. Furthermore, these generalized entanglement wedges also satisfy strong subadditivity [216], monogamy of mutual information [236] and the rest of the inequalities of the static holographic entropy cone [237].

It is thus of great interest to derive the BP proposal from the gravitational path integral. This issue has been discussed in Ref. [238], who motivated their proposal using edge modes. Since edge modes in gravity are not well understood,³ we will instead discuss a different method motivated by tensor networks and implemented directly using the gravitational path integral. Our analysis allows us to do explicit calculations in various examples, providing a clear understanding of the gravitational path integral derivation. We also discover novel conceptual aspects of the BP proposal that we outline below.

We will derive the BP proposal in time-reflection symmetric situations by first explaining what the proposal means in a random tensor network (RTN) in Sec. 4.2. RTNs, reviewed in Sec. 4.2.1, are useful toy models for holography [120] that have provided interesting insights into the entanglement structure of holographic states [242–246]. In Sec. 4.2.2, we define a generalized entanglement wedge associated to a bulk region in an RTN by first deleting the tensors in that given region and generating a new “hollowed” RTN. We then compute the entropy of the legs that were cut open in the process to define the generalized entanglement wedge.

In the hollowed RTN, there is no qualitative difference between the open legs at the original boundary and the new open legs in the hollowed region. Thus, we can use the known technique to compute the n th Rényi entropy in an RTN using its effective description as an Ising-like model on the permutation group acting on the n replica copies [120]. By doing so, we confirm that the entanglement wedge of the hollowed region matches the generalized entanglement wedge defined by the BP proposal.

The advantage of our RTN discussion is that it provides us with a concrete Hilbert space interpretation of the BP proposal. Using this, we clarify how the BP proposal should work when multiple bulk regions are involved. In particular, in Sec. 4.2.3, we propose restricted entanglement wedges as a natural candidate for computing entropies of spacelike separated bulk regions. This leads us to a new version of the holographic inequalities, which we call restricted entropy inequalities. The unrestricted entropy inequalities previously considered required stronger independence conditions and follow from our restricted inequalities in the special case when these independence conditions are satisfied. This unifies the discussion of

² Ref. [235] also addressed this question in a slightly different way by proposing an answer for the effective entropy of bulk quantum fields in a gravitating region. Our analysis helps clarify the difference between these two proposals.

³ E.g., in JT gravity, they have subtle issues related to the Plancherel measure [239–241].

strong subadditivity (SSA), which only required spacelike separated subregions [203, 216], with the other holographic inequalities.

Having understood the BP proposal in RTNs, we can then translate our results to gravity by exploiting the connection between RTNs and fixed-geometry states in gravity [19, 244, 247–250]. In Sec. 4.3, we review how this works for the purpose of computing the entropy of boundary subregions in holographic states. Of course, holographic states in gravitational theories are usually prepared using a Euclidean path integral with specified asymptotic boundary conditions, and thus are quite unlike fixed-geometry states. They are more analogous to coherent states, while the fixed-geometry states are analogous to position eigenstates. Nevertheless, a semiclassical state can be expressed as a superposition over fixed-geometry states with the wavefunction computed using the gravitational path integral [16, 251–253]. By then employing a diagonal approximation that restricts to replica-symmetric saddles, one can reproduce the standard Lewkowycz-Maldacena derivation of the Renyi entropy [104, 254].⁴

By employing the same idea for bulk subregions, we propose a gravitational path integral derivation of the Bousso-Penington proposal in Sec. 4.4. In particular, for general holographic states prepared using a Euclidean path integral, the saddles we obtain include backreaction effects that change the geometry as a function of the Renyi index n . Our derivation thus generalizes the Lewkowycz-Maldacena computation for boundary regions [104].

While our analysis could also directly be phrased using the gravitational path integral, we find it useful to motivate the derivation using RTNs since it provides a clear formulation of what entropy it is that we are computing in the underlying theory. In particular, this gives a clear motivation for why the generalized entanglement wedge is restricted to move outward from the bulk region of interest.

An important point that we would like to emphasize is that an input to our computation is a gauge-invariant definition of the bulk region which allows us to define what we mean by the specified region in arbitrary geometries. The saddles computing the Rényi entropy end up depending on the precise way in which the region is specified, but the dependence drops out in the $n \rightarrow 1$ limit. Thus, we find that the generalized entanglement wedge proposed by BP is recovered independent of the details of how the bulk region is gauge-invariantly defined.

Finally, in order to illustrate our derivation, we provide various examples in Sec. 4.5. In Sec. 4.5.1, we first consider the bulk region to be a finite interval in Jackiw-Teitelboim (JT) gravity. We compute the Rényi entropy using the gravitational path integral for different choices of gauge-invariant specification of such a region. In Sec. 4.5.2, we then do a similar computation in higher dimensional Einstein gravity for an annulus outside a black hole horizon using a minisuperspace approximation, restricting to spherically symmetric configurations.

In Sec. 4.6, we discuss generalizations of our work. In Sec. 4.6.1, we explain how one can use the gravitational path integral to compute the min- and max-entanglement wedges

⁴ In fact, the diagonal approximation method gives the correct answer for the $n < 1$ Renyi entropy while the Lewkowycz-Maldacena method leads to the wrong answer [252].

for boundary subregions [16, 245]. The same idea can then be used for bulk regions as well using our idea of hollowing out the region. Finally, in Sec. 4.6.2, we discuss the main open problem of extending our method to time-dependent situations.

4.2 BP Proposal in RTNs

In this section, we review the definition of RTNs and present our proposal for generalized entanglement wedges within these models, confirming the BP proposal. Based on this analysis, we clarify the notion of independence of bulk subregions.

4.2.1 Review of RTNs

Here we briefly review random tensor networks; for a more detailed exposition, we refer the reader to Ref. [120]. A tensor network is defined by starting with a graph $G = \{\mathcal{V}, \mathcal{E}\}$ consisting of a set of vertices $\mathcal{V}$ and edges $\mathcal{E}$. In the context of holography, this graph can be interpreted as a coarse-grained version of a smooth spatial geometry. At each vertex $x \in \mathcal{V}$, we place a rank- k tensor $T_{x, \mu_1 \mu_2 \dots \mu_k}$, where the indices μ_i correspond to the edges connected to x . Each edge is associated with a Hilbert space $\mathcal{H}_i$, spanned by basis vectors $|\mu_i\rangle$. The tensor T_x then defines a state

$$|T_x\rangle = T_{x, \mu_1 \mu_2 \dots \mu_k} |\mu_1\rangle |\mu_2\rangle \dots |\mu_k\rangle \quad (4.2.1)$$

on the product Hilbert space $\otimes_i \mathcal{H}_i$, where the product runs over all the edges leaving the vertex x . For simplicity, all edge Hilbert spaces are taken to have the same bond dimension χ , and we will be interested in the regime $\chi \gg 1$.

To define a bulk-boundary tensor network state, we contract the tensors along their shared edges to obtain:

$$\left(\bigotimes_{(xy) \in \mathcal{E}} \langle xy| \right) \left(\bigotimes_{x \in \mathcal{V}} |T_x\rangle \right), \quad (4.2.2)$$

where $|xy\rangle = \sum_{\mu} |\mu\rangle_x |\mu\rangle_y$ is a maximally entangled state defined on the Hilbert space associated with the edge (xy) and such contracted legs are called in-plane legs. The contraction is implemented by projecting onto these maximally entangled states, with the appropriate index of tensor T_x contracted with that of tensor T_y . The resulting state then lives in the Hilbert space of the remaining open legs, i.e., $\mathcal{H}_{\text{bulk}} \otimes \mathcal{H}_{\partial}$, where $\mathcal{H}_{\partial}$ is associated with the legs on the boundary of the tensor network, and $\mathcal{H}_{\text{bulk}}$ corresponds to the bulk dangling legs. This bulk-boundary tensor network state can also equivalently be interpreted as a map from bulk to boundary, i.e., from $\mathcal{H}_{\text{bulk}}$ to $\mathcal{H}_{\partial}$.

In general, the graph G can be arbitrary, but for concreteness, the reader may focus on a model inspired by AdS/CFT in which the tensor network is defined on a triangulation of hyperbolic space representing a fixed time slice of the AdS spacetime. Then, projecting the bulk legs onto a bulk state $|\psi^{\text{bulk}}\rangle$, one obtains a resulting boundary state which is a pure

state in the Hilbert space $\mathcal{H}_\partial$:

$$|\psi\rangle = \frac{1}{N} \left(\langle \psi^{\text{bulk}} | \bigotimes_{\{xy\} \in \mathcal{E} \setminus \partial} \langle xy | \right) \left(\bigotimes_{x \in \mathcal{V}} |T_x\rangle \right) \in \mathcal{H}_\partial , \quad (4.2.3)$$

where the normalization constant N ensures that $\langle \psi | \psi \rangle = 1$.

In an RTN, each tensor $|T_x\rangle$ is drawn at random using the Haar measure. This can be implemented by acting on a fixed reference state $|0_x\rangle$ with a Haar random unitary U_x . The advantage of this choice is that the statistical properties of Haar random unitaries let us compute various averaged quantities such as the entropy. In particular, by applying Schur's lemma [255], one has:

$$\overline{(|T_x\rangle \langle T_x|)^{\otimes n}} = \int [dU_x] (U_x |0_x\rangle \langle 0_x| U_x^\dagger)^{\otimes n} \propto \sum_{g_x \in \mathcal{S}_n} g_x , \quad (4.2.4)$$

where g_x is an element of the symmetric group $\mathcal{S}_n$ which acts by permuting the ket edges of the n copies of the state.

This leads us to the computation of the entropy of a boundary subregion. Let the boundary Hilbert space be factorized as $\mathcal{H}_\partial = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$, where A is the subregion of interest. The n -th Rényi entropy can be computed using a replica trick as the expectation value of a twist operator, which cyclically permutes the n copies of $\mathcal{H}_A$ according to the permutation $\tau_n \in \mathcal{S}_n$. Using Eq. (4.2.4), the calculation in the RTN boils down to the partition function of a ferromagnetic Ising-like model with a sum over permutations at each vertex dictating the gluing pattern of the various replicas and the energy cost controlled by the Cayley distance between neighbouring permutations. The boundary conditions for this partition function are set by the replica trick, i.e., $g_x = \tau_n$ for $x \in A \subseteq \partial$, and $g_x = e$ (the identity element) for $x \in \bar{A} = \partial \setminus A$.

For the purposes of modeling AdS/CFT, we are interested in the large χ limit, which corresponds to a semiclassical $G \rightarrow 0$ limit. In this limit, the entropy is self-averaging and thus can be approximately computed via the above replica trick. In this limit, the dominant contribution to the Rényi entropy arises from the ground state of the Ising-like model: a replica-symmetric configuration consisting of domains labeled by group elements τ_n and e , separated by a domain wall. In particular, the location of the domain wall is fixed by energy minimization. Thus, in this limit, we find:

$$\overline{\text{Tr}(\rho_A^n)} = e^{-(n-1)|\gamma_A| \ln \chi} \text{Tr} \left[(\rho_{E(A)}^{\text{bulk}})^n \right] , \quad (4.2.5)$$

where $|\gamma_A|$ is the number of legs cut by the domain wall, and $\rho_{E(A)}^{\text{bulk}}$ is the partial density matrix associated with the bulk state $|\psi^{\text{bulk}}\rangle$ restricted to the τ_n domain in the ground state, labelled $E(A)$ since it will turn out to be the entanglement wedge of A . This yields:

$$S_n(A) \equiv \frac{1}{1-n} \log [\text{Tr}(\rho_A^n)] = |\gamma_A| \ln \chi + S_n(\rho_{E(A)}^{\text{bulk}}) , \quad (4.2.6)$$

where due to the self-averaging nature of this calculation, the above statement is true for almost all instances of the random tensors. The term proportional to $\ln \chi$ in Eq. (4.2.6)

plays the role of the area term, while the second term corresponds to the matter entropy in the entanglement wedge. In the $n \rightarrow 1$ limit in particular, we obtain the RTN version of the Ryu-Takayanagi (RT) formula:

$$S(A) = |\gamma_A| \ln \chi + S(\rho_{E(A)}^{\text{bulk}}) = S_{\text{gen}}(E(a)), \quad (4.2.7)$$

where we have defined an RTN version of the generalized entropy S_{gen} associated to the generalized entanglement wedge.

4.2.2 Generalized Entanglement Wedges

Having reviewed the computation of the boundary entropy in RTNs, we now propose a framework to define entropies associated to bulk regions. This defines the generalized entanglement wedge in RTNs and we show that it agrees with the BP proposal.

Definition 4.1. (*Bulk Input Regions*) Given an RTN, an input region a is defined as a subset of vertices, $a \subset \mathcal{V}$. The complement of a is denoted by $\bar{a} = \mathcal{V} \setminus a$.

The internal edges of a , denoted $\mathcal{E}_a$, consist of all edges that connect vertices within a to each other. The boundary of a , denoted by ∂a , includes all in-plane edges connecting vertices in a to $\bar{a}$. We denote the number of such boundary edges by $|\partial a|$. See Fig. 4.1.

Definition 4.2. (*Hollow-graphic State*) Given an RTN and an input region a , the generalized entanglement entropy associated with a is defined using the hollow-graphic state:

$$|\psi_a\rangle = \frac{1}{N_a} \left(\langle \psi^{\text{bulk}} | \bigotimes_{\{xy\} \in \mathcal{E} \setminus \partial a \setminus \mathcal{E}_a} \langle xy | \right) \left(\bigotimes_{x \in \mathcal{V} \setminus a} |T_x\rangle \right) \left(\bigotimes_{x \in a} |\phi_x^{\text{bulk}}\rangle \right), \quad (4.2.8)$$

where N_a is a normalization constant and $|\phi_x^{\text{bulk}}\rangle$ is a maximally entangled state on the bulk dangling leg at vertex x and an auxiliary copy of it.

Let us now explain how the hollow-graphic state is algorithmically constructed from the original RTN projected onto a bulk state. By removing tensors in a , as well as all its internal edges, we generate a new hollowed graph with open in-plane legs on ∂a as well as open bulk legs in the region a . We add in an auxiliary copy for each bulk leg opened this way and put in a maximally entangled state $|\phi_x^{\text{bulk}}\rangle$ for each $x \in a$. This is a trick that allows us to flip bras coming from postselection on the bulk dangling legs in a to kets.⁵ The hollow-graphic state $|\psi_a\rangle$ thus obtained using Eq. (4.2.8) is now defined on a larger Hilbert space $\mathcal{H}_{\partial} \otimes \mathcal{H}_{\partial a} \otimes \mathcal{H}_{a,\text{bulk}}$, where $\mathcal{H}_{\partial a}$ and $\mathcal{H}_{a,\text{bulk}}$ are the Hilbert spaces corresponding to the open in-plane legs on the boundary of a and open bulk dangling legs in a respectively.

⁵ There is an ambiguity in the precise choice of this state but it doesn't affect the entropies we are interested in computing.

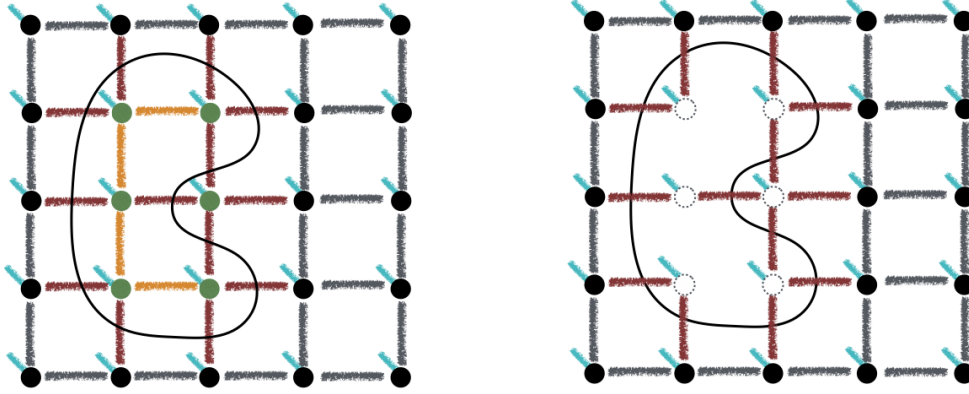


Figure 4.1: A bulk input region in a tensor network. The green dots represent the vertices comprising the input region a . The yellow edges correspond to the internal edges $\mathcal{E}_a$, while the red edges represent the boundary ∂a and the turquoise legs indicate the bulk legs. The right-hand panel shows the diagrammatic representation of the hollow-graphic state used to compute the entropy associated with a .

See Fig. 4.1. A similar construction has been proposed by Ref. [256].⁶ Note that this is a different construction from that provided in Ref. [235], where one detaches the bulk legs but does not remove the tensors or in-plane legs resulting in a proposal for the effective entropy of bulk quantum fields.

Since in an RTN, there is no qualitative difference between open legs deep in the bulk and those at the boundary, we can use the previous calculation of the entropy as is. As in the case of boundary subregions, in the limit of large bond dimension, the dominant configuration corresponds to a domain wall homologous to the newly introduced boundary ∂a , resulting from the removal of a . The Rényi entropy of this state can be computed analogously with the boundary conditions imposed by setting $g_x = \tau_n$ for $x \in \partial a$, and $g_x = e$ for $x \in \partial \setminus \partial a$:

$$\overline{\text{Tr}(\rho_a^n)} = e^{-(n-1)|\gamma_a| \ln \chi} \text{Tr} \left[(\rho_{E(a)}^{\text{bulk}})^n \right], \quad (4.2.9)$$

where $|\gamma_a|$ is the number of legs crossed by the energy minimizing domain wall, and $\rho_{E(a)}^{\text{bulk}}$ is the partial density matrix associated with the original bulk state $|\psi^{\text{bulk}}\rangle$ on the cyclically glued domain $E(a)$ of the tensor network, which will get identified with the generalized entanglement wedge of a . The resulting Rényi entropy is then

$$S_n(a) \equiv \frac{1}{1-n} \log [\text{Tr}(\rho_a^n)] = |\gamma_a| \ln(\chi) + S_n(\rho_{E(a)}^{\text{bulk}}). \quad (4.2.10)$$

Again in the $n \rightarrow 1$ limit, we obtain

$$S(a) = |\gamma_a| \ln \chi + S(\rho_{E(a)}^{\text{bulk}}) = S_{\text{gen}}(E(a)). \quad (4.2.11)$$

⁶ We thank Chris Akers for discussions on this point.

This is the analog of the RT formula for bulk subregions in RTNs.

As anticipated above, the region $E(a)$ is precisely the *generalized entanglement wedge*, providing an RTN realization of the BP proposal. In cases where the boundary of a intersects the outer boundary of the tensor network, $\partial a \cap \partial \neq \emptyset$, the imposed boundary conditions ensure that $\partial E(a) \cap \partial = \partial a \cap \partial$.

Remark 4.3. *The case where the input region is a boundary subregion, i.e., $\partial a = A$ and $\mathcal{E}_a = \emptyset$, is a special instance of this prescription.*

Our analysis above also guarantees that all standard results for RTNs follow for the hollow-graphic state as well. There is not only an RT formula, but also a version of subregion duality, which involves acting on the open legs in a in the hollow-graphic state to affect the semiclassical bulk degrees of freedom in $E(a)$.

Finally, let us comment on the map taking the original RTN state to the hollow-graphic state. Let V_a denote the map corresponding to the removal of subregion a from the tensor network, mapping the state $|\psi\rangle$ defined in Eq. (4.2.3) to the state $|\psi_a\rangle$ defined in Eq. (4.2.8). Since the state $|\psi\rangle$ resides in a smaller Hilbert space compared to that of $|\psi_a\rangle$, the map V_a is an embedding. The inverse of this map is a projection from $\mathcal{H}_\partial \otimes \mathcal{H}_{\partial a} \otimes \mathcal{H}_{a,\text{bulk}}$ back to $\mathcal{H}_\partial$. This corresponds to “sewing back in” the subregion a into the tensor network by projecting onto an appropriate state in $\mathcal{H}_{\partial a} \otimes \mathcal{H}_{a,\text{bulk}}$. We denote this projection by Π_a . This acts on $|\psi_a\rangle$ as

$$\Pi_a |\psi_a\rangle = \Pi_a (V_a |\psi\rangle) = |\psi\rangle . \quad (4.2.12)$$

These maps will play an important role in our following discussion on the independence of bulk subregions.

4.2.3 Independence Relations and Entropy Inequalities

Having understood how to make sense of generalized entanglement wedges for bulk regions, we can now discuss how generalized entanglement wedges work for multiple bulk regions, a prerequisite to discuss entropy inequalities like strong subadditivity (SSA). In quantum mechanics, SSA is stated as follows: given a state ρ on a Hilbert space that can be factorized as $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C$, the following inequality holds:

$$S(\rho^{AB}) + S(\rho^{BC}) \geq S(\rho^B) + S(\rho^{ABC}), \quad (4.2.13)$$

where the density matrices $\rho^A = \text{Tr}_{BC} \rho$ are defined through the partial trace. However, since the hollow-graphic state used to define the generalized entanglement wedge is subregion-dependent, as can be seen from Eq. (4.2.8), more care is required to achieve such a factorization.

In particular, we propose to define SSA in holography as follows: start by considering the hollow-graphic state $|\psi_{abc}\rangle$ for regions a, b, c that are spacelike separated. In this hollowed RTN, it is manifest that the Hilbert spaces associated with the open legs of regions a, b and c factorize. The statement of SSA expressed using the RT formula then takes the form:

$$S_{\text{gen}}(E(ab)|_{c'}) + S_{\text{gen}}(E(bc)|_{a'}) \geq S_{\text{gen}}(E(b)|_{(ac)'}) + S_{\text{gen}}(E(abc)), \quad (4.2.14)$$

where $E(p)|_{q'}$ denotes the *restricted entanglement wedge* of the bulk region p , defined as the entanglement wedge of p computed in the spacetime q' , where the boundary of q is treated analogously to the asymptotic boundary and the generalized entanglement wedge is not permitted to include any portion of q . Restricted entanglement wedges were first introduced in Ref. [257] in the context of the complementarity of bulk holograms, and they arise naturally in our framework as well.

Notably Eq. (4.2.14) is different from the unrestricted version of SSA considered in previous work:

$$S_{\text{gen}}(E(ab)) + S_{\text{gen}}(E(bc)) \geq S_{\text{gen}}(E(b)) + S_{\text{gen}}(E(abc)), \quad (4.2.15)$$

where one follows the standard BP proposal and doesn't impose any additional homology constraints. Nevertheless, if the following independence condition is satisfied for a collection of input regions a_i :⁷

$$a_i \subset \left[E \left(\bigcup_{j \neq i} a_j \right) \right]', \quad (4.2.16)$$

then it is guaranteed that $E(p)|_{q'} = E(p)$ since imposing the homology constraint doesn't affect the generalized entanglement wedge of p . Given this condition, Eq. (4.2.14) in fact reduces to Eq. (4.2.15).

Similarly, we note that all the holographic entropy cone inequalities can be proved with the weaker assumption that all the a_i are spacelike separated as long as the holographic entropy cone inequalities are expressed in terms of the areas of restricted entanglement wedges, as in Eq. (4.2.14). Our framework suggests that this formulation is more natural compared to the unrestricted version previously considered since the generalized entropies of restricted entanglement wedges correspond to subregion entropies of a single state, rather than entropies of subregions in different states that are related through post-selection. Moreover, the unrestricted holographic inequalities previously considered follow as a special case of our restricted inequalities when the appropriate independence conditions Eq. (4.2.16) are satisfied.

Now, Eq. (4.2.16) was indeed imposed as an independence condition for all the unrestricted higher holographic entropy cone inequalities⁸ in Refs. [236, 237]. However, a puzzling feature of the BP proposal is that, in order for bulk subregions to satisfy strong subadditivity (SSA) in the unrestricted form of Eq. (4.2.15), only a weaker notion of independence was necessary. In both time-reflection symmetric and time-dependent settings, it suffices

⁷ This independence condition is identical to the one used in the proof of the no-cloning theorem in Ref. [203]. That argument, as well as the time-dependent proofs of strong subadditivity [203] and the monogamy of mutual information in [236], relies on the assumption that the minimal and maximal entanglement wedges coincide for arbitrary unions of input subregions, i.e., $e_{\min} = e_{\max}$. This assumption is crucial: when $e_{\min} \neq e_{\max}$, the bulk dual of the von Neumann entropy is not simply given by the RT formula [16].

⁸ Note that these higher holographic entropy cone inequalities hold only at leading order in $1/G$, when one considers only the area contribution to the generalized entropy. One can prove these inequalities when including bulk matter contributions only if further entropy inequalities are imposed on the bulk matter itself [258].

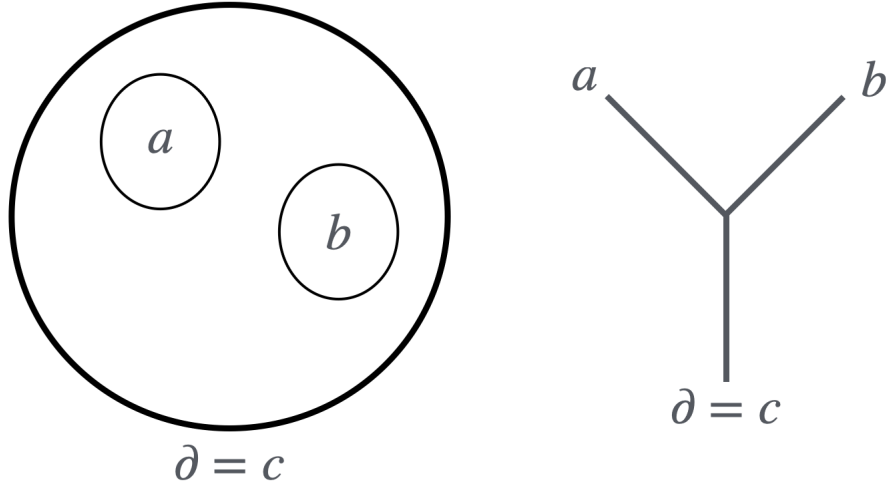


Figure 4.2: Schematic representation of the GHZ example. On the left, we see the tensor network analogue of the type of state we consider. The GHZ state is analogous to the state with subregions a and b both removed. The right side schematically represents the GHZ-type entanglement between respective subregions.

for input regions to be spacelike separated to ensure the unrestricted form of SSA. This is surprising since we have argued that Eq. (4.2.14) is the natural statement of SSA, which does not reduce to Eq. (4.2.15) under this weaker independence condition.

To resolve this confusion, our proposal here is that the fact that SSA holds under a weaker condition appears to be a special feature of semiclassical states with a good $G \rightarrow 0$ limit. Given regions a , b , and c in an RTN, the statement of Eq. (4.2.15) in general situations amounts to a statement of SSA under postselection, which is not generally true in quantum mechanics. For an explicit example, consider the GHZ state (see Fig. 4.2). In this case, we find $S(\rho^{ab}) = \ln(2)$. To compute $S(\rho^b)$ or $S(\rho^{bc})$, we first apply the postselection Π_a which we take to be postselection onto $|0\rangle$. After this projection, we find $S(\rho^b) = 0$, and since the subsystem bc is then in a pure state, we have $S(\rho^{bc}) = 0$. Similarly, $S(\rho^{abc}) = 0$. Putting this together, we see that the unrestricted version of SSA is violated as expected.

Now that we have seen why a stronger notion of independence is generally required, we still need to explain why, in the context of generalized entanglement wedges, a weaker notion of independence suffices specifically for strong subadditivity (SSA). This is a rather special and seemingly coincidental feature of SSA.

As noted above, in both the static case [216] and the time-dependent case [203], at leading order in G (where entropy is given by the area of the entanglement wedge in Planck units) no independence condition is required to prove SSA except that the input regions are spacelike. In the static setting, SSA follows from basic properties of Euclidean geometry, while in the time-dependent setting, the proof, though more intricate, remains purely geometric. The

spacelike condition ensures that, in the intermediate step of the cut-and-paste construction, the resulting regions satisfy the appropriate homology constraints.⁹

When including quantum corrections, care needs to be taken that the quantum state has a good $G \rightarrow 0$ limit, or equivalently a good large bond dimension limit in the RTN. This is a restrictive set of states which appear to satisfy the unrestricted version of SSA, but our claim is that this is only a special feature of semiclassical holographic states and not a universal statement of quantum mechanics.

It is also important to note however that unrestricted SSA need not be an indicator of a holographic dual: similar behavior can be found in theories that are not known to possess a holographic dual. E.g., in many low-energy states of general theories where the entanglement entropy of a subregion follows an area law, SSA continues to hold even when the input regions significantly overlap.

4.3 Holographic Rényi Entropy via Fixed-Geometry States

In this section, we review the calculation of Rényi entropy of a boundary subregion using the gravitational path integral. In particular, we will focus on a derivation where we decompose a holographic state into fixed-geometry states. In Refs. [252, 253], this method was shown to agree with the computation by Ref. [104] for $n > 1$ and rectify it for $n < 1$. It will also serve as a natural formalism that allows us to generalize to the case of interest, i.e., the entropy of bulk regions.

Consider a time-reflection symmetric holographic state $|\psi\rangle$ which is prepared by a standard Euclidean gravitational path integral with asymptotic boundary conditions.¹⁰ We denote the maximal volume slice anchored to the time-reflection symmetric slice on the boundary by Σ . We will assume that the leading spacetime geometry corresponding to $|\psi\rangle$ preserves time-reflection symmetry and thus, Σ is the time-reflection symmetric slice in this spacetime. This state can be decomposed into a basis of states $|h\rangle$ that have fixed geometry on the slice Σ as¹¹

$$|\psi\rangle = \sum_h \psi(h) |h\rangle. \quad (4.3.1)$$

The sum is over all possible gauge-inequivalent induced metrics h on the time-reflection

⁹ We thank Matthew Headrick for discussions that helped clarify this point.

¹⁰ For concreteness, we will consider setups with an asymptotic boundary relevant to AdS/CFT, but our analysis doesn't rely strongly on this as long as one can define the gravitational path integral in a sensible manner and choose appropriate gauge fixing conditions to specify the Cauchy slice on which we fix the geometry.

¹¹ This is true in the approximation where we ignore wormhole effects, which in fact can become important when exponentially long times are involved. See, e.g., Ref. [259] for the difference in JT gravity between the inner product in the disk approximation versus when wormhole effects are included.

