

QUANTUM ASPECTS OF BLACK HOLES: THE BAGS OF GOLD AND MONOGAMY PARADOXES

A thesis
submitted to the



Tata Institute of Fundamental Research
Mumbai, India - 400005

for the degree of

Doctor of Philosophy
in Physics

by

Joydeep Chakravarty

(Advisor: Suvrat Raju)



International Centre for Theoretical Sciences
Tata Institute of Fundamental Research
Bengaluru, India - 560089

May, 2022

(Final Version Submitted in September, 2022)

DECLARATION

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort has been made to indicate this clearly, with due reference to literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Prof. Suvrat Raju, at the International Center for Theoretical Sciences of the Tata Institute for Fundamental Research, Bengaluru.



Joydeep Chakravarty

Date: 13/09/2022

In my capacity as the supervisor of Joydeep's thesis, I certify that the above statements are true to the best of my knowledge.



Prof. Suvrat Raju

Date: 13/09/2022

List of publications

Publications relevant to the thesis:

- **Overcounting of interior excitations: A resolution to the bags of gold paradox in AdS** [1]
Joydeep Chakravarty
[arXiv: 2010.03575 \[hep-th\]](#), [JHEP 02, 027 \(2021\)](#)
- **Monogamy paradox: A toy model in flat space** [2]
T. Chakraborty, Joydeep Chakravarty, P. Paul
[arXiv: 2107.06919 \[hep-th\]](#)

Other publications:

- **Nonlinear Langevin dynamics via holography** [3]
B. Chakraborty, Joydeep Chakravarty, S. Chaudhuri, C. Jana, R. Loganayagam, A. Sivakumar
[arXiv: 1906.07762 \[hep-th\]](#), [JHEP 01, 165 \(2020\)](#)
- **Critical exponents for higher order phase transitions: Landau theory and RG flow** [4]
Joydeep Chakravarty, D. Jain
[arXiv: 2102.08398 \[cond-mat.stat-mech\]](#), [J. Stat. Mech. 09, 3204 \(2021\)](#)
- **Small corrections to semiclassical gravity and their role in information paradoxes** [5]
Joydeep Chakravarty
[arXiv: 2105.09924 \[gr-qc\]](#)
- **Normalization of D instanton amplitudes in two dimensional type 0B string theory** [6]
Joydeep Chakravarty, A. Sen
[arXiv: 2207.07138 \[hep-th\]](#)
- **On the positivity of Coon amplitude in $D=4$** [7]
Joydeep Chakravarty, P. Maity, A. Mishra
[arXiv: 2208.02735 \[hep-th\]](#)

Acknowledgements

आचार्यात् पादमादत्ते पादं शिष्यः स्वमेधया ।
सब्रह्मचारिभ्यः पादं पादं कालक्रमेण च ॥

Translation: One gains quarter of their knowledge from their acharyas (mentors), quarter from their own self-study and intellect, quarter from their peers and the remaining quarter is gained with time.

Regarding mentors, I am deeply grateful to Suvrat for his help and guidance during my graduate studies, for valuable suggestions and assistance which have shaped my research career till date, and which will keep shaping the same in the future. I am also deeply grateful to Loga from whom I have learnt a lot about various aspects of physics, and for various important discussions that shaped my research philosophy. Both of them, through different means, encouraged me to develop a better understanding of physics from various viewpoints, and use the same to analyse new developments and explore new connections. I am also thankful to Spenta for sharing his deep physical insights and valuable assistance, and also to Ashoke for teaching me various aspects of string theory.

I am also indebted to my peers, some of whom are my collaborators as well. I have learnt many different aspects, both academic and non-academic from them. In particular I deeply thank Diksha and Akhil, many discussions with whom have enriched my understanding of high energy physics, and from whom I have probably learnt the most among my peers. Similarly, I am also highly grateful to Chandan, Arnab, Tuneer, Priyadarshi, Omkar and Victor, each of whom have helped me understand different strands of string theory and quantum field theory. I also thank Varun, Manisha, Kohinoor, Santhosh, Basudeb, Parita, Sowrabh, Kasi, Sugan, Prashant, Ajit, Srikanth, Uddepta, Bhanu, Souvik, Shivam, Pronobesh, Siddharth, Chandramouli, Amiya and Aditya; and other members of the string group and the football group for various academic and non-academic discussions. I sincerely apologize if I have missed out on any names due to my short-sighted memory.

I thank my parents for always encouraging me to pursue higher education inspite of our humble situation early on, which they still consider as a duty. I also thank Diksha for her generosity to help at any point of time, and being a constant source of excitement in life. I am also grateful to my undergraduate friends Abhinav and Love, whose impact on my outlook is way too "based" to be described, as well as Bhati, Akshay and Takki. Lastly, I acknowledge gratitude to the people of India for their steady and generous support to research in basic sciences.

Abstract

In this work, we address some puzzles about the black hole interior from the bulk perspective. In the first part, we discuss how black holes in AdS can host an enormous number of semiclassical excitations in their interior, which is seemingly not reflected in the Bekenstein-Hawking entropy. In addition to the paradox in the entropy, we argue that the treatment of such excitations using effective field theory also violates black holes' expected spectral properties. We propose that these mysteries are resolved because apparently orthogonal semiclassical bulk excitations have small inner products between them; and consequently, a vast number of semiclassical excitations can be constructed using the Hilbert space which describes black hole's interior. Further we verify our resolution using consistency checks using the dual CFT. In the second part, we discuss a toy model in empty flat space which captures the essential features of the monogamy paradox for old flat space black holes within a clean calculational framework. The chief non-trivialities here are a consequence of flat space gravity having a unique vacuum and infrared structure, say in contrast to AdS. We formulate the paradox in terms of monogamy of CHSH correlations, which we use to quantify the monogamy of entanglement, and use it to write a sharp statement of the violation of the monogamy of entanglement. We argue that the resolution of the paradox is that the Gauss constraint is not properly taken into account while posing the monogamy paradox.

Contents

1	Introduction	1
1.1	The bags of gold paradox	3
1.1.1	The paradox	3
1.1.2	Proposed Resolution	5
1.2	Monogamy paradox in flat space	6
1.2.1	The paradox	6
1.2.2	The toy model in flat space	7
1.2.3	Proposed resolution	9
1.3	Organization of the thesis	9
2	Brief review of relevant topics	11
2.1	QFT in curved spacetime	11
2.1.1	Bogoliubov transformations	12
2.1.2	Rindler / Unruh modes and Bogoliubov coefficients	14
2.1.3	Bogoliubov coefficients from the universality of two-point function	18
2.2	Black holes	19
2.2.1	Schwarzschild black hole	19
2.2.2	Penrose diagrams	21
2.2.3	Astrophysical black holes	22
2.3	Focussing, light rays and black hole horizons	23
2.3.1	Lie transport	23
2.3.2	Raychaudhuri equation	25
2.3.3	Key properties of black hole horizons	27
2.4	Laws of black hole thermodynamics	29
2.4.1	The area Law	29
2.4.2	First law of black hole thermodynamics	29
2.4.3	Zeroth and the third laws	31
2.5	Black hole evaporation and Hawking's paradox	32
2.5.1	Hawking radiation	33
2.5.2	Hawking's information paradox	34

2.6	Statistical properties of typical states	35
2.7	The cloning and strong subadditivity paradoxes	37
2.7.1	Cloning paradox	37
2.7.2	Strong subadditivity paradox	39
2.7.3	Resolutions	40
2.8	Black hole complementarity	41
2.8.1	Thought experiments involving complementarity	41
2.8.2	Loss of locality in high-point correlators	43
2.9	Interior operators in the state-dependent formalism	44
3	The bags of gold paradox	47
3.1	The bags of gold paradox for the eternal black hole	48
3.1.1	Maximum volume slices in the interior	49
3.1.2	Placing semiclassical excitations on the nice slice	50
3.1.3	The paradox in the bulk	52
3.2	The resolution: Overestimation of the Hilbert space's dimensionality	52
3.2.1	How many bulk excitations can we possibly have?	55
3.3	Resolution of the paradox from the boundary perspective	57
3.4	The nature of the excitations and the initial bulk wavefunction	60
3.5	The paradox for single sided black holes	62
3.5.1	The single-sided paradox and its resolution	63
3.5.2	Why interior bulk states are non-orthogonal in the CFT Hilbert space?	64
3.6	Spectral properties of bags of gold spacetimes: Contradictions and Resolution	65
3.6.1	Spectral observables in random matrix theory and discrete systems	66
3.6.2	Spectral properties of bags of gold excitations	68
3.6.3	Resolution of spectral puzzles using overcounting	72
3.7	Study of the paradox using toy matrix models	73
3.7.1	Toy Model I: A $(0 + 1)$ -d two matrix model	74
3.7.2	Toy Model II: A $(3 + 1)$ -d CFT on $S_3 \times \mathbb{R}$	78
3.8	General properties of systems with overcounted Hilbert spaces	81
4	Monogamy paradox in flat space	85
4.1	Outline of our work	85
4.2	CHSH inequalities in quantum mechanics	85
4.2.1	CHSH operator and monogamy of entanglement	85
4.2.2	Baby example: Bell operators using simple harmonic oscillators	86
4.3	CHSH inequalities in local quantum field theory	87
4.3.1	Basic conventions and choice of operators	88

4.3.2	Vacuum projector and the most general two-point correlator	89
4.3.3	CHSH correlation between regions A and B in field theory	90
4.3.4	Summary of this section	95
4.4	The monogamy paradox in flat space	95
4.4.1	Gravity in asymptotically flat spacetime	96
4.4.2	CHSH correlation between regions A and C	99
4.4.3	The paradox and generalization to higher dimensions	106
4.4.4	Resolution of the paradox	107
5	Summary and discussion	109
5.1	The bags of gold paradox	109
5.2	The monogamy paradox	112
6	Appendices	115
6.1	Explicit examples of overcounting in small vector spaces	115
6.2	Partition function of the $U(N)$ two-oscillator model	117
6.3	Maximum volume slices for the AdS black hole	121
6.4	Technique used to accommodate vectors on the unit sphere in Hilbert space	125
6.5	Projectors onto smeared modes' vacua	126
6.6	Explicit commutator of smeared Rindler modes	128
6.7	Computation of $\langle \mathcal{G} \rangle$	128
6.8	Bogoliubov coefficients and $\langle C_{AB} \rangle \geq 2$	130
6.9	Proof of $\langle A_j C_i \rangle = \langle A_j B_i \rangle + O(\sqrt{G_N})$ and boundedness of C_i of §4.4.2 .	133
6.10	Proof of existence of C_i of §4.4.2	135
6.10.1	Positive Fourier mode reconstruction	137

Chapter 1

Introduction

A major, if not the most significant, development in 20th-century physics has been the development of the formalism of quantum field theory, which is the adequate formalism to study different aspects of high energy physics. In particular, a significant revelation in the development of this formalism was the advent of effective field theory, which enables us to describe physics using a description conveniently (e.g., a path integral description) that is valid in a chosen energy range.

Another significant development in 20th-century physics was discovering and further studying the general theory of relativity, which works very nicely as a classical theory. However, providing a consistent formulation wherein the classical general relativity theory has an underlying quantum mechanical description has proven tricky, with only a few solid promising avenues. There are many issues related to quantum gravity, and some of them are apparent at the level of quantum field theory on background spacetime itself. An important subclass of these problems arises when the background spacetime is a black hole.

In this regard, although originating from a different perspective of bootstrapping the S-matrix, string theory emerged as a strong candidate that describes gravity and low energy effective field theory in terms of underlying strings and other extended objects. In addition, string theory has provided insights into intricate relationships between entirely different theories, referred to as dualities.

Despite the tremendous progress that string theory has offered, we cannot completely understand the consequences of a stringy description when formulated in arbitrary backgrounds. Nevertheless, we can exploit the resulting dualities and other state-of-the-art tools better to understand traditional problems and newer aspects of quantum gravity.

In this regard, AdS/CFT [8–10] has provided us a robust framework to decipher features of black holes and quantum gravity in AdS in terms of boundary non-gravitational observables. The hope is that many of the issues arising in quantum gravity will be clarified with progress in fundamental aspects of string theory.

Regarding quantum gravity, black hole information paradoxes [11–13] have tradition-

ally served as beacons in the dark regarding physicists' quest to understand quantum aspects of black holes. Such paradoxes constitute an inter-related web of puzzles that arise due to the existence of the event horizon. Some important works discussing the black hole interior are [14–33].

In particular, the study of paradoxes related to the black hole interior is a promising avenue for understanding quantum mechanics in the black hole interior and for a precise realization of how quantum information is encoded in gravity. With this motivation, the chief objective of this work is geared towards understanding certain aspects of quantum gravity, guided by insights from the AdS/CFT correspondence and canonical gravity. In this light, the central theme of this thesis is to pose two such paradoxes: the bags of gold paradox [34] and the monogamy paradox [12, 35] using bulk effective field theory, and to precisely understand what goes wrong while using such a description. Much of the motivation and intuition behind such modeling comes from the setup of AdS/CFT. These paradoxes serve as valuable frames of reference illustrating the interwoven web of mysteries regarding the black hole's interior.

A very brief overview

We will briefly summarize these aspects now, which we will motivate and further delve into in the rest of the introduction. In [36], we investigated how single-sided and eternal black holes in AdS can host an enormous number of semiclassical excitations in their interior, naively leading to an entropy much larger than the Bekenstein-Hawking entropy. This paradox in entropy is known as *the bags of gold paradox* and is visually represented in Fig. 1.1. We argued that the EFT treatment of such excitations also violates the spectral properties expected of black holes.

Using state-dependent maps [33] from the dual boundary theory to the bulk gravitational theory, we concluded that these mysteries are resolved because seemingly orthogonal EFT bulk excitations have small inner products between them. Consequently, we can construct many excitations using the EFT Hilbert space describing the black hole's interior, giving rise to a far greater entropy than the Bekenstein-Hawking entropy. As a test, we also demonstrated our proposed resolution in the context of small N toy matrix models.

Apart from the bags of gold paradox, which small corrections to the inner products can potentially cure, we also studied a paradox that cannot be resolved using small gravitational corrections (as in the case of small inner products) but requires corrections of order one. In [37], we investigated a toy model in empty flat space, which captures essential features of the *monogamy of entanglement paradox* [12, 20] for old flat space black holes within a clean calculational framework. We formulated the paradox in terms of the monogamy of Bell correlations, which we used to quantify the monogamy of entanglement.

Within EFT, we showed that the entanglement of smeared bulk modes just outside a light cone with modes just inside the light cone and with modes situated at the past of future null infinity gives rise to an $O(1)$ violation in the monogamy of entanglement. These regions mentioned above are denoted in Fig. 1.2. We argued that the resolution of the paradox is that spatially separated observables probe the same underlying degrees of freedom, i.e., such observables act on a non-factorized Hilbert space.

In our work, we will see that studying puzzles regarding the black hole's interior allows us to better understand specific issues related to gravity. We hope that a study of such inconsistencies leads to an enhanced understanding of quantum aspects of gravity and quantum mechanics of black holes.

In the later part of this chapter, we will describe the paradoxes first, stating the underlying motivations and the problem statements. We will also briefly motivate our resolutions to the paradoxes while not going too deeply into technical aspects, so that reader can skim through this trailer and expect what lies ahead in the rest of the work. Later on, we will give a technical overview of some of the pre-requisites in §2 and analyze the paradoxes in significantly more detail afterward in §3 and §4.

1.1 The bags of gold paradox

The Bekenstein-Hawking entropy is a coarse-grained thermodynamic measure that states that the entropy of the black hole is proportional to its area [38, 39]. It tells us that a microscopic description of the black hole exists, with the number of the constituent microstates being the exponential of the entropy.

$$S_{BH} = \frac{A}{4} \tag{1.1}$$

A thorough understanding of black hole microstates' features is a fundamental question in itself with important implications for quantum gravity.

1.1.1 The paradox

We will briefly discuss the paradox now, which is essentially a question about the correct counting of these microstates. Specific spacelike slices which go inside the black hole interior become very large in volume for a choice of boundary time. Therefore these slices can host a considerable number of semiclassical excitations far higher than what the Bekenstein-Hawking entropy suggests, which leads to the paradox. We select these excitations such that they live far apart from each other on the Cauchy slices, thereby having zero spatial overlaps. The central question raised by the paradox is: How do we understand these states in the interior, given that they are seemingly not reflected in the

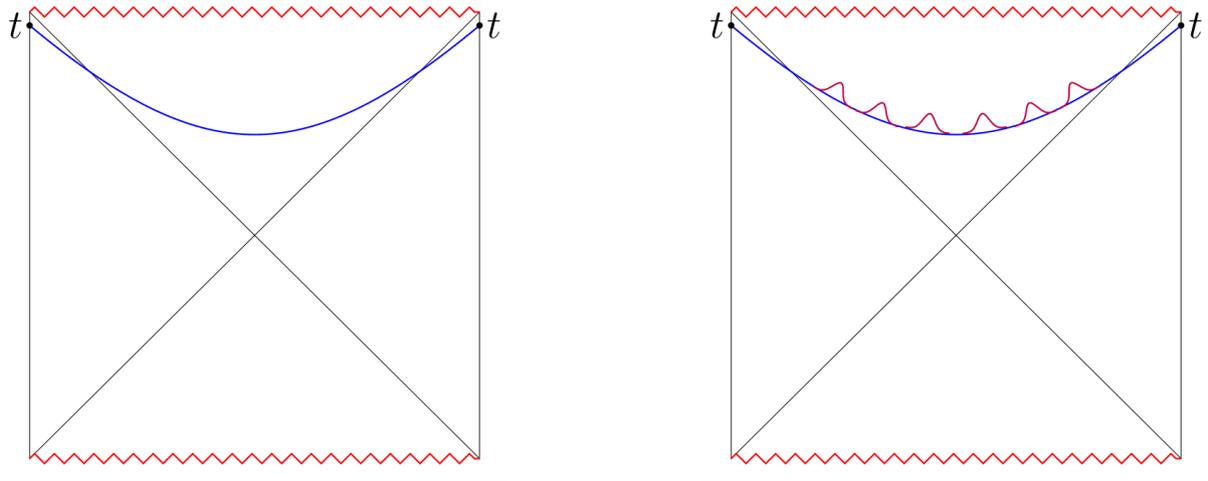


Figure 1.1: The left figure displays the maximal volume slices for the eternal AdS Schwarzschild black hole. The volumes of these slices increasingly grow with boundary time t thereby becoming very large at late times. The right figure demonstrates the bags of gold paradox for the eternal black hole on the maximal volume slices. We can accommodate an increasingly vast number of excitations placed far apart from each other on these slices which leads to the paradox.

Bekenstein-Hawking entropy?