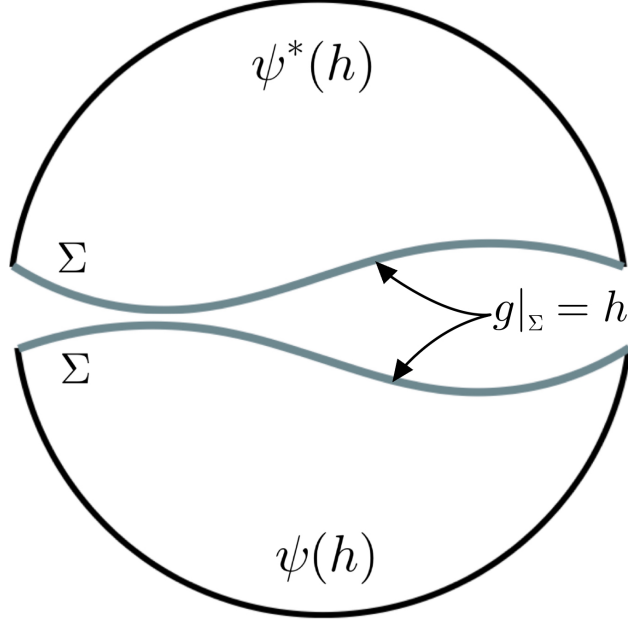


Figure 4.3: $\psi(h)$ is computed via a Euclidean gravitational path integral whose boundary conditions consist of the fixed geometry h on the bulk slice Σ and asymptotic AdS boundary conditions with sources. Gluing two copies along Σ computes the norm $|\psi(h)|^2$ and the Euclidean geometry g_h generically has a jump in extrinsic curvature across Σ . Integrating over h computes the norm of the original state $\langle\psi|\psi\rangle = \int dh |\psi(h)|^2$, resulting in a smooth saddle g_ψ .

symmetric slice Σ and $\psi(h)$ is a wavefunction that encodes the probability distribution of this superposition.

The basis state $|h\rangle$ may be prepared by a Euclidean path integral with identical asymptotic boundary conditions to $|\psi\rangle$, but with an additional postselection that fixes the bulk spatial geometry on Σ to be h . The postselection onto h can be implemented by imposing a boundary condition on Σ when computing the gravitational path integral, see Fig. 4.3. Since we have imposed a boundary condition on Σ , the Einstein equations need not be satisfied on Σ and will generically involve jumps in extrinsic curvature.

The probability distribution $p(h) = |\psi(h)|^2$ can be calculated using the gravitational path integral and is given by

$$|\psi(h)|^2 = \frac{\langle\psi|\Pi_h|\psi\rangle}{\langle\psi|\psi\rangle} = \exp(I[g_\psi] - I[g_h]), \quad (4.3.2)$$

where Π_h is a projector which fixes the geometry on Σ to be h and I is the bulk gravitational action. g_ψ is the smooth Euclidean saddle computing the norm of the state $|\psi\rangle$, while g_h is the Euclidean saddle computing the norm of the fixed-geometry state $|h\rangle$ which will generally have discontinuities in the extrinsic curvature at Σ . For semiclassical states, $\psi(h)$ is very sharply peaked around h_ψ , the induced metric on Σ in the Euclidean saddle g_ψ . We give an explicit example of this decomposition for JT gravity in Sec. 4.5.

The reduced density matrix on a boundary subregion A in this state is

$$\rho_A = \sum_{h,h'} \psi(h) \psi^*(h') \text{Tr}_{\partial\Sigma/A} (|h\rangle \langle h'|). \quad (4.3.3)$$

Thus, $\text{Tr}(\rho_A^n)$ will receive both diagonal contributions where all $h = h'$ and off-diagonal contributions where some $h \neq h'$. For a semiclassical state with a sharply peaked wavefunction around a saddle point geometry, it was argued in Ref. [252] and shown in Ref. [253] that one can ignore the off-diagonal terms for the purpose of computing the Rényi entropy. Since only the diagonal terms lead to replica-symmetric saddles in the path integral, this also agrees with the standard assumption of replica symmetry.¹² Thus, using the diagonal approximation we have

$$\text{Tr}(\rho_A^n) = \sum_h |\psi(h)|^{2n} \text{Tr}(\rho_A(h)^n), \quad (4.3.4)$$

where $\rho_A(h) = \text{Tr}_{\partial\Sigma/A} (|h\rangle \langle h|)$ is the reduced density matrix on subregion A in the fixed-geometry state $|h\rangle$.

The advantage of this method is that the computation of Rényi entropy in $|h\rangle$ is much easier than the problem of computing the Rényi entropy in $|\psi\rangle$.¹³ To compute $\text{Tr}(\rho_A(h)^n)$, we perform the replica trick. The boundary conditions involve n copies of the asymptotic boundary glued together in the standard fashion with subregion A glued cyclically across the n copies, as well as a fixed-geometry boundary condition in the bulk on each of the copies. As usual, the states prepared by the Euclidean path integral are not normalized and thus, we must fix the normalization as:

$$\text{Tr}(\rho_A(h)^n) = \frac{Z_n(A, h)}{Z_1(h)^n}, \quad (4.3.5)$$

where $Z_n(A, h)$ is the replica partition function for the fixed-geometry state $|h\rangle$ and $Z_1(h) = \langle h|h\rangle$.

In the semiclassical limit, the replica partition function can be evaluated using a saddle point approximation. For general states, this is challenging due to the difficulty in finding the corresponding saddles. However, the advantage of using fixed-geometry states is that the saddles take a simpler form. Since the bulk geometry on Σ in each copy is already measured, all one needs to do is take n copies of the bulk geometry computing Z_1 and glue together the fixed-geometry slices of the n bras and n kets.

The choice of gluing corresponds to selecting a map from the bulk spatial geometry to the permutation group $\mathcal{S}_n$. This map divides the bulk geometry into domains $\mathcal{D}(g)$ for $g \in \mathcal{S}_n$, with the n copies of bulk regions being glued according to permutation g in the region

¹² This assumption of replica symmetry can be relaxed if, by assumption, there is a unique extremal surface that is homologous to the boundary subregion. See Ref. [260] for more details.

¹³ While this is true for integer n , taking the $n \rightarrow 1$ limit for fixed-geometry states is subtle since it does not commute with the $G \rightarrow 0$ limit [261]. Our interest here is to use this as an intermediate step for more general semiclassical holographic states, and in that case, we expect the limits to commute as is usually assumed in the literature.

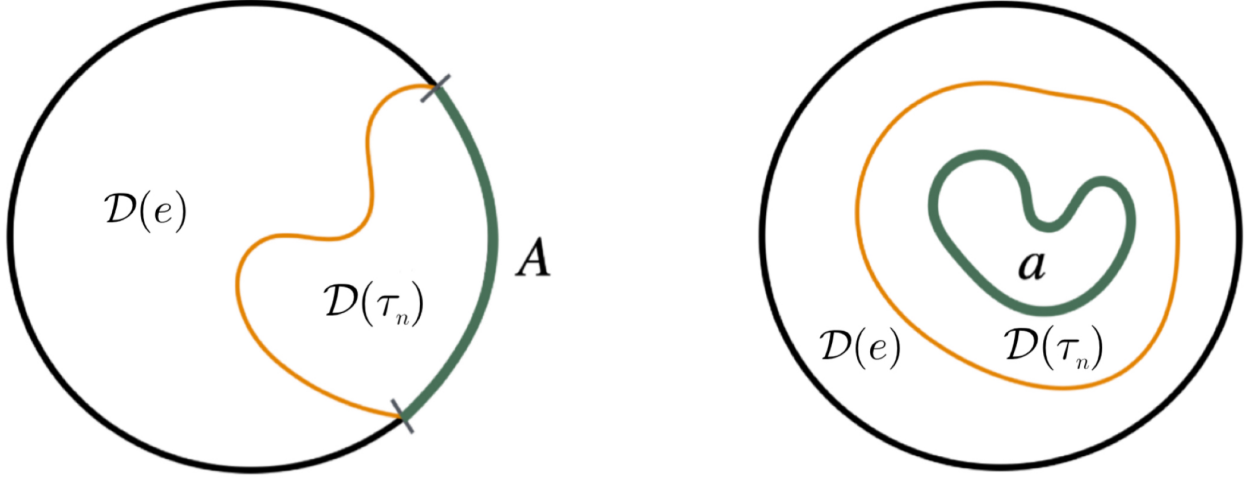


Figure 4.4: A figure representing the domain wall configurations contributing to the action when the input is a boundary subregion (left) and a bulk subregion (right).

$\mathcal{D}(g)$, subject to the appropriate boundary conditions. Moreover, the asymptotic boundary conditions are fixed by the replica trick. The boundary subregion A is assigned the cyclic permutation τ_n in the boundary replica trick and thus, the bulk region $\mathcal{D}(\tau_n)$ will have A as its conformal boundary. Similarly, the boundary subregion $\bar{A}$ is assigned the identity permutation e . It should now be clear that this is identical to the configurations showing up in the RTN calculation, see Fig. 4.4.

To find the saddle points, we must identify the configuration that minimizes the total action, which receives contributions from both the matter fields and the gravitational sector. The gravitational action itself has two components. The first arises from the domain walls separating the different domains $\mathcal{D}(g)$ on the fixed-geometry slice, while the second comes from everywhere else.

In regions away from the domain walls, the contribution to the action is given by n times the action of the saddle point that computes $Z_1(h) = \langle h|h \rangle$. This part of the action cancels against the normalization factor appearing in Eq. (4.3.5).

The remaining contribution to the gravitational action arises from the conical singularities on the domain walls. In a fashion similar to the RTN calculation, the only domains that appear in the minimal action configuration are e and τ_n , consistent with the standard assumption of replica symmetry in holography [104]. See Fig. 4.5 for the corresponding Euclidean geometry. The action associated with this conical singularity is given by

$$I_{\text{con}} = (n-1) \frac{\mathcal{A}(\partial\mathcal{D}(\tau_n), h)}{4G_N} \quad (4.3.6)$$

where $\mathcal{A}(\partial\mathcal{D}(\tau_n), h)$ is the area of the domain wall between $\mathcal{D}(e)$ and $\mathcal{D}(\tau_n)$ in geometry h . For more details on this derivation, see Refs. [19, 247, 248]. The matter contribution just comes from computing the n th moment of the reduced bulk density matrix on $\mathcal{D}(\tau_n)$, similar to the RTN calculation.

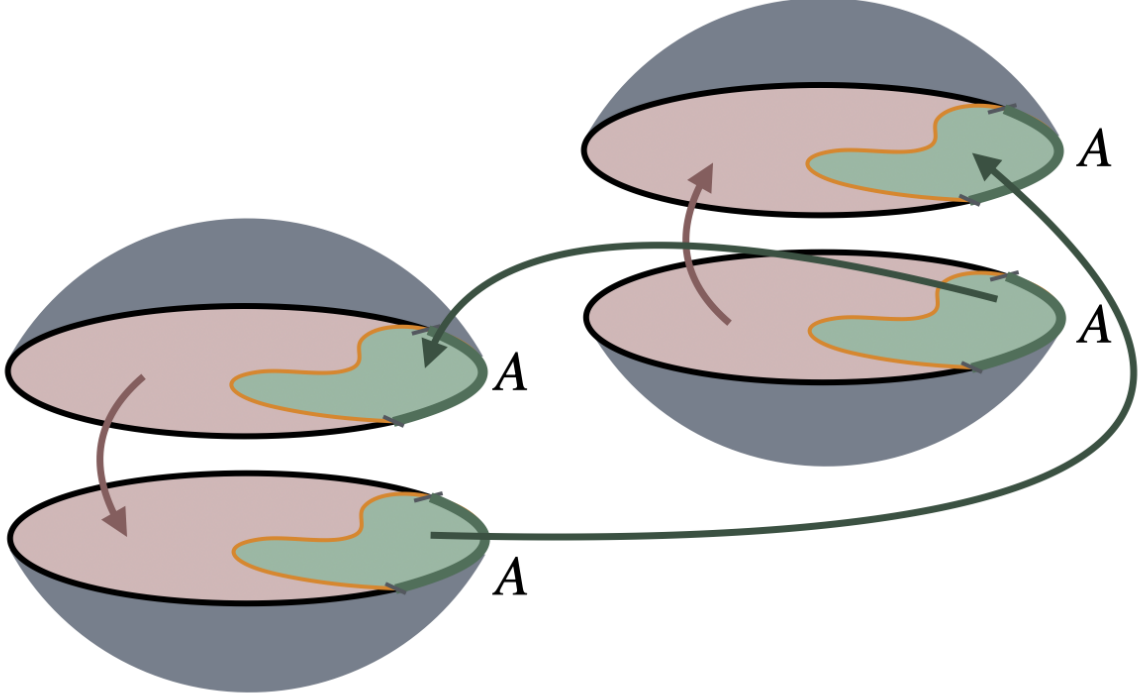


Figure 4.5: The dominant contribution to the replica path integral comes from geometries in which the regions of the bulk slice Σ within the domain wall (sharing a conformal boundary with A) are glued cyclically, while the regions outside the domain wall are glued identically.

Combining the two contributions and minimizing the resulting action, we have

$$\text{Tr}(\rho_A(h)^n) = e^{-(n-1)\frac{\mathcal{A}(\gamma_A, h)}{4G}} \text{Tr}\left[\left(\rho_{E(A)}^{\text{bulk}}(h)\right)^n\right], \quad (4.3.7)$$

where γ_A is the RT surface of A in geometry h , $E(A)$ is the corresponding entanglement wedge and $\rho^{\text{bulk}}(h)$ is the bulk density matrix of the matter fields in the fixed-geometry state. We have explicitly expressed the area of the RT surface as a function of h to clarify its dependence on the induced metric on Σ .

Using Eq. (4.3.7) in Eq. (4.3.4), we obtain:

$$\text{Tr}(\rho_A^n) = \sum_h |\psi(h)|^{2n} e^{-(n-1)\frac{\mathcal{A}(\gamma_A, h)}{4G}} \text{Tr}\left[\left(\rho_{E(A)}^{\text{bulk}}(h)\right)^n\right]. \quad (4.3.8)$$

Finally, we can perform a saddle point approximation over h to obtain

$$\text{Tr}(\rho_A^n) \approx |\psi(h_n)|^{2n} e^{-(n-1)\frac{\mathcal{A}(\gamma_A, h_n)}{4G}} \text{Tr}\left[\left(\rho_{E(A)}^{\text{bulk}}(h_n)\right)^n\right], \quad (4.3.9)$$

where h_n is the saddle point geometry of Eq. (4.3.8). Note that $h_1 = h_\psi$ and the shift in saddle as a function of n reproduces the expected backreaction for Rényi entropies [252].

The Rényi entropy is given by

$$S_n(A) = \frac{2n}{n-1} \log \left(\frac{|\psi(h_\psi)|}{|\psi(h_n)|} \right) + \frac{\mathcal{A}(\gamma_A, h_n)}{4G} + S_n \left(\rho_{E(A)}^{\text{bulk}}(h_n) \right). \quad (4.3.10)$$

The von Neumann entropy is obtained by taking the $n \rightarrow 1$ limit of the preceding expression, yielding the RT formula

$$S(A) = \lim_{n \rightarrow 1} S_n(A) = \frac{\mathcal{A}(\gamma_A, h_\psi)}{4G} + S \left(\rho_{E(A)}^{\text{bulk}}(h_\psi) \right), \quad (4.3.11)$$

where we have made the dependence on the bulk metric explicit.

4.4 BP Proposal from the Gravitational Path Integral

In order to define the entropy of bulk regions, we will now follow a strategy similar to the one employed above for boundary regions. The important new ingredient will be to propose a definition of the hollow-graphic state using the gravitational path integral in the basis of fixed-geometry states. Once we do that, the calculation of the Rényi entropy follows by employing a diagonal approximation. In the $n \rightarrow 1$ limit, we will show that this gives us the BP proposal.

Before we start, it is useful to discuss an important conceptual issue. To define a bulk subregion, we must specify it in a gauge-invariant manner. In particular, to define a bulk subregion in a general holographic state $|\psi\rangle$, which is a superposition of fixed-geometry states $|h\rangle$, we must provide a gauge-invariant definition of the subregion on all geometries h . *A priori*, there are various ways to define such regions gauge-invariantly. Our analysis in the following will remain agnostic about the particular way in which the region is specified, especially because we will recover the BP proposal in the $n \rightarrow 1$ limit independently of these details. For explicit examples of different ways to gauge-fix a subregion, see Sec. 4.5.

Let a be an input wedge specified in some gauge-invariant manner and $|h\rangle$ a fixed-geometry state. To define the entropy of this bulk region, we will take inspiration from the RTN and construct a hollow-graphic state. In the case of the RTN, the hollowing map V_a was defined by removing the tensors in that region. In the case of gravitational states, we propose to hollow it out by removing a tiny slit around the region a to introduce new boundaries a^+ and a^- , see Fig. 4.6. These new boundaries are additional locations where boundary conditions can be specified and thus describe an additional Hilbert space factor.¹⁴

In general, the induced metric on a^-, a^+ can be specified independent of the original h . When we choose the induced metric on a^-, a^+ to agree with the original h , we obtain the proposed hollow-graphic state $|h_a\rangle$. This geometric construction defines the map V_a :

$$|h_a\rangle = V_a |h\rangle. \quad (4.4.1)$$

¹⁴ A similar construction was proposed in the context of cosmology by Ref. [262].

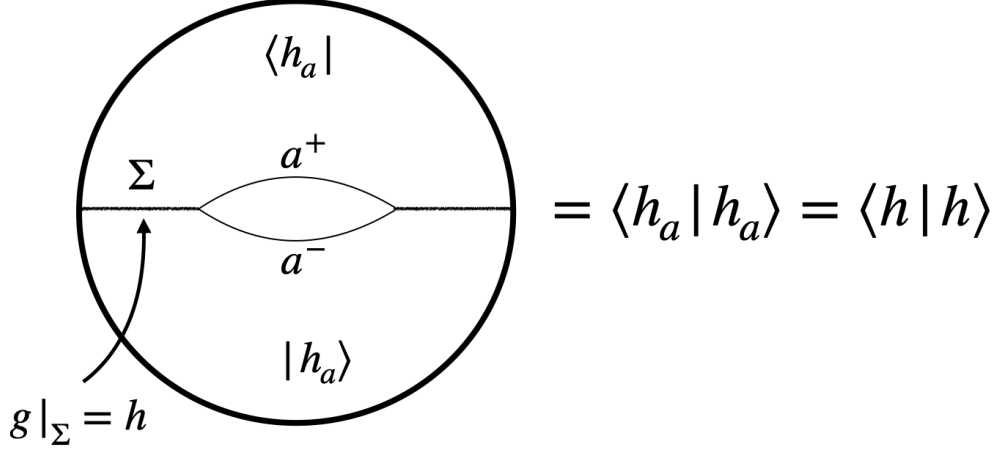


Figure 4.6: A schematic depiction of the Euclidean path integral with additional boundaries a^- and a^+ . When the induced metric on a^- , a^+ is chosen to agree with h on region a , the Euclidean spacetime is the original fixed-geometry saddle with an infinitesimal-width cut around the subregion a . The hollow-graphic map is norm-preserving.

From Fig. 4.6, it is also clear that the Euclidean action remains unchanged under removal of this infinitesimal slit and thus, the norm is preserved, i.e., $\langle h_a | h_a \rangle = \langle h | h \rangle$.

We can then linearly extend the action of V_a to a general holographic state by decomposing it in the basis of fixed-geometry states:

$$|\psi_a\rangle = V_a |\psi\rangle = \sum_h \psi(h) |h_a\rangle, \quad (4.4.2)$$

where we have used the decomposition obtained in Eq. (4.3.1). We remind the reader that this map is well-defined and unique as long as the states $|h\rangle$ must form a basis of the Hilbert space. Although this is not strictly true due to wormhole corrections, we will work in a regime where they provide an approximate basis as we demonstrate in our examples in Sec. 4.5.1.

We can now obtain the reduced density matrix for subregion a by tracing out its complement on $\partial\Sigma$:

$$\rho_a = \text{Tr}_{\partial\Sigma \setminus \partial a} (V_a \rho V_a) = \sum_{h, h'} \psi(h) \psi^*(h') \text{Tr}_{\partial\Sigma \setminus \partial a} (|h_a\rangle \langle h'_a|) . \quad (4.4.3)$$

Using this, we can compute $\text{Tr}(\rho_a^n)$ and just as in the boundary subregion case, there are both diagonal and off-diagonal terms that contribute to it. Motivated by replica symmetry, we will again assume a diagonal approximation and ignore the off-diagonal terms to obtain

$$\text{Tr}(\rho_a^n) = \sum_h |\psi(h)|^{2n} \text{Tr}(\rho_a(h)^n), \quad (4.4.4)$$

where $\rho_a(h)$ is the reduced density matrix on a in the fixed-geometry hollow-graphic state $|h_a\rangle = V_a |h\rangle$, i.e.,

$$\rho_a(h) = \text{Tr}_{\partial\Sigma \setminus \partial a} [|h_a\rangle \langle h_a|] . \quad (4.4.5)$$

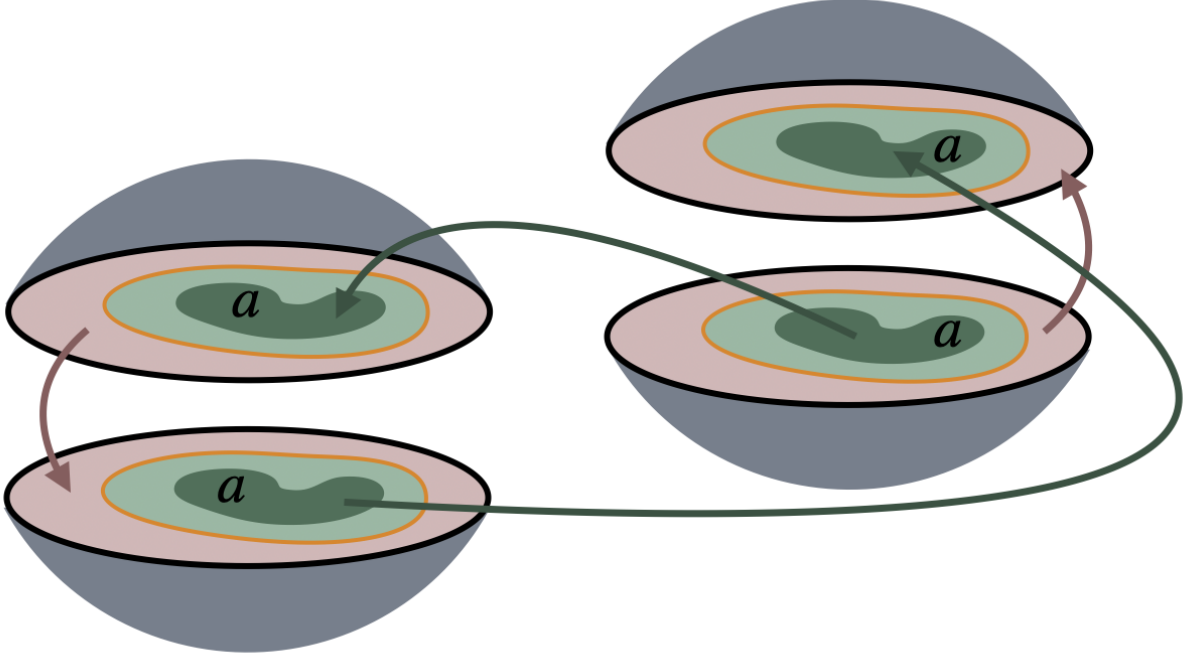


Figure 4.7: The dominant contribution to the replica path integral comes from geometries in which the regions of the bulk slice Σ within the domain wall containing a are glued cyclically, while the regions outside the domain wall are glued identically.

As in the boundary subregion case, $\text{Tr}(\rho_a(h)^n)$ is computed using the gravitational path with gluing rules analogous to those in the tensor network calculation. Namely, we divide the bulk slice Σ into domains $\mathcal{D}(g)$ associated with permutations $g \in \mathcal{S}_n$ and glue n copies of Σ with appropriate boundary conditions. As in the tensor network calculation, the boundary ∂a is assigned the cyclic permutation τ_n , and $\partial\Sigma/\partial a$ is assigned the identity permutation e . See Fig. 4.4 and Fig. 4.7.

As before, the saddle-point configuration preserves replica symmetry and consists only of two domains, $\mathcal{D}(\tau_n)$ and $\mathcal{D}(e)$. The action of the saddle-point configuration is given by the conical singularity along the domain wall $\partial\mathcal{D}(\tau_n)$. Including the contribution from the bulk density matrix, we find

$$\text{Tr}(\rho_a^n) = \sum_h |\psi(h)|^{2n} e^{-(n-1)\frac{\mathcal{A}(\gamma_a, h)}{4G}} \text{Tr}[(\rho_{E(a)}^{\text{bulk}}(h))^n], \quad (4.4.6)$$

where $\mathcal{A}(\gamma_a, h)$ is the area of the surface homologous to ∂a in the geometry h that minimizes the generalized entropy of the region $E(a)$ bounded by it.

The dominant saddle of Eq. (4.4.6) is denoted h_n , with the special case $h_1 = h_\psi$, the leading geometry of the state $|\psi\rangle$. From the form of the action, it is clear that the precise location of h_n depends on n . Moreover, it also depends on how the subregion a is gauge-

invariantly specified since this affects the action functional being optimized. Thus, we have

$$\mathrm{Tr}(\rho_a^n) \approx |\psi(h_n)|^{2n} e^{-(n-1)\frac{\mathcal{A}(\gamma_a, h_n)}{4G}} \mathrm{Tr}\left[\left(\rho_{E(a)}^{\mathrm{bulk}}(h_n)\right)^n\right], \quad (4.4.7)$$

and the Rényi entropy becomes

$$S_n(a) = \frac{2n}{n-1} \log\left(\frac{|\psi(h_\psi)|}{|\psi(h_n)|}\right) + \frac{\mathcal{A}(\gamma_a, h_n)}{4G} + S_n\left(\rho_{E(a)}^{\mathrm{bulk}}(h_n)\right), \quad (4.4.8)$$

where h_n carries implicit n -dependence.

Finally, taking the $n \rightarrow 1$ limit, and recalling that $h_1 = h_\psi$, we have

$$S(a) = \lim_{n \rightarrow 1} S_n(a) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} + S\left(\rho_{E(a)}^{\mathrm{bulk}}(h_\psi)\right). \quad (4.4.9)$$

Importantly, the dependence on the precise way in which the bulk region is defined gauge-invariantly now drops out since all such prescriptions are chosen to agree on the leading geometry h_ψ . Thus, we have derived the BP proposal Eq. (4.4.9) from the gravitational path integral. Moreover, we see that the entanglement wedge $E(a)$ in the geometry h_ψ is independent of the specific gauge-invariant prescription used to extend the definition of a to other geometries. This is consistent with the fact that Refs. [203, 216] didn't need to ever specify a particular gauge-invariant definition of bulk regions.

4.5 Examples

In this section, we will demonstrate how our formalism works in concrete examples where we compute the Rényi entropies for bulk regions on the time-symmetric slice in the thermofield double state. In particular, we will consider JT gravity, and higher-dimensional Einstein gravity in the minisuperspace approximation where we restrict to spherically symmetric configurations. A brief outline of the procedure is as follows:

- i) We begin with the semiclassical state $|\psi\rangle = |TFD(\beta)\rangle$ corresponding to a two-sided black hole of temperature $1/\beta$. The dominant geometry on the time-reflection symmetric slice Σ is denoted h_ψ . We then consider a region a on Σ .
- ii) We expand $|\psi\rangle$ in an (approximate) basis of fixed-geometry states $|h\rangle$. In JT gravity, we use fixed-dilaton states $|\Phi_h\rangle$, while in Einstein gravity, we work with states of fixed horizon area $|\mathcal{A}_h\rangle$.¹⁵ In either case, the wavefunction $\psi(h)$ is given by the Hartle-Hawking wavefunction in the relevant basis.
- iii) We then gauge-invariantly specify the region a in every geometry h . For each h , we use the gravitational path integral to compute the Rényi entropy of a .

¹⁵ In JT gravity, the entire geometry and dilaton profile is determined by the value of Φ at the horizon due to the constraint equations, see Eq. (A.1.4). Similarly, in the spherically symmetric minisuperspace approximation to Einstein gravity, fixing the horizon area fixes the entire geometry

- iv) Finally, we compute the saddle-point in the integral over geometries in Eq. (4.4.6) to obtain the Rényi entropy $S_n(a)$ in the semiclassical state $|\psi\rangle$. As $n \rightarrow 1$, this reproduces the BP proposal: $S(a) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G}$, where $\mathcal{A}(\gamma_a, h_\psi)$ is the minimal area surface homologous to a in the eternal black hole geometry.

By providing various different methods of specifying the region a , we will show explicitly how, for $n > 1$, the Rényi entropies depend on this choice, but as $n \rightarrow 1$, the dependence vanishes giving an entanglement entropy consistent with the BP proposal.

4.5.1 JT Gravity

To address (i) and (ii) above, we now briefly present the Hartle-Hawking wavefunction $\psi(h)$ for the two-sided geometry. For details of the computation, largely following Ref. [263], see A.1.2. Then we will address (iii) and (iv) by providing different gauge-invariant specifications for an interval in one of the exterior regions, and by computing the corresponding saddles for the Rényi entropy.

We begin with the thermofield double state prepared by a Euclidean evolution $\beta/2$, labelled $|\psi_{\beta/2}\rangle$. This Hartle-Hawking state can be expanded in the fixed-dilaton basis as $|\psi_{\beta/2}\rangle = \sum_{\Phi_h} \psi_{\beta/2}(\Phi_h) |\Phi_h\rangle$, where Φ_h denotes the value of the dilaton at the horizon. This provides a complete basis for the Hilbert space in the approximation where we ignore worm-hole contributions. The wavefunction $\psi_{\beta/2}(\Phi_h) = \langle \Phi_h | \psi_{\beta/2} \rangle$ is constructed using the Euclidean gravitational path integral with an asymptotic AdS boundary ∂M of length $\beta/2$, and an additional codimension-1 boundary in the bulk given by a slice Σ with the specified dilaton value at the horizon. See figure 4.3. The saddle-point geometry g_{Φ_h} of this action is a section of the two-sided black hole geometry with a kink of angle $\theta = \frac{\beta}{2}\Phi_h$ at the horizon, depicted in Fig. 4.8. Evaluating the action, the wavefunction is found to be

$$\psi_{\beta/2}(\Phi_h) \approx e^{-I_E[g_{\Phi_h}]} = \exp \left\{ \pi(\Phi_0 + \Phi_h) - \frac{\beta}{4} \Phi_h^2 \right\}. \quad (4.5.1)$$

Having decomposed $|\psi_{\beta/2}\rangle$ into fixed-dilaton states $|\Phi_h\rangle$, we are now interested in computing the entropy of an interval on Σ . To discuss this, it is useful to write down the metric and dilaton profile on Σ for each choice of Φ_h :

$$ds^2 = \frac{dx^2}{1+x^2} \quad \Phi = \Phi_h \sqrt{1+x^2}, \quad (4.5.2)$$

where the slice Σ is parametrized by coordinate $x \in (-\infty, \infty)$ and the horizon is at $x = 0$.

Let A and B denote the spacetime points at the boundaries of a , so that $\partial a = A \cup B$. Generically, for any gauge-invariant specification of the endpoints, there are two cases to consider:

- I) The endpoints A and B both lie on the same side of the horizon, for example $x_B > x_A > 0$.

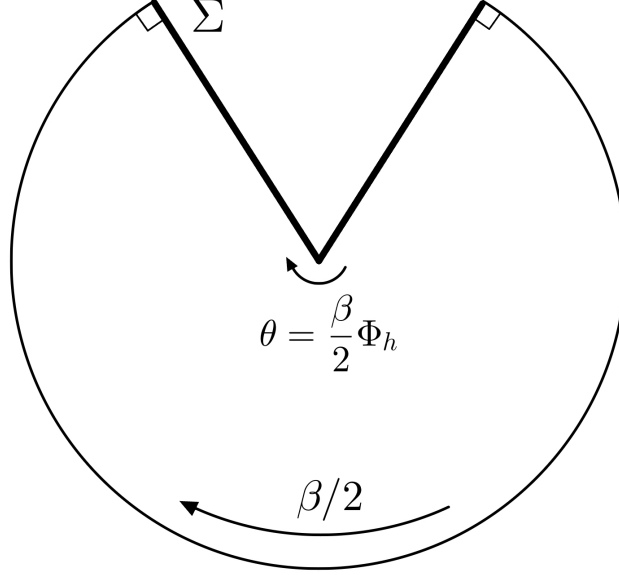


Figure 4.8: The Hartle-Hawking wavefunction is computed using a saddle-point geometry g_{Φ_h} with an asymptotic boundary of length $\frac{\beta}{2}$ and a boundary Σ with the specified dilaton value Φ_h at the horizon. At generic values of Φ_h , there is a kink in the slice Σ .

- II) The endpoints lie on opposite sides of the horizon, so that a contains the bifurcation point. For example, we may have $x_A < 0 < x_B$.

One also needs to be able to specify which side of the horizon the points A and B lie on because of the exact symmetry $\mathbb{Z}_2$ in this problem. This can be done by using the asymptotic boundaries as a reference.¹⁶ Due to the radial monotonicity of $\Phi(x)$, in case I, the entanglement wedge of a extends to the horizon, and properly contains a , i.e., $e(a) \supsetneq a$ while in case II, they are equal, i.e., $e(a) = a$. See Fig. 4.9.

In what follows we consider case I, and present the result for case II in Appendix A.2. In any given basis state labelled by Φ_h , the extremal surface homologous to a is given by $\gamma_a = H \cup B$ where H denotes the bifurcation point. Its area functional is

$$\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi(\Phi_h + \Phi_B), \quad (4.5.3)$$

where we note that $2\pi\Phi$ plays the role of $\frac{A}{4G}$ in JT gravity. Note that Φ_B in Eq. (4.5.3) is also implicitly a function of Φ_h depending on how we specify the subregion. With this setup in mind, we will now use a few different methods of gauge-invariantly specifying $a \subset \Sigma$ not only in the dominant semiclassical geometry, but in all fixed-dilaton states.

¹⁶ To illustrate this point, we note that specifying the value of Φ determines $|x|$, but not its sign. The sign can be additionally determined by specifying whether or not the endpoint is on the same side as a chosen boundary.

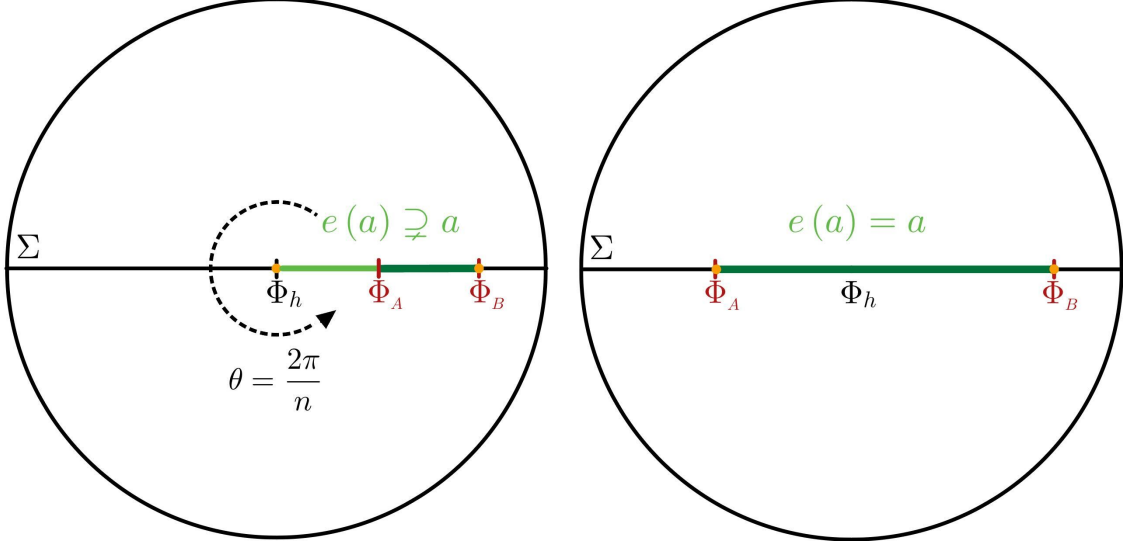


Figure 4.9: Left: For a bulk interval on a single side of the horizon, the entanglement wedge extends to the horizon. In the case where the value of the dilaton at the endpoints is specified, the opening angle at the horizon is given by $\frac{2\pi}{n}$ so that the parent spacetime is smooth. Right: For a bulk interval which crosses the horizon, the entanglement wedge is the region a itself.

Fixing the Dilaton

The first and simplest gauge-invariant specification for $a \in \Sigma$ is to directly fix the values of the dilaton at its endpoints, along with which side of the horizon the endpoints lie. Suppose that $\Phi = \Phi_A$ and $\Phi = \Phi_B$ are both fixed on the right with $\Phi_B > \Phi_A > 0$. See Fig. 4.10.

To compute the Rényi entropy for the bulk region a , we substitute our expression for $\mathcal{A}(\gamma_a, h)$ into our master formula Eq. (4.4.6) to obtain¹⁷

$$\text{Tr}(\rho_a^n) = \sum_h |\psi(\Phi_h)|^{2n} \exp\left\{-(n-1) \frac{\mathcal{A}(\gamma_a, h)}{4G}\right\} \quad (4.5.4)$$

$$= \sum_{\Phi_h \leq \Phi_A} \exp\left\{2\pi n(\Phi_0 + \Phi_h) - \frac{n\beta}{2} \Phi_h^2 - (n-1)2\pi(\Phi_0 + \Phi_h + \Phi_B)\right\} \quad (4.5.5)$$

Upon extremizing over values of Φ_h , the saddle-point geometry h_n is given by

$$\Phi_{h_n} = \frac{2\pi}{n\beta} \quad (4.5.6)$$

Using Eq. (4.4.8), we have

$$S_n(a) = 2\pi\Phi_0 + 2\pi\left(\frac{2\pi}{\beta} + \Phi_B\right) + \frac{1-n}{n} \frac{2\pi^2}{\beta} \quad (4.5.7)$$

¹⁷ We must restrict to geometries where $\Phi_h \leq \Phi_A$, so that the subregion a is well-defined.

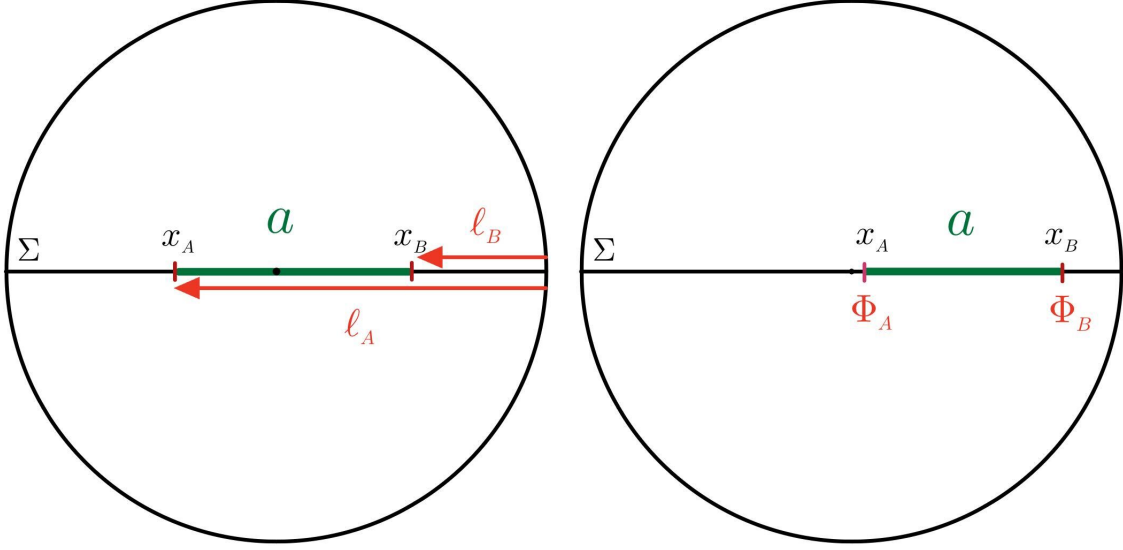


Figure 4.10: Two gauge-invariant specifications of region a : On the left, we specify the value of the dilaton at its endpoints, along with whether the points lie on the same or opposite sides of the horizon from a chosen asymptotic boundary. On the right, we specify the geodesic distance of its endpoints from a chosen asymptotic boundary.

$$S(a) = \lim_{n \rightarrow 1} S_n(a) = 2\pi\Phi_0 + 2\pi\left(\frac{2\pi}{\beta} + \Phi_B\right) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G}, \quad (4.5.8)$$

where the last line is consistent with the BP proposal.

We note that the saddle-point in this example is related to the semiclassical geometry via $\beta \rightarrow n\beta$. This indicates an opening angle of $\theta = 2\pi/n$. This is interesting since it demonstrates that the full spacetime is smooth once we glue together n copies around the horizon. This occurs since specifying the dilaton value at the endpoints is a local specification and thus, no singularities are introduced in regions outside a . This will not be the case generically as we show in other examples.

Fixing the proper distance from the asymptotic boundary

We now consider another method for gauge-invariantly specifying subregions a . Rather than directly fixing the value of the dilaton on ∂a , we fix the geodesic distance between ∂a and an asymptotic boundary. The geodesic length L from the boundary (say, at cutoff x_c) to a point x_0 is $L = \int_{x_0}^{x_c} \frac{1}{\sqrt{1+x^2}} dx = \sinh^{-1} x_c - \sinh^{-1} x_0$. This diverges as $x_c \rightarrow \infty$. Therefore, we define a renormalized geodesic distance as:

$$\ell \equiv L - \log(2\Phi|_{\partial\Sigma}) \quad (4.5.9)$$

where $\Phi|_{\partial\Sigma} = \Phi_h \sqrt{1+x_c^2}$. This gives

$$\ell \xrightarrow{x_c \rightarrow \infty} = -\sinh^{-1} x_0 - \log(2\Phi_h) = -\log\left(\Phi_h\left(x_0 + \sqrt{1 + x_0^2}\right)\right) \quad (4.5.10)$$

$$= -\log\left(\sqrt{\Phi(x_0)^2 - \Phi_h^2} + \Phi(x_0)\right) \quad (4.5.11)$$

or equivalently

$$x_0 = -\sinh\left(\ell + \log \Phi_h\right) \quad (4.5.12)$$

Moreover, the value of the dilaton at x_0 is determined by ℓ and Φ_h as

$$\Phi(x_0) = \Phi_h \cosh\left(\ell + \log \Phi_h\right) \quad (4.5.13)$$

Therefore, we can uniquely determine the location of $\partial a = A \cup B$ by specifying the renormalized geodesic distances ℓ_A, ℓ_B of the endpoints from a chosen asymptotic boundary. See Fig. 4.10. We assume without loss of generality that $\ell_A < \ell_B$. With these fixed, the extremal area becomes

$$\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi(\Phi_h + \Phi_B) = 2\pi\Phi_0 + 2\pi(\Phi_h + \Phi_h \cosh(\ell_B + \log \Phi_h)). \quad (4.5.14)$$

The saddles for $\text{Tr}(\rho_a^n)$ are given by

$$\Phi_{h_n} = \frac{2\pi/\beta}{\frac{2\pi}{\beta}e^{\ell_B}(n-1) + n} = \frac{2\pi}{\beta} - \frac{\pi}{\beta} \left(\frac{2\pi}{\beta} e^{\ell_B} + 1 \right) (n-1) + O((n-1)^2), \quad (4.5.15)$$

where we have expanded the exact formula near $n = 1$. The corresponding entanglement entropies can then be checked to be

$$S(a) = 2\pi\Phi_0 + 2\pi(\Phi_{h_\psi} + \Phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (4.5.16)$$

in agreement with the BP proposal as expected.

An important point in the above calculation is that the parent spacetime obtained by gluing together n copies will not be smooth at the horizon, unlike the previous example. This is allowed since the specification of the subregion is more non-local and thus, can backreact on the spacetime.

Fixing the dilaton relative to the horizon

Instead of fixing the value of the dilaton at the boundary of the subregion, we can instead fix its value relative to the value at the horizon, in a given geometry. For example, we may specify $\Phi_A/\Phi_h = \phi_A$ and $\Phi_B/\Phi_h = \phi_B$. This again determines the location of the endpoints x_A and x_B up to which side of the geometry they lie on.¹⁸ With $\phi_B > \phi_A > 0$, we have

$$\mathcal{A}(\gamma_a, h) = 2\pi\Phi_0 + 2\pi\Phi_h(1 + \phi_B) \quad (4.5.17)$$

¹⁸ Unlike fixing the value of the dilaton directly, in this case we need not project onto the subspace of states with $\Phi_h < \Phi_A$ to have a well-defined subregion. Instead, for $\phi_A, \phi_B > 0$, the subregion is guaranteed to exist for all choices of Φ_h .

This gives the following saddle for $\text{Tr}(\rho_a^n)$:

$$\Phi_{h_n} = \frac{2\pi}{\beta} - \frac{(n-1)}{n} \frac{2\pi}{\beta} (1 + \phi_B) \quad (4.5.18)$$

The corresponding Rényi and Von Neumann entropies are

$$S_n(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (1 + \phi_B) - \left(\frac{n-1}{n}\right) \frac{2\pi^2}{\beta} (1 + \phi_B)^2 \quad (4.5.19)$$

$$S(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (1 + \phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G}. \quad (4.5.20)$$

Fixing the proper distance from the horizon

Instead of fixing the proper distance between ∂a and the asymptotic boundary, we can also fix its distance from the horizon. Since this is a finite quantity, we need not use the renormalized geodesic distance, and we can fix the proper distance directly. Since proper distance is symmetric across the horizon, we must also specify, as before, which side of the geometry the endpoints lie on. To illustrate this, note that the distance from the horizon to a point x_0 is

$$L = \int_0^{x_0} \frac{1}{\sqrt{1+x^2}} |dx| = \sinh^{-1} |x_0| \quad (4.5.21)$$

Thus $x_0 = \pm \sinh(L)$. The value of the dilaton at x_0 is determined from L and ϕ_h as

$$\Phi(x_0) = \Phi_h \cosh L \quad (4.5.22)$$

Without loss of generality, suppose $L_B > L_A$. Then,

$$\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi(\Phi_h + \Phi_B) = 2\pi\Phi_0 + 2\pi\Phi_h(1 + \cosh L_B) \quad (4.5.23)$$

The saddles of $\text{Tr}(\rho_a^n)$ occur at

$$\Phi_{h_n} = \frac{2\pi}{\beta} - \frac{(n-1)}{n} \frac{2\pi}{\beta} (1 + \cosh L_B) \quad (4.5.24)$$

which gives the Rényi and Von Neumann entropies

$$S_n(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (1 + \cosh L_B) - \left(\frac{n-1}{n}\right) \frac{2\pi^2}{\beta} (1 + \cosh L_B)^2 \quad (4.5.25)$$

$$S(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (1 + \cosh L_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (4.5.26)$$

Again, the last line is in agreement with the BP proposal.

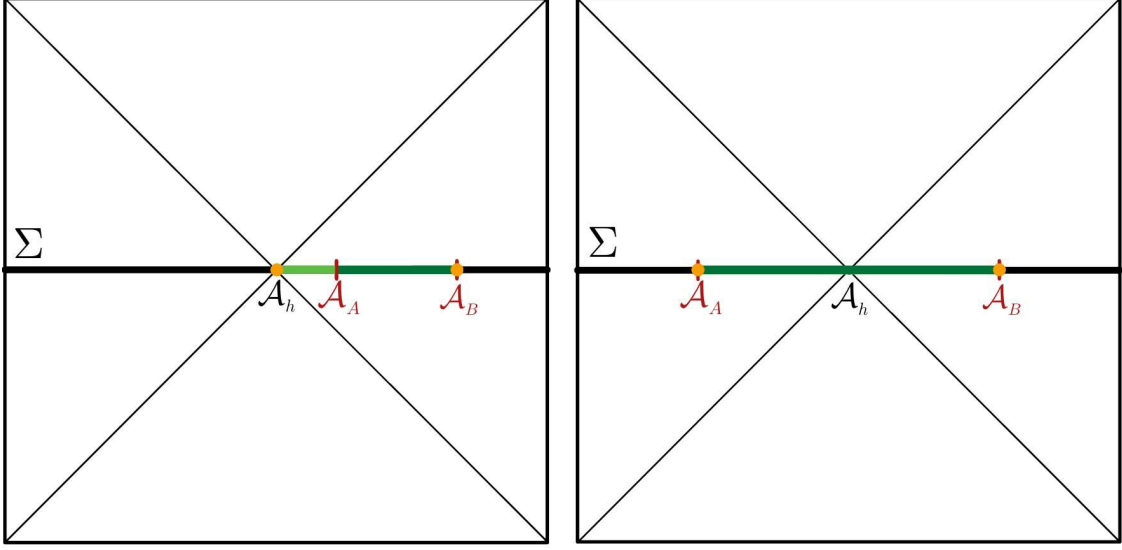


Figure 4.11: The region we choose is an annulus bounded by spheres with areas $\mathcal{A}_A$ and $\mathcal{A}_B$. On the left, the annulus is on the same exterior region in which case, the entanglement wedge extends out to the horizon. On the right, the annulus includes the horizon in which case, the entanglement wedge is the same as the region itself.

4.5.2 Einstein Gravity : Minisuperspace Approximation

Similar to the calculation in JT gravity, we can construct the Hartle-Hawking state for $(d+1)$ -dimensional black holes in the minisuperspace approximation to Einstein gravity where we assume spherical symmetry [264]. The state is likewise prepared via the Euclidean gravitational path integral with boundary conditions consisting of an asymptotic AdS boundary of length $\beta/2$, and a bulk timeslice with fixed geometry h . In the minisuperspace approximation, the geometry of slice Σ can be uniquely determined by its energy, or equivalently by its horizon area. We find that in the basis of states with fixed horizon area, denoted $|\mathcal{A}_h\rangle$, the wavefunction is given by

$$\psi_{\beta/2}(\mathcal{A}_h) = \langle \mathcal{A}_h | \psi_{\beta/2} \rangle = e^{-I[g_h]} = \exp\left\{\frac{1}{2}S(\mathcal{A}_h) - \frac{\beta}{2}E(\mathcal{A}_h)\right\} \quad (4.5.27)$$

where $S(\mathcal{A}_h) = \frac{\mathcal{A}_h}{4G}$ is the entropy, and $E(\mathcal{A}_h)$ is the energy of the slice Σ in geometry h , determined through

$$E = \frac{(d-1)\Omega_{d-1}}{16\pi G_N} r_h^{d-2} (1 + r_h^2) \quad (4.5.28)$$

where r_h is the horizon radius, e.g. $\mathcal{A}_h = \Omega_{d-1} r_h^{d-1}$. See appendix A.1.2 for the details.

We consider simply-connected regions which are bounded by two spheres, say A and B , which are centered around the origin. That is, $\partial a = A \cup B$. As in the case of dilaton gravity, we generically have two cases, where A and B bound a region on a single side of the geometry,

or a region spanning the horizon, see Fig. 4.11. We again leave the latter for appendix A.2. Since the areas of spheres monotonically increase with their radius, in any geometry h , the extremal surface homologous to ∂a is

$$\gamma_a = H \cup B \quad (4.5.29)$$

where A and B are no longer spacetime points, but are $d - 1$ dimensional spheres in Σ centered around the horizon H . See Fig. 4.11. Its area in geometry h is

$$\mathcal{A}(\gamma_a, h) = \mathcal{A}(H, h) + \mathcal{A}(B, h) \quad (4.5.30)$$

The areas of these surfaces depend on how we gauge-invariantly specify the region ∂a . Analogous to fixing the value of the dilaton in JT gravity, the simplest specification is to directly fix the area of its bounding surfaces, along with which side of the geometry A and B lie on. Specifically, we fix $\mathcal{A}(A, h) \equiv \mathcal{A}_A$ and $\mathcal{A}(B, h) \equiv \mathcal{A}_B$, and specify whether A and B lie on the same or opposite sides of chosen asymptotic boundary. Note that $\mathcal{A}(H, h) = \mathcal{A}_h$ by definition. With this choice of specification, the Rényi entropies are computed as follows.

$$\text{Tr}(\rho_a^n) = \sum_{\mathcal{A}_h} |\psi(\mathcal{A}_h)|^{2n} \exp\left\{\frac{1}{2}S(\mathcal{A}_h) - \frac{\beta}{2}E(\mathcal{A}_h) - (n-1)\frac{\mathcal{A}(\gamma_a, h)}{4G}\right\} \quad (4.5.31)$$

The saddle-point value of the horizon area $\mathcal{A}_{h_n}$ is given by

$$\mathcal{A}_{h_n} = \Omega_{d-1} \left(\frac{2\pi}{dn\beta} + \sqrt{\frac{4\pi^2}{d^2 n^2 \beta^2} - \left(\frac{d-2}{d}\right)} \right)^{d-1} \quad (4.5.32)$$

The resulting Rényi and Von Neumann entropies are

$$S_n(a) = \frac{1}{1-n} \log \left(\frac{\text{Tr}(\rho_a^n)}{(\text{Tr}\rho_a)^n} \right) = \frac{\mathcal{A}_{h_n} + \mathcal{A}_B}{4G} + O(1-n) \quad (4.5.33)$$

$$S(a) = \frac{\mathcal{A}_{h_\psi} + \mathcal{A}_B}{4G} = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (4.5.34)$$

in agreement with the BP proposal.

Finally, we also observe that the saddle for $\langle \psi | \psi \rangle$, namely $\mathcal{A}_\psi$, is related to $\mathcal{A}_{h_n}$ by $\beta \rightarrow n\beta$, indicating an opening angle of $\theta = \frac{2\pi}{n}$ around the horizon, as expected from the local specification of the subregion. This is a generalization of the standard Lewkowycz-Maldacena saddle where there is a conical defect of opening angle $2\pi n$ at B , but the geometry is smooth everywhere else.

4.6 Discussion

4.6.1 Min and Max Entanglement Wedges

In this paper, we restricted our attention to semiclassical states for which entropies can be computed under the assumption of replica symmetry. However, Ref. [16] showed that this

assumption breaks down for a class of states known as *incompressible states*, which can be constructed straightforwardly even in static spacetimes. In such cases, one must invoke tools from one-shot state merging protocols. This framework requires distinguishing between *max entanglement wedges*, where the probability of successful reconstruction is close to 1, and *min entanglement wedges*, where reconstruction is greater than some infinitesimal but nonzero ϵ . Here, we briefly comment on how such min and max entanglement wedges may be obtained via the Petz map reconstruction.

In Ref. [19], the Petz map was explicitly computed using the gravitational path integral to determine operator reconstructability.¹⁹ The authors identified two distinct regimes: one in which an operator is fully reconstructible and another in which it is entirely unreconstructible. In particular, they showed that the matrix elements of the boundary operator O_R implements the following in the bulk:

$$\langle a|O_R|b\rangle = c_1 \langle a|O|b\rangle + c_2 \delta_{ab} \text{Tr } O \ , \quad (4.6.1)$$

where $\langle a|O|b\rangle$ is the matrix elements of the operator in the bulk region. They estimated the coefficients c_1 and c_2 to quantify the quality of reconstruction.

Although Ref. [19] found a sharp transition between full reconstructability and non-reconstructability, this transition can be made more gradual by increasing the code space dimension. This allows for intermediate regimes in which reconstruction is possible with a finite $O(1)$ probability. In principle, for a fixed input bulk region, one could similarly examine reconstructability of operators localized at different points in spacetime by computing the Petz map and estimating the coefficients c_1 and c_2 . The reconstruction success probability is then given by $1 - c_2$. If this probability is sufficiently high, i.e., for $0 < c_2 < \epsilon$, the operator can be considered to be part of the max entanglement wedge. On the other hand, if this probability is low but non-zero, i.e., for $\epsilon < c_2 < 1$, it belongs to the min wedge.

While this method is not practical for general computations, it provides an *in principle* path integral prescription for determining min and max entanglement wedges.

4.6.2 Time Dependence

An important direction for future work is the open problem of extending our methods to time-dependent settings. Our primary inspiration, the tensor network picture, does not straightforwardly generalize beyond static space-times. This limitation is compounded by the fact that the distinction between min and max entanglement wedges already arises at the classical level. This suggests that the path integral construction used to compute entanglement wedges in general time-dependent spacetimes must go beyond the standard assumption of replica symmetry. Although Ref. [16] demonstrated how to identify min and max entanglement wedges of boundary subregions in static settings using fixed-area states, it remains poorly understood how to define or compute such wedges in time-dependent geometries, even for boundary input regions. Despite these challenges, progress can still

¹⁹ A similar calculation was done for RTNs in Ref. [265].

be made in certain restricted scenarios. We will explore some of these developments in forthcoming work [266].

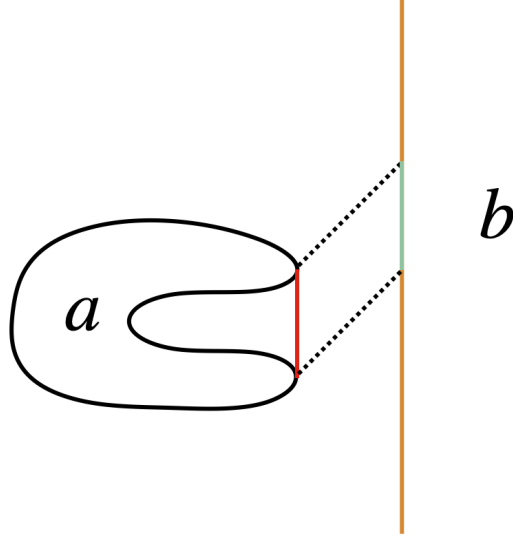


Figure 4.12: An example configuration illustrating subtleties that arise in time-dependent settings. For simplicity, we work at leading order in G and choose wedges a and b such that we have $e_{\max} = e_{\min} = e$. In this case, $e(a)$ is given by a restricted maximin procedure (see Ref. [266] for a proof of this statement). Consider a subregion a that forms a horseshoe-shaped wedge on a Cauchy slice in asymptotically flat spacetime. The portion of the boundary of $e(a)$ that does not coincide with the boundary of a is shown in red. Subregion b is chosen such that a portion of its boundary (shown in green) is null like to the red segment of $e(a)$. Generically, this green portion of ∂b will have smaller area than the red portion of $e(a)$. We also choose b such that it is normal everywhere such that $e(b) = b$. Thus a and b satisfy $a \subset e(b)'$ and $b \subset e(a)'$. Since computing the restricted entanglement wedge $e(a)|_{b'}$ requires all candidate Cauchy slices to include the boundary of b , the surface $e(a)$ can no longer be a valid maximin surface in b' . Instead, a new surface—comprising the green portion of ∂b along with the null surfaces that connect it to region a —will have smaller area and be present on all Cauchy slices. This surface therefore becomes $e(a)|_{b'}$.

Apart from the gravitational path integral derivation, there are also questions to answer about entropy inequalities and independence conditions. An interesting subtlety that arises in the context of independence conditions is that, in time-dependent settings, an independence condition of the form given in Eq. (4.2.16) is not sufficient to guarantee that the restricted entanglement wedges coincide with their unrestricted counterparts. For an explicit example, see Fig. 4.12. In this light, it may seem surprising that Ref. [236] was able to use such an independence condition to prove the monogamy of mutual information (MMI). A more detailed investigation of independence conditions in time-dependent scenarios is an

important direction for future work.

A refinement of the BP proposal was recently put forward by Ref. [257], which modifies the boundary conditions in both time-dependent and time-independent settings by introducing the fundamental complement of a bulk subregion. We note that the construction in Ref. [257] is formulated entirely in Lorentzian signature. The essential features of this approach—including the role of the fundamental complement as a boundary condition—are expected to arise naturally within a fully Lorentzian path integral framework and are not anticipated to emerge in the Euclidean path integral setting employed in this work. We leave a detailed investigation of this direction for future work.

Chapter 5

A Timelike Quantum Focusing Conjecture

5.1 Introduction

Entanglement entropy and complexity have come to be recognized as two important measures of quantum information in gravitational systems. The significance of entropy was first recognized in the context of black hole thermodynamics, where it was observed that in order for the second law of thermodynamics to be preserved, the area of causal horizons must bound the thermal entropy contained behind them [1, 2, 31]. This led naturally to the concept of generalized entropy as the sum of the areas of all horizons plus the entropy of the matter on their exteriors, as well as to a generalized second law [3]. More recently, the concept of generalized entropy has become best understood for arbitrary codimension-2 surfaces as the sum of their area (divided by $4G\hbar$) and the entanglement entropy of the semiclassical matter fields on a slice through their interior [127]. In the context of holography, generalized entropy is central to understanding the flow of information in evaporating black hole systems and appears in the QES prescription as the AdS bulk dual to the entanglement entropy of CFT subregions [15, 18, 55]. It is now believed that these methods can be extended to include the von-Neumann entropy of bulk subregions [216, 267, 268]. Beyond the context of holography, focusing arguments have been used to understand the dynamics of generalized entropy for general codimension-2 surfaces. In particular, simple assumptions about the quantum expansion of surfaces have been used to understand entropy bounds, such as the covariant entropy bound, and quantum energy conditions such as the qNEC, and used to predict the existence of quantum singularities [127, 219, 269].

Nevertheless, it is well known that the entanglement dynamics are not sufficient to account for all of the known gravitational dynamics [22]. Computational complexity was first introduced in the context of gravity to understand the growth of the volume of the black hole interior after its entropy has saturated [25, 270]. In that context, it was proposed that the holographic dual of CFT complexity is the volume of the maximal slice through the black

hole spacetime. Over the last decade, holographic complexity has become better understood, culminating in many different refinements of the original complexity/volume proposal. These include generalizations to subregions of the boundary CFT [27, 28, 271], generalizations to a broader class of candidate bulk duals beyond the volume [173, 272–274], and even an analysis of quantum corrections for different candidate complexity measures [275, 276]. Owing partially to the inherent ambiguities in quantifying the computational complexity for field theory state (e.g. choice of fundamental gate sets, cost functions, reference states), and partially to the relative recency of its introduction into quantum gravity, complexity lacks the robust theoretical understanding enjoyed by entanglement entropy. For example, it is not known which (if any) of the broad class of existing candidate complexity duals is most relevant to holography. This is largely due to our lack of a unified framework for computing holographic complexity, such as the path integral techniques used for computing holographic entanglement entropy.