In our analysis, we work with black holes in AdS, where we formulate the paradox on "nice spacelike slices" of AdS black holes. These slices stay away from singularities and significant curvature invariants everywhere. We pose the bags of gold problem in this spacetime, allowing us to utilize the AdS-CFT machinery to dissect the problem. We consider the eternal black hole [40] first, where we will demonstrate the paradox to its greatest extent by considering slices which possess the largest volumes for a given boundary. Maximizing the spacelike volume for a given value of the boundary time constructs the aforesaid nice slice. The salient feature of such a slice is that its volume in the interior becomes increasingly large as the boundary time grows. Consequently, at late times we have slices with gigantic volumes. On these late time slices, we will fit in a high number of semiclassical bulk excitations placed spatially far apart from each other such that they have zero spatial overlap, and consequently are independent of each other. The number of such excitations is much more extensive than what is stated by the Bekenstein-Hawking entropy, which leads to our paradox. Figure 1.1 displays the physical picture of the paradox. We are thus led to the question: Given that the Bekenstein-Hawking entropy is the area divided by 4, how do we account for the ever-increasing number of bulk excitations? Stated differently, does the entropy in equation (1.1) correctly count all these excitations or not?

In addition to the standard formulation of the bags of gold paradox as described above, we also argue that the effective field-theoretic description of the bags of gold excitations is inconsistent with the late time description of black holes using random matrix ensembles [19, 22, 41–49]. We will study spectral observables such as the energy

level spacing distribution and the spectral form factor, which we expect to behave in specific fashions for Gaussian unitary ensembles [41, 50–52]. We will show that an EFT description of bags of gold excitations will violate these observables’ expected features either qualitatively or quantitatively or in both fashions, thus leading to inconsistencies.

1.1.2 Proposed Resolution

Our proposed resolution to the above paradoxes is that we have tremendously overcounted the bulk states in the interior. Semiclassical bulk states placed far apart from each other in the interior are seemingly orthogonal. However, these states have small and significant inner products between them, which deviates from the semiclassical expectation of zero inner products. This is because in gravity, two coherent states corresponding to even vastly different classical configurations have a small non-vanishing inner product. In other theories such as a theory of electrodynamic interactions, two such coherent states can have a vanishingly small inner product. In contrast, the inner product between coherent states in gravity does not go to zero but saturates to a number that is larger than $O(e^{-\frac{S}{2}})$. The non-vanishing of inner products between two sufficiently distinct coherent states is the primary reason leading to overcounting.

More generally, we will show that the maximum number of vectors with small inner products that can be accommodated in a Hilbert space is exponentially larger than the dimension of the Hilbert space. This kinematical statement justifies the existence of an enormous number of interior bulk excitations leading to our paradox. As an example, if the bulk Hilbert space’s true dimensionality is e^S and the inner products between bulk excitations are of order $e^{-\frac{S}{4}}$, i.e. greater than $O(e^{-\frac{S}{2}})$; then the maximum number of bulk excitations (m) with such small inner products is a vast number given by ¹:

$$m \sim e^S \times \exp\left\{\frac{e^{\frac{S}{2}}}{2}\right\}. \quad (1.2)$$

If we consider even a small system with dimension $e^S = 3.6 \times 10^5$ with inner products of the order $e^{-\frac{S}{4}}$ then we can fit in up to $10^5 \times e^{300}$ vectors in the Hilbert space which is a huge number, far more sizeable than the number of atoms in our known observable universe ($\sim 10^{78} - 10^{82}$).

We will pause here to discuss some points. A natural extension to our present discussion is to study the paradox for single-sided black holes. We will do so using a setting similar to the eternal black holes, which we display in Figure 3.3. We advocate the same resolution for the single-sided paradox as we have for the eternal case. Another point is that there exists an entirely different way to arrive at this paradox. The paradox also arises if we glue inflating or FLRW regions inside the interior by using junction conditions

¹§3.2 gives the details of this calculation.

[53–57]. These glueings result in similar spacelike slices which have huge volumes in the interior. As a consequence of the paradox, it is also argued that the CFT does not contain the interior states. In our work, we assume that a state-dependent map reconstructs the black hole interior, thus describing the states behind the horizon [29–33]. Thus, by definition, the CFT captures our bulk interior excitations. On a related note, [58, 59] also discuss various subtleties regarding the problem of large interior volumes and advocate a similar resolution.

1.2 Monogamy paradox in flat space

Next, we move on to the monogamy paradox in flat space. Originally proposed in [12], the paradox was extensively discussed in [13, 20, 21, 29, 30, 33, 35, 60–78]. In particular, Raju [60] showed that the essential features of the monogamy paradox for old flat space black holes could be modeled using a setup in empty AdS, leading to a violation in the monogamy of entanglement there. The salient point of this setup was that it did not require the existence of a horizon, in contrast to the previous discussions of the paradox. A physically interesting question is whether such a violation in the monogamy of entanglement can be described within empty flat space, which resembles our observable universe to a good approximation. Our present work deals with addressing this question.

1.2.1 The paradox

We will briefly discuss the paradox below. Consider an old evaporating black hole in flat space at time t , such that $t > t_{\text{Page}}^2$. The outgoing near-horizon Hawking modes are strongly entangled with the near-horizon interior modes. For the final state to contain all information about the initial state, the near-horizon outgoing modes must also be entangled with Hawking modes that came out of the black hole at early times. However, this situation points to a violation of the monogamy of entanglement, which is an unavoidable consequence of quantum mechanics. This paradox is also closely related to the cloning paradox, which states that within effective field theory, a nice slice can capture both a diary thrown into a black hole and a reconstructed copy of the diary from the outgoing Hawking radiation, thereby violating the no-cloning theorem.

The physical picture portrayed by both the cloning and the monogamy paradoxes is that the interior should contain a *copy* of the exterior to resolve contradictions with basic assumptions of quantum mechanics (more precisely, with quantum information theorems).

²For a black hole with initial entropy given by S which has evaporated to S' , the number of states in the *exterior* is approximately $e^{S-S'}$. We thus pose the paradox for old black holes when the exterior contains enough degrees of freedom, i.e., $S' < \frac{S}{2}$. Later on, in line with the principle of holography of information [79–81], we will argue why we do not need to necessarily go beyond the Page time in order to set up the paradox, as is done in the standard case.

This picture is reflected within the important idea of black hole complementarity as explored in [14, 15, 82–84].

A key idea here is that the monogamy paradox for flat space black holes depends only on the entanglement of near-horizon exterior modes with near-horizon interior modes and also with modes far outside the horizon (e.g., at past of the future null infinity, i.e., \mathcal{I}_-^+). Consider the simple situation of a radially outgoing light cone at $r = r_0$ in an empty flat space, as shown in Figure 1.2. A monogamy-type paradox arises here also if we study the entanglement of the modes smeared just inside the light cone (region A) with modes smeared just outside the light cone (region B) and with another spacelike separated region (region C) (See Fig. 1.2).

1.2.2 The toy model in flat space

We will extend the construction of [60] which investigated the monogamy paradox in asymptotically AdS using Bell inequalities to asymptotically flat spacetime to understand the case of old flat space black holes. As done there, we will formulate the paradox using CHSH inequalities [85], a convenient restatement of Bell inequalities [86]. This formalism allows us to make quantitative statements regarding the monogamy of entanglement [87, 88], in particular, it allows us to rephrase statements about the monogamy of entanglement in terms of statements regarding the monogamy of CHSH correlations.

An essential ingredient that facilitates calculations in this setup compared to the original paradox is that the Hamiltonian of gravity is a boundary term [89], and thus can be used to construct a projector that projects onto the degenerate subspace of vacua labelled by supertranslations.

Important features unique to our toy model in flat space

The vacuum and low energy structure of the Fock space of canonical gravity in flat space is completely different from the same in AdS, due to the presence of supertranslations and the absence of a mass gap. In particular, AdS has a unique vacuum, while the flat space vacua span a degenerate subspace, and should be specified by their value in the supertranslation sector as well. We build upon previous works [79, 90–93] which have clarified the vacuum structure of flat space, and our definition of relevant operators and their subsequent representation in terms of the Fock space rests on the same. Here, supertranslations are not crucial to setting up the monogamy paradox in flat space but necessarily complicate the rather straightforward analysis in AdS since they introduce an additional vacuum structure.

Note that in the treatment for the AdS case in [60], there exists a natural cutoff scale set by the cosmological constant. However the issue for flat space gravity in $d = 4$ is more complicated, and one needs to specify an infrared cutoff in order to properly define the

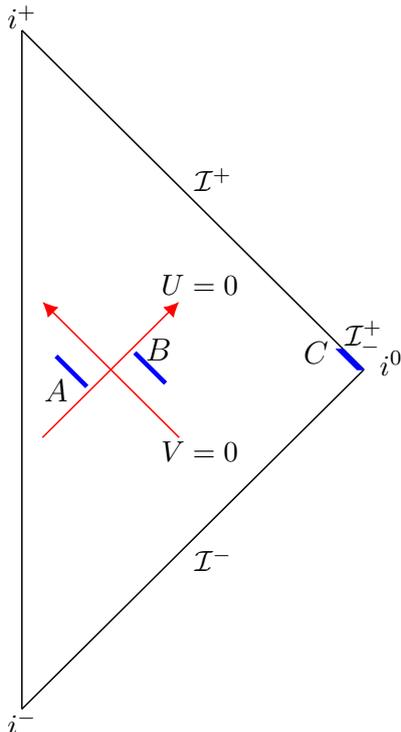


Figure 1.2: Here the red line $U = 0$ denotes a radially outgoing light shell. Regions A , B , and C are marked in blue. We will study the entanglement of modes smeared over region A with modes smeared over B and C and arrive at a paradox.

theory. Therefore in our work, we have introduced a new physically motivated projector P_δ , which projects onto energy scales below an infrared cutoff denoted by δ , and utilize the same to construct relevant operators which demonstrate the violation in monogamy of entanglement in our toy model. Physically this means that in practice, we do not work with operators that project onto the vacuum exactly but project onto states with very low energies below an IR cutoff, say $E < \delta$. This also generalizes the more abstract projector onto the vacua introduced in the context of AdS [84] and in flat space [79], and their subsequent role in how information is stored at the boundary [80].

While we can set up the monogamy paradox in flat space using the abstract projector introduced in [79] (as we demonstrate), the main thrust of our work is to utilize our physically motivated projector P_δ and use it to set up the monogamy paradox. In our work, given the infrared issues, we firstly demonstrate how operators C_i living in Region C in Fig. 1.2 which have almost the exact correlation with operators A_i in A as operators B_i defined over B have with operators A_i up to $O(\sqrt{G_N})$. Afterwards we then construct them using our physically motivated projector P_δ . This requires certain conditions on the smearing functions of relevant observables as we will qualitatively as well as rigorously explain in detail in our work.

Since these operators in C are constructed such that the AC system has almost identical CHSH correlators as the AB system and consequently the same entanglement,

we have a quantum information-theoretic contradiction.

Other details about our toy model

Given that we define our operators A_i , B_i and C_i supported in regions A , B and C respectively, the reader may ask what we mean by operator insertions in a theory of gravity. As opposed to local quantum field theory, there exists no definition of local gauge-invariant operators in a theory of quantum gravity. However, we will work with *approximately local* operators in our case, which involves taking an operator and smearing it over a small spatial interval. One way of thinking about such approximately local operators is to regard them as gauge fixing (i.e., up to small diffeomorphisms) in the bulk. We will not work with explicitly diffeomorphism invariant operators constructed using gravitational dressing, like ones formed by attaching geodesics from the boundary. This is because the paradox necessarily requires us to look at local bulk observables and such constructions are by definition non-local.

1.2.3 Proposed resolution

We note that the violation in the monogamy is $O(1)$, and cannot be removed by including minor corrections of $O(\sqrt{G_N})$, which are essential to the resolution of Hawking's original paradox and the bags of gold paradox [1, 5, 28, 40, 58, 94–96]. As we discuss below, this is an important observation that strongly hints towards a resolution of the paradox via complementarity, i.e., the interior degrees of freedom are complicated polynomials of the exterior degrees of freedom. The commutators are of $O(1)$ because in principle, we are acting upon the *same degrees of freedom* in the interior and the exterior, and complicated enough exterior operators can probe the information contained in the interior. The resolution is not surprising given that the Hilbert space does not factorize upon spatially partitioning a given manifold in a theory of gravity. This simple fact follows from the Gauss constraint of gravity. An implicit ingredient in our local quantum field theoretic construction is that the system partitioned into regions A , B , and C has factorized Hilbert spaces. However, the monogamy statement here is not violated since upon turning on gravity, the Hilbert space does not factorize, and consequently, the interior and exterior degrees of freedom are not independent degrees of freedom. This is the primary origin of the $O(1)$ violation, demonstrating why local quantum field theory is not a good framework for dealing with questions regarding quantum information and entanglement in gravity.

1.3 Organization of the thesis

After this initial motivation, we will discuss some essential aspects of quantum aspects of black holes in §2. In particular, we will stress how QFT in curved spacetime leads

to puzzles about quantum information and use it to study some standard information puzzles regarding the black hole interior. In the same chapter, we will also develop necessary tools relevant to understanding later discussions regarding the paradoxes.

In §3, we will discuss the bags of gold paradox in detail, while in §4, we will discuss the monogamy paradox in flat space. We will first set up the paradox in detail in both cases and then discuss our resolution. In the case of the bags of gold, we will perform various consistency checks that demonstrate our proposal.

Finally, in §5, we will summarize the thesis, discuss various aspects, and outline related directions. The Appendices contain results and calculations which are too detailed to be included in the main body of our work.

Chapter 2

Brief review of relevant topics

This section will first develop the necessary tools to understand quantum field theory in curved spacetime. In particular, we will stress how the regularity of correlators across horizons plays an essential role in understanding the physics of horizons. We then discuss black holes, the properties of black hole horizons, and, consequently, the laws of black hole thermodynamics. We then formulate QFT in black hole spacetime and deduce interesting properties for black holes like Hawking radiation and black hole evaporation.

In the later part, we use these essential tools to probe the black hole's interior and analyze related puzzles. In particular, we demonstrate Hawking's paradox and use statistical properties of pure states to demonstrate why the paradox is not on a firm footing. We then introduce two related paradoxes: the cloning and the strong subadditivity paradox, and briefly argue their resolutions. The notion of black hole complementarity is then introduced to discuss these issues. Finally, we discuss how to write interior operators using state-dependent formalism.

These tools serve as primary foundations, building upon which we will discuss our thesis work on the bags of gold and the monogamy paradoxes.

2.1 QFT in curved spacetime

First, let us review some basic tools involving quantum field theory in curved spacetime, which is the leading order contribution to quantum gravity. The basic idea here is that of a fixed background upon which quantum fields can fluctuate without changing the spacetime geometry. These fields include not only matter but also gravitons. The Lagrangian density of a minimally coupled real scalar field is given by

$$L = \frac{\sqrt{-g}}{2} \left(g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{m^2 \phi^2}{2} \right) \quad (2.1)$$

where $g_{\mu\nu}$ is the metric and $g = \det(g_{\mu\nu})$. The action is the integral over the Lagrangian density

$$S = \int d^d x L. \quad (2.2)$$

The corresponding equations of motion are given by

$$\frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} g^{\mu\nu} \partial_\nu \phi + m^2 \phi = 0. \quad (2.3)$$

As in any quantum field theory the key observables are the vacuum correlators,

$$\langle 0 | \phi(x_1) \phi(x_2) \dots \phi(x_n) | 0 \rangle \quad (2.4)$$

The critical feature that distinguishes QFT on a generic curved spacetime compared to QFT in Minkowski space is that the vacuum is not invariant, i.e., not all observers moving in the geometry will agree on a common vacuum. Consequently, the particle number is not an invariant measure as well. This is a consequence of the fact that there is no canonical choice of time.

We can quantize the scalar field system by imposing equal time commutation relations. Since there is no canonical choice of time, we make a choice of time by hand, i.e. we fix a spacelike slice upon which the canonical commutation relations are imposed. The metric then splits up into the so-called ADM form

$$ds^2 = N^2 dt^2 - h_{ij} (dx^i + N^i dt) (dx^j + N^j dt), \quad (2.5)$$

where we have $\sqrt{-g} = N\sqrt{h}$. The commutation relations on a fixed spacelike slice labelled by t are given by

$$[\phi(x, t), \Pi(x', t)] = i\delta(x - x') \quad (2.6)$$

2.1.1 Bogoliubov transformations

We write the mode expansion of the solution to the field equations in (2.3) as follows

$$\phi = \sum_i a_i f_i(t, x) + a_i^\dagger f_i^*(t, x), \quad (2.7)$$

where i labels the quantum number over which the sum takes place, such as, the energy. Substituting (2.7) in the canonical commutation relations gives us the commutators for the creation and annihilation operators

$$[a_i, a_j^\dagger] = \delta_{ij} \quad (2.8)$$

provided the mode functions satisfy the following relation:

$$\sum_i \sqrt{-g} \left[f_i(t, x) g^{0\mu} \partial_\mu f_i^*(t, x) - f_i^*(t, x) g^{0\mu} \partial_\mu f_i(t, x) \right] = \delta(x - x') \quad (2.9)$$

One can now use the creation and annihilation operators to construct the Fock space. The important point here is that one can perform the mode expansion of equation (2.3) in terms of a basis of completely different modes as well, i.e.

$$\phi = \sum_i b_i g_i(t, x) + b_i^\dagger g_i^*(t, x). \quad (2.10)$$

However since ϕ is the solution of a linear differential equation, therefore both mode decompositions are just a simple change of basis related by a linear transformation.

$$\begin{aligned} a_i &= \sum_j \alpha_{ij} b_j + \beta_{ij}^* b_j^\dagger \\ a_j^\dagger &= \sum_j \beta_{ij} b_j + \alpha_{ij}^* b_j^\dagger \end{aligned} \quad (2.11)$$

and consequently the mode functions similarly satisfy

$$\begin{aligned} g_i &= \sum_j \alpha_{ij} f_j + \beta_{ij} f_j^* \\ g_j^* &= \sum_j \beta_{ij}^* f_j + \alpha_{ij}^* f_j^* \end{aligned} \quad (2.12)$$

This implies that the positive frequency modes in one basis are just a linear combination of positive and negative frequency modes in another basis. Let us define the different vacua as

$$a_i |\Omega\rangle = 0; \quad b_i |0\rangle = 0. \quad (2.13)$$

By imposing the canonical commutation relations in the other basis, one can show that the vacuum is not invariant, but rather there is a relation between the two vacua. In particular one vacuum written down in terms of oscillators acting on another vacuum resembles a squeezed state:

$$|\Omega\rangle = \exp\left(\frac{1}{2} b_j^\dagger c_{jk} b_k^\dagger\right) |0\rangle \quad (2.14)$$

where the matrix c_{mj} is defined as

$$c_{mj} = - \sum_i \beta_{mi}^* \gamma_{ij} \quad (2.15)$$

such that the matrix γ_{ij} represents the inverse of the matrix α_{ij} :

$$\sum_i \alpha_{ji} \gamma_{ik} = \delta_{jk}. \quad (2.16)$$

Note, however, that the correlators defined in equation (2.4) are still the invariant quantities since the field operators remain unchanged under the Bogoliubov transformations, which compensate for the transformation of the state.