It is widely understood that quantum corrections to the geometric part of any holographic complexity must be included to fully account for the dynamics beyond the classical limit. The need for such corrections strongly suggests that complexity is dual not purely to geometric volume (or other proposed measures), but must include corrections from the complexity of the semiclassical bulk quantum fields. The authors of [277] argue that a notion of generalized complexity may be necessary to quantify the complexity dynamics of evaporating black holes. They demonstrate that a generalization of the CV proposal which includes bulk quantum corrections may be used to understand the appearance of the quantum extremal island, and they argue that the resulting complexity growth agrees with expectations from finite-dimensional qubit models. This leads naturally to the question of whether such a notion of generalized complexity may be applicable to generic spacetime subregions.

In what follows, we introduce a notion of generalized complexity for arbitrary partial slices Σ as the sum of their volume (in appropriate units) plus corrections from the semiclassical QFT complexity of the fields associated with Σ . In order to study the dynamics of this generalized complexity, we introduce a notion of timelike quantum expansion, as the variation of the generalized complexity under local deformations of Σ along an orthogonal timelike congruence through Σ . We find that simple focusing assumptions lead to a complexity-based quantum strong energy condition, as well as a covariant bound on the complexity of fields by the geometric volume of their support. We demonstrate that the focusing construction requires certain consistency conditions on the measure of QFT complexity. For example, it must be codimension-0, and it must obey strong subadditivity. We regard these conditions as a method for narrowing down large class of existing complexity measures, in the absence of reined mathematical control, to those natural for semiclassical gravity.

In section 2, we introduce our definition of generalized complexity, and define a timelike quantum expansion. In section 3, we investigate the consequences of the timelike quantum focusing assumption, and show that it entails a quantum energy condition, a complexity bound, and that its consistency requires strong-subadditivity of QFT complexity. For clarity of presentation, in sections 2 and 3 we adapt the notion of the original QFC paper [127] to the context of timelike congruences. In section 4, we discuss a classical analogue of the derived entropy bound, an application to diagnosing quantum singularities, and then explore the relationship between the quantum Strong Energy Condition, its classical counterpart, and

the cosmological constant.

5.2 Definitions

5.2.1 Generalized Complexity

[277] proposes that an appropriate extension of CV for incorporating islands is the following. Given a boundary region R , the holographic complexity of this subregion is

$$C(R) = \max_{\Sigma \in e(R)} \left(\frac{\text{Vol}(\Sigma)}{\hbar G \ell} + C_{QFT}(\Sigma) \right) \quad (5.2.1)$$

where $\Sigma \in e(R)$ is a slice through the bulk component of the entanglement wedge of R , and $C_{QFT}(\Sigma)$ is the computational complexity of the semiclassical fields propagating in $e(R)$. This motivates the introduction of a notion of generalized complexity for arbitrary subregions:

Given a partial Cauchy slice Σ , we define its generalized complexity

$$C_{gen}(\Sigma) = \frac{\text{Vol}(\Sigma)}{\hbar G \ell} + C_{QFT}(\Sigma) \quad (5.2.2)$$

so that Eq. (5.2.1) takes a form analogous to the QES prescription, but with a maximization over codimension-1 slices, rather than a minimization over codimension-2 surfaces:

$$C(R) = \max_{\Sigma \in e(R)} C_{gen}(\Sigma) \quad (5.2.3)$$

We do not yet make any specific claim about which complexity measure is used to quantify $C_{QFT}(\Sigma)$, nor the length scale ℓ . Many proposed complexity measures exist throughout the literature [278–281], but there is no reason any of them should be excluded at this step. We take the point of view that there exists a measure (potentially many) for the complexity of the full quantum gravitational degrees of freedom associated with the slice Σ , and that this measure can be decomposed into a geometric piece and a piece associated with the semiclassical fields propagating on Σ . For any given complexity measure, the geometric and QFT terms will be separately cutoff dependent, but we expect that their sum is cutoff independent. Any rigorous justification will require a specific complexity measure. The cutoff dependence of certain measures have been studied, for example, in [279, 280, 282–285].

5.2.2 Timelike Quantum Expansion

Consider a congruence T of timelike geodesics which is torsionless, and which is hypersurface orthogonal to some partial Cauchy slice Σ_0 . We can introduce coordinates near the surface Σ_0 given by (τ, x) such that Σ_0 is parameterized by $(0, x) \in \Sigma_0$. That is, τ is the coordinate

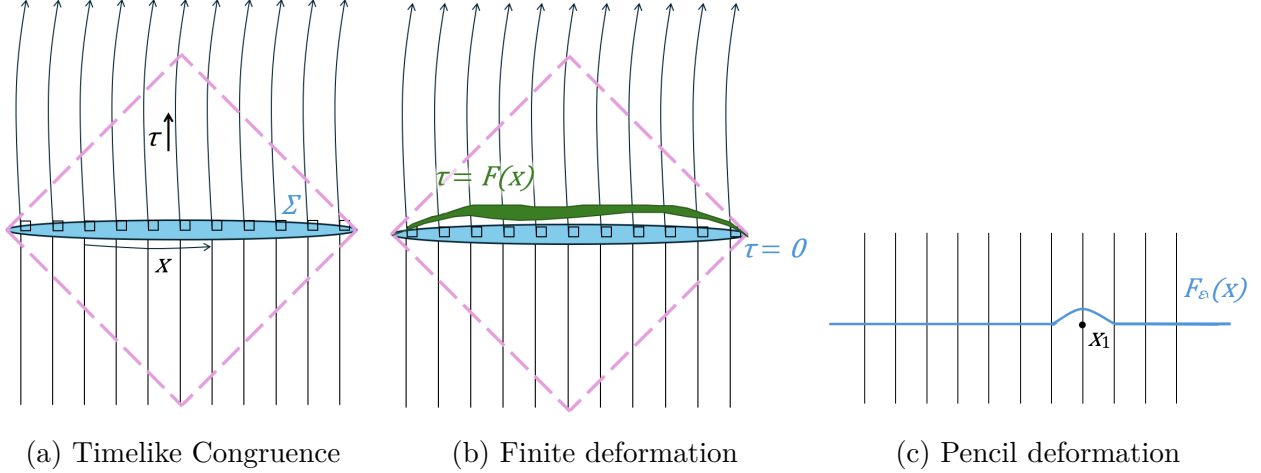


Figure 5.1: Panel (a) depicts a hypersurface orthogonal timelike congruence generated by a slice Σ_0 . In panel (b), deformations of the slice are defined by surfaces $\tau = F(x)$ where τ is the proper time along the timelike generator through $x \in \Sigma_0$. Panel (c) depicts a local deformation along a “pencil” of generators near a point x_0 .

orthogonal to the surface Σ_0 , and labels the proper time along the geodesic that intersects Σ_0 at $(0, x)$. See Fig. 5.1.

Other partial Cauchy slices through the domain of dependence of Σ_0 (i.e. those which share a boundary with Σ_0) can be obtained using a positive definite function $\tau = F(x)$, with the condition that $F(x) = 0$ for $x \in \partial\Sigma_0$. To ensure that the surface $\tau = F(x)$ is spacelike, we also restrict to the class of functions satisfying $|\nabla F(x)|^2 < 1$. This condition ensures that the normal vector to Σ remains timelike. By defining partial Cauchy slices in this way, we can regard the generalized complexity as a functional of the embedding $F(x)$:

$$C[F(x)] = \frac{\text{Vol}[F(x)]}{\hbar G \ell} + C_{QFT}[F(x)] \quad (5.2.4)$$

We may now consider deforming the slice in a neighborhood of a single generator through x_1 . That is, we consider a nearby slice of T which differs from Σ only in a neighborhood near x_1 having infinitesimal volume $\mathcal{V}$. Such a deformation is described by a function $F_\epsilon(x) = F(x) + \epsilon \Delta_{x_1}(x)$ where $\Delta_{x_1} = 1$ in a neighborhood of volume $\mathcal{V}$ around x_1 and is zero elsewhere¹. We define the local deformation of $C_{gen}[F(x)]$ as follows:

$$\left. \frac{dC_{gen}}{d\epsilon} \right|_{x_1} \equiv \lim_{\epsilon \rightarrow 0} \frac{C_{gen}[F_\epsilon(x)] - C_{gen}[F(x)]}{\epsilon} \quad (5.2.5)$$

¹ Note that $F_\epsilon(x)$ remains in our restricted class of functions, since ϵ may be chosen sufficiently small so that $F_\epsilon(x) = 0 = F(x)$ for $x \in \partial\Sigma_0$. Additionally, to first order in ϵ , $|\nabla F_\epsilon(x)|^2 < 1$ so long as the condition holds for $F(x)$.

We then define the timelike quantum expansion as

$$\Xi[F(x); x_1] \equiv \lim_{\mathcal{V} \rightarrow 0} \frac{G\hbar\ell}{\mathcal{V}} \frac{dC_{gen}}{d\epsilon} \Big|_{x_1} \quad (5.2.6)$$

Equivalently, we may write the timelike quantum expansion as a functional derivative:

$$\Xi[F(x); x_1] \equiv \frac{G\hbar\ell}{\sqrt{g(x_1)}} \frac{\delta C_{gen}}{\delta F(x_1)} \quad (5.2.7)$$

where $\sqrt{g}$ is the finite volume element of the metric restricted to Σ , and ensures that the functional derivative is taken per unit of coordinate-independent geometric volume.

The factor of $\hbar G\ell$ is chosen so that in the limit that $\hbar \rightarrow 0$ the geometric term dominates the generalized complexity, and the quantum expansion reduces to the classical timelike expansion:

$$\lim_{\hbar \rightarrow 0} \Xi[F(x); x_1] = \lim_{\hbar \rightarrow 0} \frac{G\hbar\ell}{\sqrt{g(x_1)}} \frac{\delta C_{gen}}{\delta F(x_1)} = \frac{1}{\sqrt{g(x_1)}} \frac{\delta \text{Vol}[F(x)]}{\delta F(x_1)} = \vartheta[F(x); x_1] \quad (5.2.8)$$

5.3 Timelike Focusing and its Consequences

The timelike quantum focusing conjecture (TQFC) is the statement that the functional variation of the timelike quantum expansion is everywhere negative:

$$\frac{\delta}{\delta F(x_2)} \Xi[F(x); x_1] \leq 0 \quad (5.3.1)$$

In what follows, we investigate three aspects of the TQFC inequality.

1. First, we consider the diagonal components of the TQFC, where $x_2 = x_1$, and we demonstrate that it implies a quantum strong energy condition.
2. Second, we consider the off-diagonal case, $x_2 \neq x_1$, and we argue that for a suitable definition of C_{QFT} , the TQFC follows from strong subadditivity of C_{QFT} .
3. Third, we integrate the quantum expansion along non-expanding generators to obtain a complexity/volume bound which is analogous to the covariant entropy bound.

5.3.1 A Quantum Strong Energy Condition

The TQFC is a constraint on the second functional derivative of the generalized complexity:

$$\frac{\delta}{\delta F(x_2)} \frac{G\hbar\ell}{\sqrt{g(x_1)}} \frac{\delta C_{gen}}{\delta F(x_1)} \leq 0 \quad (5.3.2)$$

Consider first the diagonal case, where $x_2 = x_1$. The timelike quantum expansion can be written, using a shorthand notation, as

$$\Xi = \vartheta + \frac{\hbar G \ell}{\mathcal{V}} \dot{C}_{QFT} \quad (5.3.3)$$

where the “ $\dot{\cdot}$ ” indicates a first variation along the (timelike) generator at x_1^2 . The statement of the (diagonal) TQFC is that $\dot{\Xi} \leq 0$, which gives

$$0 \geq \dot{\vartheta} + \frac{G \hbar \ell}{\mathcal{V}} (\ddot{C}_{QFT} - \dot{C}_{QFT} \vartheta) \quad (5.3.4)$$

The classical timelike expansion obeys the Raychaudhuri equation [286]

$$\dot{\vartheta} = -\frac{1}{3} \vartheta^2 - \sigma^2 + \omega^2 - R_{\tau\tau} \quad (5.3.5)$$

where $R_{\tau\tau} = R_{\mu\nu} u^\mu u^\nu$ and u^μ is the (unit) timelike vector field normal to $\tau = F(x)$. σ and ω denote the contracted shear and torsion tensors. Then, using the trace-reversed form of the semiclassical Einstein equation³ and substituting into the TQFC inequality, we obtain

$$0 \geq -\frac{1}{3} \vartheta^2 - \sigma^2 - \omega^2 - 8\pi G \langle T_{\tau\tau} \rangle - 4\pi G \langle T \rangle + \frac{G \hbar \ell}{\mathcal{V}} (\ddot{C}_{QFT} - \dot{C}_{QFT} \vartheta) \quad (5.3.6)$$

Note that we are free to choose a congruence such that at the point x_1 the shear and torsion vanish, as well as the classical expansion: $\sigma^2 = \omega^2 = \vartheta^2 = 0$. With this special choice of congruence, the $\hbar \rightarrow 0$ limit yields the usual strong energy condition:

$$\langle T_{\tau\tau} \rangle + \frac{1}{2} \langle T \rangle \geq 0 \quad (5.3.7)$$

For nonzero $\hbar$, the same special choice of congruence gives an inequality which we refer to as the quantum strong energy condition (qSEC):

$$\langle T_{\tau\tau} \rangle + \frac{1}{2} \langle T \rangle \geq \frac{\hbar \ell}{8\pi \mathcal{V}} \ddot{C}_{QFT} \quad (5.3.8)$$

This inequality is desirable for cases (such as AdS) in which the strong energy condition $T_{\tau\tau} + \frac{1}{2}T \geq 0$ is obeyed classically, but becomes locally violated once negative energy quantum fluctuations are allowed [287, 288], yielding $\langle T_{\tau\tau} \rangle_\Psi + \frac{1}{2} \langle T \rangle_\Psi < 0$. The bound Eq. (5.3.8), if it holds, allows for negative energy fluctuations so long as they don't cause $\langle T_{\tau\tau} \rangle_\Psi + \frac{1}{2} \langle T \rangle_\Psi$ to become more negative than the second variation of the complexity of the state Ψ . We also point out that, like the QNEC, the energy condition does not depend on the gravitational constant G , despite being derived from focusing conditions on geodesic congruences. This suggests that the qSEC is a purely quantum inequality, and can be investigated without worrying about gravitational effects. The inequality does however depend on the length parameter ℓ , which we regard as a free parameter in our definition of quantum complexity. In Sec. 5.4 we discuss the plausibility of this inequality in general, and in cases where the strong energy condition is violated classically, such as de Sitter.

² We use a dot rather than a prime to emphasize that the variation plays the role of a time-derivative

³ With this form of the Einstein equation, we have $R_{\mu\nu} = 8\pi(\langle T_{\mu\nu} \rangle - \frac{1}{2} \langle T \rangle g_{\mu\nu})$. For a unit timelike vector u^μ , we have $\langle T_{\mu\nu} \rangle u^\mu u^\nu - \frac{1}{2} \langle T \rangle g_{\mu\nu} u^\mu u^\nu = \langle T_{\tau\tau} \rangle - \frac{1}{2} \langle T \rangle (-1) = \langle T_{\tau\tau} \rangle + \frac{1}{2} \langle T \rangle$.

5.3.2 Off-diagonal TQFC from SSA of Complexity

In the previous section, the diagonal form of the TQFC was used to arrive at the qSEC. In what follows, we show that the off-diagonal part, i.e. the case where $x_1 \neq x_2$, follows from the strong sub-additivity (SSA) of complexity, once we restrict to a suitable class of C_{QFT} measures. In the context of entanglement entropy, SSA is the following property [289]. Suppose a state ρ is supported on a partial Cauchy slice Σ which can be decomposed into (possibly overlapping) subsystems Σ_A and Σ_B supporting states ρ_A and ρ_B . Then,

$$S(A \cup B) - S(A) - S(B) + S(A \cap B) \leq 0 \quad (5.3.9)$$

Equivalently, for a tripartition of Σ into (non-overlapping subregions) $\Sigma_A \cup \Sigma_B \cup \Sigma_C$,

$$S(ABC) - S(AB) - S(BC) + S(C) \leq 0 \quad (5.3.10)$$

For measures of complexity that are slice-dependent, i.e., they depend only on the operators/fields acting on a slice Σ , it is unclear how the off-diagonal TQFC is related to SSA, if at all, since the functional variations vary the surface Σ itself, rather than varying subregions of Σ . However, we show below that if C_{QFT} is taken to be in a class of codimension-0 measures, the off-diagonal QFC follows from SSA for codimension-0 subregions. To demonstrate this, we examine the form of the second C_{gen} variation.

Consider the case where $x_2 \neq x_1$. Since $\text{Vol}[F(x)]$ is the integral of a local functional of $F(x)$, and the factor $\sqrt{g}$ is evaluated at x_1 , it follows that the off-diagonal second variation only receives a contribution from C_{QFT} . Specifically,

$$\frac{\delta}{\delta F(x_2)} \frac{G\hbar\ell}{\sqrt{g(x_1)}} \frac{\delta C_{gen}}{\delta F(x_1)} = \frac{G\hbar\ell}{\sqrt{g(x_1)}} \frac{\delta^2 C_{QFT}}{\delta F(x_2)\delta F(x_1)} \quad (5.3.11)$$

The second functional derivative can be written as follows:

$$\frac{\delta^2 C_{QFT}}{\delta F(x_2)\delta F(x_1)} = \quad (5.3.12)$$

$$\lim_{\mathcal{V}_1, \mathcal{V}_2 \rightarrow 0} \frac{\sqrt{g(x_1)g(x_2)}}{\mathcal{V}_1 \mathcal{V}_2} \lim_{\epsilon_1, \epsilon_2 \rightarrow 0} \frac{C_{QFT}[F_{\epsilon_1, \epsilon_2}(x)] - C_{QFT}[F_{\epsilon_1}(x)] - C_{QFT}[F_{\epsilon_2}(x)] + C_{QFT}[F(x)]}{\epsilon_1 \epsilon_2} \quad (5.3.13)$$

In the above expression, $\mathcal{V}_{1,2}$ are the volume elements at x_1 and x_2 , and $\epsilon_{1,2}$ are the deformation parameters for the surface along the generators at x_1 and x_2 . The functions $F_\epsilon(x)$ are defined as before, e.g.

$$F_{\epsilon_1, \epsilon_2}(x) = F(x) + \epsilon_1 \Delta_{x_1}(x) + \epsilon_2 \Delta_{x_2}(x) \quad (5.3.14)$$

where the Δ_{x_i} are supported on $\mathcal{V}_i$, and are assumed to be non-overlapping.

In the context of ordinary (null) quantum focusing, an identical expression to Eq. (5.3.14) appears, but with S_{out} replacing C_{QFT} , and with deformations taken along null generators rather than timelike ones. A key insight was that the subregions of the (codimension-1)

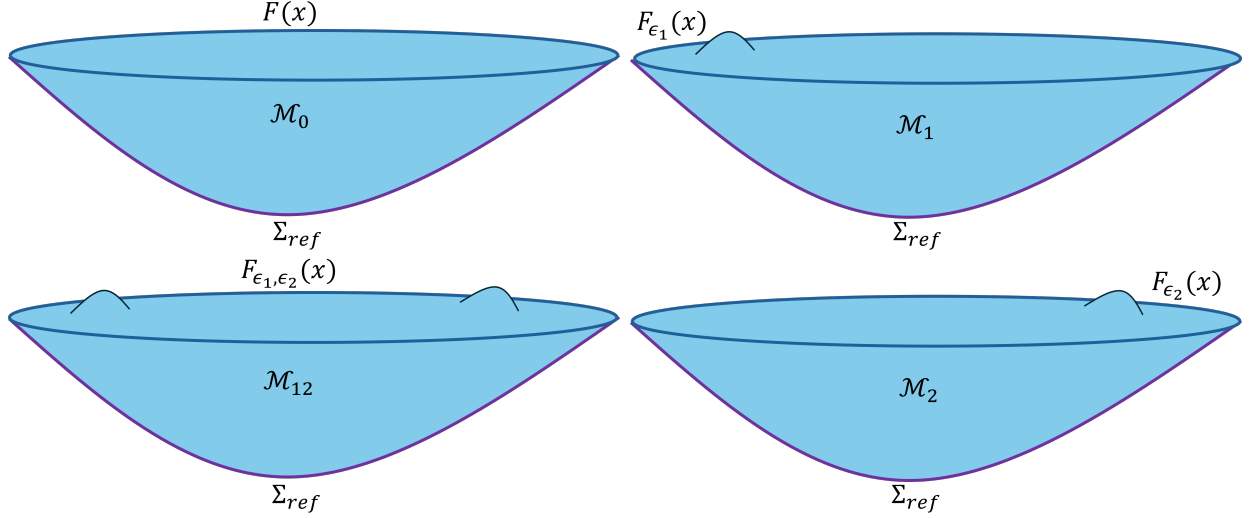


Figure 5.2: Depicted are the codimension-0 regions between the chosen reference slice Σ_{ref} and the local deformations of the slice Σ_0 . The undeformed slice is defined by $\tau = F[x]$ (top left), and the deformations at x_1 and x_2 (right, bottom left) are implemented by $F_{\epsilon_1}(x)$ and $F_{\epsilon_2}(x)$. The terms appearing in the second variation correspond to the complexities associated with these codimension-0 regions.

lightsheet above the slices deformed by the $\epsilon_i \Delta_{x_i}$ s form a tripartition of the lightsheet. This allows one to identify the quantities $S_{out}[F_{\epsilon_i}(y)]$ with the entanglement entropies of different lightsheet subregions. In doing so, the non-positivity of the second variation becomes precisely the statement of strong subadditivity. This suggests that $C_{QFT}[F(x)]$ ought to be defined in such a way that its variations allow us to form a tripartition of a (codimension-0) portion of the timelike congruence.

A natural way to proceed is to identify $C_{QFT}[F(x)]$ not as the complexity of constructing the state ρ_Σ on Σ using operators living on Σ , but as the complexity of preparing ρ_Σ on the slice $\tau = F(x)$ from a reference state ρ_{ref} on some earlier slice $\tau = F_{ref}(x)$. Indeed, complexity is fundamentally a count of operators and any operator which moves a state from one slice to another may be represented as a product of operators with support in the intermediate codimension-0 region. Optimizing over such preparation procedures allows us to associate $C_{QFT}[F(x)]$ with the codimension-0 region between Σ_{ref} and Σ rather than with Σ alone⁴. Moreover, such a notion allows us to associate the terms in Eq. (5.3.14) with subregions of the congruence beneath Σ and its deformations, the union and intersection of which form a tripartition of a codimension-0 region.

Let us explicitly consider the terms in Eq. (5.3.14). Denote the singly deformed slices $\tau = F_{\epsilon_i}(x)$ by Σ_i , the double deformed slice $\tau = F_{\epsilon_1, \epsilon_2}(x)$ by Σ_{12} , and the undeformed slice

⁴ A simple example of such a procedure is the circuit associated with Hamiltonian evolution. For a general reference state ρ_{ref} additional operators may be necessary to prepare ρ_Σ . In general, whether the Hamiltonian circuit is optimal depends on the precise details of the complexity measure, cost functions, etc.

$\tau = F(x)$ by Σ_0 . Then, if we introduce some past reference state Σ_{ref} , we can denote the codimension-0 subregions of the congruence between Σ_{ref} and Σ_i by $M_i \equiv D^-(\Sigma_i) \cap D^+(\Sigma_{ref})$, $\mathcal{M}_{12} \equiv D^-(\Sigma_{12}) \cap D^-(\Sigma_{ref})$, etc. See Fig. 5.2. Explicitly, then, we have

$$C_{QFT}[F(x)] \equiv C(\mathcal{M}_0) = C\left(\text{img}\right), \quad (5.3.15)$$

$$C_{QFT}[F_{\epsilon_1}(x)] \equiv C(\mathcal{M}_1) = C\left(\text{img}\right), \quad (5.3.16)$$

$$C_{QFT}[F_{\epsilon_2}(x)] \equiv C(\mathcal{M}_2) = C\left(\text{img}\right), \quad (5.3.17)$$

$$C_{QFT}[F_{\epsilon_1, \epsilon_2}(x)] \equiv C(\mathcal{M}_{12}) = C\left(\text{img}\right) \quad (5.3.18)$$

Finally, since $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_{12}$ are all obtained from $\mathcal{M}_0$ by future-directed deformations, it follows that $\mathcal{M}_0 = \mathcal{M}_1 \cap \mathcal{M}_2$ and $M_{12} = \mathcal{M}_1 \cup \mathcal{M}_2$, and the terms in Eq. (5.3.14) become

$$C_{QFT}[F_{\epsilon_1, \epsilon_2}(x)] - C_{QFT}[F_{\epsilon_1}(x)] - C_{QFT}[F_{\epsilon_2}(x)] + C_{QFT}[F(x)] \quad (5.3.19)$$

$$= C(\mathcal{M}_1 \cup \mathcal{M}_2) - C(\mathcal{M}_1) - C(\mathcal{M}_2) + C(\mathcal{M}_1 \cap \mathcal{M}_2) \quad (5.3.20)$$

Imposing nonpositivity on the last expression yields precisely the statement of strong subadditivity for the complexity associated with the overlapping regions $\mathcal{M}_1$ and $\mathcal{M}_2$. The above expression can be written in a more familiar form if we tri-partition the subregion $\mathcal{M}_1 \cup \mathcal{M}_2$ as follows:

$$\mathcal{A} \equiv \mathcal{M}_1 \cap (\mathcal{M}_1 \cap \mathcal{M}_2) = \text{img} = \mathcal{B} \equiv \mathcal{M}_2 \cap (\mathcal{M}_1 \cap \mathcal{M}_2) \quad (5.3.21)$$

$$\mathcal{C} \equiv \mathcal{M}_1 \cap \mathcal{M}_2 = \text{img} \quad (5.3.22)$$

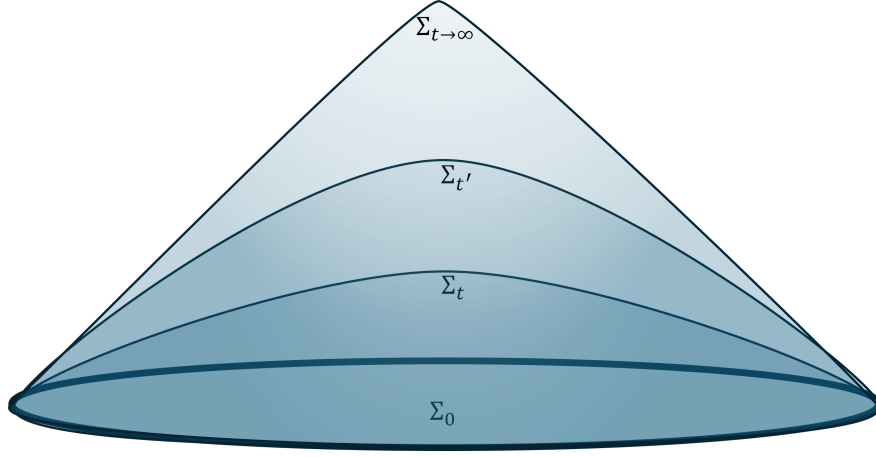
Then, the non-positivity of Eq. (5.3.20) can be stated as

$$C(\mathcal{ABC}) - C(\mathcal{AC}) - C(\mathcal{BC}) + C(\mathcal{C}) \leq 0 \quad (5.3.23)$$

In either case, we find that the TQFC condition Eq. (5.3.14) follows from the SSA inequality

$$C\left(\text{img}\right) - C\left(\text{img}\right) - C\left(\text{img}\right) + C\left(\text{img}\right) \leq 0 \quad (5.3.24)$$

This result reveals that the off-diagonal part of the focusing condition Eq. (5.3.1) will hold in instances where the candidate codimension-0 complexity measure has the property of strong subadditivity. There are many examples of codimension-0 complexity measures throughout the literature, including the complexity=action proposal [173], path-integral optimization [290, 291], and the infinite class of measures constructed in [292]. Different complexity measures obey various additivity inequalities, and it is not expected that every codimension-0 measure obeys SSA [293–298]. We therefore regard the subadditivity condition as a useful guide for discerning which complexity measures are most relevant for focusing considerations.



$$0 = C_{QFT}(\Sigma_0) \leq C_{QFT}(\Sigma_t) \leq C_{QFT}(\Sigma_{t'}) \leq C_{QFT}(\Sigma_\infty) \leq \text{Vol}(\Sigma_0)$$

Figure 5.3: If the timelike quantum expansion is everywhere negative throughout the congruence, then the complexity associated with all future slices Σ_t is bounded above by the volume of the initial slice Σ_0 .

5.3.3 A Covariant Complexity Bound

We now investigate the consequence of integrating the timelike quantum expansion along non-expanding timelike generators, and arrive at a bound which we refer to as the covariant complexity bound (CCB).

Consider an initial slice Σ defined by $\tau = 0$, and some later slice Σ' defined by $\tau = F(x)$ which is related to Σ by deforming only along generators for which the timelike quantum expansion is negative. In particular, we require that $F(x) = 0$ for all points $x \in \Sigma$ where $\Xi[F(x); x] < 0$. Otherwise, $F(x) \geq 0$. Then, since the nonpositivity of Ξ along a generator x implies that the generalized complexity is non-increasing along the generator, we see that the generalized complexity of Σ' must be less than that of Σ . That is, for this choice of $F(x)$, we have

$$C_{gen}[F(x)] \leq C_{gen}[0], \quad (5.3.25)$$

or equivalently,

$$C_{QFT}[\Sigma'] + \frac{\text{Vol}(\Sigma')}{\hbar G \ell} \leq C_{QFT}[\Sigma] + \frac{\text{Vol}(\Sigma)}{\hbar G \ell}, \quad (5.3.26)$$

and moreover

$$C_{QFT}[\Sigma'] - C_{QFT}[\Sigma] \leq \frac{\text{Vol}(\Sigma) - \text{Vol}(\Sigma')}{\hbar G \ell} \quad (5.3.27)$$

The above inequality is most illuminating in the following limit. Suppose that we define C_{QFT} as the complexity associated with Hamiltonian evolution through a codimension-0 region, and consider Σ itself to be the reference slice supporting the initial field data. Then, we have $C_{QFT}(\Sigma) = 0$. If Σ' is chosen to be the future-most time-slice in any foliation of $D(\Sigma)$, then Σ' approaches the future null boundary of $D(\Sigma)$, and its volume vanishes, i.e.