2.1.2 Rindler / Unruh modes and Bogoliubov coefficients

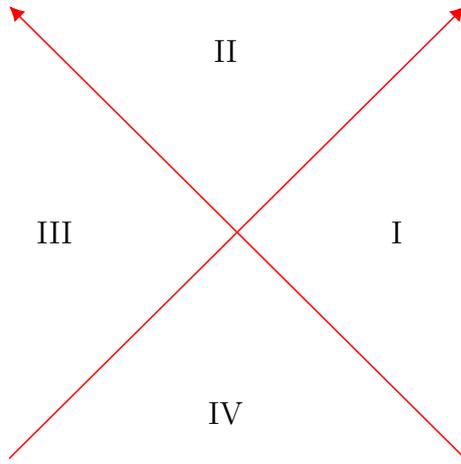
To initially define the Rindler and Unruh modes, we first work in flat space in two dimensions where the metric is simply given by

$$ds^2 = dt^2 - dx^2 \quad (2.17)$$

Here we can define the coordinates $U = t - x$, and $V = t + x$. In the region $x > |t|$, we define the Rindler coordinates

$$\begin{aligned} U &= -\frac{1}{a} \exp(-a u_r) \\ V &= \frac{1}{a} \exp(a v_r). \end{aligned} \quad (2.18)$$

$$U = 0, u_r = \infty, t_r = \infty, x_r = -\infty$$



$$V = 0, v_r = -\infty, t_r = -\infty, x_r = -\infty$$

Figure 2.1: The Rindler coordinates

We call this region quadrant I which is depicted in Figure 2.1. The metric in these new coordinates takes the conformal form

$$ds^2 = \frac{\exp(2ax_r)}{a^2} (dt_r^2 - dx_r^2), \quad (2.19)$$

where $u_r = t_r - x_r$, $v_r = t_r + x_r$. One can define other quadrants similarly by changing the signs in front of the exponential, e.g. in quadrant III given by $x < -|t|$, we can define

the coordinates

$$\begin{aligned} U &= \frac{1}{a} \exp(-a u_r) \\ V &= -\frac{1}{a} \exp(a v_r), \end{aligned} \tag{2.20}$$

while in region II given by $t > |x|$, we have

$$\begin{aligned} U &= \frac{1}{a} \exp(-a u_r) \\ V &= \frac{1}{a} \exp(a v_r), \end{aligned} \tag{2.21}$$

while correspondingly in region IV given by $t < -|x|$, we have

$$\begin{aligned} U &= -\frac{1}{a} \exp(-a u_r) \\ V &= -\frac{1}{a} \exp(a v_r). \end{aligned} \tag{2.22}$$

Rindler mode expansions

Since the modes are conformally flat, we can define a mode expansion similar to the Minkowski expansion. For region I, this takes the form:

$$\phi_I = \int \frac{d\omega}{\sqrt{\omega}} \left(a_\omega e^{-i\omega u_r} + b_\omega e^{-i\omega v_r} + \text{h.c.} \right). \tag{2.23}$$

The mode expansion takes a similar form for Region III, which is given by

$$\phi_{III} = \int \frac{d\omega}{\sqrt{\omega}} \left(\tilde{a}_\omega e^{i\omega u_r} + \tilde{b}_\omega e^{i\omega v_r} + \text{h.c.} \right). \tag{2.24}$$

Note that the exponential signs here are different since the Rindler time t_r runs in the opposite direction in Region III compared to Region I. Consequently, on the slice at $t = 0$, we have the following mode expansion:

$$\phi = \int \frac{d\omega}{\sqrt{\omega}} \left(a_\omega e^{-i\omega u_r} + b_\omega e^{-i\omega v_r} + \tilde{a}_\omega e^{i\omega u_r} + \tilde{b}_\omega e^{i\omega v_r} + \text{h.c.} \right). \tag{2.25}$$

Similarly in global modes, the left moving and the right moving modes of the Minkowski space can be written as:

$$\phi = \int \frac{d\omega}{\sqrt{\omega}} \left(c_\omega e^{-i\omega(t-x)} + d_\omega e^{-i\omega(t+x)} + \text{h.c.} \right). \tag{2.26}$$

Unruh modes

In general, finding Bogoliubov coefficients denoting the change of basis from Minkowski to Rindler modes is cumbersome. However, one can define Unruh modes and use them to find the Bogoliubov transformation conveniently between the Rindler and the Unruh modes.

Then the task of calculating Minkowski to Rindler Bogoliubov coefficients simplifies since the Unruh and Minkowski vacuum is essentially the same. We will not compute the Minkowski to Rindler coefficients explicitly but will outline the construction of these modes.

The Unruh mode is defined as follows:

$$\begin{aligned} U_U(u_r) &= e^{-i\omega u_r}, & \text{In Region III} \\ &= e^{\frac{\pi\omega}{a}} e^{-i\omega u_r}, & \text{In Region I} \end{aligned} \quad (2.27)$$

which can be essentially written as

$$U_U(u_r) = a^{\frac{i\omega}{a}} U^{\frac{i\omega}{a}}. \quad (2.28)$$

Note that $U^{\frac{i\omega}{a}}$ has a branch cut, which we can choose in the upper half plane. Apart from this U does not have any singularities in the lower half plane. Consequently one can analytically continue U in the lower half plane from Region III to Region I. This gives rise to the extra factor of $e^{\frac{\pi\omega}{a}}$ in eqn (2.27).

Since these modes are analytic in the lower half-plane, it is easy to see that the modes satisfy the following property

$$\int_{-\infty}^{\infty} dU U^{\frac{i\omega}{a}} e^{-i\omega' U} = 0, \quad \text{for } \omega' > 0 \quad (2.29)$$

since the integral can be continued analytically in the lower half-plane and is zero. Thus the Unruh mode has only positive Minkowski frequencies, and since there is no mixing of positive and negative frequencies, both the Unruh vacuum and the Minkowski vacuum are the same.

Using the u_r coordinate one can similarly define the other Unruh $\tilde{U}_U(u_r)$ mode as

$$\begin{aligned} \tilde{U}_U(u_r) &= e^{i\omega u_r}, & \text{In Region III} \\ &= e^{-\frac{\pi\omega}{a}} e^{i\omega u_r}, & \text{In Region I} \end{aligned} \quad (2.30)$$

We can similarly define the corresponding Unruh mode using the v_r Rindler mode,

which is given by

$$\begin{aligned} V_U(v_r) &= e^{-i\omega v_r}, & \text{In Region III} \\ &= e^{\frac{\pi\omega}{a}} e^{-i\omega v_r}, & \text{In Region I} \end{aligned} \quad (2.31)$$

while the other Unruh mode $\tilde{V}_U(v_r)$ is given by

$$\begin{aligned} \tilde{V}_U(v_r) &= e^{i\omega v_r}, & \text{In Region III} \\ &= e^{-\frac{\pi\omega}{a}} e^{i\omega v_r}, & \text{In Region I} \end{aligned} \quad (2.32)$$

The mode expansion of the scalar field in the Unruh modes is given by

$$\phi = \int \frac{d\omega}{\sqrt{\omega}} \left(e_\omega U_U(u_r) + \tilde{e}_\omega \tilde{U}_U(u_r) + f_\omega V_U(v_r) + \tilde{f}_\omega \tilde{V}_U(v_r) + \text{h.c.} \right). \quad (2.33)$$

From the Rindler oscillators defined in eqn (2.25), we can simply read off the Rindler to Unruh coefficients:

$$e_\omega = \frac{a_\omega - e^{-\frac{\pi\omega}{a}} \tilde{a}_\omega^\dagger}{e^{\frac{\pi\omega}{a}} - e^{-\frac{\pi\omega}{a}}} \quad \text{and} \quad \tilde{e}_\omega = \frac{\tilde{a}_\omega - e^{-\frac{\pi\omega}{a}} a_\omega^\dagger}{1 - e^{-2\frac{\pi\omega}{a}}} \quad (2.34)$$

Using these Bogoliubov coefficients and equation (2.14), one can write the Minkowski vacuum in terms of the Rindler vacuum as

$$|\Omega_M\rangle = \exp \left(\sum_\omega e^{-\frac{\pi\omega}{a}} a_\omega^\dagger \tilde{a}_\omega^\dagger + \text{b-terms} \right) |\Omega_{I,III}\rangle \quad (2.35)$$

One can expand this state which results in the well-known property of Minkowski vacuum that it looks like a thermofield doubled state when written in terms of Rindler modes:

$$|\Omega_M\rangle = \sum_E e^{-\frac{\beta E}{2}} |E, E\rangle \quad (2.36)$$

where $\beta = \frac{2\pi}{a}$ and $E = \sum \omega n_\omega$ is the energy of the state. Note that if we sum over one of either the regions I or III, we get the canonical density matrix for the other region, i.e.

$$\rho(E) = \frac{e^{-\beta E}}{\sum_E e^{-\beta E}} \quad (2.37)$$

2.1.3 Bogoliubov coefficients from the universality of two-point function

Consider a $d + 1$ dimensional spacetime where we introduce the coordinates $U = t - r$ and $V = t + r$, such that the metric takes the form

$$ds^2 = -dU dV + \dots, \quad (2.38)$$

where dots denote the contribution from $d - 1$ transverse directions. Then any two-point function on this spacetime has a universal part, which nicely reflects itself at short enough distances. This arises from the fact that zooming into a region of spacetime reveals a universal Minkowski-like behaviour at the leading order. In particular, given a state $|\psi\rangle$ dual to a given geometry, the two-point correlator at short distances takes the universal form:

$$\langle \Psi | \phi(x_1) \phi(x_2) | \Psi \rangle = \frac{C}{|x_1 - x_2|^{\frac{d-1}{2}}} (1 + \mathcal{O}(|x_1 - x_2|)) \quad (2.39)$$

where C is a constant that can be fixed using the free scalar theory, i.e.

$$C = \frac{\Gamma(d-1)}{2^d \pi^{\frac{d}{2}} \Gamma\left(\frac{d}{2}\right)} \quad (2.40)$$

Extraction of modes and Bogoliubov coefficients

An interesting property follows from here, which is relevant to our later discussion of the monogamy paradox. Consider the regions A and B , which are situated just inside and outside an outgoing light cone as shown in Fig. 1.2. Given the scalar field, one can construct any general correlator by extracting the oscillator creation/annihilation operator.

We will consider the case of Rindler smearing. To perform this, we introduce a tuning function $\mathcal{T}(U)$ such that it is supported only on the small bounded regions and smoothly dies off. Thus we define the smeared operators on the regions A and B by

$$\begin{aligned} \alpha_A &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2} \Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) \Phi(t_A(U), r_A(U), \Omega) \\ \alpha_B &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2} \Omega r_B^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}(U) \Phi(t_B(U), r_B(U), \Omega) \\ \alpha_A^\dagger &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2} \Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}^*(U) \Phi(t_A(U), r_A(U), \Omega) \\ \alpha_B^\dagger &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2} \Omega r_B^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}^*(U) \Phi(t_B(U), r_B(U), \Omega) \end{aligned} \quad (2.41)$$

Here r_s and t_s , where $s = A, B$ denote the coordinates on the regions A and B . The

key idea here is that the smearing function oscillates increasingly as we tend to go near $U = 0$, and thus even a small interval very close to $U = 0$ is useful to extract out the Rindler modes.

One can use the above oscillators to calculate the two-point function involving the creation and annihilation operators. This ultimately gives us the Bogoliubov coefficients as derived in equation (2.34). This is demonstrated in detail in [80].

The above statement is quite general, and one can use such correlations across a null surface to extract details of the Bogoliubov coefficients, with the key feature here being the fact that there is a universal part of the correlator that, in some instances, can be extracted out by smearing over a bounded region. We will discuss further details of this construction in §4.

2.2 Black holes

We will now move on to a brief review of black holes and their properties.

2.2.1 Schwarzschild black hole

The metric of the Schwarzschild black hole in four dimensional asymptotically flat space is given by

$$ds^2 = - \left(1 - \frac{2m}{r}\right) dt^2 + \frac{dr^2}{1 - \frac{2m}{r}} + r^2 d\Omega^2 \quad (2.42)$$

On inspecting the metric, we observe two different singularities, one at $r = 0$ and another at $r = r_h$. The singularity at $r = r_h$ is not a real singularity but a coordinate singularity; however, the one at $r = 0$ is a real singularity. This can be seen because scalar curvature invariants diverge at $r = 0$ but not at $r = r_h$.

We can also use other convenient coordinate systems to see that the singularity is essentially due to our choice of coordinate system. One such coordinate system is the Tortoise coordinate system.

In addition, we can calculate the proper time it takes for an observer to fall into a black hole. Using this, we can see that it takes a finite proper time to pass through the horizon as well as the observer reaches the singularity in a finite proper time. Hence we can see that the proper time analysis indicates that the observer should eventually cross the horizon, even though by computing the redshift, they never appear to cross the horizon.

As argued before, the coordinate singularity can be replaced using an appropriate coordinate redefinition such as the Tortoise coordinate system. These coordinates can be

defined so that the metric looks like

$$ds^2 = f(r_*) \left[-dt^2 + dr_*^2 \right] + g(r_*) d\Omega^2. \quad (2.43)$$

To do this, we need to define

$$\left(1 - \frac{2m}{r}\right) dr_*^2 = \frac{dr^2}{\left(1 - \frac{2m}{r}\right)}; \quad \text{i.e.} \quad dr_* = \frac{dr}{\left(1 - \frac{2m}{r}\right)} \quad (2.44)$$

which leads to

$$r_* = r + 2m \log \left| \frac{r - 2m}{2m} \right|. \quad (2.45)$$

This has the following falloff for the Tortoise coordinate: as $r \rightarrow 2m$, then $r_* \rightarrow 0$. Next, we define the following Rindler like coordinates:

$$U = -\exp \alpha (r_* - t), \quad V = \exp \alpha (r_* + t) \quad (2.46)$$

Using these coordinates, the metric takes the form

$$dU dV = -\alpha^2 (dr_* - dt) (dr_* + dt) \exp 2\alpha r_* \quad (2.47)$$

Now near the horizon the exponential part, i.e., $\exp 2\alpha r_*$ can be expanded as

$$\exp 2\alpha r_* = \exp 2\alpha r_h \left(\frac{r - 2m}{2m} \right)^{4m\alpha} \quad (2.48)$$

Setting $4m\alpha = 1$, the near horizon metric becomes proportional to $dU dV$. The exact metric then becomes

$$ds^2 = -dU dV \left(\frac{32M^3 e^{-\frac{r}{2m}}}{r} \right) + r^2 d\Omega^2 \quad (2.49)$$

which is non-singular as we take the limit $r \rightarrow r_h$. Here the horizon is at $U = 0$, and we can smoothly extend the metric into the region $U > 0$, reminiscent of Rindler space as discussed earlier. The true singularity at $r = 0$ cannot be removed by any coordinate redefinition and can be seen from the divergence of scalar curvature invariants.

Once again, as in the case of flat space, we can define coordinates in different regions. We can similarly define coordinates in "Region IV" where ($U < 0, V < 0$):

$$U = -\exp \alpha (r_* - t), \quad V = -\exp \alpha (r_* + t). \quad (2.50)$$

Here surfaces of constant t are given by fixing $\frac{U}{V}$, while surfaces of constant r are given by fixing UV . So constant r surfaces are hyperboloids in the $U - V$ plane while constant t surfaces are straight lines passing through the origin.

In terms of the Tortoise coordinate, the singularity $r = 0$ is same as $r_* = 0$ and

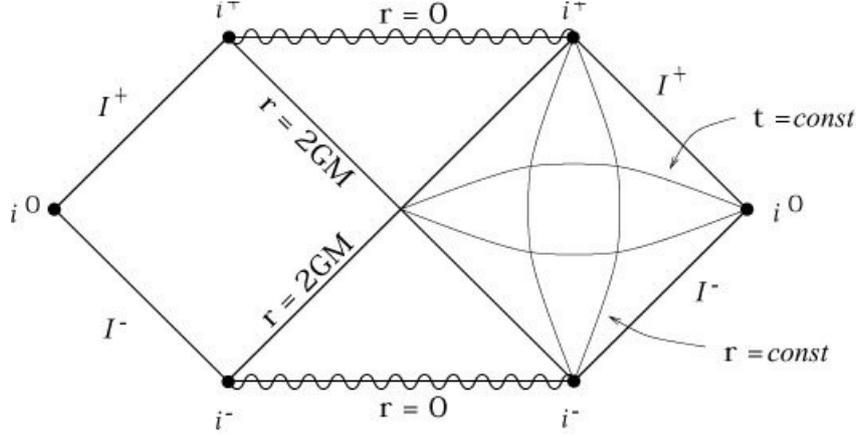


Figure 2.2: Penrose diagram for Schwarzschild black hole in asymptotic flat spacetime. Image taken from Google.

$UV = 1$.

2.2.2 Penrose diagrams

In our work, we will use Penrose diagrams to study various spacetimes. These diagrams are very useful to study classical solutions. They help visualize the causal structure by compactifying the spacetime, especially the boundary structure. As an example, for the case of the Schwarzschild black hole, we can define

$$x = \arctan U; \quad \xi = \arctan V \quad (2.51)$$

while compactifies the spacetime to the region $x \in \left[\frac{\pi}{2}, -\frac{\pi}{2}\right]$ and similarly $\xi \in \left[\frac{\pi}{2}, -\frac{\pi}{2}\right]$, except the cutoff at $UV = 1$ or $\tan x \tan \xi = 1$.

We can now rotate the resultant diagram by 45 degrees in the anti-clockwise direction in order to make null rays at 45 degrees. This gives us the diagram in Figure 2.2.

Flat Schwarzschild black hole

Let us briefly discuss Figure 2.2. The rightmost region is the exterior of the flat Schwarzschild black hole. The lightlike asymptotic boundaries of this region are called the future null infinity \mathcal{I}^+ and the past null infinity \mathcal{I}^- , while the spacelike boundary is denoted by i^0 . The timelike boundaries are denoted by i^\pm . Note here that the causal structure of light rays is preserved as can be seen from the horizons, i.e., the light rays are either at 45 degrees or 135 degrees.

The leftmost region and the bottom part are the Kruskal extensions of this Penrose diagram. This Kruskal extension corresponds to a wormhole-like geometry, indicated by a common shared interior region connecting two different exterior boundaries. As indicated in the figure, this region is causally disconnected from the exterior in the right

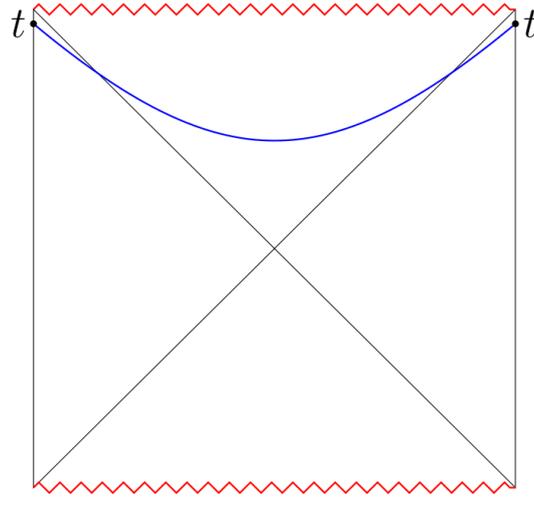


Figure 2.3: Penrose diagram of eternal black hole in asymptotically AdS, where the blue slice is a nice slice, i.e. a spacelike Cauchy slice that stays away from singularities.

region.

Eternal AdS Schwarzschild black hole

The Penrose diagram for the eternal black hole in AdS is similar apart from the asymptotic boundary structure, as shown in Figure 2.3. This again has the wormhole connecting two different exterior regions. The only difference from the Kruskal extension of the flat space two-sided Schwarzschild black hole is in the presence of timelike boundaries in AdS, which is different from the asymptotic structure of flat space.