$\text{Vol}(\Sigma') = 0$. In this limit, the inequality becomes

$$C_{QFT}(\Sigma') \leq \frac{\text{Vol}(\Sigma)}{\hbar G \ell} \quad (5.3.28)$$

Physically, this inequality states that if the fields on some compact slice Σ are evolved as far as is causally allowed using the local Hamiltonian, then the complexity of the final state cannot exceed the volume of the slice Σ . If we further make the reasonable assumption that C_{QFT} is monotonically increasing along a foliation of slices Σ_t , with $\Sigma = \Sigma_0$, i.e. we assume that $t < t' \Rightarrow C_{QFT}(\Sigma_t) < C_{QFT}(\Sigma_{t'})$, then the inequality implies that

$$C_{QFT}(\Sigma_t) \leq \frac{\text{Vol}(\Sigma_0)}{\hbar G \ell} \quad \text{for all } t > 0 \quad (5.3.29)$$

That is, the volume of a slice Σ bounds the complexity of all future states obtainable from it by Hamiltonian evolution. We discuss implications of this result in Sec. 5.4.

5.4 Discussion

We have demonstrated that a notion of generalized complexity can be used to define a complexity-based timelike quantum expansion, and that focusing conditions on this quantum expansion imply a complexity-based quantum strong energy condition, as well as a new complexity bound. We have also shown that the focusing condition (for non-overlapping deformations) follows from the strong subadditivity of complexity for codimension-0 complexity measures.

We regard the focusing condition as a tool for narrowing the broad class of proposed complexity measures. That is to say, we propose that the physically relevant complexity measure for semiclassical fields is one which obeys the necessary focusing conditions, e.g. SSA. We leave a detailed investigation of the focusing properties of particular complexity measure for future investigation.

Below, we discuss the plausibility of the main consequences of focusing, as well as potential applications. We provide evidence for the complexity bound Eq. (5.3.28) which doesn't rely on any notion of generalized complexity, and discuss how the complexity bound may be used to diagnose quantum singularities. We then argue for the plausibility of the qSEC, and discuss how it may be preserved in spacetimes with positive cosmological constant where the SEC is violated.

5.4.1 Classical Analogue of the Covariant Complexity Bound

In Sec. 5.3.3, we showed that the complexity of causally evolved fields from a slice Σ is bounded by the volume of Σ , if the TQFC is assumed. We now show that this bound is a covariant version of a complexity bound which can be deduced without any focusing considerations.

Consider a region of space Σ enclosed by a sphere of size l . The amount of energy that can fit in this region without collapsing it into a black hole is bounded by ⁵

$$E \lesssim \frac{l}{G}. \quad (5.4.1)$$

Lloyd's bound [299] places a fundamental upper bound on the rate of complexity growth,

$$\dot{C} \lesssim \frac{E}{\hbar} \quad (5.4.2)$$

where we have suppressed order one constants. Given that the region Σ has size l , the longest (future-directed) timelike geodesic contained in Σ 's domain of dependence has proper time of order l , i.e. $\Delta\tau \lesssim l$. Then, for sufficiently small l (or for sufficiently slowly varying $\dot{C}$), Lloyd's bound can be integrated to obtain

$$\Delta C \lesssim \frac{E}{\hbar} \cdot l \lesssim \frac{l^2}{G\hbar} \quad (5.4.3)$$

Since in $d = 3$ the volume of Σ is $O(l^3)$, the above inequality can be written as

$$\Delta C \lesssim \frac{\text{Vol}(\Sigma)}{\hbar G l} \quad (5.4.4)$$

We regard the bound Eq. (5.3.28) as the covariant generalization of this statement, potentially indicating that the length parameter ℓ in the generalized complexity Eq. (5.2.2) is state-dependent.

5.4.2 Diagnosing Singularities with Complexity Bounds

The covariant complexity bound (5.3.28) may be used to signal the existence of quantum singularities, in a manner analogous to [219, 269]. To outline how this may be done, begin with a surface Σ_0 and consider future timelike deformations of Σ_0 which bring it to another surface Σ in its domain of dependence. Let's assume that the orthogonal congruence from Σ_0 used to perform deformations is future timelike trapped, or that everywhere in the congruence we have

$$\vartheta < 0 \quad (5.4.5)$$

Let us also assume deformations along the congruence only increase the quantum complexity, $C_{\text{QFT}}(\Sigma)$. We will assume that by evolving along the slice we can bound the growth of complexity below by

$$\frac{d}{d\tau} C_{\text{QFT}}(\Sigma) \geq \gamma \quad (5.4.6)$$

⁵ For example, the mass contained in a sphere of radius r must be less than that of a Schwarzschild black hole of the same radius, i.e. $M \leq r/2G$.

for some positive γ . Integrating this inequality gives

$$C_{\text{QFT}}(\Sigma) \geq \gamma\tau + C_{\text{QFT}}(\Sigma_0) \quad (5.4.7)$$

Comparing this to Eq. (5.3.28) we get

$$\frac{\text{Vol}(\Sigma_0)}{\hbar Gl} \geq \gamma\tau + C_{\text{QFT}}(\Sigma_0) \quad (5.4.8)$$

or equivalently

$$\tau \leq \frac{1}{\gamma} \left(\frac{\text{Vol}(\Sigma_0)}{\hbar Gl} - C_{\text{QFT}}(\Sigma_0) \right) \quad (5.4.9)$$

This inequality cannot hold beyond a finite proper time. If the future development of Σ_0 were regular and timelike geodesically complete, one could continue the orthogonal flow to arbitrarily large τ , contradicting the above inequality. Physically, this says that finite regions cannot supply an infinite amount of complexity growth. The singularity appears when the resources of a finite computational budget have run out.

5.4.3 The qSEC and the cosmological constant

In spacetimes with a nonzero cosmological constant Λ , the strong energy condition is the statement that for any unit timelike vector field u^μ

$$\langle T_{\mu\nu} \rangle u^\mu u^\nu + \frac{1}{2} \langle T \rangle - \frac{\Lambda}{8\pi G} \geq 0 \quad (5.4.10)$$

For $\Lambda < 0$, the SEC holds in the vacuum (i.e. AdS), but can be violated if quantum fluctuations cause the energy expectation value to drop too low, i.e. if they cause $\langle T_{\mu\nu} \rangle u^\mu u^\nu + \frac{1}{2} \langle T \rangle \leq \frac{\Lambda}{8\pi G} < 0$. However, the qSEC Eq. (5.3.8) becomes (upon including the cosmological constant),

$$\langle T_{\tau\tau} \rangle + \frac{1}{2} \langle T \rangle - \frac{\Lambda}{8\pi G} \geq \frac{\hbar \ell}{8\pi \mathcal{V}} \ddot{C}_{\text{QFT}} \quad (5.4.11)$$

This inequality may still be preserved, provided that $\ddot{C}_{\text{QFT}}$ is sufficiently negative for states with SEC-violating energy fluctuations. That is, if $\ddot{C}_{\text{QFT}} < 0$, the qSEC allows for energy fluctuations to be more negative than the SEC allows for. Moreover, since C_{QFT} is explicitly state-dependent, it is plausible that $\ddot{C}_{\text{QFT}}$ provides a weaker (more negative) lower bound for states with greater negative energy fluctuations.

For spacetimes with $\Lambda > 0$, such as vacuum de Sitter, the SEC is violated even at the classical level, since $T_{\mu\nu} = 0$ in the vacuum. This reduces the SEC inequality to $-\frac{\Lambda}{8\pi G} \geq 0$, which is explicitly violated in cases where $\Lambda > 0$. In order for the qSEC to hold while the SEC is violated, it must be the case that $\ddot{C}_{\text{QFT}} \lesssim -\Lambda$ even in the classical limit. We speculate that this may be possible due to the cutoff dependence of C_{QFT} , and the natural IR cutoff inherent to observers in $\Lambda > 0$ spacetimes.

Below, we sketch a framework for quantifying semiclassical complexity using a coarse-grained modular flow. Within this framework, $\tilde{C}_{QFT}$ satisfies the negativity requirements outlined above, and scales with Λ in a manner consistent with the requirements of the qSEC when $\Lambda > 0$.

Coarse-grained modular channels

Consider a reference state on a slice through a domain of dependence D in a generic space-time. An observer in a static patch will have a limit set on what observables they are actually able to measure. We thus specify an algebra A_λ of observables in the diamond below a cutoff λ which keeps only observables with modular frequencies $|\omega| \leq \lambda$. We also introduce a tolerance $\varepsilon > 0$ within which nearby states are identified in an appropriate norm. The resulting Hilbert space is finite-dimensional, and we suppose that there is a finite set of gate operators $G \subset A_\lambda$ which can be used to construct arbitrary semiclassical states with accuracy ε .

This allows us to define a coarse grained modular channel as follows. Let $U_D = \exp\{-iKs\}$ denote the modular operator, where K is the modular hamiltonian and s is the modular parameter. An operator $X \in A_\lambda$ evolves under modular evolution via $X_s = U_D(-s)XU_D(s)$. We define its evolution in the coarse-grained modular channel as

$$X_s^{(\lambda)} = P_{A_\lambda}(U_D(-s)XU_D(s)), \quad X \in A_\lambda \quad (5.4.12)$$

where P_{A_λ} is an operator that projects only on observables below the cutoff λ . This is necessary since exact modular evolution may evolve X_s outside of A_λ . We can then define an effective channel complexity as the minimal number of gates needed to approximately construct $X^{(\lambda)}$ from X within the ε tolerance. Formally, if $W = g_1 \cdots g_N$ is a channel constructed of gates $g_i \in G$ such that $|X_s^{(\lambda)} - W^\dagger X W| \leq \varepsilon$, then the effective channel complexity is the minimal N over all such channels W :

$$C_c(s, \varepsilon, \lambda) = \min_W N(W) \quad (5.4.13)$$

Once the effective Hilbert space is finite dimensional and one works at fixed nonzero tolerance ε , the set of distinguishable channels is compact up to an ε -net, so the approximate complexity is bounded above by some finite C_{max} depending on the regulator, gate set, and tolerance [300, 301]. This means that

$$0 \leq C_c(s) \leq C_{max} < \infty \quad (5.4.14)$$

Following other examples in the literature [302, 303], we introduce an “uncomplexity reservoir” as the difference between the complexity and the maximum complexity

$$u(s) \equiv C_{max} - C_c(s) \quad (5.4.15)$$

which is always positive. Since we have coarse grained, we do not track microscopic recurrences and only ask if the next increment of modular flow is taking the retained observables into a new channel class or into one already explored. When $u(s)$ is large, nearly every step

explores a new channel. Late in the evolution, the growth slows down and the system saturates. We can model the evolution of the coarse-grained complexity with a simple saturation ansatz

$$\frac{d}{ds}C_c(s) = \Gamma u(s) = \Gamma(C_{max} - C_c(s)) \quad (5.4.16)$$

where Γ is some coefficient that describes the effective mixing rate. Generically, Γ will may depend on the length scale of the chosen geometry, as well as the cutoff scale λ . Solving this differential equation gives

$$\dot{C}_c(s) = \Gamma(C_{max} - C_c(0))e^{-\Gamma s} \quad (5.4.17)$$

and therefore

$$\ddot{C}_c(s) = -\Gamma^2(C_{max} - C_c(0))e^{-\Gamma s} < 0 \quad (5.4.18)$$

The classical limit in AdS and dS

If one assumes that the timelike congruence can be chosen so that $\frac{d^2}{d\tau^2}$ and $\frac{d^2}{ds^2}$ only differ by a positive factor, the above inequality gives the negative concavity required by the qSEC. In the case of AdS, if the mixing rate Γ goes like λ as we would expect, i.e. $\Gamma \sim \lambda$, then in the limit $\lambda \rightarrow 0$ where all QFT modes are coarse-grained over, we have $\Gamma \rightarrow 0$ and hence $\ddot{C} \rightarrow 0$, thus reducing the qSEC to the SEC.

In the case of the dS static patch vacuum, it remains true that $\ddot{C} \leq 0$ under the earlier assumptions. However, flowing the observer's cutoff scale into the IR corresponds to $\lambda \rightarrow \ell_{dS}^{-1}$, rather than $\lambda \rightarrow 0$, since the static patch radius is the largest possible length scale in de Sitter. In this limit, we thus have that $\Gamma \sim 1/\ell_{dS}$. To estimate the saturation complexity, we take the effective number of fundamental qubits N_{eff} needed to describe the static patch to scale with the dS entropy.

$$N_{\text{eff}} \sim S_{dS} = \frac{A_{dS}}{\hbar G} = 4\pi \frac{\ell_{dS}^2}{\hbar G} \quad (5.4.19)$$

The static patch volume is

$$V_{\text{patch}} \sim \frac{4}{3}\pi \ell_{dS}^3, \quad (5.4.20)$$

so the corresponding maximum coarse-grained complexity density in the IR scales like

$$c_{max} = \frac{C_{max}}{V_{\text{patch}}} \sim \frac{N_{\text{eff}}}{V_{\text{patch}}} \sim \frac{3}{\hbar G \ell_{dS}} \quad (5.4.21)$$

Plugging this into our saturation model from before gives

$$\ddot{c} \sim -\Gamma^2 c_{max} = -\frac{3}{\hbar G \ell_{dS}^3} \quad (5.4.22)$$

Finally,

$$\frac{\hbar \ell_{dS}}{8\pi V} \ddot{C} \sim -\frac{3\hbar \ell_{dS}}{8\pi \hbar G \ell_{dS}^3} = -\frac{3}{8\pi G \ell_{dS}^2} = -\frac{3\Lambda}{8\pi G} = -\rho_\Lambda \quad (5.4.23)$$

where ρ_Λ is precisely the vacuum energy density in dS. Comparing this with the discussion below Eq. (5.4.11) reveals that the semiclassical complexity is at precisely the scale needed

for the qSEC to hold in vacuum de Sitter. This suggests that $\ddot{C}$ may remain negative in the classical limit and may exist as a classical observable. We leave a detailed analysis of the relationship between effective QFT complexity and the cosmological constant to future work.

These conclusion are possible because $C_{gen}(\Sigma) = \frac{\text{Vol}(\Sigma)}{\hbar G \ell} + C_{QFT}(\Sigma)$ is intrinsically a semiclassical object. In particular, for the qSEC to make sense in de Sitter-like spacetimes, the renormalized theory must contain finite infrared complexity dynamics. Since the exact vacuum state in a causal diamond is stationary under exact modular flow, this finite contribution cannot naturally be interpreted as the second derivative of exact state complexity. Exact modular flow leaves the reduced vacuum state invariant, whereas coarse-grained modular flow can display relaxation and saturation, making a negative second derivative of an effective complexity observable possible. Therefore, it should be associated with this renormalized and coarse-grained notion of modular complexity, or an equivalent effective complexity observable attached to the observer algebra of the static patch. If such an observable is nonzero, then the static patch exhibits infrared complexity dynamics even in the vacuum. This suggests a maximum finite coarse-grained complexity capacity for the patch, which is naturally compatible with the presence of a horizon, a temperature, and a finite de Sitter entropy. A similar interpretation should hold in black-hole spacetimes, where the relevant observer algebra is again associated with a thermal horizon subsystem rather than a globally pure vacuum state.

Appendix A

A.1 The Hartle-Hawking Wavefunction

A.1.1 JT Gravity

The action of interest is the Euclidean JT gravity action:¹

$$I_E = -\frac{1}{2} \int_M d^2x \sqrt{-g} \left(\Phi_0 R + \Phi(R+2) \right) - \int_{\partial M} dt \sqrt{|\gamma|} \left(\Phi_0 K_{\partial M} + \Phi(K_{\partial M} - 1) \right) \quad (\text{A.1.1})$$

$$- \int_{\Sigma} dt \sqrt{|\gamma|} \left(\Phi_0 K_{\Sigma} + \Phi(K_{\Sigma} - 1) \right) \quad (\text{A.1.2})$$

where $K_{\partial M}$ and K_{Σ} are the extrinsic curvatures on the AdS and bulk boundaries. We take $K_{\partial M} = \gamma^{\mu\nu} \nabla_{\mu} r_{\nu}$ where r^{μ} is the normal vector in ∂M , $\gamma^{\mu\nu}$ is the induced metric and Φ is the dilaton field. The asymptotic boundary conditions on ∂M are chosen to be

$$\gamma_{t_E t_E} \Big|_{\partial M} = \frac{1}{r_c^2}$$

$$\Phi \Big|_{\partial M} = r_c$$

with $r_c \rightarrow \infty$. The boundary conditions on the timeslice Σ are taken to be

$$n_{\mu} \partial^{\mu} \Phi_{\Sigma} = 0 \quad (\text{A.1.3})$$

$$\Phi_{\Sigma} = \Phi_h \sqrt{1+x^2} \quad (\text{A.1.4})$$

Thus, Σ is the time-reflection symmetric slice, and the value of the dilaton at the bifurcation point is Φ_h . These conditions amount to projecting $|\psi_{\beta/2}\rangle$ onto the fixed-dilaton state $|\Phi_h\rangle$.

The saddle point of the Euclidean action is simply the section of the Euclidean Schwarzschild solution with boundary length $r_c \beta/2$, i.e. with $t_E \in (0, \beta/2)$. See Fig. 4.8. Therefore, the saddle-point geometry covers only a fraction of the entire black hole geometry and has an opening angle determined by the ratio $\frac{\beta/2}{\beta_h}$. In particular, the opening angle is $\theta = \pi \frac{\beta}{\beta_h} = \frac{\beta}{2} \Phi_h$.

¹ Note we remove a relative factor of 2 compared to Ref. [263] to match the normalization used for JT gravity in other literature.

In Schwarzschild coordinates, the horizon is located at $r_h = \Phi_h$. Since we do *not* impose smoothness at the horizon, $\Phi_h \neq 2\pi/\beta$. Rather, there is a conical defect at the horizon of angle $\theta = \beta\Phi_h/2$. The contribution to the action from this corner² is

$$I_{\text{corner}} = -(\Phi_0 + \Phi_h)(\pi - \theta)$$

The contribution away from the corners is simply a fraction of the action for the full (smooth) geometry:

$$I[g_h] = \frac{\beta/2}{2\pi/\Phi_h} I_{\text{tot}} \quad (\text{A.1.5})$$

Evaluating the Euclidean action including the defect contribution gives [263]

$$I_E[g_h] = -\pi(\Phi_0 + \Phi_h) + \frac{\beta}{4}\Phi_h^2 \quad (\text{A.1.6})$$

and therefore

$$\psi(\Phi_h) \approx e^{-I_E[g_h]} = \exp\left\{\pi(\Phi_0 + \Phi_h) - \frac{\beta}{4}\Phi_h^2\right\} \quad (\text{A.1.7})$$

A.1.2 Einstein Gravity

Here, we primarily follow the construction of Ref. [264]. With a $d+1$ dimensional bulk $\mathcal{M}$, we impose the asymptotic AdS boundary conditions

$$\gamma_{ab}dy^a dy^b = r_c^2(d\tau^2 + d\Omega_{d-1}^2) \quad (\text{A.1.8})$$

where $r_c \rightarrow \infty$. The geometry of the bulk slice Σ is taken to be that of an eternal black hole with fixed horizon area $\mathcal{A}_h$, or equivalently, a fixed energy $E(\mathcal{A}_h)$ ³. Specifically,

$$h_{ij}dx^i dx^j = d\rho^2 + r_h^2 \cosh^2 \rho d\Omega_{d-2}^2 \quad (\text{A.1.9})$$

where ρ is the radial coordinate. The bulk slice Σ meets the AdS boundary at $\rho_c = \pm \log \frac{r_c}{r_h}$, and r_h is the horizon radius, related to E through

$$E(\mathcal{A}_h) = \frac{(d-1)\Omega_{d-1}}{16\pi G_N} r_h^{d-2} (1 + r_h^2) \quad (\text{A.1.10})$$

with r_h defined through $\mathcal{A}_h = \Omega_{d-1} r_h^{d-1}$

The Euclidean action of interest is

² There are also corners of angle $\pi/2$ where the AdS boundary meets the bulk slice Σ , but these have a vanishing contribution to the action.

³ In this way, we may identify states $|E\rangle$ of energy with states of fixed horizon area $|\mathcal{A}_h\rangle$, analogous to the fixed dilaton states used in Sec. 4.5.

$$I_E = -\frac{1}{16\pi G} \int_{\mathcal{M}} d^{d+1}x \sqrt{g} (R - 2\Lambda) - \frac{1}{8\pi G} \int_{\partial\mathcal{M}} d^d y \sqrt{\gamma} K_{\partial\mathcal{M}} \quad (\text{A.1.11})$$

$$- \frac{1}{8\pi G} \int_{\Sigma} d^d y \sqrt{\gamma} K_{\Sigma} \quad (\text{A.1.12})$$

where $K_{\partial\mathcal{M}}$ and K_{Σ} are the extrinsic curvatures on the asymptotic and bulk boundaries. The unnormalized Hartle-Hawking wavefunction $\psi_{\beta/2}$ is again given simply by the path integral in Eq. (4.3.2)

$$\psi_{\beta/2}(E) \approx e^{-I_E[g_h]}, \quad (\text{A.1.13})$$

where g_h is the saddle point of the action with fixed geometry h on Σ , corresponding to horizon radius r_h . The saddle-point geometry, as in JT gravity, is given by a slice of the eternal black hole geometry with an asymptotic boundary of length $\beta/2$, with a conical defect at the horizon (as in Fig. 4.8). At fixed energy E , the black hole geometry has an inverse temperature

$$\beta_h = 4\pi \frac{r_h}{r_h^2 d + d - 2} \quad (\text{A.1.14})$$

Therefore the saddle geometry covers only a fraction of the entire black hole geometry, and has opening angle determined by the ratio $\frac{\beta/2}{\beta_h}$. In particular, the opening angle is $\theta = \pi \frac{\beta}{\beta_h}$.

The saddle-point action has two contributions, coming from the fraction of the full black hole geometry, and from the corners at the horizon, and where Σ meets $\partial\mathcal{M}$. The full black hole geometry has action

$$I_{tot}[g_h] = \beta_h E(\mathcal{A}_h) - S(\mathcal{A}_h) \quad (\text{A.1.15})$$

So, including the contribution from the corners, the action of the saddlepoint geometry evaluates to

$$I[g_h] = \frac{\beta/2}{\beta_h} I_{tot}[g_h] + I_{\text{corners}} = -\frac{1}{2} S(\mathcal{A}_h) + \frac{\beta}{2} E(\mathcal{A}_h) \quad (\text{A.1.16})$$

. This gives

$$\psi_{\beta/2}(\mathcal{A}_h) = \exp\left\{\frac{1}{2} S(\mathcal{A}_h) - \frac{\beta}{2} E(\mathcal{A}_h)\right\} \quad (\text{A.1.17})$$

A.2 Bulk Subregions Spanning the Horizon

We briefly state the results for the Rényi entropies of bulk regions which span the horizon, using our four different gauge-invariant specifications. Specifically, we consider the case where the endpoints A and B lie on opposite sides of the horizon, so that $e(a) = a$, and hence $\gamma_a = A \cup B$ with area

$$\mathcal{A}(\gamma_a, h) = \mathcal{A}(A, h) + \mathcal{A}(B, h) \quad (\text{A.2.1})$$

where $\mathcal{A}(A, h)$ and $\mathcal{A}(B, h)$ are determined by our gauge-invariant subregion specification.

Fixed Dilaton Value

In this case, we have $\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi(\Phi_A + \Phi_B)$ which gives

$$\Phi_{h_n} = \frac{2\pi}{\beta} \quad (\text{A.2.2})$$

$$S_n(a) = 2\pi\Phi_0 + 2\pi(\Phi_A + \Phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.3})$$

$$S(a) = 2\pi\Phi_0 + 2\pi(\Phi_A + \Phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.4})$$

Fixed Renormalized Distance

In this case, we have

$$\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi\Phi_h (\cosh(\ell_A + \log \Phi_h) + \cosh(\ell_B + \log \Phi_h)) \quad (\text{A.2.5})$$

which gives

$$\Phi_{h_n} = \frac{2\pi/\beta}{\frac{2\pi}{n\beta}(e^{\ell_A} + e^{\ell_B})(n-1) + 1} = \frac{2\pi}{\beta} - \frac{4\pi^2}{\beta^2}(e^{\ell_A} + e^{\ell_B})(n-1) + O((n-1)^2) \quad (\text{A.2.6})$$

$$S_n(a) = 2\pi\Phi_0 + 2\pi\frac{2\pi}{\beta} \cosh(\ell_A + \log \frac{2\pi}{\beta}) + 2\pi\frac{2\pi}{\beta} \cosh(\ell_B + \log \frac{2\pi}{\beta}) + O(n-1) \quad (\text{A.2.7})$$

$$S(a) = 2\pi\Phi_0 + 2\pi(\Phi_A + \Phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.8})$$

Fixed Relative Dilaton Value

In this case, $\mathcal{A}(\gamma_a, h) = 2\pi\Phi_0 + 2\pi\Phi_h(\phi_A + \phi_B)$ which gives

$$\Phi_{h_n} = \frac{2\pi}{\beta} - \frac{(n-1)\pi}{n\beta}(\phi_A + \phi_B) \quad (\text{A.2.9})$$

$$S_n(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta}(\phi_A + \phi_B) - \left(\frac{n-1}{n}\right)\frac{2\pi^2}{\beta}(\phi_A + \phi_B)^2 \quad (\text{A.2.10})$$

$$S(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta}(\phi_A + \phi_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.11})$$

Fixed Distance from Horizon

With $\frac{\mathcal{A}(\gamma_a, h)}{4G} = 2\pi\Phi_0 + 2\pi\Phi_h(\cosh L_A + \cosh L_B)$, we have

$$\Phi_{h_n} = \frac{2\pi}{\beta} - \frac{(n-1)}{n} \frac{2\pi}{\beta} (\cosh L_A + \cosh L_B) \quad (\text{A.2.12})$$

$$S_n(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (\cosh L_A + \cosh L_B) - \left(\frac{n-1}{n}\right) \frac{2\pi^2}{\beta} (\cosh L_A + \cosh L_B)^2 \quad (\text{A.2.13})$$

$$S(a) = 2\pi\Phi_0 + \frac{4\pi^2}{\beta} (\cosh L_A + \cosh L_B) = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.14})$$

Fixed Horizon Area in Einstein Gravity

In this case, we have $\mathcal{A}(\gamma_a, h) = \mathcal{A}_A + \mathcal{A}_B$. This gives a saddlepoint geometry with horizon area

$$\mathcal{A}_{h_n} = \Omega_{d-1} \left(\frac{2\pi}{d\beta} + \sqrt{\frac{4\pi^2}{d^2\beta^2} - \left(\frac{d-2}{d}\right)} \right)^{d-1} \quad (\text{A.2.15})$$

Note that, as in the fixed-dilaton specification, the saddlepoint has trivial n dependence, and agrees with the expected semiclassical horizon area $\mathcal{A}_{h_\psi}$. This arises from the fact that γ_a is independent of h . Similarly, the spectrum is flat. The Rényi and Von Neumann entropies are

$$S_n(a) = \frac{\mathcal{A}_A + \mathcal{A}_B}{4G} = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.16})$$

$$S(a) = \frac{\mathcal{A}_A + \mathcal{A}_B}{4G} = \frac{\mathcal{A}(\gamma_a, h_\psi)}{4G} \quad (\text{A.2.17})$$

Bibliography

- [1] Jacob D. Bekenstein. “Black holes and entropy”. In: *Phys. Rev. D* 7 (1973), pp. 2333–2346. DOI: [10.1103/PhysRevD.7.2333](https://doi.org/10.1103/PhysRevD.7.2333).
- [2] James M. Bardeen, Brandon Carter, and Stephen W. Hawking. “The four laws of black hole mechanics”. In: *Commun. Math. Phys.* 31 (1973), p. 161. DOI: [10.1007/BF01645742](https://doi.org/10.1007/BF01645742).
- [3] Jacob D. Bekenstein. “Generalized second law of thermodynamics in black hole physics”. In: *Phys. Rev. D* 9 (1974), p. 3292. DOI: [10.1103/PhysRevD.9.3292](https://doi.org/10.1103/PhysRevD.9.3292).
- [4] Jacob D. Bekenstein. “A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems”. In: *Phys. Rev. D* 23 (1981), p. 287. DOI: [10.1103/PhysRevD.23.287](https://doi.org/10.1103/PhysRevD.23.287).
- [5] Ted Jacobson. “Thermodynamics of space-time: The Einstein equation of state”. In: *Phys. Rev. Lett.* 75 (1995), pp. 1260–1263. DOI: [10.1103/PhysRevLett.75.1260](https://doi.org/10.1103/PhysRevLett.75.1260). arXiv: [gr-qc/9504004](https://arxiv.org/abs/gr-qc/9504004).
- [6] Raphael Bousso. “A covariant entropy conjecture”. In: *JHEP* 07 (1999), p. 004. DOI: [10.1088/1126-6708/1999/07/004](https://doi.org/10.1088/1126-6708/1999/07/004). arXiv: [hep-th/9905177](https://arxiv.org/abs/hep-th/9905177).
- [7] Gerard 't Hooft. “Dimensional reduction in quantum gravity”. In: *Conf. Proc. C* 930308 (1993), pp. 284–296. arXiv: [gr-qc/9310026](https://arxiv.org/abs/gr-qc/9310026).
- [8] Leonard Susskind. “The world as a hologram”. In: *J. Math. Phys.* 36 (1995), pp. 6377–6396. DOI: [10.1063/1.531249](https://doi.org/10.1063/1.531249). arXiv: [hep-th/9409089](https://arxiv.org/abs/hep-th/9409089).
- [9] Raphael Bousso. “The holographic principle”. In: *Rev. Mod. Phys.* 74 (2002), pp. 825–874. DOI: [10.1103/RevModPhys.74.825](https://doi.org/10.1103/RevModPhys.74.825). arXiv: [hep-th/0203101](https://arxiv.org/abs/hep-th/0203101).
- [10] Juan Martin Maldacena. “The Large N limit of superconformal field theories and supergravity”. In: *Adv. Theor. Math. Phys.* 2 (1998), pp. 231–252. DOI: [10.4310/ATMP.1998.v2.n2.a1](https://doi.org/10.4310/ATMP.1998.v2.n2.a1). arXiv: [hep-th/9711200](https://arxiv.org/abs/hep-th/9711200).