Limitations of Penrose diagrams

While Penrose diagrams are handy for studying classical geometries, they may be misleading when one tries to visualize certain features of gravity, especially quantum effects. One crucial effect in black hole spacetimes is the infinite blue shifts that emerge when we backtrace light rays to the past, which naively is not apparent in the Penrose diagram.

Another important example is an evaporating black hole, where the final state after evaporation becomes a flat spacetime with propagating gravitons and matter. One may try to modify the Penrose diagram here by adding an extra wedge to denote the final state of the evaporating black hole. However, this conflicts with how information is localized in a theory of quantum gravity [80].

2.2.3 Astrophysical black holes

If we restrict ourselves to classical general relativity, then the Schwarzschild black hole in flat space is eternal. However, even in classical general relativity, the eternal black hole

is not a very physical scenario since most astrophysical black holes are formed from an initial state of gravitationally collapsing matter.

A very general class of such solutions is the Oppenheimer-Snyder-Datt solution. A crucial ingredient leading to this solution is the Birkhoff theorem: any spherically symmetric solution of Einstein field equations is static, i.e., the Schwarzschild solution is the unique solution in the absence of matter.

Now consider two regions, the collapsing matter, and the external region with no matter. As Birkhoff pointed out, the external solution is just the Schwarzschild solution. Regarding the solution in the region of constant matter, we can write the candidate metric as

$$ds^2 = -d\tau^2 + A dr^2 + X d\Omega^2 \quad (2.52)$$

One can systematically solve for this metric, i.e., the coefficients A and X using a particular form for the stress-energy $T_{\mu\nu}$, and glue it appropriately with the exterior solution to extract the physics of the solution. The solution's causal structure involves forming an apparent horizon that merges into the Schwarzschild event horizon towards the end of gravitational collapse.

As we will see later, by incorporating quantum effects, the late time state of a black hole is not a stable black hole solution but is a final state involving Hawking quanta.

2.3 Focussing, light rays and black hole horizons

In this section, we will first develop the necessary mathematical tools leading to the derivation of the Raychaudhuri equation. Using the Raychaudhuri equation, we will describe the focussing effect of light rays given certain restrictions on the energy conditions. Finally, we will apply our understanding to black hole horizons and use it to prove theorems regarding the same.

2.3.1 Lie transport

Let us start with the well-known notion of parallel transport. Recall that parallel transport is necessary to define geodesics, i.e. to implement $\xi^a \nabla_a \xi^b = 0$ where ξ denotes the velocity field. The reason is clear when we write the geodesic equation

$$\frac{d^2 x^b}{d\tau^2} + \Gamma_{ac}^b \frac{dx^a}{d\tau} \frac{dx^c}{d\tau} = 0 \implies \frac{dx^a}{d\tau} \cdot \nabla_a \frac{dx^b}{d\tau} = 0, \quad (2.53)$$

Now if we set $\frac{dx^a}{d\tau} = \xi^a$, then it follows that $(\xi^a \nabla_a \xi^b) = 0$.

We can similarly parallel transport a vector T^b around a curve $\xi^a = \frac{dx^a}{d\tau}$ by $\xi^a \nabla_a T^b = 0$. The implication here is that the T^b 's angle with the curve ξ^a will not be the same but depend on the manifold's curvature.

However, consider the following physical situation. Consider a fixed observer at a certain altitude over the earth (not the geosynchronous orbit, but say the observer is kept afloat by some rockets). The observer wants to keep their telescope oriented directly outward. How can the vector describing their telescope's direction be transported along the observer's worldline? In other words, how is the telescope vector T shifted along the observer's vector field O ?

The Lie transport $L_O T = 0$ is designed to describe this situation. Consider a curve C which is parametrized by x^μ such that the tangent vector is given by

$$u^\mu = \frac{dx^\mu}{d\lambda}. \quad (2.54)$$

Then the Lie derivative is given by

$$L_u v^\mu = u^\nu \nabla_\nu v^\mu - v^\nu \nabla_\nu u^\mu \quad (2.55)$$

This definition can be extended to tensors as well.

More precisely, Lie transport preserves the observer's angle with the telescope, and the Lie derivative is the generator of the telescope's transport along the observer's curve.

Another way to understand the same is that parallel transport depends on the manifold's intrinsic properties like connection, curvature, etc., while Lie transport depends on the curves only and is independent of the manifold's properties.

Isometries

The defining property of a Killing vector field ζ is that it satisfies the Killing equation, i.e., the equation where the Lie transport is precisely zero.

$$L_\zeta g_{\mu\nu} = \nabla_\mu \zeta_\nu + \nabla_\nu \zeta_\mu = 0 \quad (2.56)$$

One can study the transport of metric along Killing vector fields and use it to define spacetime isometries. Spacetimes in which the above equation has a maximal number of solutions to the Killing equation, i.e., $\frac{d(d+1)}{2}$ Killing vectors are called maximally symmetric spacetimes. Important examples of such spacetimes are flat space (where the Poincare group constitutes the isometries), Anti-de Sitter space, and de Sitter space.

Static and stationery spacetimes

Stationary spacetime means a spacetime with an asymptotically timelike Killing vector. Static means a stationary spacetime in which the asymptotically timelike Killing vector is hypersurface orthogonal, and therefore the metric components $g_{0\mu}$ vanish.

2.3.2 Raychaudhuri equation

Consider a congruence of null geodesics, which is defined by each geodesic passing through a point through which no other geodesic passes. Therefore one can define a notion of distance between the geodesics, which is a vector field measuring the deviation between the geodesics. We will call the deviation vector field η .¹

Let the tangent vector field be denoted by ξ^a . We can define the following tensor using the tangent vector field

$$B_{ab} = \nabla_a \xi_b,$$

which tells us how the deviation vector field η is parallel transported, as we shortly see. Since the deviation vector field η is Lie transported along a geodesic, the Lie derivative with respect to the tangent vector field ξ is zero.

$$L_\xi \eta = 0 \implies \xi_a \nabla^a \eta_b - \eta_a \nabla^a \xi_b = 0 \implies \xi_a \nabla^a \eta_b = \eta_a \nabla^a \xi_b = \eta_a B_b^a \quad (2.57)$$

The above equation demonstrates that B tells us how the deviation is parallel transported along a geodesic ξ .

Note here that the deviation vector field is not unique. One can have different notions of the deviation vector, and the above equation is valid for all.

Now consider the following expression

$$\xi^a \nabla_a (\xi^b \eta_b) = (\xi^a \nabla_a \xi^b) \eta_b + (\xi^a \nabla_a \eta_b) \xi^b \quad (2.58)$$

Here we have used the fact that ξ^a satisfies the geodesic equation given by $(\xi^a \nabla_a \xi^b) = 0$. The second term in the above equation is zero as well. One way is to choose the deviation vectors such that $\xi \cdot \eta = 0$.

Another way to see why $\xi^a \nabla_a (\xi^b \eta_b)$ vanishes is to note that

$$\xi^a \nabla_a (\xi^b \eta_b) = (\xi^a \nabla_a \eta_b) \xi^b - (\eta^a \nabla_a \xi^b) \xi_b + (\eta^a \nabla_a \xi^b) \xi_b = (L_\xi \eta)^b \xi_b + \eta^a \nabla_a \left(\frac{\xi^b \xi_b}{2} \right) = 0. \quad (2.59)$$

We have thus shown that the transport of the non-orthogonal part of η (i.e., $\eta \cdot \xi$) along the vector field ξ is uninteresting as it is constant. We now focus when η is orthogonal to ξ , i.e. $\eta \cdot \xi = 0$. Consider the case when we have the null geodesics

$$\xi \cdot \xi = 0 \quad (2.60)$$

Given this class of geodesics, we consider the equivalence class of the deviation fields

¹Note that the choice of the deviation is not unique, as we will shortly demonstrate later on.

given by shifting the deviation along ξ and parametrized by λ

$$\eta \rightarrow \eta + \lambda\xi, \quad (2.61)$$

which leads us to consider the two-dimensional vector space spanned by η and ξ . This space is convenient to consider because $(\eta, \xi) = 0$. Let us now consider the induced metric on this two-dimensional space h_{ab} .

Note here that picking up different representations of η does not lead to two different answers. This is because of the fact

$$h_{ab}\eta^a(\eta^b + \lambda\xi^b) = 0, \quad (2.62)$$

since $(\eta, \xi) = 0$. The tensor B also has a well defined action on this two dimensional space because

$$(B_b^a \eta_a) \xi^b = 0. \quad (2.63)$$

The action is well defined because taking $\eta \rightarrow \eta + \lambda\xi$ doesn't change $(B_b^a \eta_a) \xi^b = 0$. We now want to decompose the tensor B_{ab} in the following fashion

$$B_{ab} = \frac{1}{2}\theta h_{ab} + \sigma_{ab} + \omega_{ab} \quad (2.64)$$

which resembles the standard stress energy decomposition of a fluid. Here the different stress energy components have the following relations:

$$\sigma_{ab} = \sigma_{ba} \quad \omega_{ab} = -\omega_{ba} \quad \sigma_{ab} h^{ab} = 0 \quad (2.65)$$

Given the decomposition, we can multiply the same with h^{ab} on both sides, and write down θ in terms of B_{ab} and h^{ab} .

$$\theta = B_{ab} h^{ab}. \quad (2.66)$$

Our goal is to study the evolution of θ as we move along the geodesic flow. More precisely, we want to study how $\frac{d\theta}{d\lambda}$ changes along the vector field ξ . Substituting the form for B_b^a in eqn. (2.66) and taking appropriate covariant derivatives, we get

$$\xi^a \nabla_a \theta = \xi^c \nabla_c \nabla_a \xi_b h^{ab}. \quad (2.67)$$

Let us now analyse the term given by $\xi^c \nabla_c \nabla_a \xi_b$, which takes the form

$$\xi^c \nabla_c \nabla_a \xi_b = \xi^c \nabla_a \nabla_c \xi_b + R_{cab}^d \xi_d \xi^c = -(\nabla_a \xi^c)(\nabla_c \xi^b) + R_{cab}^d \xi_d \xi^c = R_{cab}^d \xi_d \xi^c - X_a^c X_{cb}. \quad (2.68)$$

If we take the trace after substituting the above simplification in eqn (2.67), we get the

following expression for change of θ along the affine parameter

$$\frac{d\theta}{d\lambda} = \xi^a \nabla_a \theta = -R_{cd} \xi_d \xi^c - X_a^c X_c^a = -\frac{1}{2} \theta^2 - \sigma_{ab} \sigma^{ab} + \omega_{ab} \omega^{ab} - R_{cd} \xi_d \xi^c. \quad (2.69)$$

Next we consider geodesic congruences which are hypersurface orthogonal. In other words, we restrict ourselves to the class where $\xi = fdg$. Since $d^2 = 0$, acting with the differential operator on ξ gives us

$$d\xi = df \wedge dg = \frac{df}{f} \wedge \xi \implies \nabla_a \xi_b - \nabla_b \xi_a = V_a \xi_b - V_b \xi_a. \quad (2.70)$$

This sets $\omega_{ab} = 0$, as now we have

$$\nabla_{[a} \xi_{b]} n_1^a n_2^b = 0. \quad (2.71)$$

We will now enforce the **weak null energy condition**, which is a classically reasonable assumption²

$$R_{cd} \xi^c \xi^d = G_{cd} \xi^c \xi^d = 8\pi T_{cd} \xi^c \xi^d > 0. \quad (2.72)$$

Consequently we arrive at the following equation:

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2} \theta^2. \quad (2.73)$$

Integrating the same with respect to the affine parameter λ gives us an useful result for focussing of light rays,

$$\frac{1}{\theta_f} - \frac{1}{\theta_i} \leq \frac{1}{2} (\lambda_f - \lambda_i). \quad (2.74)$$

This implies that θ always decreases. Once you have a negative value of $\theta_i = -\theta_0$, therefore in affine time $\lambda_f = 2\theta_0$, $\theta_f \rightarrow -\infty$.

2.3.3 Key properties of black hole horizons

Recall that in a black hole, the horizon bounds the interior region where gravity is so strong that massless signals from the interior cannot escape to its complementary region, especially to the asymptotics. In other words, say in asymptotically flat space, a black hole horizon in asymptotically flat spacetime is the boundary of the causal past of the future null infinity. This implies that the horizon is a global aspect of general relativity.

We will now list some theorems about the black hole horizons in asymptotically flat space and sketch rough proofs. Note that all theorems are generalizable to other spacetimes if we replace the asymptotic structure of flat spacetime with another asymptotic region. In particular, we can replace the spacetime boundary given by future null infinity

²but not quantum mechanically, as we will see for the case of Hawking radiation

(\mathcal{I}^+) using a timelike boundary such as that of Anti-de Sitter. The focus is more on the physical aspects of the proofs and not on mathematical rigor.

- **Theorem 1:** Let the horizon be denoted by H . If events $P_1 \in H$ and $P_2 \in H$, then P_1 cannot be in the causal past of P_2 .

Proof: If P_1 is in past of P_2 , then there is a timelike curve connecting the two events. Then by continuity, in the neighbourhood of P_2 there are timelike curves from P_1 which connect to P_2 . By definition the neighbourhood of P_2 is in the causal past of \mathcal{J}^+ . Therefore there exist timelike curves from $P_1 \in H$ to \mathcal{J}^+ , thereby contradicting the statement that $P_1 \in H$.

- **Theorem 2:** A horizon cannot be an entirely spacelike surface.

Proof: Consider a point P' in neighbourhood of $P \in H$. From the point P' , there are timelike curves from P' to \mathcal{J}^+ . As we move P' to P , the timelike curves have no choice but to become null, as null curves are the boundary between timelike and spacelike. Therefore the horizon is composed of spacelike and null curves.

Note that one could use Theorem 2 as a starting point to prove Theorem 1 as well.

- **Theorem 3:** Null curves cannot intersect. Consider two null curves, curve A passing through points A, O, and C; and curve B passing through the points B, O, and D, which pass through the common point O. Then there exists a timelike curve from Curve A to Curve B.

Proof: If A,B,C,D are infinitesimally close together, then the segments AO, OD are future directed null vectors. So $(AO + OD)^2 < 0$.

Therefore the null generators of the horizon do not intersect.

- **Theorem 4:** Null generators are hypersurface orthogonal.

Proof: Let ξ be the null tangent vector to the horizon. Let us consider a spacelike cut of the horizon. If ξ is not orthogonal to this cut, we have

$$\xi = h + O \tag{2.75}$$

where h lies along the horizon and $\langle O, h \rangle = 0$. Then h must be spacelike as the horizon can have only spacelike and null tangents, and a null tangent can be formed only from the addition of two spacelike. Therefore $O = \xi - h$ is also tangent to the horizon, and since it is on the horizon, it is also null. But then we have

$$\xi^2 = (h + O)^2 = \langle h, h \rangle + \langle O, O \rangle + 2 \langle h, O \rangle > 0. \tag{2.76}$$

which contradicts the fact that ξ is a null vector.

Therefore the tangent vectors are hypersurface orthogonal. There exists no component h of ξ that lies along the horizon, which has a non-zero overlap with the spacelike cut of the horizon, and hence the null generators are orthogonal to any spacelike slicing of the horizon.

2.4 Laws of black hole thermodynamics

This section will briefly describe how our discussion leads to the laws of black hole thermodynamics. For simplicity, we will start with the second law of black hole thermodynamics and then move on to the other laws.

2.4.1 The area Law

The area law is the second law of black hole thermodynamics. Recall that the horizon is made up of hypersurface orthogonal null geodesics which obey the following properties:

1. As shown previously, null rays constituting the same can not intersect by theorem 3.
2. The null rays cannot hit a singularity by assumption.

Keeping in mind these properties, we are led to demand $\theta \geq 0$ in the Raychaudhuri equation (see (2.69) and (2.73)), which essentially implies that

$$\frac{dA}{d\lambda} \geq 0, \tag{2.77}$$

i.e., the horizon's area always increases classically, provided weak energy condition is satisfied.

2.4.2 First law of black hole thermodynamics

We will now move on to the first law of black hole thermodynamics. In order to go into the same, we first outline the notion of surface gravity.

Surface gravity and Killing/Affine parametrization

Consider a stationary spacetime with a timelike Killing vector χ at asymptotic infinity. This vector χ becomes null at the horizon for the stationary black hole. We normalize at asymptotic infinity

$$\langle \chi, \frac{\partial}{\partial t} \rangle \rightarrow -1. \tag{2.78}$$

Since the Killing vector becomes a null vector at horizon, $\langle \chi, \chi \rangle = 0$. In general, if a surface is given by $f(x) = 0$, the normal to it is $\nabla_\mu f(x) = \nabla_\mu (\chi \cdot \chi)$. Then the defining equation for surface gravity is given by

$$\nabla_\mu (\chi \cdot \chi) = -2\kappa \chi_\mu. \quad (2.79)$$

Using the Killing equation, the above relation becomes

$$\chi^a \nabla_a \chi_\mu = \kappa \chi_\mu. \quad (2.80)$$

Parametrizing geodesics on the horizon using an affine parameter would have given us a vanishing result on the right-hand side of the above equation. Therefore κ measures the failure of the affine parameter to be a Killing parameter.

Now let $\chi = \frac{dx}{dv}$, where v is the Killing parameter. Set $\frac{d\lambda}{dv} = e^{\kappa v}$, and a new vector $k^\mu = \frac{dx^\mu}{d\lambda}$. Therefore we have $\lambda = \frac{1}{\kappa} e^{\kappa v}$ from $\frac{d\lambda}{dv} = e^{\kappa v}$. Using this we can see that

$$\chi^\mu = \frac{d\lambda}{dv} k^\mu = e^{\kappa v} k^\mu = \kappa \lambda k^\mu, \quad (2.81)$$

which is the standard relation between the Killing vector on the horizon with the Killing parameter and the null vector with the affine parameter. One can also show that

$$k^\mu \nabla_\mu k^\nu = 0, \quad (2.82)$$

which is expected of a null geodesic.

1st law of black hole thermodynamics

In order to focus on the physical aspects concerning black hole thermodynamics, we consider non-rotating black holes first. Let us begin by throwing some matter into the black hole. The energy current is given by

$$j_\mu = T_\mu^\nu \chi_\nu \quad (2.83)$$

and therefore the increase in mass is given by

$$\delta E = \int j_\mu dx^\mu dA = \int j_\mu k^\mu d\lambda dA = \int T_\mu^\nu \chi_\nu k^\mu d\lambda dA = \int T_{\mu\nu} (\kappa \lambda) k^\nu k^\mu d\lambda dA, \quad (2.84)$$

where the last line follows from the earlier definitions. Using the Einstein equation, we have

$$\delta E = \frac{\kappa}{8\pi G} \int R_{\mu\nu} (\kappa \lambda) k^\nu k^\mu d\lambda dA, \quad (2.85)$$

Consider the stress-energy that is thrown in in the form of a shockwave. One can locally make the shear, and the strain vanishes up to the first order. Therefore we have

$$\frac{d\theta}{d\lambda} \approx R_{cd}k^ck^d. \quad (2.86)$$

This gives us the first law of black hole thermodynamics

$$\delta E = -\frac{\kappa}{8\pi G} \int \frac{d\theta}{d\lambda} \lambda d\lambda dA = \frac{\kappa}{8\pi G} \int \theta d\lambda dA = \frac{\kappa}{8\pi G} dA \quad (2.87)$$

where in the second step the boundary terms were dropped as $\theta = 0$ at the beginning and the end of the process.

We can similarly generalize the first law to the case of rotating black holes, which we will not perform here, but the proof is straightforward.³

2.4.3 Zeroth and the third laws

In addition, there are two other laws of black hole thermodynamics as well. We will list them below with brief motivation:

- **Zeroth law:** For a stationary black hole, the surface gravity κ is constant throughout the horizon. This is the thermodynamic analog of constant temperature throughout a body.
- **Third law:** Naively, one would guess that the statement for the third law takes the form $S \rightarrow 0$ as $T \rightarrow 0$. However, this definition does not hold for a specific class of black holes. In particular, the entropy is finite for charged black holes, i.e., Reissner-Nordstrom black holes, as the temperature goes to zero, i.e., $T \rightarrow 0$, since the black hole has a finite area for extremal solutions.⁴

Therefore a better formulation of the third law is that one cannot reach $T \rightarrow 0$ through a finite process. Again this mirrors the third law of thermodynamics.

Further details of these laws and the proof of the first law for rotating black holes are given in [97].

³A rotating black hole in asymptotically flat space, or the Kerr solution, has an angular momentum J about an axis and charge M . The presence of two horizons characterizes such a solution. In addition, there exists a region called the ergosphere, where an observer is dragged so strongly that it cannot rotate around the black hole in a direction opposite to the angular momentum of the black hole. Note that there is no analog of the Birkhoff theorem for such solutions.

⁴A Reissner-Nordstrom black hole is a black hole with a charge Q and mass M . The crucial property of this black hole is that it has two horizons; as we take the limit $Q \rightarrow M$, both horizons coincide while the temperature goes to zero. The entropy, however, is still finite.

2.5 Black hole evaporation and Hawking's paradox

We will calculate the temperature of the Schwarzschild black hole here and show that a black hole solution is not a static solution but evaporates after a finite time due to quantum effects.

Originally used by Hawking, a convenient and physically illuminating way to derive the temperature of the black hole is through ray tracing. We will briefly sketch the picture here. Consider a black hole formed from the gravitational collapse of a shell. Then the evolution of the state in the exterior can be understood as the evolution from \mathcal{I}^- to $H \cup \mathcal{I}^+$, where H denotes the horizon. Let us start with the following mode expansion on \mathcal{I}^-

$$\phi = \int \frac{a_\omega}{\sqrt{\omega}} \frac{e^{ikV}}{r} Y_m(\Omega) + \text{h.c.} \quad (2.88)$$

The universe is given in the initial state $|\Omega\rangle$ such that

$$a_\omega |\Omega\rangle = 0. \quad (2.89)$$

Let the coordinate V_0 denote the last ray that gets stuck at the horizon rather than escaping from the black hole. Now modes that escape to \mathcal{I}^+ can be related to modes with support just before the last ray at V_0 by ray tracing. Appropriately mapping such rays gives rise to the ray-tracing formula

$$U^{\text{out}} = 4m + \delta V + 4m \log \frac{\delta V}{4m} - V_s \quad (2.90)$$

where V_s is the coordinate of the collapsing shell, and δV denotes the distance from the last ray at V_0 . Very close to the last ray, the relation approximately becomes

$$U^{\text{out}} = 4m \log \delta V \quad (2.91)$$

Therefore a solution of the form $e^{i\omega U}$ on \mathcal{I}^+ can be backtraced to

$$\frac{(V_0 - V)^{4i\omega M}}{4m}, \quad V < V_0 \quad (2.92)$$

using the ray tracing formula on \mathcal{I}^- . Then at late times, we can consider the mode expansion

$$\begin{aligned} \phi &= \int \frac{b_\omega}{\sqrt{\omega}} \frac{(V_0 - V)^{4i\omega M}}{4m} Y_m(\Omega) + \text{h.c.} & V < V_0 \\ &= \int \frac{c_\omega}{\sqrt{\omega}} \frac{(V - V_0)^{4i\omega M}}{4m} Y_m(\Omega) + \text{h.c.} & V > V_0 \end{aligned} \quad (2.93)$$

where we demand that the final state is in the $b - c$ vacuum. Then the transformation

between the initial and the final state is just the Minkowski to Rindler transformation which we discussed in detail before. Consequently an observer at \mathcal{I}^+ who observes only the b modes finds the state $|\Omega\rangle$ to be appearing as a thermal density matrix, whose temperature is

$$\beta = 8\pi M; \quad (2.94)$$

which is the Hawking temperature of the Schwarzschild black hole.

Note that there is a caveat to this analysis: from the ray tracing formula, an infinite region near \mathcal{I}^+ is filled from a small region near V_0 on \mathcal{I}^- . This leads to the blueshifting of rays at $O(1)$ frequencies at \mathcal{I}^+ when traced back to \mathcal{I}^- . This potentially takes us outside the domain of effective field theory.

A better derivation of the temperature is by using universal properties of the two point correlator to argue that the expectation value of the number operator is exactly thermal. More precisely, the number operator of b modes have to satisfy

$$\langle b_\omega^\dagger b_\omega \rangle = \frac{e^{-\beta\omega}}{1 - e^{-\beta\omega}} \quad (2.95)$$

in order to reproduce the correct near horizon properties of the correlator

$$\langle \phi(t, r, \Omega) \phi(t', r', \Omega') \rangle. \quad (2.96)$$

Note that a failure to reproduce the correct near horizon properties implies that the geometry is not regular near the horizon, which is not the case. Hence we infer that the Hawking temperature of the black hole is $\beta = 8\pi M$.

2.5.1 Hawking radiation

One can further use the properties of near horizon correlators to find pair production across the horizon and, consequently, show that radiation is emitted from the horizon. In order to understand this, we need to look at the correlation of right moving operators d_ω behind the horizon and right moving operators b_ω outside the horizon. Note that the annihilation operators d_ω in the black hole interior carry the opposite energy due to the flipping of space and time.

Again in order to get the correct behaviour of the two point correlator $\langle \phi(t, r, \Omega) \phi(t', r', \Omega') \rangle$ across the horizon, one needs to impose

$$\begin{aligned} \langle b_\omega d_\omega \rangle &= \frac{e^{-\frac{\beta\omega}{2}}}{1 - e^{-\beta\omega}} \\ \langle b_\omega^\dagger d_\omega^\dagger \rangle &= \frac{e^{-\frac{\beta\omega}{2}}}{1 - e^{-\beta\omega}} \end{aligned} \quad (2.97)$$

Therefore, we develop a physical understanding of Hawking radiation from the above equations. There is a pair of positive and negative particles at the horizon, and the positive particle escapes to infinity. In contrast, the negative particle falls into the black hole and decreases the mass of the black hole. Note that both particles are entangled in this description. This is the basic process through which black holes evaporate.

This statement can also be understood as the fact that there is a positive outward flux of energy at asymptotic infinity, while there is an inward flux of particles at the horizon.

Note that the quantum effects leading to Hawking radiation due to pair production violate the weak energy condition that we used to derive the second law of black hole thermodynamics. Therefore the weak energy condition used in the Raychaudhuri equation to obtain the focussing effect is no longer valid, and as a consequence, we can violate the second law of black hole thermodynamics in the presence of quantum effects, i.e., the area decreases as Hawking quanta escape the black hole.

2.5.2 Hawking's information paradox

Let us now look at a puzzle related to the black hole interior, which arises due to black hole evaporation. From the two-point function, the black hole emits information just like a blackbody⁵. Therefore the black hole loses mass at the rate

$$\frac{dM}{dt} = -\sigma AT^4 \quad (2.98)$$

Since the area of the horizon A and the temperature T of the black hole go respectively as $A \propto M^2$ and $T \propto \frac{1}{M}$, we find that the lifetime of black holes is of the order

$$t_{\text{life}} \propto M^3. \quad (2.99)$$

Now consider a black hole formed from gravitational collapse, which is a pure state. Our analysis of Hawking radiation suggests that the quanta that went out constitute a thermal bath at a temperature given by $\beta = 8\pi M$. This implies that the escaped Hawking quanta obey a thermal ensemble of uncorrelated particles, and the final state on \mathcal{I}^+ is a thermal density matrix. This leads to a picture where there is an evolution from an initial pure state to a mixed state described by a thermal density matrix.

However, this paradox is not a paradox since a large class of pure states can mimic mixed states to exponential accuracy, as described in the next section. Consequently, the above analysis of Hawking quanta is not sufficient to distinguish between a final pure state and a final mixed state.

⁵Strictly speaking greybody factors may also be involved.

2.6 Statistical properties of typical states

This section will attempt to precisely understand how certain states in a statistical ensemble, called typical states, mimic thermal behaviour. We will revisit some properties of typical states within a general formalism.

We consider a n dimensional Hilbert space such that $n = \exp S$. Let us work with an orthonormal basis of states $|\psi_i\rangle$ where $i \in [1, n]$. We will now write down the most general pure state living on this Hilbert space:

$$|\psi\rangle = \sum_{i=1}^n c_i |\psi_i\rangle \quad \text{where} \quad \sum_{i=1}^n |c_i|^2 = 1. \quad (2.100)$$

Equation (2.100) describes a sphere S^{2n-1} where all these vectors live, which we previously encountered in §3.2.1. We define the Haar measure on the pure states which guarantees that each pure state is equally likely.

$$d\mu = a dc_1 dc_1^* dc_2 dc_2^* \dots dc_n dc_n^* \delta\left(1 - \sum_{i=1}^n |c_i|^2\right) \quad (2.101)$$

Here a is fixed using the following condition:

$$\int d\mu = 1. \quad (2.102)$$

The measure $d\mu$ is invariant under independent rotations of phases $c_i \rightarrow e^{i\alpha_i} c_i$. Now consider a linear operator O acting on this Hilbert space \mathcal{H} . We want to study the properties of the expectation value $\langle\psi|O|\psi\rangle$ and also how it depends on c_i 's. We argue that for most choices of c_i , the expectation value is independent of the typical state ψ provided that n is very large. Firstly the average $\overline{\langle\psi|O|\psi\rangle}$ over all states is given by

$$\overline{\langle\psi|O|\psi\rangle} = \int d\mu \langle\psi|O|\psi\rangle = \sum_{i,j} O_{ij} \int d\mu c_i^* c_j. \quad (2.103)$$

This integral is non-zero only if $i = j$ due to invariance of $d\mu$ under independent rotations of phases $c_i \rightarrow e^{i\alpha_i} c_i$. Therefore we write

$$\int d\mu c_i^* c_j = A_i \delta_{ij}, \quad \text{where} \quad A_i = \int d\mu |c_i|^2 \quad (2.104)$$

Since all c_i 's enter the measure in an equivalent way and $d\mu$ is independent under permutations of c_i 's, the index i on A_i is redundant. Therefore $A_i \equiv A = \int d\mu |c_i|^2$. In order to evaluate A we now sum over all c_i 's in equation (2.104).

$$\overline{A} = \int d\mu \mathbf{1} = \frac{1}{n}. \quad (2.105)$$

We use the value of \bar{A} to imply that:

$$\int d\mu c_i^* c_j = \frac{1}{n} \delta_{ij} \implies \overline{\langle \psi | O | \psi \rangle} = \sum_{i,j} O_{ij} \int d\mu c_i^* c_j = \frac{1}{n} O_{ij} = \text{Tr} [\rho_\Omega O], \quad (2.106)$$

where $\rho_\Omega = \frac{1}{n}$ is the microcanonical density matrix. We thus conclude that the the average of the expectation value of operators over all typical states is that of the maximally mixed state. We now want to understand how close is the expectation value of an operator is to the maximally mixed state. In order to do this we need to look at the variance which can be similarly calculated in the following equation:

$$\overline{[\langle \psi | O | \psi \rangle - \text{Tr} (\rho_\Omega O)]^2} = \frac{1}{e^S + 1} [\text{Tr} (\rho_\Omega O^2) - \text{Tr} (\rho_\Omega O)^2] \quad (2.107)$$

We see that the variance is exponentially small in entropy. Therefore we conclude that most pure states must look exponentially close to the mixed state, or else we will obtain a more significant number in the variance. Therefore almost all states mimic thermal behaviour, which justifies our claim that almost all states $|\psi\rangle$ are typical |TYP>. As a result, we write for almost all such states:

$$\langle \text{TYP} | O | \text{TYP} \rangle = \text{Tr} [\rho_\Omega O] + \mathcal{O} \left(\frac{1}{e^S} \right). \quad (2.108)$$

An important assumption that goes into calculating the variance is that the degree of the operator k is small compared to the dimension of the Hilbert space $k \ll n$. A violation of this property leads to a more substantial variance. Therefore, we demand that the number of operator insertions is much smaller than the Hilbert space's dimension. The degree k being small provides a statistical basis for imposing this condition on operators in the small algebra and is the boundary counterpart of demanding that the backreaction due to operator insertions is small.

Now consider the case where we are looking at energy eigenkets spread over ΔE such that $\Delta E \ll E$. In this limit, the canonical density matrix approaches the microcanonical density matrix. Each typical state therefore satisfies the following property:

$$\langle \text{TYP} | O_1 O_2 \dots O_k | \text{TYP} \rangle = \frac{1}{Z(\beta)} \text{Tr} [e^{-\beta H} O_1 O_2 \dots O_k] + \mathcal{O} \left(\frac{1}{S} \right) \quad (2.109)$$

where we have replaced the microcanonical ensemble with the canonical ensemble. Here again the number of operators k is much smaller than the entropy of the state, $k \ll n$. (2.109) essentially says that k -point correlators on the typical state are indistinguishable from thermal k -point correlators up to $\mathcal{O} \left(\frac{1}{S} \right)$ corrections. This kinematical statement about the correlators is quite surprising; any typical state exhibits such behaviour.

We will clarify a physical question here. How does the typical state know about the

inverse temperature β ? This information is contained in the number of energy eigenstates comprising the typical state and the energy interval where the states live. Hence the system knows about the temperature.

The analysis in this section also addresses the question of Hawking's paradox. Since we have established that typical states are exponentially close to mixed states in an ensemble and constitute most of the states in a statistical ensemble, Hawking's calculation is not precise enough to determine that the final state is a thermal state.

2.7 The cloning and strong subadditivity paradoxes

This section will look at two closely related puzzles regarding the black hole interior: the no-cloning and the strong subadditivity paradoxes [12].

2.7.1 Cloning paradox

Consider a nice slice in a black hole spacetime that stays away from singularities and is spacelike everywhere. We call such a slice a nice slice. Now consider an old black hole that has evaporated past the Page time.⁶

We are concerned with slices that asymptote to substantially late times in the exterior. This captures almost all the outgoing Hawking radiation. Now consider some infalling matter, whose state is given by $|\psi\rangle$. Therefore, this slice has two copies of information from different regions, one from the infalling matter itself and another reconstructed from the Hawking radiation. In other words, it seems that the Hawking process leads to the following situation with two copies of the state on the nice slice:

$$|\psi\rangle \rightarrow |\psi\rangle \otimes |\psi\rangle. \quad (2.110)$$

However, the above scenario contradicts the no-cloning theorem, which follows from a fundamental assumption of quantum mechanics, i.e., the unitarity assumption. We will now describe the proof of the no-cloning theorem in quantum mechanics.

The no-cloning theorem in quantum mechanics

Consider a factorized Hilbert space $H = H_A \otimes H_B$ denoting two subsystems A and B , such that $H_A = H_B$. Using unitary operations, we want to copy a generic state $|\phi\rangle_A$ over an "ancilla" state of the subsystem B given by $|e\rangle_B$ to obtain $|\phi\rangle_B$. In other words, we

⁶The Page time as defined here is half of the black hole evaporation time. The original motivation of Page was to treat a black hole as a factorized system where we have separate interior and exterior subsystems and quantify the flow of information between these subsystems. However, as we shall argue later on, this assumption conflicts with the Gauss constraint of gravity.

want to start with the initial composite system given by

$$|X\rangle_{in} = |\phi\rangle_A \otimes |e\rangle_B \quad (2.111)$$

and show that there exists no unitary operator U ⁷ that takes us from the above initial state to the following final state

$$|X\rangle_{out} = |\phi\rangle_A \otimes |\phi\rangle_B. \quad (2.112)$$

To demonstrate the proof, we consider two arbitrary states $|\phi\rangle$ and $|\psi\rangle$ in the Hilbert space H_A . Then the unitary operator U operates on the initial state and takes it to the final state

$$|\phi\rangle_A \otimes |e\rangle_B \rightarrow |\phi\rangle_A \otimes |\phi\rangle_B \quad (2.113)$$

and similarly

$$|\psi\rangle_A \otimes |e\rangle_B \rightarrow |\psi\rangle_A \otimes |\psi\rangle_B \quad (2.114)$$

Since the operator is unitary the inner product between the two initial states and the two final states must be the same. This implies

$$|\langle\phi|\psi\rangle|^2 = |\langle\phi|\psi\rangle| \quad (2.115)$$

which implies that either $|\langle\phi|\psi\rangle| = 1$ or $|\langle\phi|\psi\rangle| = 0$. But this implies that the states $|\phi\rangle$ and $|\psi\rangle$ in the Hilbert space H_A are specially selected and not arbitrary. Hence such an operator U does not exist for a generic state.

One can also write down the representation of the unitary operator on the two-dimensional subspace of initial and final states, and using unitarity show that such an operator has a particular action and does not exist for a generic state.

Why the paradox may be ill posed

Note that this paradox cannot be resolved through small corrections of $O(\sqrt{G_N})$ and is an $O(1)$ paradox.

In this regard, one might state the following reasons why the paradox can potentially be ill-defined:

1. The paradox is difficult to observe since one needs to collect all the escaped Hawking radiation and then perform a quantum computation of the radiation. Then in order for the infalling observer to violate the no-cloning theorem, the details of this quantum computation have to be sent to the infalling observer, which then has to perform a measurement that demonstrates the violation of quantum mechanics.

⁷which in our case is the unitary time evolution operator

Such a violation is seemingly hard in practice.

2. There exists a firewall at the horizon.

Even though the measurement of violation may be hard, a copy of the information on the nice Cauchy slice still exists. In addition, firewalls imply a breakdown of effective field theory, especially the two-point correlator across the horizon, which a priori seems unnecessary. We will not further look into the firewall story in our work for these reasons.

2.7.2 Strong subadditivity paradox

This paradox is closely related to the cloning paradox, as described earlier. We will first explain the statement of strong subadditivity, which is a fundamental theorem involving Von Neumann entropies of subsystems.

Consider a tripartite system's Hilbert space denoted by

$$H = H_A \otimes H_B \otimes H_C. \quad (2.116)$$

The strong subadditivity theorem states that

$$S_{ABC} + S_B \leq S_{AB} + S_{BC} \quad (2.117)$$

Stated a different way by adding and removing a system D on this tripartite description, one has the following relation

$$S_{AB} + S_{BC} \geq S_A + S_C. \quad (2.118)$$

We will be concerned with this above statement of strong subadditivity in our work.