- [11] Ahmed Almheiri, Xi Dong, and Daniel Harlow. “Bulk locality and quantum error correction in AdS/CFT”. In: *JHEP* 04 (2015), p. 163. DOI: [10.1007/JHEP04\(2015\)163](https://doi.org/10.1007/JHEP04(2015)163). arXiv: [1411.7041](https://arxiv.org/abs/1411.7041) [hep-th].
- [12] Xi Dong, Daniel Harlow, and Aron C. Wall. “Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality”. In: *Phys. Rev. Lett.* 117 (2016), p. 021601. DOI: [10.1103/PhysRevLett.117.021601](https://doi.org/10.1103/PhysRevLett.117.021601). arXiv: [1601.05416](https://arxiv.org/abs/1601.05416) [hep-th].
- [13] Shinsei Ryu and Tadashi Takayanagi. “Holographic derivation of entanglement entropy from AdS/CFT”. In: *Phys. Rev. Lett.* 96 (2006), p. 181602. DOI: [10.1103/PhysRevLett.96.181602](https://doi.org/10.1103/PhysRevLett.96.181602). arXiv: [hep-th/0603001](https://arxiv.org/abs/hep-th/0603001).
- [14] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. “Quantum corrections to holographic entanglement entropy”. In: *JHEP* 11 (2013), p. 074. DOI: [10.1007/JHEP11\(2013\)074](https://doi.org/10.1007/JHEP11(2013)074). arXiv: [1307.2892](https://arxiv.org/abs/1307.2892) [hep-th].
- [15] Netta Engelhardt and Aron C. Wall. “Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime”. In: *JHEP* 01 (2015), p. 073. DOI: [10.1007/JHEP01\(2015\)073](https://doi.org/10.1007/JHEP01(2015)073). arXiv: [1408.3203](https://arxiv.org/abs/1408.3203) [hep-th].
- [16] Chris Akers and Geoff Penington. “Leading order corrections to the quantum extremal surface prescription”. In: *JHEP* 04 (2021), p. 062. DOI: [10.1007/JHEP04\(2021\)062](https://doi.org/10.1007/JHEP04(2021)062). arXiv: [2008.03319](https://arxiv.org/abs/2008.03319) [hep-th].
- [17] Stephen W. Hawking. “Breakdown of predictability in gravitational collapse”. In: *Phys. Rev. D* 14 (1976), pp. 2460–2473. DOI: [10.1103/PhysRevD.14.2460](https://doi.org/10.1103/PhysRevD.14.2460).
- [18] Geoff Penington. “Entanglement wedge reconstruction and the information paradox”. In: *JHEP* 09 (2020), p. 002. DOI: [10.1007/JHEP09\(2020\)002](https://doi.org/10.1007/JHEP09(2020)002). arXiv: [1905.08255](https://arxiv.org/abs/1905.08255) [hep-th].
- [19] Geoff Penington et al. “Replica wormholes and the black hole interior”. In: *JHEP* 03 (2022), p. 205. DOI: [10.1007/JHEP03\(2022\)205](https://doi.org/10.1007/JHEP03(2022)205). arXiv: [1911.11977](https://arxiv.org/abs/1911.11977) [hep-th].
- [20] Ahmed Almheiri et al. “Replica wormholes and the entropy of Hawking radiation”. In: *JHEP* 05 (2020), p. 013. DOI: [10.1007/JHEP05\(2020\)013](https://doi.org/10.1007/JHEP05(2020)013). arXiv: [1911.12333](https://arxiv.org/abs/1911.12333) [hep-th].
- [21] Daniel Harlow and Patrick Hayden. “Quantum computation vs. firewalls”. In: *JHEP* 06 (2013), p. 085. DOI: [10.1007/JHEP06\(2013\)085](https://doi.org/10.1007/JHEP06(2013)085). arXiv: [1301.4504](https://arxiv.org/abs/1301.4504) [hep-th].
- [22] Leonard Susskind. “Entanglement is not enough”. In: *Fortsch. Phys.* 64 (2016), pp. 49–71. DOI: [10.1002/prop.201500095](https://doi.org/10.1002/prop.201500095). arXiv: [1411.0690](https://arxiv.org/abs/1411.0690) [hep-th].

- [23] Adam R. Brown et al. “The python’s lunch: geometric obstructions to decoding Hawking radiation”. In: *JHEP* 08 (2020), p. 121. DOI: [10.1007/JHEP08\(2020\)121](https://doi.org/10.1007/JHEP08(2020)121). arXiv: [1912.00228](https://arxiv.org/abs/1912.00228) [hep-th].
- [24] Netta Engelhardt et al. “Cryptographic Censorship”. In: *JHEP* 01 (2025), p. 122. DOI: [10.1007/JHEP01\(2025\)122](https://doi.org/10.1007/JHEP01(2025)122). arXiv: [2402.03425](https://arxiv.org/abs/2402.03425) [hep-th].
- [25] Leonard Susskind. “Computational complexity and black hole horizons”. In: *Fortsch. Phys.* 64 (2016), p. 24. DOI: [10.1002/prop.201500092](https://doi.org/10.1002/prop.201500092). arXiv: [1402.5674](https://arxiv.org/abs/1402.5674) [hep-th].
- [26] Leonard Susskind. *Three Lectures on Complexity and Black Holes*. 2018. arXiv: [1810.11563](https://arxiv.org/abs/1810.11563) [hep-th].
- [27] Mohsen Alishahiha. “Holographic Complexity”. In: *Phys. Rev. D* 92.12 (2015), p. 126009. DOI: [10.1103/PhysRevD.92.126009](https://doi.org/10.1103/PhysRevD.92.126009). arXiv: [1509.06614](https://arxiv.org/abs/1509.06614) [hep-th].
- [28] Yale Fan et al. “Sharp Transitions for Subsystem Complexity”. In: (Oct. 2025). arXiv: [2510.18832](https://arxiv.org/abs/2510.18832) [hep-th].
- [29] Samir D. Mathur. “The information paradox: A pedagogical introduction”. In: *Class. Quant. Grav.* 26 (2009), p. 224001. DOI: [10.1088/0264-9381/26/22/224001](https://doi.org/10.1088/0264-9381/26/22/224001). arXiv: [0909.1038](https://arxiv.org/abs/0909.1038) [hep-th].
- [30] Ahmed Almheiri et al. “Black holes: complementarity or firewalls?” In: *JHEP* 02 (2013), p. 062. DOI: [10.1007/JHEP02\(2013\)062](https://doi.org/10.1007/JHEP02(2013)062). arXiv: [1207.3123](https://arxiv.org/abs/1207.3123) [hep-th].
- [31] Stephen W. Hawking. “Particle creation by black holes”. In: *Commun. Math. Phys.* 43 (1975), pp. 199–220. DOI: [10.1007/BF02345020](https://doi.org/10.1007/BF02345020). Erratum: *Commun. Math. Phys.* 46, 206 (1976).
- [32] Tom Banks et al. “M theory as a matrix model: A conjecture”. In: *Phys. Rev. D* 55 (1997), pp. 5112–5128. DOI: [10.1103/PhysRevD.55.5112](https://doi.org/10.1103/PhysRevD.55.5112). arXiv: [hep-th/9610043](https://arxiv.org/abs/hep-th/9610043).
- [33] Alex Hamilton et al. “Local bulk operators in AdS/CFT: a Boundary view of horizons and locality”. In: *Phys. Rev. D* 73 (2006), p. 086003. DOI: [10.1103/PhysRevD.73.086003](https://doi.org/10.1103/PhysRevD.73.086003). arXiv: [hep-th/0506118](https://arxiv.org/abs/hep-th/0506118) [hep-th].
- [34] Alex Hamilton et al. “Holographic representation of local bulk operators”. In: *Phys. Rev. D* 74 (2006), p. 066009. DOI: [10.1103/PhysRevD.74.066009](https://doi.org/10.1103/PhysRevD.74.066009). arXiv: [hep-th/0606141](https://arxiv.org/abs/hep-th/0606141) [hep-th].
- [35] Yasunori Nomura, Pratik Rath, and Nico Salzetta. “Pulling the boundary into the bulk”. In: *Phys. Rev. D* 98 (2018), p. 026010. DOI: [10.1103/PhysRevD.98.026010](https://doi.org/10.1103/PhysRevD.98.026010). arXiv: [1805.00523](https://arxiv.org/abs/1805.00523) [hep-th].

- [36] Chitraang Murdia, Yasunori Nomura, and Pratik Rath. “Coarse-graining holographic states: a semiclassical flow in general spacetimes”. In: *Phys. Rev. D* 102 (2020), p. 086001. DOI: [10.1103/PhysRevD.102.086001](https://doi.org/10.1103/PhysRevD.102.086001). arXiv: [2008.01755](https://arxiv.org/abs/2008.01755) [hep-th].
- [37] Netta Engelhardt, Geoff Penington, and Arvin Shahbazi-Moghaddam. “A world without pythons would be so simple”. In: *Class. Quant. Grav.* 38 (2021), p. 234001. DOI: [10.1088/1361-6382/ac2de5](https://doi.org/10.1088/1361-6382/ac2de5). arXiv: [2102.07774](https://arxiv.org/abs/2102.07774) [hep-th].
- [38] Gerard 't Hooft. “On the quantum structure of a black hole”. In: *Nucl. Phys. B* 256 (1985), p. 727. DOI: [10.1016/0550-3213\(85\)90418-3](https://doi.org/10.1016/0550-3213(85)90418-3).
- [39] Gerard 't Hooft. “The black hole interpretation of string theory”. In: *Nucl. Phys. B* 335 (1990), p. 138. DOI: [10.1016/0550-3213\(90\)90174-C](https://doi.org/10.1016/0550-3213(90)90174-C).
- [40] Leonard Susskind, Larus Thorlacius, and John Uglum. “The stretched horizon and black hole complementarity”. In: *Phys. Rev. D* 48 (1993), p. 3743. DOI: [10.1103/PhysRevD.48.3743](https://doi.org/10.1103/PhysRevD.48.3743). arXiv: [hep-th/9306069](https://arxiv.org/abs/hep-th/9306069) [hep-th].
- [41] Don N. Page. “Information in black hole radiation”. In: *Phys. Rev. Lett.* 71 (1993), p. 3743. DOI: [10.1103/PhysRevLett.71.3743](https://doi.org/10.1103/PhysRevLett.71.3743). arXiv: [hep-th/9306083](https://arxiv.org/abs/hep-th/9306083) [hep-th].
- [42] Yasunori Nomura. “Reanalyzing an evaporating black hole”. In: *Phys. Rev. D* 99 (2019), p. 086004. DOI: [10.1103/PhysRevD.99.086004](https://doi.org/10.1103/PhysRevD.99.086004). arXiv: [1810.09453](https://arxiv.org/abs/1810.09453) [hep-th].
- [43] Yasunori Nomura. “Spacetime and universal soft modes — black holes and beyond”. In: *Phys. Rev. D* 101 (2020), p. 066024. DOI: [10.1103/PhysRevD.101.066024](https://doi.org/10.1103/PhysRevD.101.066024). arXiv: [1908.05728](https://arxiv.org/abs/1908.05728) [hep-th].
- [44] Yasunori Nomura. “The interior of a unitarily evaporating black hole”. In: *Phys. Rev. D* 102 (2020), p. 026001. DOI: [10.1103/PhysRevD.102.026001](https://doi.org/10.1103/PhysRevD.102.026001). arXiv: [1911.13120](https://arxiv.org/abs/1911.13120) [hep-th].
- [45] Kevin Langhoff and Yasunori Nomura. “Ensemble from coarse graining: reconstructing the interior of an evaporating black hole”. In: *Phys. Rev. D* 102 (2020), p. 086021. DOI: [10.1103/PhysRevD.102.086021](https://doi.org/10.1103/PhysRevD.102.086021). arXiv: [2008.04202](https://arxiv.org/abs/2008.04202) [hep-th].
- [46] Yasunori Nomura. “Black hole interior in unitary gauge construction”. In: *Phys. Rev. D* 103 (2021), p. 066011. DOI: [10.1103/PhysRevD.103.066011](https://doi.org/10.1103/PhysRevD.103.066011). arXiv: [2010.15827](https://arxiv.org/abs/2010.15827) [hep-th].
- [47] Kyriakos Papadodimas and Suvrat Raju. “An infalling observer in AdS/CFT”. In: *JHEP* 10 (2013), p. 212. DOI: [10.1007/JHEP10\(2013\)212](https://doi.org/10.1007/JHEP10(2013)212). arXiv: [1211.6767](https://arxiv.org/abs/1211.6767) [hep-th].

- [48] Erik P. Verlinde and Herman L. Verlinde. “Black hole entanglement and quantum error correction”. In: *JHEP* 10 (2013), p. 107. DOI: [10.1007/JHEP10\(2013\)107](https://doi.org/10.1007/JHEP10(2013)107). arXiv: [1211.6913](https://arxiv.org/abs/1211.6913) [hep-th].
- [49] Yasunori Nomura and Jaime Varela. “A note on (no) firewalls: the entropy argument”. In: *JHEP* 07 (2013), p. 124. DOI: [10.1007/JHEP07\(2013\)124](https://doi.org/10.1007/JHEP07(2013)124). arXiv: [1211.7033](https://arxiv.org/abs/1211.7033) [hep-th].
- [50] Juan Martin Maldacena and Leonard Susskind. “Cool horizons for entangled black holes”. In: *Fortsch. Phys.* 61 (2013), p. 781. DOI: [10.1002/prop.201300020](https://doi.org/10.1002/prop.201300020). arXiv: [1306.0533](https://arxiv.org/abs/1306.0533) [hep-th].
- [51] Patrick Hayden and Geoff Penington. “Learning the alpha-bits of black holes”. In: *JHEP* 12 (2019), p. 007. DOI: [10.1007/JHEP12\(2019\)007](https://doi.org/10.1007/JHEP12(2019)007). arXiv: [1807.06041](https://arxiv.org/abs/1807.06041) [hep-th].
- [52] Kyriakos Papadodimas and Suvrat Raju. “State-dependent bulk-boundary maps and black hole complementarity”. In: *Phys. Rev. D* 89 (2014), p. 086010. DOI: [10.1103/PhysRevD.89.086010](https://doi.org/10.1103/PhysRevD.89.086010). arXiv: [1310.6335](https://arxiv.org/abs/1310.6335) [hep-th].
- [53] Kyriakos Papadodimas and Suvrat Raju. “Remarks on the necessity and implications of state-dependence in the black hole interior”. In: *Phys. Rev. D* 93 (2016), p. 084049. DOI: [10.1103/PhysRevD.93.084049](https://doi.org/10.1103/PhysRevD.93.084049). arXiv: [1503.08825](https://arxiv.org/abs/1503.08825) [hep-th].
- [54] Ahmed Almheiri et al. “The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole”. In: *JHEP* 12 (2019), p. 063. DOI: [10.1007/JHEP12\(2019\)063](https://doi.org/10.1007/JHEP12(2019)063). arXiv: [1905.08762](https://arxiv.org/abs/1905.08762) [hep-th].
- [55] Ahmed Almheiri et al. “The Page curve of Hawking radiation from semiclassical geometry”. In: *JHEP* 03 (2020), p. 149. DOI: [10.1007/JHEP03\(2020\)149](https://doi.org/10.1007/JHEP03(2020)149). arXiv: [1908.10996](https://arxiv.org/abs/1908.10996) [hep-th].
- [56] Yasunori Nomura. “From the black hole conundrum to the structure of quantum gravity”. In: *Mod. Phys. Lett. A* 36 (2021), p. 2130007. DOI: [10.1142/S021773232130007X](https://doi.org/10.1142/S021773232130007X). arXiv: [2011.08707](https://arxiv.org/abs/2011.08707) [hep-th].
- [57] Gary W. Gibbons and Stephen W. Hawking. “Cosmological event horizons, thermodynamics, and particle creation”. In: *Phys. Rev. D* 15 (1977), p. 2738. DOI: [10.1103/PhysRevD.15.2738](https://doi.org/10.1103/PhysRevD.15.2738).
- [58] Yasunori Nomura, Pratik Rath, and Nico Salzetta. “Spacetime from unentanglement”. In: *Phys. Rev. D* 97 (2018), p. 106010. DOI: [10.1103/PhysRevD.97.106010](https://doi.org/10.1103/PhysRevD.97.106010). arXiv: [1711.05263](https://arxiv.org/abs/1711.05263) [hep-th].

- [59] Fabio Sanches and Sean J. Weinberg. “Holographic entanglement entropy conjecture for general spacetimes”. In: *Phys. Rev. D* 94 (2016), p. 084034. DOI: [10.1103/PhysRevD.94.084034](https://doi.org/10.1103/PhysRevD.94.084034). arXiv: [1603.05250](https://arxiv.org/abs/1603.05250) [hep-th].
- [60] Yasunori Nomura et al. “Toward a holographic theory for general spacetimes”. In: *Phys. Rev. D* 95 (2017), p. 086002. DOI: [10.1103/PhysRevD.95.086002](https://doi.org/10.1103/PhysRevD.95.086002). arXiv: [1611.02702](https://arxiv.org/abs/1611.02702) [hep-th].
- [61] Mohsen Alishahiha et al. “The dS/dS correspondence”. In: *AIP Conf. Proc.* 743 (2004), p. 393. DOI: [10.1063/1.1848341](https://doi.org/10.1063/1.1848341). arXiv: [hep-th/0407125](https://arxiv.org/abs/hep-th/0407125) [hep-th].
- [62] Mohsen Alishahiha, Andreas Karch, and Eva Silverstein. “Hologravity”. In: *JHEP* 06 (2005), p. 028. DOI: [10.1088/1126-6708/2005/06/028](https://doi.org/10.1088/1126-6708/2005/06/028). arXiv: [hep-th/0504056](https://arxiv.org/abs/hep-th/0504056) [hep-th].
- [63] Leonard Susskind. “Black holes hint towards de Sitter-matrix theory”. In: (2021). arXiv: [2109.01322](https://arxiv.org/abs/2109.01322) [hep-th].
- [64] Leonard Susskind. “Entanglement and chaos in de Sitter space holography: an SYK example”. In: *JHAP* 1 (2021), p. 1. DOI: [10.22128/jhap.2021.455.1005](https://doi.org/10.22128/jhap.2021.455.1005). arXiv: [2109.14104](https://arxiv.org/abs/2109.14104) [hep-th].
- [65] Edgar Shaghoulian. “The central dogma and cosmological horizons”. In: *JHEP* 01 (2022), p. 132. DOI: [10.1007/JHEP01\(2022\)132](https://doi.org/10.1007/JHEP01(2022)132). arXiv: [2110.13210](https://arxiv.org/abs/2110.13210) [hep-th].
- [66] Edgar Shaghoulian and Leonard Susskind. “Entanglement in de Sitter space”. In: *JHEP* 08 (2022), p. 198. DOI: [10.1007/JHEP08\(2022\)198](https://doi.org/10.1007/JHEP08(2022)198). arXiv: [2201.03603](https://arxiv.org/abs/2201.03603) [hep-th].
- [67] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. “A bound on chaos”. In: *JHEP* 08 (2016), p. 106. DOI: [10.1007/JHEP08\(2016\)106](https://doi.org/10.1007/JHEP08(2016)106). arXiv: [1503.01409](https://arxiv.org/abs/1503.01409) [hep-th].
- [68] Patrick Hayden and John Preskill. “Black holes as mirrors: quantum information in random subsystems”. In: *JHEP* 09 (2007), p. 120. DOI: [10.1088/1126-6708/2007/09/120](https://doi.org/10.1088/1126-6708/2007/09/120). arXiv: [0708.4025](https://arxiv.org/abs/0708.4025) [hep-th].
- [69] Yasuhiro Sekino and Leonard Susskind. “Fast scramblers”. In: *JHEP* 10 (2008), p. 065. DOI: [10.1088/1126-6708/2008/10/065](https://doi.org/10.1088/1126-6708/2008/10/065). arXiv: [0808.2096](https://arxiv.org/abs/0808.2096) [hep-th].
- [70] William G. Unruh. “Notes on black hole evaporation”. In: *Phys. Rev. D* 14 (1976), p. 870. DOI: [10.1103/PhysRevD.14.870](https://doi.org/10.1103/PhysRevD.14.870).
- [71] Werner Israel. “Thermo field dynamics of black holes”. In: *Phys. Lett. A* 57 (1976), pp. 107–110. DOI: [10.1016/0375-9601\(76\)90178-X](https://doi.org/10.1016/0375-9601(76)90178-X).

- [72] Tom Banks. “Cosmological breaking of supersymmetry?” In: *Int. J. Mod. Phys. A* 16 (2001), p. 910. DOI: [10.1142/S0217751X01003998](https://doi.org/10.1142/S0217751X01003998). arXiv: [hep-th/0007146](https://arxiv.org/abs/hep-th/0007146).
- [73] Edward Witten. “Quantum gravity in de Sitter space”. In: *Strings 2001*. 2001. arXiv: [hep-th/0106109](https://arxiv.org/abs/hep-th/0106109).
- [74] Leonard Susskind. *Addendum to fast scramblers*. 2011. arXiv: [1101.6048](https://arxiv.org/abs/1101.6048) [[hep-th](#)].
- [75] Nima Goheer, Matthew Kleban, and Leonard Susskind. “The trouble with de Sitter space”. In: *JHEP* 07 (2003), p. 056. DOI: [10.1088/1126-6708/2003/07/056](https://doi.org/10.1088/1126-6708/2003/07/056). arXiv: [hep-th/0212209](https://arxiv.org/abs/hep-th/0212209).
- [76] Jason Pollack et al. “Eigenstate thermalization and disorder averaging in gravity”. In: *Phys. Rev. Lett.* 125 (2020), p. 021601. DOI: [10.1103/PhysRevLett.125.021601](https://doi.org/10.1103/PhysRevLett.125.021601). arXiv: [2002.02971](https://arxiv.org/abs/2002.02971) [[hep-th](#)].
- [77] Alexandre Belin and Jan de Boer. “Random statistics of OPE coefficients and Euclidean wormholes”. In: *Class. Quant. Grav.* 38 (2021), p. 164001. DOI: [10.1088/1361-6382/ac1082](https://doi.org/10.1088/1361-6382/ac1082). arXiv: [2006.05499](https://arxiv.org/abs/2006.05499) [[hep-th](#)].
- [78] Hong Liu and Sagar Vardhan. “Entanglement entropies of equilibrated pure states in quantum many-body systems and gravity”. In: *PRX Quantum* 2 (2021), p. 010344. DOI: [10.1103/PRXQuantum.2.010344](https://doi.org/10.1103/PRXQuantum.2.010344). arXiv: [2008.01089](https://arxiv.org/abs/2008.01089) [[hep-th](#)].
- [79] Ben Freivogel, Dora Nikolakopoulou, and Antonio F. Rotundo. “Wormholes from averaging over states”. In: *SciPost Phys.* 14.3 (2023), p. 026. DOI: [10.21468/SciPostPhys.14.3.026](https://doi.org/10.21468/SciPostPhys.14.3.026). arXiv: [2105.12771](https://arxiv.org/abs/2105.12771) [[hep-th](#)].
- [80] Jordan Cotler and Kristan Jensen. “A precision test of averaging in AdS/CFT”. In: *JHEP* 11 (2022), p. 070. DOI: [10.1007/JHEP11\(2022\)070](https://doi.org/10.1007/JHEP11(2022)070). arXiv: [2205.12968](https://arxiv.org/abs/2205.12968) [[hep-th](#)].
- [81] Raphael Bousso and Elizabeth Wildenhain. “Gravity/ensemble duality”. In: *Phys. Rev. D* 102 (2020), p. 066005. DOI: [10.1103/PhysRevD.102.066005](https://doi.org/10.1103/PhysRevD.102.066005). arXiv: [2006.16289](https://arxiv.org/abs/2006.16289) [[hep-th](#)].
- [82] Netta Engelhardt, Sebastian Fischetti, and Alexander Maloney. “Free energy from replica wormholes”. In: *Phys. Rev. D* 103 (2021), p. 046021. DOI: [10.1103/PhysRevD.103.046021](https://doi.org/10.1103/PhysRevD.103.046021). arXiv: [2007.07444](https://arxiv.org/abs/2007.07444) [[hep-th](#)].
- [83] Renato Renner and Jie Wang. *The black hole information puzzle and the quantum de Finetti theorem*. 2021. arXiv: [2110.14653](https://arxiv.org/abs/2110.14653) [[hep-th](#)].

- [84] Xiao-Liang Qi, Shangnan Zhou, and Zhenbin Yang. “Holevo information and ensemble theory of gravity”. In: *JHEP* 02 (2022), p. 056. DOI: [10.1007/JHEP02\(2022\)056](https://doi.org/10.1007/JHEP02(2022)056). arXiv: [2111.05355](https://arxiv.org/abs/2111.05355) [hep-th].
- [85] Ahmed Almheiri and Henry W. Lin. “The entanglement wedge of unknown couplings”. In: *JHEP* 08 (2022), p. 062. DOI: [10.1007/JHEP08\(2022\)062](https://doi.org/10.1007/JHEP08(2022)062). arXiv: [2111.06298](https://arxiv.org/abs/2111.06298) [hep-th].
- [86] Andreas Blommaert, Luca V. Iliesiu, and Jorrit Kruthoff. “Gravity factorized”. In: *JHEP* 09 (2022), p. 080. DOI: [10.1007/JHEP09\(2022\)080](https://doi.org/10.1007/JHEP09(2022)080). arXiv: [2111.07863](https://arxiv.org/abs/2111.07863) [hep-th].
- [87] Jeevan Chandra and Thomas Hartman. “Coarse graining pure states in AdS/CFT”. In: *JHEP* 10 (2023), p. 030. DOI: [10.1007/JHEP10\(2023\)030](https://doi.org/10.1007/JHEP10(2023)030). arXiv: [2206.03414](https://arxiv.org/abs/2206.03414) [hep-th].
- [88] Henry Maxfield and Gustavo J. Turiaci. “The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral”. In: *JHEP* 01 (2021), p. 118. DOI: [10.1007/JHEP01\(2021\)118](https://doi.org/10.1007/JHEP01(2021)118). arXiv: [2006.11317](https://arxiv.org/abs/2006.11317) [hep-th].
- [89] Matthew Heydeman et al. “The statistical mechanics of near-BPS black holes”. In: *J. Phys. A* 55 (2022), p. 014004. DOI: [10.1088/1751-8121/ac3be9](https://doi.org/10.1088/1751-8121/ac3be9). arXiv: [2011.01953](https://arxiv.org/abs/2011.01953) [hep-th].
- [90] Tullio Regge and John A. Wheeler. “Stability of a Schwarzschild singularity”. In: *Phys. Rev.* 108 (1957), p. 1063. DOI: [10.1103/PhysRev.108.1063](https://doi.org/10.1103/PhysRev.108.1063).
- [91] Richard H. Price. “Nonspherical perturbations of relativistic gravitational collapse. I. scalar and gravitational perturbations”. In: *Phys. Rev. D* 5 (1972), p. 2419. DOI: [10.1103/PhysRevD.5.2419](https://doi.org/10.1103/PhysRevD.5.2419).
- [92] Saul A. Teukolsky. “Perturbations of a rotating black hole. I. Fundamental equations for gravitational electromagnetic and neutrino field perturbations”. In: *Astrophys. J.* 185 (1973), p. 635. DOI: [10.1086/152444](https://doi.org/10.1086/152444).
- [93] Yasuyuki Hatsuda. “An alternative to the Teukolsky equation”. In: *Gen. Rel. Grav.* 53 (2021), p. 93. DOI: [10.1007/s10714-021-02866-4](https://doi.org/10.1007/s10714-021-02866-4). arXiv: [2007.07906](https://arxiv.org/abs/2007.07906) [gr-qc].
- [94] Jacob D. Bekenstein. “Nonexistence of baryon number for static black holes”. In: *Phys. Rev. D* 5 (1972), p. 1239. DOI: [10.1103/PhysRevD.5.1239](https://doi.org/10.1103/PhysRevD.5.1239).
- [95] Ya. B. Zeldovich. “A novel type of radioactive decay: gravitational baryon annihilation”. In: *Zh. Eksp. Teor. Fiz.* 72 (1977), p. 18.

- [96] Tom Banks and Nathan Seiberg. “Symmetries and strings in field theory and gravity”. In: *Phys. Rev. D* 83 (2011), p. 084019. DOI: [10.1103/PhysRevD.83.084019](#). arXiv: [1011.5120 \[hep-th\]](#).
- [97] Daniel Harlow and Hiroshi Ooguri. “Constraints on symmetries from holography”. In: *Phys. Rev. Lett.* 122 (2019), p. 191601. DOI: [10.1103/PhysRevLett.122.191601](#). arXiv: [1810.05337 \[hep-th\]](#).
- [98] Gia Dvali. “Black holes and large N species solution to the hierarchy problem”. In: *Fortsch. Phys.* 58 (2010), p. 528. DOI: [10.1002/prop.201000009](#). arXiv: [0706.2050 \[hep-th\]](#).
- [99] Don N. Page. “Is black hole evaporation predictable?” In: *Phys. Rev. Lett.* 44 (1980), p. 301. DOI: [10.1103/PhysRevLett.44.301](#).
- [100] Yasunori Nomura, Jaime Varela, and Sean J. Weinberg. “Black holes, information, and Hilbert space for quantum gravity”. In: *Phys. Rev. D* 87 (2013), p. 084050. DOI: [10.1103/PhysRevD.87.084050](#). arXiv: [1210.6348 \[hep-th\]](#).
- [101] Sinya Aoki, Tetsuya Onogi, and Shuichi Yokoyama. “Conserved charges in general relativity”. In: *Int. J. Mod. Phys. A* 36 (2021), p. 2150098. DOI: [10.1142/S0217751X21500986](#). arXiv: [2005.13233 \[gr-qc\]](#).
- [102] Yasunori Nomura and Sean J. Weinberg. “Entropy of a vacuum: what does the covariant entropy count?” In: *Phys. Rev. D* 90 (2014), p. 104003. DOI: [10.1103/PhysRevD.90.104003](#). arXiv: [1310.7564 \[hep-th\]](#).
- [103] Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. “A covariant holographic entanglement entropy proposal”. In: *JHEP* 07 (2007), p. 062. DOI: [10.1088/1126-6708/2007/07/062](#). arXiv: [0705.0016 \[hep-th\]](#).
- [104] Aitor Lewkowycz and Juan Maldacena. “Generalized gravitational entropy”. In: *JHEP* 08 (2013), p. 090. DOI: [10.1007/JHEP08\(2013\)090](#). arXiv: [1304.4926 \[hep-th\]](#).
- [105] Xi Dong, Aitor Lewkowycz, and Mukund Rangamani. “Deriving covariant holographic entanglement”. In: *JHEP* 11 (2016), p. 028. DOI: [10.1007/JHEP11\(2016\)028](#). arXiv: [1607.07506 \[hep-th\]](#).
- [106] Ahmed Almheiri, Xi Dong, and Brian Swingle. “Linearity of holographic entanglement entropy”. In: *JHEP* 02 (2017), p. 074. DOI: [10.1007/JHEP02\(2017\)074](#). arXiv: [1606.04537 \[hep-th\]](#).