The paradox

We look at three different regions given a black hole.

1. Region C just inside the horizon.
2. Region B just outside the horizon
3. Region A which is the rest of the slice outside the horizon.

In our discussion, we will neglect the entropy of the vacuum and will subtract this leading correction. For our discussion, we assume that the entropy of region A follows the Page curve, i.e., it attains a maximum value at the Page time t_{Page} and gradually reduces afterward till the black hole completely evaporates. We consider the situation where we are after the Page time.

All the three regions A, B, and C have Hawking quanta in them, which dynamically shifts between these three regions, i.e., the quanta in Region B move out to region A after a certain time interval. As a result, we can write

$$S_A(t + \delta t) = S_{AB} \quad (2.119)$$

where δt is the light crossing time for Region B. Then it trivially follows that

$$S_{AB} < S_A. \quad (2.120)$$

Next the Hawking radiation in Region B and C are highly entangled. This means that

$$S_{BC} < S_B \approx S_C. \quad (2.121)$$

Both B and C are approximately thermal when viewed independently but are almost pure when viewed together. As a consequence, adding S_{AB} and S_{BC} up we obtain

$$S_{AB} + S_{BC} \leq S_A + S_C. \quad (2.122)$$

This indicates a violation of strong subadditivity as quantitatively posed in equation (2.118). The violation of strong subadditivity is a statement about the violation of monogamy of entanglement: If Region B is entangled with Region A, then it cannot be entangled with Region C. Here Region B is entangled with Region A due to unitary evolution and with Region C due to the smoothness of the horizon.

Just as in cloning paradox, the strong subadditivity violation is an $O(1)$ violation and is closely related to it as well. The fact that B is entangled with A and C is similar to the paradox that quantum information is cloned on A and C.

2.7.3 Resolutions

One trivial way these paradoxes are resolved is due to the appearance of structure at the horizon, i.e., fuzzball/ firewall kind of proposals. However, we focus on a more plausible aspect that does not deal with the breakdown of effective field theory at the horizon.

The essential feature here which resolves the paradox is the fact that in gravity, the regions A, B, and C are not independent, i.e. the decomposition assumption

$$H = H_A \otimes H_B \otimes H_C. \quad (2.123)$$

is incorrect. This essentially gives rise to the above-mentioned $O(1)$ violation.

2.8 Black hole complementarity

The key idea of black hole complementarity is that different observers observe physics differently, with none spotting a violation, i.e., one can have two alternative descriptions of the black hole [14, 15]. These descriptions can be listed as follows, and the goal of this section is to explain these statements in detail:

1. An outside observer who can think of a "membrane" at the stretched horizon and consider unitary evolution. Here the stretched horizon is defined as

$$r = r_h + \frac{l_P^2}{r_h}. \quad (2.124)$$

The above equation implies that the stretched horizon area is l_P^2 larger than the event horizon.

2. An infalling observer who sees a smooth horizon but cannot verify unitarity since they perish into the singularity before being able to perform any such experiment.

As we will show later on, one can formulate a stronger statement of black hole complementarity using the non-factorization mentioned above of the Hilbert space while considering high point correlators.

2.8.1 Thought experiments involving complementarity

We will briefly review the approach of [14, 15] here, who considered thought experiments to show that no observer can detect complementarity without knowing about Planck scale physics.

Existence of a stretched membrane

Consider an infalling observer who sends messages reporting the smoothness of the horizon to an asymptotic observer. The redshift factor here is given by

$$\omega_\infty = \sqrt{1 - \frac{r_h}{r}} \omega_{in} \quad (2.125)$$

which at the stretched horizon becomes

$$\omega_\infty = \sqrt{1 - \frac{r_h}{r_h + \frac{l_P^2}{r_h}}} \omega_{in} = \frac{l_P}{r_h} \omega_{in} \quad (2.126)$$

The asymptotic observer is in a bath of radiation at temperature $T \propto M^{-1}$. To be able to read a message the outside observer requires a frequency

$$\omega_\infty > \frac{1}{r_h} \implies \omega_{in} > \frac{1}{l_P}. \quad (2.127)$$

Consequently, beyond the stretched horizon, the infalling observer has to use Planckian frequencies to communicate the absence of a membrane to distinguish the signal from the bath.

Alternatively, one can say that the acceleration required by an observer to go up to the stretched horizon and then escape outside is of the order

$$a \sim \frac{1}{l_P} \quad (2.128)$$

which again requires the knowledge of Planckian physics to confirm the existence of a stretched membrane.

Is cloning possible?

Let us study whether the infalling and the exterior observer who falls in later can detect information cloning.

Consider a situation where the first infalling observer crosses the horizon at $V = 1$. The second observer waits for a time longer than the Page time, decodes the Hawking radiation, and then jumps in at a late enough time, say of the order

$$V = \exp\left(\frac{t_{Page}}{2r_h}\right) = \exp r_h K \left(\frac{r_h}{l_P}\right) \quad (2.129)$$

where K is a numerical constant. Recall that the singularity is at $UV = 1$. Consequently the first observer has to send a message before

$$U = \exp\left(-\frac{t_{Page}}{2r_h}\right) \quad (2.130)$$

in order for the second observer to receive the same. It can be shown that the maximum proper time for the first observer to send this signal is

$$r_h \exp\left(-\frac{r_h^2}{l_P^2}\right) \quad (2.131)$$

which is simply the fact that the first observer has to use Planckian frequencies to send the signal. In general, as long as the later observer jumps after

$$t = r_h \log \frac{r_h}{l_P} \quad (2.132)$$

the first observer has to use Planckian frequencies.

Preskill Hayden time

Consider the following question: Suppose a black hole has evaporated past the Page time, and someone collects all the radiation till that time. Now, if you throw new information into the black hole, how long will it take before it emits the same?

It was argued in [19] that the black hole will emit this radiation in the time

$$t = r_h \log \frac{r_h}{l_P}. \quad (2.133)$$

The key idea in [19] is that the time it takes for the black hole to collect the information about the diary is the same time that it takes for the black hole to scramble the information of the diary over its degrees of freedom, which is a much shorter time scale as compared to the evaporation time scale.

This time scale is reminiscent of the time scale encountered in the cloning process and, more generally, is the threshold bound for black hole complementarity. In a certain sense, the existence of this time scale demonstrates that it is impossible to obtain the cloning of quantum information without going to Planckian frequencies.

2.8.2 Loss of locality in high-point correlators

The description here follows the analysis of [29–33]. Previously, the focus was not on how to read information in Hawking radiation using local observables. The focus of this subsection is that the exterior and interior regions are not independent. More precisely, on a nice slice, the statement of complementarity that resolves cloning paradox (and also has implications for the strong subadditivity paradox) is given by

$$\phi(x^{in}) = P(\phi(x_1^{out}), \phi(x_1^{out}), \dots, \phi(x_1^{out})) \quad (2.134)$$

Note that in ordinary quantum field theories such as scalar field theory such a relation is manifestly wrong since causality implies that for spacelike points x_i and x_j , with $x_i \neq x_j$ on a slice

$$[\phi(x_i), \phi(x_j)] = 0. \quad (2.135)$$

The roots of this statement lie in the fact that there is a Gauss constraint in gravity. A qualitatively similar statement can also be displayed in the case of gauge theories. We will further look into such issues in §4.

2.9 Interior operators in the state-dependent formalism

In this section, we review the construction of state-dependent operators describing modes behind the horizon [29–33]. This construction is similar to both pure and eternal black holes. For eternal black holes, the CFT observables mean the right CFT’s observables, which are our objects of interest.

We start with the black hole state $|\psi\rangle$ whose average energy is given by E . Firstly an algebra \mathcal{A} ⁸ is generated by *simple operators* which are defined as operator polynomials of degree n such that $n \ll \mathcal{N}$, where \mathcal{N} is the central charge of CFT ($\mathcal{N} = N^2$). We will call this algebra the *small algebra*. The small algebra is associated with the state $|\psi\rangle$ and omits the Hamiltonian. We exclude the Hamiltonian because we do not want to include any annihilation operators in the algebra, and the Hamiltonian annihilates the state, i.e., $(H - E)|\psi\rangle = 0$. However, we want the algebra to be approximately closed under time evolution.

Keeping these aspects in mind, we construct and work in the *small Hilbert space*, which is obtained by the action of these simple operators on the state $|\psi\rangle$.

$$\mathcal{H}_{|\psi\rangle} := \mathcal{A}|\psi\rangle \quad (2.136)$$

By doing this, we have laid out the basic framework in order to derive various axiomatic/algebraic QFT results. Note that this construction does not involve a doubling of the Hilbert space. In particular, provided the algebra \mathcal{A} is a Von Neumann algebra, we can derive the Tomita-Takesaki theorem, which constructs a commutant algebra \mathcal{A}' for us.

To state the theorem, we firstly define the following antilinear map from $S : H_{|\psi\rangle} \rightarrow H_{|\psi\rangle}$ and $O \in \mathcal{A}$.

$$S O |\psi\rangle = O^\dagger |\psi\rangle \quad (2.137)$$

We now decompose the operator S as $S = J\Delta^{1/2}$, where J is an anti-unitary operator and Δ is Hermitian. Consequently we have $S^\dagger S = \Delta$. The Tomita-Takesaki theorem says that there exists a commutant algebra $\mathcal{A}' \equiv J\mathcal{A}J$, with the property that operators $\tilde{O} \in \mathcal{A}'$ defined by $\tilde{O} = JOJ$ commutes with all elements $O' \in \mathcal{A}$

$$[\tilde{O}, O'] = 0. \quad (2.138)$$

⁸The picture to keep in mind is that of a Von Neumann algebra, which is basically a set of bounded operators closed under certain operation. More precisely, in quantum gravity, we do not deal with Von Neumann algebras, but rather a subspace of bounded operators (constructed using smearing), which approximates a Von Neumann algebra to a good degree. This is because the space of bounded operators is not closed under multiplication of operators, and one can go out of this space using repeated multiplication.

Since Δ is Hermitian, we express it as $\Delta = \exp\{-K\}$, where K is defined as the modular Hamiltonian for the algebras \mathcal{A} and \mathcal{A}' generating the Hilbert space $\mathcal{H}_{|\psi\rangle}$, and is expressed in terms of the antilinear operator S as:

$$K = -\log S^\dagger S. \quad (2.139)$$

Our job now is to construct the precise form of the modular Hamiltonian and the tilde operators. To construct these, we will apply the above construction to a system with Hamiltonian H acting on the state $|\psi\rangle$. As given in [98], the modular Hamiltonian up to the leading order in N takes the form:

$$K = \beta(H - E) + \mathcal{O}\left(\frac{1}{N}\right) \quad (2.140)$$

where E is the average energy of the state ψ on which it is acting. Therefore to the leading order in N one can give a precise form for the \tilde{O} operators. Using the definition given in equation (2.137) with the definitions given by $S^\dagger S = \Delta$ and $\tilde{O} = J O J$, we can write down the action of the \tilde{O} operators on the Hilbert space:

$$\tilde{O}(\omega) O' |\psi\rangle = O' e^{-\beta\omega} O^\dagger(\omega) |\psi\rangle \quad \& \quad [H, \tilde{O}(\omega)] O' |\psi\rangle = \omega \tilde{O}(\omega) O' |\psi\rangle. \quad (2.141)$$

The state-dependent operators \tilde{O} describing the right moving modes in the interior are thus constructed in the above fashion. The commutant algebra allows us to impose causality and locality between the interior and the exterior operators. The role of modular operators is to push the excitations behind the horizon. The unique feature of this construction is that this follows naturally for any well-defined quantum field theory, provided the algebra of simple operators satisfies the requirements mentioned above.

Chapter 3

The bags of gold paradox

Overview of results

Following our initial motivation and basic tools, we will now overview our results quickly. §3.1 poses the paradox discussed above for eternal black holes in detail. §3.2 discusses our proposed resolution, where we also determine the maximum number of vectors that can be fit inside a Hilbert space with small inner products. In §3.3, we demonstrate that the paradox does not show up in the fine-grained entropy of the CFT. From the CFT perspective, the action of state-dependent operators on the state of the black hole generates the interior bulk states in our construction. We show that the bulk state produced by the action of interior operators on the thermofield double state [40, 99, 100] does not lead to any change in the Von Neumann entropy of the CFT. We also calculate the fine-grained entropy using quantum extremal surfaces for the eternal black hole. These surfaces do not enter the black hole interior and, therefore, do not capture our interior excitations. Consequently, there is no paradox in the dual CFT. These observations strongly support our claim that the interior states arise due to overcounting and are not independent excitations in quantum gravity.

In the bulk description, it is essential to understand the behaviour of the interior excitations. From the CFT perspective, our excitations appear to be in equilibrium when probed using simple operators in the right side CFT. However, they are out of equilibrium when probed with operators belonging to the complement of the *small algebra* of *simple operators*¹. These properties of the excitations lead to a bulk picture of the excitations arising from the left-past horizon of the eternal black hole and travelling through the left side. Afterwards, they fall into the left future horizon where they go on and intersect the nice slices. The above nature of the excitations physically demonstrates the paradox in Figure 3.1 where different excitations come out of the left horizon at particular times governed by the unitary operator $U(t)$. The initial bulk state of the excitations on the

¹See §2.9 for the definition of simple operators, the small algebra and its complement

black hole is a Euclidean black hole glued to the Lorentzian geometry [101, 102]. Here the excitations are generated using operators at the Euclidean AdS boundary (See Figure 3.2).

We estimate that two excitations placed far apart on the nice slices of single-sided black holes have an overlap larger than $O(e^{-\frac{S}{2}})$. Such an overlap strongly backs our resolution involving small inner products and is the topic of §3.5. We discuss how the treatment of bags of gold excitations using effective field theory violates black holes' expected spectral properties in §3.6. We provide some toy examples of bags of gold configurations there, which violate the qualitative and quantitative features of spectral form factor and energy level spacing distribution. We also argue how our resolution fixes these issues. Next, we explicitly demonstrate that there can be a large number of excitations living in the black hole interior using toy models in §3.7. These toy models are small N matrix models in which we first construct a typical state [103–105] in order to model single-sided black holes. We then use the typical state and the small algebra to construct the *small Hilbert space*² describing interior bulk excitations. Random combinations of operators living on this small Hilbert space gives rise to smeared bulk excitations. We see that the small Hilbert space can embed a large number of states having small inner products with each other. We then construct states resembling excitations placed far apart from each other on the Cauchy slice in these matrix models. These states have small inner products, thereby confirming our resolution discussed in §3.5.

It is natural to ask why such an overcounting does not occur for quantum statistical systems and is special to black holes. Consider a statistical system which has a Hilbert space of dimension m . One can apply our resolution to this system and ask whether this system has a much smaller dimension n , with $m \gg n$. While we can kinematically pose such a statement, such a situation leads to discrepancies in thermodynamic observables. Another consequence of such a modelling is that forbidden quantum state transfers can occur in the larger system modelled with n vectors. We discuss these issues in §3.8.

3.1 The bags of gold paradox for the eternal black hole

This section will outline the construction of the maximal volume surfaces. Afterwards, we will place excitations on these slices. Lastly, we will pose and discuss the paradox in detail.

²See §2.9 for a review of the state-dependent formalism and associated definitions. Loosely speaking this is also known as the code subspace.

3.1.1 Maximum volume slices in the interior

Consider an eternal black hole at boundary time t as in Figure 1.1. We want to construct nice slices which stay away from singularity everywhere and possess the maximum volume for a given boundary time t . We will work with the AdS Schwarzschild metric in $d + 1$ dimensions is given by

$$ds^2 = -\frac{4f(r)}{f'(r_h)^2} e^{-f'(r_h)r^*} du_k dv_k + r^2 d\Omega_{d-1}^2, \quad (3.1)$$

where $f(r) = r^2 + 1 - \frac{C}{r^{d-2}}$, r_h is the black hole horizon and r^* is the tortoise coordinate. The subscript k denotes Kruskal coordinates. Our goal is to show is that the interior's volume grows as we increase the boundary time t .

Since the paradox involves only the interior, a demonstration of the growth of the interior volume will be sufficient for our purposes. Instead of parametrizing the slices with the boundary time, we will parametrize them using the Kruskal coordinates $(u_0, 0)$ on the left horizon and $(0, v_0)$ on the right horizon, as shown in Figure 3.1. Thus we change our problem to a similar one where we compute the maximum volume of slices which end at $(u_0, 0)$ on the left horizon and $(0, v_0)$ on the right horizon. This problem has two advantages. We see the first advantage of calculating the maximal volume surfaces in the case of single-sided black holes in §3.5. These black holes possess the entire interior region but do not have a boundary time on the left. Therefore we can utilize this construction of maximum volume slices for the single-sided case. This problem also overcomes the problem of infinite exterior volumes³.

We set $u_0 = v_0$ using the isometry of AdS spacetime. It is convenient to use the infalling Eddington-Finkelstein coordinate $v = t + r^*$ in order to calculate the maximum volume surfaces.

$$ds^2 = -f(r) dv^2 + 2dr dv + r^2 d\Omega_{d-1}^2 \quad (3.2)$$

Note that v here is different from the Kruskal coordinate v_k . We define an affine spacelike parameter σ to parametrize the nice slice. We now need to extremize the following volume integral to obtain the maximum volume of these surfaces.

$$V = V_{d-1} \int d\sigma r^{d-1} \left(-f(r) \dot{v}^2 + 2\dot{r} \dot{v} \right)^{\frac{1}{2}}, \quad (3.3)$$

where V_{d-1} is the volume of the $(d - 1)$ spherical ball. We end up with the following expression for the volume [106–108]:

$$V = \frac{\beta A(r_{min})}{2\pi} \log u_0 + O(1) \quad (3.4)$$

³though this can also be tamed by introducing a boundary cutoff.

where $A(r_{min})$ and $O(1)$ are terms of order one which do not grow with u_0 . In equation (3.4), r_{min} is determined using

$$f(r_{min}) r_{min}^{2(d-1)} + E^2 = 0, \quad (3.5)$$

where E is a conserved quantity $E = -\frac{\partial L}{\partial \dot{v}}$ with L denoting the integrand of equation 3.3. The volume extremization, derivation of the resulting equation (3.4) and $A(r_{min})$ are calculated in Appendix 6.3. The important observation here is that the interior volume of the nice slice increasingly grows with the Kruskal time. The physical reason is that the wormhole grows larger and larger with Kruskal time.