- [107] Wojciech H. Zurek. “Entropy evaporated by a black hole”. In: *Phys. Rev. Lett.* 49 (1982), p. 1683. DOI: [10.1103/PhysRevLett.49.1683](https://doi.org/10.1103/PhysRevLett.49.1683).
- [108] Don N. Page. “Comment on ‘Entropy evaporated by a black hole’”. In: *Phys. Rev. Lett.* 50 (1983), p. 1013. DOI: [10.1103/PhysRevLett.50.1013](https://doi.org/10.1103/PhysRevLett.50.1013).
- [109] James B. Hartle and Stephen W. Hawking. “Path integral derivation of black hole radiance”. In: *Phys. Rev. D* 13 (1976), p. 2188. DOI: [10.1103/PhysRevD.13.2188](https://doi.org/10.1103/PhysRevD.13.2188).
- [110] Yasunori Nomura, Fabio Sanches, and Sean J. Weinberg. “Black hole interior in quantum gravity”. In: *Phys. Rev. Lett.* 114 (2015), p. 201301. DOI: [10.1103/PhysRevLett.114.201301](https://doi.org/10.1103/PhysRevLett.114.201301). arXiv: [1412.7539 \[hep-th\]](https://arxiv.org/abs/1412.7539).
- [111] Yasunori Nomura, Fabio Sanches, and Sean J. Weinberg. “Relativeness in quantum gravity: limitations and frame dependence of semiclassical descriptions”. In: *JHEP* 04 (2015), p. 158. DOI: [10.1007/JHEP04\(2015\)158](https://doi.org/10.1007/JHEP04(2015)158). arXiv: [1412.7538 \[hep-th\]](https://arxiv.org/abs/1412.7538).
- [112] William G. Unruh. “Origin of the particles in black hole evaporation”. In: *Phys. Rev. D* 15 (1977), p. 365. DOI: [10.1103/PhysRevD.15.365](https://doi.org/10.1103/PhysRevD.15.365).
- [113] Werner Israel. *Shenanigans at the black hole horizon: pair creation or Boulware accretion?* 2015. arXiv: [1504.02419 \[gr-qc\]](https://arxiv.org/abs/1504.02419).
- [114] Steven B. Giddings. “Hawking radiation, the Stefan–Boltzmann law, and unitarization”. In: *Phys. Lett. B* 754 (2016), p. 39. DOI: [10.1016/j.physletb.2015.12.076](https://doi.org/10.1016/j.physletb.2015.12.076). arXiv: [1511.08221 \[hep-th\]](https://arxiv.org/abs/1511.08221).
- [115] William G. Unruh and Robert M. Wald. “Acceleration radiation and generalized second law of thermodynamics”. In: *Phys. Rev. D* 25 (1982), p. 942. DOI: [10.1103/PhysRevD.25.942](https://doi.org/10.1103/PhysRevD.25.942).
- [116] Adam R. Brown. “Tensile strength and the mining of black holes”. In: *Phys. Rev. Lett.* 111 (2013), p. 211301. DOI: [10.1103/PhysRevLett.111.211301](https://doi.org/10.1103/PhysRevLett.111.211301). arXiv: [1207.3342 \[gr-qc\]](https://arxiv.org/abs/1207.3342).
- [117] Don N. Page. “Particle emission rates from a black hole: massless particles from an uncharged, nonrotating hole”. In: *Phys. Rev. D* 13 (1976), p. 198. DOI: [10.1103/PhysRevD.13.198](https://doi.org/10.1103/PhysRevD.13.198).
- [118] Sudip Ghosh and Suvrat Raju. “Quantum information measures for restricted sets of observables”. In: *Phys. Rev. D* 98 (2018), p. 046005. DOI: [10.1103/PhysRevD.98.046005](https://doi.org/10.1103/PhysRevD.98.046005). arXiv: [1712.09365 \[hep-th\]](https://arxiv.org/abs/1712.09365).

- [119] Fernando Pastawski et al. “Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence”. In: *JHEP* 06 (2015), p. 149. DOI: [10.1007/JHEP06\(2015\)149](https://doi.org/10.1007/JHEP06(2015)149). arXiv: [1503.06237](https://arxiv.org/abs/1503.06237) [hep-th].
- [120] Patrick Hayden et al. “Holographic duality from random tensor networks”. In: *JHEP* 11 (2016), p. 009. DOI: [10.1007/JHEP11\(2016\)009](https://doi.org/10.1007/JHEP11(2016)009). arXiv: [1601.01694](https://arxiv.org/abs/1601.01694) [hep-th].
- [121] Sidney R. Coleman and Frank De Luccia. “Gravitational effects on and of vacuum decay”. In: *Phys. Rev. D* 21 (1980), p. 3305. DOI: [10.1103/PhysRevD.21.3305](https://doi.org/10.1103/PhysRevD.21.3305).
- [122] Ning Bao et al. “Beyond toy models: distilling tensor networks in full AdS/CFT”. In: *JHEP* 11 (2019), p. 069. DOI: [10.1007/JHEP11\(2019\)069](https://doi.org/10.1007/JHEP11(2019)069). arXiv: [1812.01171](https://arxiv.org/abs/1812.01171) [hep-th].
- [123] Brian Swingle. “Entanglement renormalization and holography”. In: *Phys. Rev. D* 86 (2012), p. 065007. DOI: [10.1103/PhysRevD.86.065007](https://doi.org/10.1103/PhysRevD.86.065007). arXiv: [0905.1317](https://arxiv.org/abs/0905.1317) [cond-mat.str-el].
- [124] Masamichi Miyaji and Tadashi Takayanagi. “Surface/state correspondence as a generalized holography”. In: *PTEP* 2015 (2015), 073B03. DOI: [10.1093/ptep/ptv089](https://doi.org/10.1093/ptep/ptv089). arXiv: [1503.03542](https://arxiv.org/abs/1503.03542) [hep-th].
- [125] Netta Engelhardt and Aron C. Wall. “Extremal surface barriers”. In: *JHEP* 03 (2014), p. 068. DOI: [10.1007/JHEP03\(2014\)068](https://doi.org/10.1007/JHEP03(2014)068). arXiv: [1312.3699](https://arxiv.org/abs/1312.3699) [hep-th].
- [126] Idse Heemskerk et al. “Bulk and transhorizon measurements in AdS/CFT”. In: *JHEP* 10 (2012), p. 165. DOI: [10.1007/JHEP10\(2012\)165](https://doi.org/10.1007/JHEP10(2012)165). arXiv: [1201.3664](https://arxiv.org/abs/1201.3664) [hep-th].
- [127] Raphael Bousso et al. “Quantum focusing conjecture”. In: *Phys. Rev. D* 93 (2016), p. 064044. DOI: [10.1103/PhysRevD.93.064044](https://doi.org/10.1103/PhysRevD.93.064044). arXiv: [1506.02669](https://arxiv.org/abs/1506.02669) [hep-th].
- [128] Raphael Bousso. “Holography in general space-times”. In: *JHEP* 06 (1999), p. 028. DOI: [10.1088/1126-6708/1999/06/028](https://doi.org/10.1088/1126-6708/1999/06/028). arXiv: [hep-th/9906022](https://arxiv.org/abs/hep-th/9906022).
- [129] Raphael Bousso and Netta Engelhardt. “Generalized second law for cosmology”. In: *Phys. Rev. D* 93 (2016), p. 024025. DOI: [10.1103/PhysRevD.93.024025](https://doi.org/10.1103/PhysRevD.93.024025). arXiv: [1510.02099](https://arxiv.org/abs/1510.02099) [hep-th].
- [130] Charles Fefferman and C. Robin Graham. “Conformal invariants”. In: *The Mathematical Heritage of Élie Cartan (Lyon, 1984)*. 1985, p. 95.
- [131] Emil T. Akhmedov. “A remark on the AdS/CFT correspondence and the renormalization group flow”. In: *Phys. Lett. B* 442 (1998), p. 152. DOI: [10.1016/S0370-2693\(98\)01270-2](https://doi.org/10.1016/S0370-2693(98)01270-2). arXiv: [hep-th/9806217](https://arxiv.org/abs/hep-th/9806217).

- [132] Enrique Álvarez and Cesar Gómez. “Geometric holography, the renormalization group and the c-theorem”. In: *Nucl. Phys. B* 541 (1999), p. 441. DOI: [10.1016/S0550-3213\(98\)00752-4](#). arXiv: [hep-th/9807226](#).
- [133] Vijay Balasubramanian and Per Kraus. “Space-time and the holographic renormalization group”. In: *Phys. Rev. Lett.* 83 (1999), p. 3605. DOI: [10.1103/PhysRevLett.83.3605](#). arXiv: [hep-th/9903190](#).
- [134] Kostas Skenderis and Paul K. Townsend. “Gravitational stability and renormalization group flow”. In: *Phys. Lett. B* 468 (1999), p. 46. DOI: [10.1016/S0370-2693\(99\)01212-5](#). arXiv: [hep-th/9909070](#).
- [135] Jan de Boer, Erik P. Verlinde, and Herman Verlinde. “On the holographic renormalization group”. In: *JHEP* 08 (2000), p. 003. DOI: [10.1088/1126-6708/2000/08/003](#). arXiv: [hep-th/9912012](#).
- [136] Hermann Bondi, M. G. J. van der Burg, and A. W. K. Metzner. “Gravitational waves in general relativity VII. Waves from axisymmetric isolated systems”. In: *Proc. Roy. Soc. Lond. A* 269 (1962), p. 21. DOI: [10.1098/rspa.1962.0161](#).
- [137] Rainer K. Sachs. “Gravitational waves in general relativity VIII. Waves in asymptotically flat space-times”. In: *Proc. Roy. Soc. Lond. A* 270 (1962), p. 103. DOI: [10.1098/rspa.1962.0206](#).
- [138] Jan de Boer and Sergey N. Solodukhin. “A holographic reduction of Minkowski space-time”. In: *Nucl. Phys. B* 665 (2003), p. 545. DOI: [10.1016/S0550-3213\(03\)00494-2](#). arXiv: [hep-th/0303006](#).
- [139] Andrew Strominger. “On BMS invariance of gravitational scattering”. In: *JHEP* 07 (2014), p. 152. DOI: [10.1007/JHEP07\(2014\)152](#). arXiv: [1312.2229 \[hep-th\]](#).
- [140] Clifford Cheung, Anton de la Fuente, and Raman Sundrum. “4D scattering amplitudes and asymptotic symmetries from 2D CFT”. In: *JHEP* 01 (2017), p. 112. DOI: [10.1007/JHEP01\(2017\)112](#). arXiv: [1609.00732 \[hep-th\]](#).
- [141] Sabrina Pasterski, Shu-Heng Shao, and Andrew Strominger. “Flat space amplitudes and conformal symmetry of the celestial sphere”. In: *Phys. Rev. D* 96 (2017), p. 065026. DOI: [10.1103/PhysRevD.96.065026](#). arXiv: [1701.00049 \[hep-th\]](#).
- [142] Alok Laddha et al. “The holographic nature of null infinity”. In: *SciPost Phys.* 10 (2021), p. 041. DOI: [10.21468/SciPostPhys.10.2.041](#). arXiv: [2002.02448 \[hep-th\]](#).

- [143] Lisa Dyson, Matthew Kleban, and Leonard Susskind. “Disturbing implications of a cosmological constant”. In: *JHEP* 10 (2002), p. 011. DOI: [10.1088/1126-6708/2002/10/011](https://doi.org/10.1088/1126-6708/2002/10/011). arXiv: [hep-th/0208013](https://arxiv.org/abs/hep-th/0208013).
- [144] Shamit Kachru et al. “de Sitter vacua in string theory”. In: *Phys. Rev. D* 68 (2003), p. 046005. DOI: [10.1103/PhysRevD.68.046005](https://doi.org/10.1103/PhysRevD.68.046005). arXiv: [hep-th/0301240](https://arxiv.org/abs/hep-th/0301240).
- [145] L. F. Abbott and Stanley Deser. “Stability of gravity with a cosmological constant”. In: *Nucl. Phys. B* 195 (1982), p. 76. DOI: [10.1016/0550-3213\(82\)90049-9](https://doi.org/10.1016/0550-3213(82)90049-9).
- [146] Yasunori Nomura and Sean J. Weinberg. “Black holes, entropies, and semiclassical spacetime in quantum gravity”. In: *JHEP* 10 (2014), p. 185. DOI: [10.1007/JHEP10\(2014\)185](https://doi.org/10.1007/JHEP10(2014)185). arXiv: [1406.1505](https://arxiv.org/abs/1406.1505) [[hep-th](#)].
- [147] P. V. Buividovich and Mikhail I. Polikarpov. “Entanglement entropy in gauge theories and the holographic principle for electric strings”. In: *Phys. Lett. B* 670 (2008), p. 141. DOI: [10.1016/j.physletb.2008.10.032](https://doi.org/10.1016/j.physletb.2008.10.032). arXiv: [0806.3376](https://arxiv.org/abs/0806.3376) [[hep-th](#)].
- [148] William Donnelly. “Decomposition of entanglement entropy in lattice gauge theory”. In: *Phys. Rev. D* 85 (2012), p. 085004. DOI: [10.1103/PhysRevD.85.085004](https://doi.org/10.1103/PhysRevD.85.085004). arXiv: [1109.0036](https://arxiv.org/abs/1109.0036) [[hep-th](#)].
- [149] Horacio Casini, Marina Huerta, and Javier A. Rosabal. “Remarks on entanglement entropy for gauge fields”. In: *Phys. Rev. D* 89 (2014), p. 085012. DOI: [10.1103/PhysRevD.89.085012](https://doi.org/10.1103/PhysRevD.89.085012). arXiv: [1312.1183](https://arxiv.org/abs/1312.1183) [[hep-th](#)].
- [150] Sudip Ghosh, Ronak M. Soni, and Sandip P. Trivedi. “On the entanglement entropy for gauge theories”. In: *JHEP* 09 (2015), p. 069. DOI: [10.1007/JHEP09\(2015\)069](https://doi.org/10.1007/JHEP09(2015)069). arXiv: [1501.02593](https://arxiv.org/abs/1501.02593) [[hep-th](#)].
- [151] William Donnelly and Laurent Freidel. “Local subsystems in gauge theory and gravity”. In: *JHEP* 09 (2016), p. 102. DOI: [10.1007/JHEP09\(2016\)102](https://doi.org/10.1007/JHEP09(2016)102). arXiv: [1601.04744](https://arxiv.org/abs/1601.04744) [[hep-th](#)].
- [152] Tadashi Takayanagi and Kotaro Tamaoka. “Gravity edges modes and Hayward term”. In: *JHEP* 02 (2020), p. 167. DOI: [10.1007/JHEP02\(2020\)167](https://doi.org/10.1007/JHEP02(2020)167). arXiv: [1912.01636](https://arxiv.org/abs/1912.01636) [[hep-th](#)].
- [153] Sheldon Goldstein et al. “Canonical typicality”. In: *Phys. Rev. Lett.* 96 (2006), p. 050403. DOI: [10.1103/PhysRevLett.96.050403](https://doi.org/10.1103/PhysRevLett.96.050403). arXiv: [cond-mat/0511091](https://arxiv.org/abs/cond-mat/0511091) [[cond-mat.stat-mech](#)].

- [154] Sandu Popescu, Anthony J. Short, and Andreas Winter. “Entanglement and the foundations of statistical mechanics”. In: *Nature Phys.* 2.11 (2006), pp. 754–758. DOI: [10.1038/nphys444](#). arXiv: [quant-ph/0511225](#).
- [155] Don N. Page. “Average entropy of a subsystem”. In: *Phys. Rev. Lett.* 71 (1993), p. 1291. DOI: [10.1103/PhysRevLett.71.1291](#). arXiv: [gr-qc/9305007](#).
- [156] Jonathan Maltz and Leonard Susskind. “de Sitter space as a resonance”. In: *Phys. Rev. Lett.* 118 (2017), p. 101602. DOI: [10.1103/PhysRevLett.118.101602](#). arXiv: [1611.00360 \[hep-th\]](#).
- [157] Raphael Bousso. “Violations of the equivalence principle by a nonlocally reconstructed vacuum at the black hole horizon”. In: *Phys. Rev. Lett.* 112 (2014), p. 041102. DOI: [10.1103/PhysRevLett.112.041102](#). arXiv: [1308.3697 \[hep-th\]](#).
- [158] Daniel L. Jafferis and Lampros Lamprou. “Inside the hologram: reconstructing the bulk observer’s experience”. In: *JHEP* 03 (2022), p. 084. DOI: [10.1007/JHEP03\(2022\)084](#). arXiv: [2009.04476 \[hep-th\]](#).
- [159] Ping Gao and Lampros Lamprou. “Seeing behind black hole horizons in SYK”. In: *JHEP* 06 (2022), p. 143. DOI: [10.1007/JHEP06\(2022\)143](#). arXiv: [2111.14010 \[hep-th\]](#).
- [160] D’enes Petz. “Sufficient subalgebras and the relative entropy of states of a von Neumann algebra”. In: *Commun. Math. Phys.* 105 (1986), p. 123. DOI: [10.1007/BF01212345](#).
- [161] D’enes Petz. “Sufficiency of channels over von Neumann algebras”. In: *Quart. J. Math. Oxford Ser.* 39 (1988), p. 97. DOI: [10.1093/qmath/39.1.97](#).
- [162] Bartłomiej Czech et al. “The gravity dual of a density matrix”. In: *Class. Quant. Grav.* 29 (2012), p. 155009. DOI: [10.1088/0264-9381/29/15/155009](#). arXiv: [1204.1330 \[hep-th\]](#).
- [163] Aron C. Wall. “Maximin surfaces, and the strong subadditivity of the covariant holographic entanglement entropy”. In: *Class. Quant. Grav.* 31 (2014), p. 225007. DOI: [10.1088/0264-9381/31/22/225007](#). arXiv: [1211.3494 \[hep-th\]](#).
- [164] Matthew Headrick et al. “Causality & holographic entanglement entropy”. In: *JHEP* 12 (2014), p. 162. DOI: [10.1007/JHEP12\(2014\)162](#). arXiv: [1408.6300 \[hep-th\]](#).
- [165] Daniel L. Jafferis et al. “Relative entropy equals bulk relative entropy”. In: *JHEP* 06 (2016), p. 004. DOI: [10.1007/JHEP06\(2016\)004](#). arXiv: [1512.06431 \[hep-th\]](#).

- [166] Jordan Cotler et al. “Entanglement wedge reconstruction via universal recovery channels”. In: *Phys. Rev. X* 9 (2019), p. 031011. DOI: [10.1103/PhysRevX.9.031011](https://doi.org/10.1103/PhysRevX.9.031011). arXiv: [1704.05839](https://arxiv.org/abs/1704.05839) [hep-th].
- [167] Joseph Polchinski, Leonard Susskind, and Nicolaos Toumbas. “Negative energy, superluminality and holography”. In: *Phys. Rev. D* 60 (1999), p. 084006. DOI: [10.1103/PhysRevD.60.084006](https://doi.org/10.1103/PhysRevD.60.084006). arXiv: [hep-th/9903228](https://arxiv.org/abs/hep-th/9903228).
- [168] Steven B. Giddings and Matthew Lippert. “Precursors, black holes, and a locality bound”. In: *Phys. Rev. D* 65 (2002), p. 024006. DOI: [10.1103/PhysRevD.65.024006](https://doi.org/10.1103/PhysRevD.65.024006). arXiv: [hep-th/0103231](https://arxiv.org/abs/hep-th/0103231).
- [169] Isaac Kim, Eugene Tang, and John Preskill. “The ghost in the radiation: robust encodings of the black hole interior”. In: *JHEP* 06 (2020), p. 031. DOI: [10.1007/JHEP06\(2020\)031](https://doi.org/10.1007/JHEP06(2020)031). arXiv: [2003.05451](https://arxiv.org/abs/2003.05451) [hep-th].
- [170] Juan Martin Maldacena. “Eternal black holes in anti-de Sitter”. In: *JHEP* 04 (2003), p. 021. DOI: [10.1088/1126-6708/2003/04/021](https://doi.org/10.1088/1126-6708/2003/04/021). arXiv: [hep-th/0106112](https://arxiv.org/abs/hep-th/0106112).
- [171] Stephen H. Shenker and Douglas Stanford. “Black holes and the butterfly effect”. In: *JHEP* 03 (2014), p. 067. DOI: [10.1007/JHEP03\(2014\)067](https://doi.org/10.1007/JHEP03(2014)067). arXiv: [1306.0622](https://arxiv.org/abs/1306.0622) [hep-th].
- [172] Kyriakos Papadodimas and Suvrat Raju. “Local operators in the eternal black hole”. In: *Phys. Rev. Lett.* 115 (2015), p. 211601. DOI: [10.1103/PhysRevLett.115.211601](https://doi.org/10.1103/PhysRevLett.115.211601). arXiv: [1502.06692](https://arxiv.org/abs/1502.06692) [hep-th].
- [173] Adam R. Brown et al. “Holographic complexity equals bulk action?” In: *Phys. Rev. Lett.* 116 (2016), p. 191301. DOI: [10.1103/PhysRevLett.116.191301](https://doi.org/10.1103/PhysRevLett.116.191301). arXiv: [1509.07876](https://arxiv.org/abs/1509.07876) [hep-th].
- [174] Ahmed Almheiri, Raghu Mahajan, and Juan Maldacena. *Islands outside the horizon*. 2019. arXiv: [1910.11077](https://arxiv.org/abs/1910.11077) [hep-th].
- [175] Adam Bouland, Bill Fefferman, and Umesh Vazirani. *Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality*. 2019. arXiv: [1910.14646](https://arxiv.org/abs/1910.14646) [quant-ph].
- [176] Dominik Neuenfeld. “Homology conditions for RT surfaces in double holography”. In: *Class. Quant. Grav.* 39 (2022), p. 075009. DOI: [10.1088/1361-6382/ac51e7](https://doi.org/10.1088/1361-6382/ac51e7). arXiv: [2105.01130](https://arxiv.org/abs/2105.01130) [hep-th].
- [177] Herman Verlinde. *Deconstructing the wormhole: factorization, entanglement and decoherence*. 2021. arXiv: [2105.02142](https://arxiv.org/abs/2105.02142) [hep-th].

- [178] Samuel Leutheusser and Hong Liu. “Causal connectability between quantum systems and the black hole interior in holographic duality”. In: *Phys. Rev. D* 108.8 (2023), p. 086019. DOI: [10.1103/PhysRevD.108.086019](https://doi.org/10.1103/PhysRevD.108.086019). arXiv: [2110.05497](https://arxiv.org/abs/2110.05497) [hep-th].
- [179] Samuel Aaron Wehlau Leutheusser and Hong Liu. “Emergent Times in Holographic Duality”. In: *Phys. Rev. D* 108.8 (2023), p. 086020. DOI: [10.1103/PhysRevD.108.086020](https://doi.org/10.1103/PhysRevD.108.086020). arXiv: [2112.12156](https://arxiv.org/abs/2112.12156) [hep-th].
- [180] Edward Witten. “Gravity and the crossed product”. In: *JHEP* 10 (2022), p. 008. DOI: [10.1007/JHEP10\(2022\)008](https://doi.org/10.1007/JHEP10(2022)008). arXiv: [2112.12828](https://arxiv.org/abs/2112.12828) [hep-th].
- [181] Andrew Strominger. “The dS/CFT correspondence”. In: *JHEP* 10 (2001), p. 034. DOI: [10.1088/1126-6708/2001/10/034](https://doi.org/10.1088/1126-6708/2001/10/034). arXiv: [hep-th/0106113](https://arxiv.org/abs/hep-th/0106113).
- [182] Lars Aalsma and Gary Shiu. “Chaos and complementarity in de Sitter space”. In: *JHEP* 05 (2020), p. 152. DOI: [10.1007/JHEP05\(2020\)152](https://doi.org/10.1007/JHEP05(2020)152). arXiv: [2002.01326](https://arxiv.org/abs/2002.01326) [hep-th].
- [183] Hao Geng. “Non-local entanglement and fast scrambling in de-Sitter holography”. In: *Annals Phys.* 426 (2021), p. 168402. DOI: [10.1016/j.aop.2021.168402](https://doi.org/10.1016/j.aop.2021.168402). arXiv: [2005.00021](https://arxiv.org/abs/2005.00021) [hep-th].
- [184] Hao Geng, Yasunori Nomura, and Hao-Yu Sun. “Information paradox and its resolution in de Sitter holography”. In: *Phys. Rev. D* 103 (2021), p. 126004. DOI: [10.1103/PhysRevD.103.126004](https://doi.org/10.1103/PhysRevD.103.126004). arXiv: [2103.07477](https://arxiv.org/abs/2103.07477) [hep-th].
- [185] Lars Aalsma et al. “Shocks and information exchange in de Sitter space”. In: *JHEP* 10 (2021), p. 104. DOI: [10.1007/JHEP10\(2021\)104](https://doi.org/10.1007/JHEP10(2021)104). arXiv: [2105.12737](https://arxiv.org/abs/2105.12737) [hep-th].
- [186] Jamie Kames-King, Erik M. H. Verheijden, and Erik P. Verlinde. “No Page curves for the de Sitter horizon”. In: *JHEP* 03 (2022), p. 040. DOI: [10.1007/JHEP03\(2022\)040](https://doi.org/10.1007/JHEP03(2022)040). arXiv: [2108.09318](https://arxiv.org/abs/2108.09318) [hep-th].
- [187] Vijay Balasubramanian et al. “Multiboundary wormholes and holographic entanglement”. In: *Class. Quant. Grav.* 31 (2014), p. 185015. DOI: [10.1088/0264-9381/31/18/185015](https://doi.org/10.1088/0264-9381/31/18/185015). arXiv: [1406.2663](https://arxiv.org/abs/1406.2663) [hep-th].
- [188] Gary W. Gibbons and Stephen W. Hawking. “Action integrals and partition functions in quantum gravity”. In: *Phys. Rev. D* 15 (1977), p. 2752. DOI: [10.1103/PhysRevD.15.2752](https://doi.org/10.1103/PhysRevD.15.2752).
- [189] Daniel Harlow and Edgar Shaghoulian. “Global symmetry, Euclidean gravity, and the black hole information problem”. In: *JHEP* 04 (2021), p. 175. DOI: [10.1007/JHEP04\(2021\)175](https://doi.org/10.1007/JHEP04(2021)175). arXiv: [2010.10539](https://arxiv.org/abs/2010.10539) [hep-th].

- [190] Luca V. Iliesiu, Mark Mezei, and Gabor Sarosi. “The volume of the black hole interior at late times”. In: *JHEP* 07 (2022), p. 073. DOI: [10.1007/JHEP07\(2022\)073](https://doi.org/10.1007/JHEP07(2022)073). arXiv: [2107.06286](https://arxiv.org/abs/2107.06286) [hep-th].
- [191] T. S. Bunch and Paul C. W. Davies. “Quantum field theory in de Sitter space: renormalization by point splitting”. In: *Proc. Roy. Soc. Lond. A* 360 (1978), p. 117. DOI: [10.1098/rspa.1978.0060](https://doi.org/10.1098/rspa.1978.0060).
- [192] Hao Geng et al. “Information transfer with a gravitating bath”. In: *SciPost Phys.* 10 (2021), p. 103. DOI: [10.21468/SciPostPhys.10.5.103](https://doi.org/10.21468/SciPostPhys.10.5.103). arXiv: [2012.04671](https://arxiv.org/abs/2012.04671) [hep-th].
- [193] Curtis G. Callan and Frank Wilczek. “On geometric entropy”. In: *Phys. Lett. B* 333 (1994), p. 55. DOI: [10.1016/0370-2693\(94\)91007-3](https://doi.org/10.1016/0370-2693(94)91007-3). arXiv: [hep-th/9401072](https://arxiv.org/abs/hep-th/9401072).
- [194] Pasquale Calabrese and John Cardy. “Entanglement entropy and quantum field theory”. In: *J. Stat. Mech.* (2004), P06002. DOI: [10.1088/1742-5468/2004/06/P06002](https://doi.org/10.1088/1742-5468/2004/06/P06002). arXiv: [hep-th/0405152](https://arxiv.org/abs/hep-th/0405152).
- [195] Xi Dong and Aitor Lewkowycz. “Entropy, extremality, Euclidean variations, and the equations of motion”. In: *JHEP* 01 (2018), p. 081. DOI: [10.1007/JHEP01\(2018\)081](https://doi.org/10.1007/JHEP01(2018)081). arXiv: [1705.08453](https://arxiv.org/abs/1705.08453) [hep-th].
- [196] Donald Marolf and Henry Maxfield. “Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information”. In: *JHEP* 08 (2020), p. 044. DOI: [10.1007/JHEP08\(2020\)044](https://doi.org/10.1007/JHEP08(2020)044). arXiv: [2002.08950](https://arxiv.org/abs/2002.08950) [hep-th].
- [197] Jyotirmoy Chakravarty. “Overcounting of interior excitations: a resolution to the bags of gold paradox in AdS”. In: *JHEP* 02 (2021), p. 027. DOI: [10.1007/JHEP02\(2021\)027](https://doi.org/10.1007/JHEP02(2021)027). arXiv: [2010.03575](https://arxiv.org/abs/2010.03575) [hep-th].
- [198] Marios Christodoulou and Carlo Rovelli. “How big is a black hole?” In: *Phys. Rev. D* 91 (2015), p. 064046. DOI: [10.1103/PhysRevD.91.064046](https://doi.org/10.1103/PhysRevD.91.064046). arXiv: [1411.2854](https://arxiv.org/abs/1411.2854) [gr-qc].
- [199] Marios Christodoulou and Tommaso De Lorenzo. “Volume inside old black holes”. In: *Phys. Rev. D* 94 (2016), p. 104002. DOI: [10.1103/PhysRevD.94.104002](https://doi.org/10.1103/PhysRevD.94.104002). arXiv: [1604.07222](https://arxiv.org/abs/1604.07222) [gr-qc].
- [200] Jacob McNamara and Cumrun Vafa. *Baby universes, holography, and the swampland*. 2020. arXiv: [2004.06738](https://arxiv.org/abs/2004.06738) [hep-th].
- [201] Daniel L. Jafferis. *Bulk reconstruction and the Hartle-Hawking wavefunction*. 2017. arXiv: [1703.01519](https://arxiv.org/abs/1703.01519) [hep-th].

- [202] Thomas Faulkner and Aitor Lewkowycz. “Bulk locality from modular flow”. In: *JHEP* 07 (2017), p. 151. DOI: [10.1007/JHEP07\(2017\)151](https://doi.org/10.1007/JHEP07(2017)151). arXiv: [1704.05464](https://arxiv.org/abs/1704.05464) [[hep-th](#)].
- [203] Raphael Bousso and Geoff Penington. “Holograms in our world”. In: *Phys. Rev. D* 108 (2023), p. 046007. DOI: [10.1103/PhysRevD.108.046007](https://doi.org/10.1103/PhysRevD.108.046007). arXiv: [2302.07892](https://arxiv.org/abs/2302.07892) [[hep-th](#)].
- [204] Chitraang Murdia, Yasunori Nomura, and Kyle Ritchie. “Black hole and de Sitter microstructures from a semiclassical perspective”. In: *Phys. Rev. D* 107 (2023), p. 026016. DOI: [10.1103/PhysRevD.107.026016](https://doi.org/10.1103/PhysRevD.107.026016). arXiv: [2207.01625](https://arxiv.org/abs/2207.01625) [[hep-th](#)].
- [205] William K. Wootters and Wojciech H. Zurek. “A single quantum cannot be cloned”. In: *Nature* 299 (1982), p. 802. DOI: [10.1038/299802a0](https://doi.org/10.1038/299802a0).
- [206] Sidney Coleman. “Black holes as red herrings: topological fluctuations and the loss of quantum coherence”. In: *Nucl. Phys. B* 307 (1988), p. 867. DOI: [10.1016/0550-3213\(88\)90110-1](https://doi.org/10.1016/0550-3213(88)90110-1).
- [207] Steven B. Giddings and Andrew Strominger. “Loss of incoherence and determination of coupling constants in quantum gravity”. In: *Nucl. Phys. B* 307 (1988), p. 854. DOI: [10.1016/0550-3213\(88\)90109-5](https://doi.org/10.1016/0550-3213(88)90109-5).
- [208] Steven B. Giddings and Andrew Strominger. “Baby universes, third quantization and the cosmological constant”. In: *Nucl. Phys. B* 321 (1989), p. 481. DOI: [10.1016/0550-3213\(89\)90353-2](https://doi.org/10.1016/0550-3213(89)90353-2).
- [209] Donald Marolf and Henry Maxfield. “Observations of Hawking radiation: the Page curve and baby universes”. In: *JHEP* 04 (2021), p. 272. DOI: [10.1007/JHEP04\(2021\)272](https://doi.org/10.1007/JHEP04(2021)272). arXiv: [2010.06602](https://arxiv.org/abs/2010.06602) [[hep-th](#)].
- [210] Joseph Polchinski and Andrew Strominger. “A possible resolution of the black hole information puzzle”. In: *Phys. Rev. D* 50 (1994), p. 7403. DOI: [10.1103/PhysRevD.50.7403](https://doi.org/10.1103/PhysRevD.50.7403). arXiv: [hep-th/9407008](https://arxiv.org/abs/hep-th/9407008).
- [211] Oleg Lunin and Samir D. Mathur. “AdS/CFT duality and the black hole information paradox”. In: *Nucl. Phys. B* 623 (2002), p. 342. DOI: [10.1016/S0550-3213\(01\)00620-4](https://doi.org/10.1016/S0550-3213(01)00620-4). arXiv: [hep-th/0109154](https://arxiv.org/abs/hep-th/0109154).
- [212] Oleg Lunin, Juan Martin Maldacena, and Liat Maoz. *Gravity solutions for the D1-D5 system with angular momentum*. 2002. arXiv: [hep-th/0212210](https://arxiv.org/abs/hep-th/0212210).
- [213] Samir D. Mathur, Ashish Saxena, and Yogesh K. Srivastava. “Constructing ‘hair’ for the three charge hole”. In: *Nucl. Phys. B* 680 (2004), p. 415. DOI: [10.1016/j.nuclphysb.2003.12.022](https://doi.org/10.1016/j.nuclphysb.2003.12.022). arXiv: [hep-th/0311092](https://arxiv.org/abs/hep-th/0311092).

- [214] Ahmed Almheiri et al. “An apologia for firewalls”. In: *JHEP* 09 (2013), p. 018. DOI: [10.1007/JHEP09\(2013\)018](https://doi.org/10.1007/JHEP09(2013)018). arXiv: [1304.6483](https://arxiv.org/abs/1304.6483) [hep-th].
- [215] Donald Marolf and Joseph Polchinski. “Gauge/gravity duality and the black hole interior”. In: *Phys. Rev. Lett.* 111 (2013), p. 171301. DOI: [10.1103/PhysRevLett.111.171301](https://doi.org/10.1103/PhysRevLett.111.171301). arXiv: [1307.4706](https://arxiv.org/abs/1307.4706) [hep-th].
- [216] Raphael Bousso and Geoff Penington. “Entanglement wedges for gravitating regions”. In: *Phys. Rev. D* 107 (2023), p. 086002. DOI: [10.1103/PhysRevD.107.086002](https://doi.org/10.1103/PhysRevD.107.086002). arXiv: [2208.04993](https://arxiv.org/abs/2208.04993) [hep-th].
- [217] David A. Lowe et al. “Black hole complementarity versus locality”. In: *Phys. Rev. D* 52 (1995), p. 6997. DOI: [10.1103/PhysRevD.52.6997](https://doi.org/10.1103/PhysRevD.52.6997). arXiv: [hep-th/9506138](https://arxiv.org/abs/hep-th/9506138).
- [218] Steven B. Giddings. “Black hole information, unitarity, and nonlocality”. In: *Phys. Rev. D* 74 (2006), p. 106005. DOI: [10.1103/PhysRevD.74.106005](https://doi.org/10.1103/PhysRevD.74.106005). arXiv: [hep-th/0605196](https://arxiv.org/abs/hep-th/0605196).
- [219] Raphael Bousso and Arvin Shahbazi-Moghaddam. “Quantum singularities”. In: *Phys. Rev. D* 107.6 (2023), p. 066002. DOI: [10.1103/PhysRevD.107.066002](https://doi.org/10.1103/PhysRevD.107.066002). arXiv: [2206.07001](https://arxiv.org/abs/2206.07001) [hep-th].
- [220] Yasunori Nomura. “Quantum mechanics, spacetime locality, and gravity”. In: *Found. Phys.* 43 (2013), p. 978. DOI: [10.1007/s10701-013-9729-1](https://doi.org/10.1007/s10701-013-9729-1). arXiv: [1110.4630](https://arxiv.org/abs/1110.4630) [hep-th].
- [221] Harold Ollivier, David Poulin, and Wojciech H. Zurek. “Objective Properties from Subjective Quantum States: Environment as a Witness”. In: *Phys. Rev. Lett.* 93.22 (2004), p. 220401. DOI: [10.1103/PhysRevLett.93.220401](https://doi.org/10.1103/PhysRevLett.93.220401). arXiv: [quant-ph/0307229](https://arxiv.org/abs/quant-ph/0307229).
- [222] Robin Blume-Kohout and Wojciech H. Zurek. “Quantum Darwinism: Entanglement, branches, and the emergent classicality of redundantly stored quantum information”. In: *Phys. Rev. A* 73.6 (2006), p. 062310. DOI: [10.1103/PhysRevA.73.062310](https://doi.org/10.1103/PhysRevA.73.062310). arXiv: [quant-ph/0505031](https://arxiv.org/abs/quant-ph/0505031).
- [223] Raphael Bousso and Leonard Susskind. “The multiverse interpretation of quantum mechanics”. In: *Phys. Rev. D* 85 (2012), p. 045007. DOI: [10.1103/PhysRevD.85.045007](https://doi.org/10.1103/PhysRevD.85.045007). arXiv: [1105.3796](https://arxiv.org/abs/1105.3796) [hep-th].
- [224] Vijay Balasubramanian et al. “The entropy of a hole in spacetime”. In: *JHEP* 10 (2013), p. 220. DOI: [10.1007/JHEP10\(2013\)220](https://doi.org/10.1007/JHEP10(2013)220). arXiv: [1305.0856](https://arxiv.org/abs/1305.0856) [hep-th].

- [225] Stephen A. Fulling. “Nonuniqueness of canonical field quantization in Riemannian space-time”. In: *Phys. Rev. D* 7 (1973), p. 2850. DOI: [10.1103/PhysRevD.7.2850](https://doi.org/10.1103/PhysRevD.7.2850).
- [226] Paul C. W. Davies. “Scalar particle production in Schwarzschild and Rindler metrics”. In: *J. Phys. A* 8 (1975), p. 609. DOI: [10.1088/0305-4470/8/4/022](https://doi.org/10.1088/0305-4470/8/4/022).
- [227] Maulik Parikh and Andrew Svesko. “Einstein’s equations from the stretched future light cone”. In: *Phys. Rev. D* 98 (2018), p. 026018. DOI: [10.1103/PhysRevD.98.026018](https://doi.org/10.1103/PhysRevD.98.026018). arXiv: [1712.08475 \[hep-th\]](https://arxiv.org/abs/1712.08475).
- [228] Phil Saad, Stephen H. Shenker, and Douglas Stanford. “JT gravity as a matrix integral”. In: (Mar. 2019). arXiv: [1903.11115 \[hep-th\]](https://arxiv.org/abs/1903.11115).
- [229] Douglas Stanford and Edward Witten. “JT gravity and the ensembles of random matrix theory”. In: *Adv. Theor. Math. Phys.* 24.6 (2020), pp. 1475–1680. DOI: [10.4310/ATMP.2020.v24.n6.a4](https://doi.org/10.4310/ATMP.2020.v24.n6.a4). arXiv: [1907.03363 \[hep-th\]](https://arxiv.org/abs/1907.03363).
- [230] Yasunori Nomura. “The Static Quantum Multiverse”. In: *Phys. Rev. D* 86 (2012), p. 083505. DOI: [10.1103/PhysRevD.86.083505](https://doi.org/10.1103/PhysRevD.86.083505). arXiv: [1205.5550 \[hep-th\]](https://arxiv.org/abs/1205.5550).
- [231] Edward Witten. “Matrix Models and Deformations of JT Gravity”. In: *Proc. Roy. Soc. Lond. A* 476.2244 (2020), p. 20200582. DOI: [10.1098/rspa.2020.0582](https://doi.org/10.1098/rspa.2020.0582). arXiv: [2006.13414 \[hep-th\]](https://arxiv.org/abs/2006.13414).
- [232] Phil Saad, Stephen H. Shenker, and Douglas Stanford. “A semiclassical ramp in SYK and in gravity”. In: (June 2018). arXiv: [1806.06840 \[hep-th\]](https://arxiv.org/abs/1806.06840).
- [233] Hao Geng et al. “Inconsistency of islands in theories with long-range gravity”. In: *JHEP* 01 (2022), p. 182. DOI: [10.1007/JHEP01\(2022\)182](https://doi.org/10.1007/JHEP01(2022)182). arXiv: [2107.03390 \[hep-th\]](https://arxiv.org/abs/2107.03390).
- [234] Suvrat Raju. “Failure of the split property in gravity and the information paradox”. In: *Class. Quant. Grav.* 39.6 (2022), p. 064002. DOI: [10.1088/1361-6382/ac482b](https://doi.org/10.1088/1361-6382/ac482b). arXiv: [2110.05470 \[hep-th\]](https://arxiv.org/abs/2110.05470).
- [235] Xi Dong et al. “Effective entropy of quantum fields coupled with gravity”. In: *JHEP* 10 (2020), p. 052. DOI: [10.1007/JHEP10\(2020\)052](https://doi.org/10.1007/JHEP10(2020)052). arXiv: [2007.02987 \[hep-th\]](https://arxiv.org/abs/2007.02987).
- [236] Raphael Bousso and Sami Kaya. “Geometric quantum states beyond the AdS/CFT correspondence”. In: *Phys. Rev. D* 110.6 (2024), p. 066017. DOI: [10.1103/PhysRevD.110.066017](https://doi.org/10.1103/PhysRevD.110.066017). arXiv: [2404.11644 \[hep-th\]](https://arxiv.org/abs/2404.11644).
- [237] Raphael Bousso and Sami Kaya. “Holographic entropy cone beyond AdS/CFT”. In: *Phys. Rev. D* 111.8 (2025), p. 086014. DOI: [10.1103/PhysRevD.111.086014](https://doi.org/10.1103/PhysRevD.111.086014). arXiv: [2502.03516 \[hep-th\]](https://arxiv.org/abs/2502.03516).

- [238] Vijay Balasubramanian and Charlie Cummings. “The entropy of finite gravitating regions”. In: (Dec. 2023). arXiv: [2312.08434 \[hep-th\]](#).
- [239] Jennifer Lin. “Ryu-Takayanagi Area as an Entanglement Edge Term”. In: (Apr. 2017). arXiv: [1704.07763 \[hep-th\]](#).
- [240] Jennifer Lin. “Entanglement entropy in Jackiw-Teitelboim Gravity”. In: (July 2018). arXiv: [1807.06575 \[hep-th\]](#).
- [241] Daniel Louis Jafferis and David K. Kolchmeyer. “Entanglement Entropy in Jackiw-Teitelboim Gravity”. In: (Nov. 2019). arXiv: [1911.10663 \[hep-th\]](#).
- [242] Xi Dong, Xiao-Liang Qi, and Michael Walter. “Holographic entanglement negativity and replica symmetry breaking”. In: *JHEP* 06 (2021), p. 024. DOI: [10.1007/JHEP06\(2021\)024](#). arXiv: [2101.11029 \[hep-th\]](#).
- [243] Chris Akers et al. “Reflected entropy in random tensor networks”. In: *JHEP* 05 (2022), p. 162. DOI: [10.1007/JHEP05\(2022\)162](#). arXiv: [2112.09122 \[hep-th\]](#).
- [244] Chris Akers et al. “Reflected entropy in random tensor networks. Part II. A topological index from canonical purification”. In: *JHEP* 01 (2023), p. 067. DOI: [10.1007/JHEP01\(2023\)067](#). arXiv: [2210.15006 \[hep-th\]](#).
- [245] Chris Akers et al. “One-shot holography”. In: *SciPost Phys.* 16.6 (2024), p. 144. DOI: [10.21468/SciPostPhys.16.6.144](#). arXiv: [2307.13032 \[hep-th\]](#).
- [246] Chris Akers et al. “Reflected entropy in random tensor networks. Part III. Triway cuts”. In: *JHEP* 12 (2024), p. 209. DOI: [10.1007/JHEP12\(2024\)209](#). arXiv: [2409.17218 \[hep-th\]](#).
- [247] Chris Akers and Pratik Rath. “Holographic Renyi Entropy from Quantum Error Correction”. In: *JHEP* 05 (2019), p. 052. DOI: [10.1007/JHEP05\(2019\)052](#). arXiv: [1811.05171 \[hep-th\]](#).
- [248] Xi Dong, Daniel Harlow, and Donald Marolf. “Flat entanglement spectra in fixed-area states of quantum gravity”. In: *JHEP* 10 (2019), p. 240. DOI: [10.1007/JHEP10\(2019\)240](#). arXiv: [1811.05382 \[hep-th\]](#).
- [249] Xi Dong and Donald Marolf. “One-loop universality of holographic codes”. In: *JHEP* 03 (2020), p. 191. DOI: [10.1007/JHEP03\(2020\)191](#). arXiv: [1910.06329 \[hep-th\]](#).
- [250] Geoff Penington, Michael Walter, and Freek Witteveen. “Fun with replicas: tripartitions in tensor networks and gravity”. In: *JHEP* 05 (2023), p. 008. DOI: [10.1007/JHEP05\(2023\)008](#). arXiv: [2211.16045 \[hep-th\]](#).

- [251] Donald Marolf, Shannon Wang, and Zhencheng Wang. “Probing phase transitions of holographic entanglement entropy with fixed area states”. In: *JHEP* 12 (2020), p. 084. DOI: [10.1007/JHEP12\(2020\)084](https://doi.org/10.1007/JHEP12(2020)084). arXiv: [2006.10089](https://arxiv.org/abs/2006.10089) [hep-th].
- [252] Xi Dong, Jonah Kudler-Flam, and Pratik Rath. “A modified cosmic brane proposal for holographic Renyi entropy”. In: *JHEP* 06 (2024), p. 120. DOI: [10.1007/JHEP06\(2024\)120](https://doi.org/10.1007/JHEP06(2024)120). arXiv: [2312.04625](https://arxiv.org/abs/2312.04625) [hep-th].
- [253] Geoff Penington and Pratik Rath. “Diagonal Approximation for Holographic Rényi Entropies”. In: *Phys. Rev. Lett.* 134.16 (2025), p. 161501. DOI: [10.1103/PhysRevLett.134.161501](https://doi.org/10.1103/PhysRevLett.134.161501). arXiv: [2412.03670](https://arxiv.org/abs/2412.03670) [hep-th].
- [254] Xi Dong. “The Gravity Dual of Renyi Entropy”. In: *Nature Commun.* 7 (2016), p. 12472. DOI: [10.1038/ncomms12472](https://doi.org/10.1038/ncomms12472). arXiv: [1601.06788](https://arxiv.org/abs/1601.06788) [hep-th].
- [255] Aram W. Harrow. *The Church of the Symmetric Subspace*. 2013. arXiv: [1308.6595](https://arxiv.org/abs/1308.6595) [quant-ph].
- [256] Chris Akers et al. “On observers in holographic maps”. In: *JHEP* 05 (2025), p. 201. DOI: [10.1007/JHEP05\(2025\)201](https://doi.org/10.1007/JHEP05(2025)201). arXiv: [2503.09681](https://arxiv.org/abs/2503.09681) [hep-th].
- [257] Raphael Bousso and Sami Kaya. “Fundamental complement of a gravitating region”. In: *Gen. Rel. Grav.* 57.8 (2025), p. 124. DOI: [10.1007/s10714-025-03462-6](https://doi.org/10.1007/s10714-025-03462-6). arXiv: [2505.15886](https://arxiv.org/abs/2505.15886) [hep-th].
- [258] Chris Akers, Sergio Hernández-Cuenca, and Pratik Rath. “Quantum Extremal Surfaces and the Holographic Entropy Cone”. In: *JHEP* 11 (2021), p. 177. DOI: [10.1007/JHEP11\(2021\)177](https://doi.org/10.1007/JHEP11(2021)177). arXiv: [2108.07280](https://arxiv.org/abs/2108.07280) [hep-th].
- [259] Luca V. Iliesiu et al. “On the non-perturbative bulk Hilbert space of JT gravity”. In: *JHEP* 10 (2024), p. 220. DOI: [10.1007/JHEP10\(2024\)220](https://doi.org/10.1007/JHEP10(2024)220). arXiv: [2403.08696](https://arxiv.org/abs/2403.08696) [hep-th].
- [260] Jesse Held et al. “Euclidean and complex geometries from real-time computations of gravitational Rényi entropies”. In: *JHEP* 02 (2025), p. 136. DOI: [10.1007/JHEP02\(2025\)136](https://doi.org/10.1007/JHEP02(2025)136). arXiv: [2409.17428](https://arxiv.org/abs/2409.17428) [hep-th].
- [261] Xi Dong, Donald Marolf, and Pratik Rath. “Constrained HRT Surfaces and their Entropic Interpretation”. In: *JHEP* 02 (2024), p. 151. DOI: [10.1007/JHEP02\(2024\)151](https://doi.org/10.1007/JHEP02(2024)151). arXiv: [2311.18290](https://arxiv.org/abs/2311.18290) [hep-th].
- [262] Victor Ivo, Yue-Zhou Li, and Juan Maldacena. “The no boundary density matrix”. In: *JHEP* 02 (2025), p. 124. DOI: [10.1007/JHEP02\(2025\)124](https://doi.org/10.1007/JHEP02(2025)124). arXiv: [2409.14218](https://arxiv.org/abs/2409.14218) [hep-th].

- [263] Daniel Harlow and Daniel Jafferis. “The Factorization Problem in Jackiw-Teitelboim Gravity”. In: *JHEP* 02 (2020), p. 177. DOI: [10.1007/JHEP02\(2020\)177](https://doi.org/10.1007/JHEP02(2020)177). arXiv: [1804.01081](https://arxiv.org/abs/1804.01081) [hep-th].
- [264] Wan Zhen Chua and Thomas Hartman. “Black hole wavefunctions and microcanonical states”. In: *JHEP* 06 (2024), p. 054. DOI: [10.1007/JHEP06\(2024\)054](https://doi.org/10.1007/JHEP06(2024)054). arXiv: [2309.05041](https://arxiv.org/abs/2309.05041) [hep-th].
- [265] Hewei Frederic Jia and Mukund Rangamani. “Petz reconstruction in random tensor networks”. In: *JHEP* 10 (2020), p. 050. DOI: [10.1007/JHEP10\(2020\)050](https://doi.org/10.1007/JHEP10(2020)050). arXiv: [2006.12601](https://arxiv.org/abs/2006.12601) [hep-th].
- [266] Sami Kaya, Pratik Rath, and Kyle Ritchie. “To appear”. In: (2026).
- [267] Kristan Jensen, Jonathan Sorce, and Antony J. Speranza. “Generalized Entropy for General Subregions in Quantum Gravity”. In: *Journal of High Energy Physics* 2023.12 (2023), p. 020. DOI: [10.1007/JHEP12\(2023\)020](https://doi.org/10.1007/JHEP12(2023)020). arXiv: [2306.01837](https://arxiv.org/abs/2306.01837) [hep-th].
- [268] Sami Kaya, Pratik Rath, and Kyle Ritchie. “Hollow-grams: generalized entanglement wedges from the gravitational path integral”. In: *JHEP* 09 (2025), p. 032. DOI: [10.1007/JHEP09\(2025\)032](https://doi.org/10.1007/JHEP09(2025)032). arXiv: [2506.10064](https://arxiv.org/abs/2506.10064) [hep-th].
- [269] Arvin Shahbazi-Moghaddam. “Restricted quantum focusing”. In: *Phys. Rev. D* 109.6 (2024), p. 066023. DOI: [10.1103/PhysRevD.109.066023](https://doi.org/10.1103/PhysRevD.109.066023). arXiv: [2212.03881](https://arxiv.org/abs/2212.03881) [hep-th].
- [270] Douglas Stanford and Leonard Susskind. “Complexity and Shock Wave Geometries”. In: *Phys. Rev. D* 90.12 (2014), p. 126007. DOI: [10.1103/PhysRevD.90.126007](https://doi.org/10.1103/PhysRevD.90.126007). arXiv: [1406.2678](https://arxiv.org/abs/1406.2678) [hep-th].
- [271] Cesar A. Agón, Matthew Headrick, and Brian Swingle. “Subsystem Complexity and Holography”. In: *JHEP* 02 (2019), p. 145. DOI: [10.1007/JHEP02\(2019\)145](https://doi.org/10.1007/JHEP02(2019)145). arXiv: [1804.01561](https://arxiv.org/abs/1804.01561) [hep-th].
- [272] Josiah Couch, Willy Fischler, and Phuc H. Nguyen. “Noether charge, black hole volume, and complexity”. In: *JHEP* 03 (2017), p. 119. DOI: [10.1007/JHEP03\(2017\)119](https://doi.org/10.1007/JHEP03(2017)119). arXiv: [1610.02038](https://arxiv.org/abs/1610.02038) [hep-th].
- [273] Juan F. Pedraza et al. “Sewing spacetime with Lorentzian threads: complexity and the emergence of time in quantum gravity”. In: *JHEP* 02 (2022), p. 093. DOI: [10.1007/JHEP02\(2022\)093](https://doi.org/10.1007/JHEP02(2022)093). arXiv: [2106.12585](https://arxiv.org/abs/2106.12585) [hep-th].

- [274] Robert C. Myers and Shan-Ming Ruan. “Complexity Equals (Almost) Anything”. In: *Gravity, Strings and Fields: A Conference in Honour of Gordon Semenoff*. Mar. 2024. DOI: [10.1007/978-3-031-91266-5_11](https://doi.org/10.1007/978-3-031-91266-5_11). arXiv: [2403.17475](https://arxiv.org/abs/2403.17475) [hep-th].
- [275] Roberto Emparan et al. “Holographic complexity of quantum black holes”. In: *JHEP* 02 (2022), p. 204. DOI: [10.1007/JHEP02\(2022\)204](https://doi.org/10.1007/JHEP02(2022)204). arXiv: [2112.04860](https://arxiv.org/abs/2112.04860) [hep-th].
- [276] Abdulrahim Al Balushi et al. “Holographic complexity of rotating black holes”. In: *JHEP* 05 (2021), p. 226. DOI: [10.1007/JHEP05\(2021\)226](https://doi.org/10.1007/JHEP05(2021)226). arXiv: [2010.11203](https://arxiv.org/abs/2010.11203) [hep-th].
- [277] Violet Concepcion et al. “Complexity for an Evaporating Black Hole”. In: *to appear* (2026).
- [278] Mark R. Dowling and Michael A. Nielsen. “The geometry of quantum computation”. In: *Quant. Inf. Comput.* 8.10 (2008), pp. 0861–0899. DOI: [10.26421/QIC8.10-1](https://doi.org/10.26421/QIC8.10-1). arXiv: [quant-ph/0701004](https://arxiv.org/abs/quant-ph/0701004).
- [279] Ro Jefferson and Robert C. Myers. “Circuit complexity in quantum field theory”. In: *JHEP* 10 (2017), p. 107. DOI: [10.1007/JHEP10\(2017\)107](https://doi.org/10.1007/JHEP10(2017)107). arXiv: [1707.08570](https://arxiv.org/abs/1707.08570) [hep-th].
- [280] Shira Chapman et al. “Toward a Definition of Complexity for Quantum Field Theory States”. In: *Phys. Rev. Lett.* 120.12 (2018), p. 121602. DOI: [10.1103/PhysRevLett.120.121602](https://doi.org/10.1103/PhysRevLett.120.121602). arXiv: [1707.08582](https://arxiv.org/abs/1707.08582) [hep-th].
- [281] Vijay Balasubramanian et al. “Quantum chaos and the complexity of spread of states”. In: *Phys. Rev. D* 106.4 (2022), p. 046007. DOI: [10.1103/PhysRevD.106.046007](https://doi.org/10.1103/PhysRevD.106.046007). arXiv: [2202.06957](https://arxiv.org/abs/2202.06957) [hep-th].
- [282] Arpan Bhattacharyya, Arvind Shekar, and Aninda Sinha. “Circuit complexity in interacting QFTs and RG flows”. In: *JHEP* 10 (2018), p. 140. DOI: [10.1007/JHEP10\(2018\)140](https://doi.org/10.1007/JHEP10(2018)140). arXiv: [1808.03105](https://arxiv.org/abs/1808.03105) [hep-th].
- [283] Hugo A. Camargo et al. “Path integral optimization as circuit complexity”. In: *Phys. Rev. Lett.* 123.1 (2019), p. 011601. DOI: [10.1103/PhysRevLett.123.011601](https://doi.org/10.1103/PhysRevLett.123.011601). arXiv: [1904.02713](https://arxiv.org/abs/1904.02713) [hep-th].
- [284] Bowen Chen, Bartłomiej Czech, and Zi-zhi Wang. “Query complexity and cutoff dependence of the CFT₂ ground state”. In: *Phys. Rev. D* 103.2 (2021), p. 026015. DOI: [10.1103/PhysRevD.103.026015](https://doi.org/10.1103/PhysRevD.103.026015). arXiv: [2004.11377](https://arxiv.org/abs/2004.11377) [hep-th].

- [285] Kiran Adhikari, Sayantan Choudhury, and Abhishek Roy. “Krylov Complexity in Quantum Field Theory”. In: *Nucl. Phys. B* 993 (2023), p. 116263. DOI: [10.1016/j.nuclphysb.2023.116263](#). arXiv: [2204.02250 \[hep-th\]](#).
- [286] Robert M. Wald. *General Relativity*. Chicago, USA: Chicago Univ. Pr., 1984. DOI: [10.7208/chicago/9780226870373.001.0001](#).
- [287] H. Epstein, V. Glaser, and A. Jaffe. “Nonpositivity of energy density in Quantized field theories”. In: *Nuovo Cim.* 36 (1965), p. 1016. DOI: [10.1007/BF02749799](#).
- [288] Christopher J. Fewster and Eleni-Alexandra Kontou. “Quantum strong energy inequalities”. In: *Phys. Rev. D* 99.4 (2019), p. 045001. DOI: [10.1103/PhysRevD.99.045001](#). arXiv: [1809.05047 \[gr-qc\]](#).
- [289] E. H. Lieb and M. B. Ruskai. “Proof of the strong subadditivity of quantum-mechanical entropy”. In: *J. Math. Phys.* 14 (1973), pp. 1938–1941. DOI: [10.1063/1.1666274](#).
- [290] Jan Boruch, Pawel Caputa, and Tadashi Takayanagi. “Path-Integral Optimization from Hartle-Hawking Wave Function”. In: *Phys. Rev. D* 103.4 (2021), p. 046017. DOI: [10.1103/PhysRevD.103.046017](#). arXiv: [2011.08188 \[hep-th\]](#).
- [291] Jan Boruch et al. “Holographic path-integral optimization”. In: *JHEP* 07 (2021). [Erratum: *JHEP* 09, 111 (2022)], p. 016. DOI: [10.1007/JHEP07\(2021\)016](#).
- [292] Alexandre Belin et al. “Complexity equals anything II”. In: *JHEP* 01 (2023), p. 154. DOI: [10.1007/JHEP01\(2023\)154](#). arXiv: [2210.09647 \[hep-th\]](#).
- [293] Elena Caceres et al. “Complexity of Mixed States in QFT and Holography”. In: *JHEP* 03 (2020), p. 012. DOI: [10.1007/JHEP03\(2020\)012](#). arXiv: [1909.10557 \[hep-th\]](#).
- [294] Elena Cáceres et al. “Holographic Purification Complexity”. In: *Phys. Rev. D* 99.8 (2019), p. 086016. DOI: [10.1103/PhysRevD.99.086016](#). arXiv: [1811.10650 \[hep-th\]](#).
- [295] Jeff Murugan and Hendrik J. R. van Zyl. “Superadditivity of Krylov Complexity for Tensor Products”. In: (Jan. 2026). arXiv: [2601.08723 \[hep-th\]](#).
- [296] Roberto Auzzi et al. “On subregion action complexity in AdS_3 and in the BTZ black hole”. In: *JHEP* 01 (2020), p. 066. DOI: [10.1007/JHEP01\(2020\)066](#). arXiv: [1910.00526 \[hep-th\]](#).
- [297] Elena Cáceres, Rafael Carrasco, and Juan F. Pedraza. “Lorentzian threads and non-local computation in holography”. In: (Dec. 2025). arXiv: [2512.07963 \[hep-th\]](#).

- [298] Tadahito Nakajima. “Holographic Subregion Complexity and Fidelity Susceptibility in Noncommutative Yang–Mills Theory”. In: (Feb. 2026). arXiv: [2602.14448 \[hep-th\]](#).
- [299] Seth Lloyd. “Ultimate physical limits to computation”. In: *Nature* 406 (2000), pp. 1047–1054. DOI: [10.1038/35023282](#). arXiv: [quant-ph/9908043](#).
- [300] Christopher M. Dawson and Michael A. Nielsen. “The Solovay-Kitaev algorithm”. In: *Quant. Inf. Comput.* 6 (2006), pp. 081–095. DOI: [10.26421/QIC6.1-6](#). arXiv: [quant-ph/0505030](#).
- [301] Michal Oszmaniec et al. “Saturation and Recurrence of Quantum Complexity in Random Local Quantum Dynamics”. In: *Phys. Rev. X* 14.4 (2024), p. 041068. DOI: [10.1103/PhysRevX.14.041068](#). arXiv: [2205.09734 \[quant-ph\]](#).
- [302] Adam R. Brown and Leonard Susskind. “The Second Law of Quantum Complexity”. In: *Physical Review D* 97.8 (2018), p. 086015. DOI: [10.1103/PhysRevD.97.086015](#). arXiv: [1701.01107 \[hep-th\]](#).
- [303] Ying Zhao. “Uncomplexity and Black Hole Geometry”. In: *Physical Review D* 97.12 (2018), p. 126007. DOI: [10.1103/PhysRevD.97.126007](#). arXiv: [1711.03125 \[hep-th\]](#).