3.1.2 Placing semiclassical excitations on the nice slice

Since the volume of the nice slice in the interior keeps increasing with the Kruskal time, the interior can accommodate an increasingly large number of semiclassical excitations far apart from each other such that their spatial overlaps with each other are zero. At late times the slice's volume goes to infinity, and therefore a high number of excitations can be placed far apart from each other. These interior excitations are created by acting with unitary operators on the right CFT in the thermofield double state. Eqn. (3.6) describes an interior excitation,

$$|\psi_1^e\rangle = C_1 e^{-\frac{\beta H_R}{2}} U_1(O(t_1)) e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle. \quad (3.6)$$

Here $U(O(t_1))$ is an unitary operator acting on the right CFT, C_1 is the normalization constant, H_R is the right CFT's Hamiltonian. Here $O(t_1)$ is can be thought of as a simple operator, i.e. a smeared single trace operator of low conformal dimension. The state $|\psi_e\rangle$ represents our excitation. These states are motivated by the state-dependent formalism, which we review in §2.9⁴. The unitary operator U_1 controls the position of these excitations on the slice. This control results due to the timelike coordinate t in the exterior becoming a spacelike coordinate in the interior. We now create another excitation in the interior:

$$|\psi_2^e\rangle = C_2 e^{iH_R T} e^{-\frac{\beta H_R}{2}} U_2(t_2) e^{\frac{\beta H_R}{2}} e^{-iH_R T} |\text{TFD}\rangle. \quad (3.7)$$

⁴Since the unitaries that we are looking at are of the form $U(O(t_0))$, indeed one can express them in the form $U(O(t_0)) = e^{iPO(t_0)}$. Since this is an exponentiated operator, it is not really a monomial constructed out of creation operators. Still in a certain sense, the above operators are *motivated* by the state dependent picture, since the modulation by factors of $e^{\pm\frac{\beta}{2}}$ pushes them inside the interior. The state dependent formalism was invoked since a similar picture is arises there as well. Otherwise they could be treated as independent excitations on their own. Thus even though these operators do not have the standard form of being a monomial of creation operators, the reason why they describe (non-equilibrium) insertions behind the horizon is essentially the same as the state dependent construction.

The action of the Rindler Hamiltonian using factors of $e^{-iH_R T}$ spatially separates this second excitation from the first one. Since the exterior timelike coordinate becomes spacelike in the interior, these excitations are placed far apart from each other if T is large enough. We now generate m number of such excitations similarly, with each excitation placed far apart from the previous one as a result of modulating with the factor e^{-iHT} , where T denotes the time difference between consequent excitations. We will discuss the nature of these excitations in more detail in §3.4. Therefore a physical picture of placing the excitations on the nice slice is as follows: Consider an excitation created at time t_1 . This excitation proceeds into the interior of the black hole and intersects the nice slice. Other excitations are created at time t_2 , and so on at t_3, t_4 up to t_m .

Physical properties of the excitations

We now demand certain physical properties which these excitations should satisfy. We generate the excitations such that the backreaction is very small as compared to the mass of the black hole. If we have m excitations each having energy of the order of E_0 , then the condition for preventing backreaction is given by

$$m E_0 \ll M_{BH}. \quad (3.8)$$

We will ensure that the density of excitations $\rho = \frac{m}{V}$ is a finite quantity in the thermodynamic limit, i.e. with m and V large. Fixing the density allows us to calculate the entropy of these excitations in the effective field theory approximation by treating the system as a "dilute gas" of excitations living on the nice slice of the black hole. We also want that the separation between any two excitations is quite more substantial than the smearing time scale δt associated with each excitation which leads to the following condition.

$$\delta t_i \ll |t_i - t_j|, \quad \forall j \neq i \quad (3.9)$$

We also impose an IR cutoff for the excitations which restricts them completely to the interior of the black hole. In the late time limit, we demand that the excitations have a length scale shorter than the volume of the black hole divided by the number of excitations, which gives rise to the following bound:

$$\frac{V}{m V_{d-1}} \gg \frac{1}{E_0}. \quad (3.10)$$

where V_{d-1} is the volume of the unit spherical ball as defined previously. Thus our construction defines a "dilute gas" of excitations living in the black hole interior, such that each of these excitations has zero spatial overlap with the others. We will clarify further details regarding the physical behaviour of the excitations in the bulk in §3.4.

3.1.3 The paradox in the bulk

We will now roughly calculate the entropy of the "dilute gas" of excitations in the bulk interior using the microcanonical ensemble, assuming that our excitations behave *classically*⁵. Let E denote the total energy of the configuration. The volume Σ_p of a shell with uncertainty Δ_E centred about E in the momentum space is given by:

$$\Sigma_p = \frac{\sqrt{m}}{(m-1)!} E^{m-1} \Delta_E. \quad (3.11)$$

Using this we calculate the volume of the phase space spanned by the gas.

$$\Omega(E, V, m) = \frac{V^m}{m!} \frac{\sqrt{m}}{(m-1)!} E^{m-1} \Delta_E. \quad (3.12)$$

The phase space volume enables us to calculate the entropy of the ensemble. We use Stirling approximation and ignore the subleading terms in m . Finally, the expression for entropy with $\rho = \frac{m}{V}$ is obtained to be

$$S(E, V, m) = m \log \frac{V E}{m^2} = m \log \frac{V E_0}{m} = V \rho \log \frac{E_0}{\rho}, \quad (3.13)$$

where $E = mE_0$, with E_0 being the average energy of a single excitation. Since we have imposed that the density ρ is a finite non-zero quantity, (3.13) indicates that the entropy scales as the volume. This scaling gives rise to the paradox that the entropy of the dilute gas is larger than the Bekenstein-Hawking entropy of the black hole at late times.

3.2 The resolution: Overestimation of the Hilbert space's dimensionality

The reason why the paradox arises is due to a colossal overcounting of the bulk Hilbert space. In our construction, we ensured that the semiclassical excitations have zero spatial overlap, which is sufficient for two different excitations to be independent in effective field theory, i.e. with a vanishingly small inner product. This section motivates why the above assertion is not correct in quantum gravity and demonstrates that we can embed many more vectors in a Hilbert space with small inner products than given by the dimension of the space. Some results in this section were also discussed in unpublished notes in [109].

We first review why semiclassical gravity predicts that the inner product between two vectors in the Hilbert space can be arbitrarily small. Afterwards, we will look at why such a prediction does not hold true in quantum gravity.

⁵By classically, here the assumption essentially means to perform such a counting while ignoring particle production and related effects which may spoil the estimation.

Inner products in semiclassical gravity

We will follow the work of [33] here in order to compute the inner product between semiclassical states. We work with a background metric $g_{\mu\nu}^0(x)$ in $d + 1$ dimensions, and consider small linearized fluctuations $g'_{\mu\nu} = g_{\mu\nu}^0 + \sqrt{8\pi G_N} \delta g_{\mu\nu}$ about it. In general these linearized fluctuations can be expressed in terms of creation and annihilation operators

$$\delta g_{\mu\nu}(x) = \sum_i \sum_k a^i(k) g_{\mu\nu}^i(k) + \text{h.c.} \quad (3.14)$$

where i denotes the $\frac{(d+1)(d-2)}{2}$ polarizations and k goes over the momenta. We choose the functions $g_{\mu\nu}^i(k)$ such that the creation and annihilation operators obey the same commutation relations for a simple harmonic oscillator. We will look at the coherent states formed by the action of the creation operators which creates the excited spacetime:

$$|\alpha\rangle = C_\alpha e^{\sum_i \sum_k a_i^\dagger(k) \alpha(k)} |0\rangle, \quad (3.15)$$

where C_α is the normalization constant and $|0\rangle$ is the vacuum such that $a_i(k) |0\rangle = 0$. The expectation value of the metric operator on a coherent state $|g^{\text{cl}}\rangle$ gives us the classical value of the metric:

$$g_{\mu\nu}^{\text{cl}} = \langle g^{\text{cl}} | \hat{g}_{\mu\nu}(x) | g^{\text{cl}} \rangle. \quad (3.16)$$

We now consider the inner product between the background spacetime and the excited spacetime, such that the two spacetimes are "distant" in the phase space. Here "distant" means a substantial classical perturbation $\delta g_{\mu\nu} \sim \frac{\Delta}{\sqrt{8\pi G_N}} = \Delta \mathcal{N}$, where \mathcal{N} is the central charge of the CFT ($\mathcal{N} = N^2$ for gauge theories with gauge group N). For small linearized fluctuations, we set $\Delta \sim \text{O}(1) \ll \text{O}(\mathcal{N})$ such that $\Delta \ll 1$, which still allows us to do linearized perturbations while not being vanishingly small. As shown in [33], the semiclassical inner product between the two bulk states is given by

$$\langle g_{\mu\nu}^0 | g_{\mu\nu}^{\text{cl}} \rangle = \exp\left\{-\mathcal{N} v(g^0, g^{\text{cl}})\right\} \quad (3.17)$$

where $v(g^0, g^{\text{cl}})$ is an $\text{O}(1)$ quantity. Thus we conclude that the inner product between two different semiclassical excitations can be arbitrarily small. This is a feature common to a QFT, coherent states corresponding to quite different classical excitations can have a vanishingly small overlap.

Inner products in quantum gravity from the CFT description

Using the dual CFT description, we will see why the analysis in the preceding subsection is misleading when the phase space "distance" between the classical configurations becomes large. Contrary to the semiclassical indication, the inner product between two different

vectors might be a small but finite number even if the classical description is completely different [32, 33, 110]. A simple example is the overlap between two factorized AdS spacetimes and the thermofield double, which are very different classical configurations. These two have an overlap given by:

$$\langle 0, 0 | \text{TFD} \rangle = \frac{1}{\sqrt{Z(\beta)}}, \quad (3.18)$$

which is small but nonvanishing. The physical basis behind this small overlap is the following: the semiclassical inner product is obeyed only up to a particular "distance" in the phase space between two different classical configurations. Beyond this distance, inner products are saturated and differ from the semi classical inner product.

An example of this saturation is given by "time-shifted states" in the CFT [32, 33], which represent different bulk configurations. Consider the time shifted state given by time evolution on the left CFT acting on the thermofield double:

$$|\psi_T\rangle = e^{iH_L T} |\text{TFD}\rangle. \quad (3.19)$$

On the thermofield double consider $m = e^S$ distinct time shifted states each shifted by a time $(T_1, T_2 \dots T_m)$. Now there exist a solution for α_i 's given in the following equation:

$$\left| |\text{TFD}\rangle - \sum_{i=1}^{e^S} \alpha_i e^{iH_L T_i} |\text{TFD}\rangle \right|^2 \geq O(e^{-N}). \quad (3.20)$$

This leads to the inner products developing a saturated "fat tail" of magnitude larger than $O(e^{-N})$ which is our primary motivation for overcounting. This shows that these bulk states are not really independent of each other.

We give another proof of the presence of small inner products from the CFT description in §3.5.2 for single-sided black holes. Given a CFT dual to a single-sided black hole, we will show that two far apart excitations have an inner product larger than $O(e^{-\frac{S}{2}})$, which serves as the basis for overcounting in the single-sided black holes.

The fundamental reason why this saturation of inner products happens in gravity is an obstruction to the lifting of classical observables living on the phase space to the Hilbert space. The d -metric and its canonical conjugate momentum in the $d + 1$ ADM decomposition cannot be naively lifted to well-defined operators on the Hilbert space, as they give rise to the semiclassical inner product. Apart from these examples, there also exist other cases where the inner product in effective field theory receives small corrections in quantum gravity. This "fat tail" is similar to the "spectral form factor" in [49]. Another striking example is the statement that two states in quantum gravity might turn out to be the same [111].

3.2.1 How many bulk excitations can we possibly have?

We saw in the preceding subsection that all distinct bulk excitations are not independent of each other. Since the inner products saturate, taking excitations far apart would not make them independent. With this motivation, it becomes a natural question to ask how many bulk excitations can we fit inside a Hilbert space of dimension n .

This question has a profound consequence: a black hole with coarse-grained entropy S_{BH} can still have a vast number of bulk excitations living on the nice slices, and hence there is no paradox.

How many vectors can we fit inside a Hilbert space of dimension n ?

We consider the following problem: In a Hilbert space \mathbf{H} of dimension n , what is the maximum number $m_n(\epsilon) \equiv m$ of vectors $\{v_i\}$ which satisfy the following relations:

$$\langle v_i | v_i \rangle = 1 \quad \& \quad |\langle v_i | v_j \rangle| \leq \epsilon, \quad i \neq j; \quad \epsilon \geq \frac{1}{\sqrt{n}}. \quad (3.21)$$

We have $m_n(0) = n$ trivially. The solution to this problem is as follows. Unit vectors in the Hilbert space live on the surface of an $(2n - 1)$ dimensional real sphere. We can fix one vector to be $|v_1\rangle = (1, 0, \dots, 0, 0)$. The remaining vectors $|v_i\rangle = (a_1, a_2, \dots, a_{n-1}, a_n)$ will satisfy the following equation,

$$|a_1|^2 + |a_2|^2 + \dots + |a_{n-1}|^2 + |a_n|^2 = 1. \quad (3.22)$$

For $i \neq 1$, (3.21) implies that $|a_1|^2 \leq \epsilon^2$. Therefore around a vector $|v\rangle$, there is an exclusion zone where there can be no other vector. The boundary of this region is given by

$$|a_2|^2 + \dots + |a_{n-1}|^2 + |a_n|^2 = 1 - \epsilon^2. \quad (3.23)$$

Since $a_k \in \mathbb{C}$, we write $a_k = c_k + id_k$, where $c_k, d_k \in \mathbb{R}$. We perform the worst-case estimate of the number of vectors by assuming all the inner products are of the order ϵ . We obtain the naive estimate for the number of vectors that satisfy the inner product bounds by dividing the surface area of the $2n - 1$ dimensional real sphere (since $|a_1|^2 \sim \epsilon^2$) with the area of the exclusion zone. The exclusion zone for each vector has the radius $\frac{\epsilon}{2}$. Therefore each sphere will have the volume given by

$$\sum_{k=2}^n c_k^2 + d_k^2 = 1 - \frac{\epsilon^2}{4}. \quad (3.24)$$

A more accurate computation would also require the packing fraction of such exclusion zones. One can then count the number of vectors and multiply it by the packing ratio to approximately get the highest number of vectors.

$$m \approx \frac{P S_{2n-1}}{V_{2n-2}} \left(\frac{1}{1 - \frac{\epsilon^2}{4}} \right)^{2n-2} = 2\pi P \left(1 - \frac{\epsilon^2}{4} \right)^{-2n+2} \quad (3.25)$$

Here S_{2n-1} is the surface area of the $(2n-1)$ dimensional sphere, the volume enclosed by the $(2n-3)$ dimensional sphere is given by V_{2n-2} and P denotes the constant of proportionality which gets contribution from the packing fraction and also takes into account small errors which may have resulted from our rough counting method. We have also used $\frac{S_{2n-1}}{V_{2n-2}} = 2\pi$. Let us have a look at the function $\left(1 - \frac{\epsilon^2}{4}\right)^{-2n+2}$. We are interested when n becomes very large. Now using the definition of the exponential function we obtain

$$\lim_{n \rightarrow \infty} \left(\frac{1}{1 - \frac{\epsilon^2}{4}} \right)^{2n-2} = \lim_{n \rightarrow \infty} \left(1 - \frac{n\epsilon^2}{4} \right)^{-2n+2} \approx \lim_{n \rightarrow \infty} \left(1 - \frac{n\epsilon^2}{4} \right)^{-2n} = e^{\frac{n\epsilon^2}{2}}. \quad (3.26)$$

Note that the above expression is valid for any value of ϵ , including our case where $n\epsilon^2 \gg 1$. We evaluate the value of m in the limit of large n to be

$$m \approx 2\pi P e^{\frac{n\epsilon^2}{2}}. \quad (3.27)$$

Since our small inner products in question are very close to zero, i.e. $\epsilon \approx 0$, we fix the proportionality constant $2\pi P$ in the case when $\epsilon = 0$, which sets $2\pi P = n$. Therefore the formula describing maximum possible vectors for small ϵ is given by

$$m \approx n e^{\frac{n\epsilon^2}{2}}. \quad (3.28)$$

We pause here to reflect upon what our formula in (3.28) tells us. With tiny inner products $\epsilon \ll 1$ such that $\epsilon > e^{-\frac{S}{2}}$, we can obtain an extremely enormous overcounting of the Hilbert space. As an example consider all inner products $\epsilon \sim e^{-\frac{S}{4}}$. The maximum number of states with such a small inner product that can be embedded in the Hilbert space of dimensionality e^S is given by:

$$m \sim e^S \times \exp \left\{ \frac{e^{\frac{S}{2}}}{2} \right\}. \quad (3.29)$$

This counting suggests that even for a small S like $S = 10^5$, m is a vast number. Thus a high number of bulk states can be embedded in the actual smaller Hilbert space with tiny inner products, which is the surprising fact underlying our resolution.

In low dimensions equation (3.28) seems to contradict our intuition, for we do not see such tremendous growth. Appendix 6.1 deals with the calculation of inner products for vectors denoting the corners of regular polyhedra in general dimensions while building up from low dimensional examples. Inner products of these corner vectors of regular

polyhedra eventually reproduce equation (3.28) when the dimensionality becomes large. This approach helps develop our intuition for large Hilbert spaces since it builds up starting from low dimensional examples.

3.3 Resolution of the paradox from the boundary perspective

As mentioned in the introduction, equation (1.1) is the coarse-grained entropy of a black hole. The origin of coarse-grained quantities like the thermodynamic entropy is due to inherent sloppiness since we measure only a small subspace of the Hilbert space. As a result, coarse-grained quantities can grow under unitary time evolution. In contrast, the fine-grained entropy or the Von Neumann entropy is a more accurate measure of the degrees of freedom. The fine-grained entropy remains invariant under unitary time evolution.

We hereby investigate the paradox from the boundary viewpoint and calculate the Von Neumann entropy on the CFT side. We will show that the calculation of the entropy of the CFT reveals the absence of any paradox because the insertion of the excitations on the thermofield double preserves the Von Neumann entropy.

Computation of the generalized entanglement entropy also demonstrates that there is no paradox in the CFT. This computation involves a choice of quantum extremal surfaces and does not depend on the precise details of the excitations.

We note an important point here: The proof that there is no paradox in the boundary does not capture the qualitative picture of the paradox in bulk. However, this indicates a crucial fact: the excitations do not increase the fine-grained entropy. The invariance of fine-grained and coarse-grained entropy along with the assumption that state-dependent operators reconstruct the black hole interior leaves us with no choice apart from overcounting of vectors to resolve this paradox.

CFT excitations: No paradox

In this subsection, we will look at the entanglement entropy of the right CFT. Consider the thermofield double state, which consists of the left and the right CFTs. Tracing over the left region gives us the reduced density matrix for the right CFT, which is the thermal density matrix ρ_T .

$$\rho_R = \text{Tr}_L |\text{TFD}\rangle \langle \text{TFD}| = \sum_i \frac{e^{-\beta E_i}}{Z(\beta)} |E_i\rangle_R \langle E_i|_R = \rho_T, \quad (3.30)$$

where ρ_T is the thermal density matrix. Equation (3.6) describes an excitation in the interior:

$$|\psi_1^e\rangle = N_1 e^{-\frac{\beta H_R}{2}} U_R^1 e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle, \quad (3.31)$$

We will define the following unitary operators for our convenience:

$$V_R^i \equiv e^{iH_R T} U_R^i e^{-iH_R T}. \quad (3.32)$$

Note that here we have included the time evolution contributions $e^{-iH_R T}$'s inside the unitary V 's since they represent unitary contributions. Till now we have worked in the semiclassical picture where we have treated the excitations as m different vectors. However from the CFT perspective, the boundary state with m interior excitations is written as the action of a single interior operator on the thermofield double state. The following expression is due to the specific form of the interior operators:

$$|\psi^e\rangle = C e^{-\frac{\beta H_R}{2}} V_R^m(t_m) V_R^{m-1}(t_{m-1}) \dots V_R^2(t_2) V_R^1(t_1) e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle = C e^{-\frac{\beta H_R}{2}} V_R e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle \quad (3.33)$$

We now calculate the reduced density matrix on the right region for this system of excitations.

$$\begin{aligned} \rho'_R &= \text{Tr}_L |\psi^e\rangle \langle \psi^e| = |C|^2 e^{-\frac{\beta H_R}{2}} V_R e^{\frac{\beta H_R}{2}} \text{Tr}_L |\text{TFD}\rangle \langle \text{TFD}| e^{\frac{\beta H_R}{2}} V_R^\dagger e^{-\frac{\beta H_R}{2}} \\ &= \frac{1}{Z(\beta)} e^{-\frac{\beta H_R}{2}} V_R e^{\frac{\beta H_R}{2}} e^{-\beta H_R} e^{\frac{\beta H_R}{2}} V_R^\dagger e^{-\frac{\beta H_R}{2}} \\ &= \rho_T. \end{aligned} \quad (3.34)$$

The above manipulations follow because V_R is a unitary operator. We expect the thermal density matrix to remain unchanged under interior operator insertions because the thermal behaviour arises due to the horizon's existence and is irrespective of insertions in the interior unless a large backreaction changes the horizon. The interior operators are defined only in the effective field theory limit, i.e. the backreaction is small, and hence the thermal density matrix remains invariant. Since the density matrix itself does not change due to the excitations, the entanglement entropy does not change as well. Therefore we see that there is no paradox in the CFT as interior excitations do not change the entanglement entropy.

Generalized entanglement entropy of the CFT

Using the generalized entanglement entropy [112–117], we can again show that there is no paradox in the CFT. Quantum extremal surfaces are defined as surfaces which extremize the sum of area and bulk entanglement entropy contributions, given a boundary subregion B . This extremized sum is the generalized entanglement entropy of B .

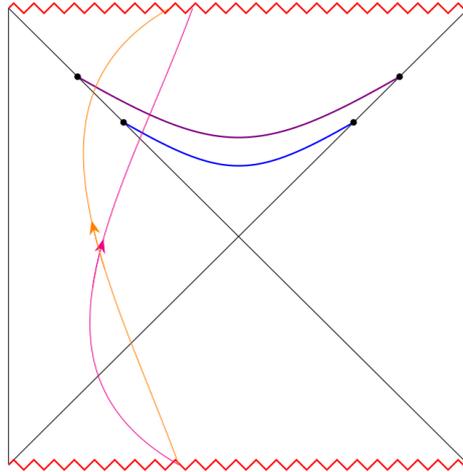


Figure 3.1: Bulk excitations denoted by orange and magenta lines arise from the left past horizon and fall into the left future horizon. These states come out of equilibrium as indicated in equation (3.38) around the time t_1 and t_2 for the orange and magenta excitations. In the bulk this out of equilibrium behaviour is indicated by how far the excitations protrude out on the left. The unitaries control the position of the excitation on the slice, and large $|t_1 - t_2|$ leads to large spatial separation. Interior excitations at late times are visible only to later slices. Consequently, we can keep accommodating more and more excitations at later and later times which leads to the paradox.

$$S_{\text{gen}}(B) = \text{Min}_X \text{Ext}_X \left[\frac{\text{Area}(X)}{4G_N} + S_{\text{bulk}}(\Sigma_B^X) \right] \quad (3.35)$$

Consider B to be the right boundary region R on which the right CFT lives. We will consider the case with no excitations living on the black hole first. Quantum extremal surfaces for this case end at the horizon, therefore the generalized entanglement entropy is given by:

$$S_{\text{gen}}(R) = \frac{\text{Area of black hole}}{4G_N} \quad (3.36)$$

Now consider a situation where the matter content due to excitations in the interior is very large, which is our case of interest. Consequently, the bulk entropy S_{bulk} in the interior of the black hole due to all the excitations is very large. In this case, the quantum extremal surfaces are no different and go only up to the horizon, thereby not capturing S_{bulk} in the interior region. As a result, the fine-grained entropy again is given by (3.36). Therefore we conclude that there is no bag of gold paradox. We note that we do not need the precise form of the excitations in order to derive this conclusion.

3.4 The nature of the excitations and the initial bulk wavefunction

In this section, we are interested in understanding the exact nature of the excitations created by interior operators as given in (3.6). The excitations' behaviour also holds the key to qualitatively understand the initial state in the bulk, which leads to the paradox. From the CFT perspective, the states given in (3.6) are non-equilibrium states [98], which we briefly describe. These states $|\psi_1^e\rangle = C_1 e^{-\frac{\beta H_R}{2}} U_1(t_1) e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle$ arise from the past left horizon and end up at the future left horizon. To see this, we first show that these excitations are invisible to the small algebra \mathcal{A}_R ⁶. Therefore the time dependence of these observables cannot be seen by probing with $O \in \mathcal{A}_R$.

$$\frac{d}{dt} \langle \psi_1^e | O(t) | \psi_1^e \rangle \sim \mathcal{O}\left(\frac{1}{S}\right) \quad (3.37)$$

However, these states are truly non-equilibrium when probed by the Hamiltonian [98]. The Hamiltonian has support on both \mathcal{A}_R and \mathcal{A}'_R and therefore can detect the excitations on the commutant \mathcal{A}'_R . Writing the state as $|\psi_1^e\rangle = W(t) |\text{TFD}\rangle$, it can be shown that

$$\frac{d}{dt} \langle \psi_1^e | O(t) H | \psi_1^e \rangle = \frac{d}{dt} \langle \text{TFD} | W^\dagger O(t) [H, W] | \text{TFD} \rangle + \mathcal{O}\left(\frac{1}{S}\right). \quad (3.38)$$

Equation (3.38) shows that the state $|\psi_1^e\rangle$ is out of equilibrium. The bulk interpretation is now clear as the operators in the right exterior of the black hole cannot detect the excitations $|\psi_1^e\rangle$. These excitations emerge from the past singularity and are short-lived. At around $t \sim t_1$ they arrive at the left part of the diagram. At a later time, they fall into the future singularity. These non equilibrium states are out of equilibrium at around $t \sim t_1$, but remain in equilibrium for $t \ll t_1$ and come back to equilibrium for $t \gg t_1$, and are therefore transient.

It is now easy to generalize from a single excitation to many excitations as given in (3.33), where as before, we include the factors $e^{-iH_R T}$'s inside the unitaries V 's.

$$|\psi^e\rangle = C e^{-\frac{\beta H_R}{2}} V_R^m(t_m) V_R^{m-1}(t_{m-1}) \dots V_R^2(t_2) V_R^1(t_1) e^{\frac{\beta H_R}{2}} |\text{TFD}\rangle \quad (3.39)$$

This state in (3.39) will be seen in equilibrium at $t \gg t_1, t_2, \dots, t_m$ and $t \ll t_1, t_2, \dots, t_m$. However when probed by the Hamiltonian at intermediate times, say at $t \sim t_1, t \sim t_2$ or at $t \sim t_m$ the state will appear out of equilibrium. The bulk picture describing out-of-equilibrium behaviour of the excitations at these intermediate times is understood as them coming out of the past left horizon and travelling in the left exterior

⁶The algebras \mathcal{A} , \mathcal{A}' and the associated small Hilbert space formed by acting with them on the thermofield double state are reviewed in §2.9.

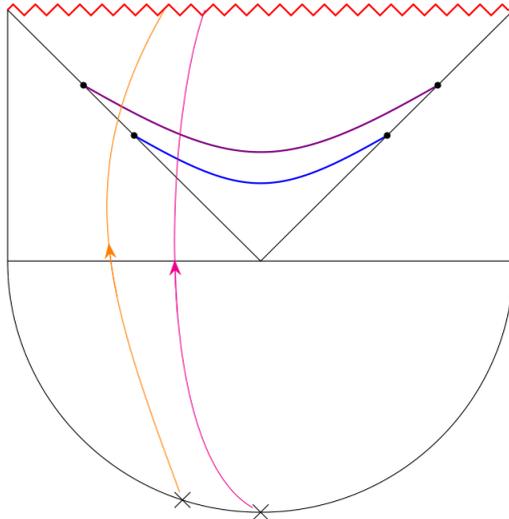


Figure 3.2: The initial state of the black hole is created by glueing the Euclidean AdS to the bottom half of the Lorentzian Penrose diagram. Boundary deformations of the Euclidean AdS create our excitations. Just after the initial time, all the bags of gold excitations are in the left exterior.

before falling into the future horizon (See Figure 3.1).

The nature of the excitations reveals the physical picture of the paradox as well. As we have argued earlier, all excitations possess an energy E_0 , where $m E_0 \ll M_{BH}$. This small energy means that the excitations cannot protrude very much outside the interior on the left-hand side, and all excitations protrude a similar distance after coming out of the past horizon before travelling and falling inside the future horizon.

Now consider early excitations governed by small t 's, e.g. $(t_1, t_2 \dots t_i)$, where $i \ll m$, which come out from the past horizon and fall into the future horizon. These excitations intersect the Cauchy slices with boundaries at earlier Kruskal times and keep intersecting future Cauchy slices at later Kruskal times as well. In contrast, the excitations which come outside the past horizon and fall inside the future horizon at late t 's will not intersect the early Kruskal time Cauchy slices. However, these excitations will intersect the late Kruskal time Cauchy slices in the interior (See Figure 3.1). These above features give rise to the physical picture of the paradox. On the late time slices, there will be more and more excitations where the number of excitations is tuned such that they constitute a dilute gas of a fixed density ρ . Therefore we have slices which have an increasingly large value of entropy at late times which becomes more substantial than the Bekenstein-Hawking entropy.

In the bulk Lorentzian description it naively seems that the excitations emerge out of the past singularity. This apparent problem is rectified by writing down an initial bulk state for the problem [101, 102]. The way we construct the initial state or the Hartle-Hawking wavefunction of the eternal Lorentzian geometry is by glueing it to a Euclidean AdS part and then performing the path integral over the Euclidean part. We can thus obtain the Hartle-Hawking state. At $t = t_0$ all excitations are in the exterior and

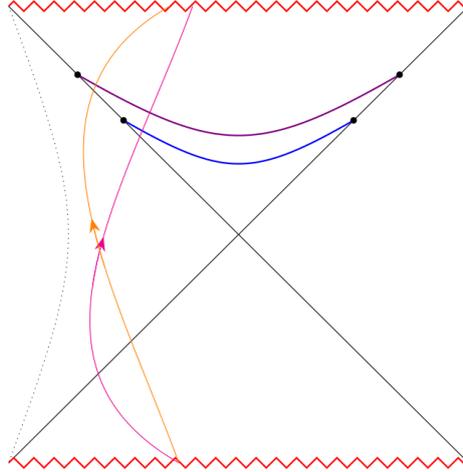


Figure 3.3: We demonstrate the paradox for pure state black holes which are single-sided. The dotted line on the left denotes the UV cutoff for the theory living on the right boundary, which prohibits us from reaching arbitrary close to the left boundary. Since the interior region is similar for both the single-sided and eternal black holes, the physical picture of the paradox and its resolution is similar.

propagate afterwards on the left side of the Penrose diagram. We write our initial state at this $t = t_0$ when all excitations are outside the horizon. Here each excitation should be treated as a small deformation of the initial wavefunction and can be generated by inserting operators at the Euclidean AdS boundary as shown in Figure 3.2. This gives us the CFT state (3.39). The initial state in the bulk is a path integral performed over this configuration of an eternal geometry plus small boundary deformations, which is given in (3.39). This path integral qualitatively resolves the problem of constructing a valid initial bulk state in order to pose the paradox.

3.5 The paradox for single sided black holes

Till now, we have discussed at length the paradox for the eternal black hole. We can also pose a similar paradox for pure state black holes. These black holes are described on the boundary by a pure state on a single-sided CFT. Using a single side CFT on the right, the bulk description of the black hole can be reconstructed using HKLL reconstruction [118] in the exterior right. The top and bottom regions of the fully extended Kruskal diagram of AdS Schwarzschild black hole can be reconstructed using state-dependent operators as given in §2.9. A part of the left bulk region can also be reconstructed; however, we cannot go too far on the left side as one needs operators with higher and higher energies to approach closer and closer to the left boundary. In other words, a UV cutoff on the right boundary CFT prevents us from going arbitrarily close to the left boundary in the bulk.

Single sided black holes are represented by typical states [103–105] on the boundary

CFT. We define these states by considering a quantum statistical system at a temperature T and average energy E . The relevant example in our case is a CFT at a temperature T . Let us consider a small interval ΔE centred about E in the CFT energy spectrum with $\Delta E \ll E$. We will be looking at n energy eigenstates $|E_i\rangle$ in the interval ΔE , each with energy E_i . The entropy is therefore given by $\Delta S = \frac{\Delta E}{T}$, while the number of states n in the interval is related to the entropy as $n = e^{\Delta S}$. We now define a state by randomly superposing the energy eigenstates:

$$|\psi_s\rangle = \sum_{i=1}^n c_i |E_i\rangle \quad (3.40)$$

such that $\sum_i |c_i|^2 = 1$. Here c_i 's are chosen at random. These $|\psi_s\rangle$ states obey a surprising property: For a quantum statistical system, "almost" all the states $|\psi_s\rangle$ mimic thermal behaviour, and we define such states which look thermal as typical states $|\text{TYP}\rangle$. Further details of the construction of single sided black holes using typical states are reviewed in §2.6.

3.5.1 The single-sided paradox and its resolution

We now state the paradox for the single-sided black holes. The construction of maximal volume surfaces in §3.1 is the same for this case because the single-sided black hole possesses the same interior region as the eternal black hole does. The excitations in the black hole interior are also similar with the difference being their action on the typical state rather than on the thermofield double state. A single excitation is given by:

$$|\psi_1^e\rangle = C_1 e^{-\frac{\beta H_R}{2}} U_1(t_1) e^{\frac{\beta H_R}{2}} |\text{TYP}\rangle \quad (3.41)$$

As before, we can place similar excitations far apart from each other on the nice slice by adjusting the unitary U to create a dilute gas of density ρ . Calculating the entropy of this semiclassical configuration again violates the coarse-grained Bekenstein-Hawking entropy at late times in the bulk. The nature of the excitations is also similar, they emerge out from the bottom interior by coming out of the left past horizon and propagate on the left side for some time, and fall into the left future horizon.

Our resolution to the bulk paradox for the single-sided black holes is the same resolution which we have proposed for the eternal case. We have hugely overcounted the excitations in this case as well due to small inner products between coherent bulk states describing the excitations. The resolution for this case is unchanged because the interior possessed by single-sided and eternal black holes is the same.

As before we see that the fine-grained entropy remains unchanged. This consistency arises as the typical state is a pure state and the entanglement entropy of this system is zero. Similar to what was derived in 3.3, insertion of multiple bulk interior excitations

on the typical state leaves the density matrix unchanged. As a result, we again conclude that there is no paradox in the CFT. Even though the fine-grained entropy of the system is zero, the coarse-grained entropy is S_{BH} . In the following §3.5.2 we justify our claim that the enormous number of semiclassical bulk excitations arise due to an overcounting of the bulk Hilbert space.

3.5.2 Why interior bulk states are non-orthogonal in the CFT Hilbert space?

We consider typical states in the CFT which are dual to the single sided black hole in the bulk and are centered about an average energy E with range $2\Delta E$:

$$|\text{TYP}\rangle = \sum_{i=E-\Delta E}^{E+\Delta E} c_i |\psi_i\rangle, \quad (3.42)$$

where $|\psi_i\rangle$ are normalized states and $\sum_i |c_i|^2 = 1$. We will denote O_ω as operators in the boundary CFT with energy ω . (3.42) is constructed by acting with a string of O 's on the ground state such that the string's total energy is E_i , which then leads to the state $|\psi_i\rangle$. We are looking at states of the form:

$$|\psi\rangle = K e^{-\frac{\beta H}{2}} U(O_\omega) e^{\frac{\beta H}{2}} |\text{TYP}\rangle \quad (3.43)$$

where K is the constant of normalization, $U(O_\omega)$ is an unitary operator creating bulk excitation generated by products of O_ω . These operator insertions do not change the energy of the typical state much, i.e. $\omega \ll O(\mathcal{N})$. Another requirement is that the number of single oscillator operator insertions in $U(O_\omega)$ is lesser than $O(\mathcal{N})$. These conditions define the small algebra of observables \mathcal{A} which act on the ground state to give the small Hilbert space. For the CFT this means that the number of operator insertions is very small as compared to the energy of the state, and the insertions don't have very high energy themselves. We now want to evaluate the inner product of the two such states in the small Hilbert space, where as previously, $V_j(O_{\omega_j}, T_j)$ includes the e^{iHt} insertions, i.e. $V_j(O_{\omega_j}, T_j) = e^{iHt} U_j(O_{\omega_j}) e^{-iHt}$.

$$|\psi_i\rangle = K_i e^{-\frac{\beta H}{2}} V_i(O_{\omega_i}, T_i) e^{\frac{\beta H}{2}} |\text{TYP}\rangle \quad \& \quad |\psi_j\rangle = K_j e^{-\frac{\beta H}{2}} V_j(O_{\omega_j}, T_j) e^{\frac{\beta H}{2}} |\text{TYP}\rangle \quad (3.44)$$

These states defined above live in the small Hilbert space and the indices i, j go over the small Hilbert space. The inner product between these states is given by

$$\begin{aligned}
 \langle \psi_j | \psi_i \rangle &= K_i K_j \langle \text{TYP} | e^{\frac{\beta H}{2}} V_j^\dagger(O_j, T_j) e^{-\beta H} V_i(O_i, T_i) e^{\frac{\beta H}{2}} | \text{TYP} \rangle \\
 &= K_i K_j \sum_{k,l,o,p,q,r} c_k^* c_n \langle E_k | e^{\frac{\beta H}{2}} | E_o \rangle \langle E_o | V_j^\dagger | E_p \rangle \langle E_p | e^{-\beta H} | E_q \rangle \langle E_q | V_i | E_r \rangle \langle E_r | e^{\frac{\beta H}{2}} | E_l \rangle \\
 &= K_i K_j \sum_{k,l,p} c_k^* c_n e^{\frac{\beta E_k}{2}} e^{\frac{\beta E_l}{2}} e^{-\beta E_p} \langle E_k | V_j^\dagger | E_p \rangle \langle E_p | V_i | E_l \rangle
 \end{aligned} \tag{3.45}$$

We see here that in general, these states are not orthogonal. This non-orthogonality arises since we are working with restricted energy operators on the typical states. Because these states are normalized and since $e^{\frac{\beta E_k}{2}} e^{\frac{\beta E_l}{2}} e^{-\beta E_p} \approx 1$ as the states lie in the small Hilbert space; we can write the inner products as

$$\langle \psi_j | \psi_i \rangle = \sum_{k,l} (c_k^j)^* c_l^i, \tag{3.46}$$

such that $\sum_l |c_l^i|^2 = 1$. Equation (3.46) gives rise to a small but finite number larger than $O(e^{-\frac{S}{2}})$, where the dimension of the Hilbert space is e^S ⁷. We thus see that the inner product between the vectors is a small number if $|\psi_j\rangle$ live in a huge dimensional Hilbert space. These small inner products naturally give rise to overcounting in CFTs.

3.6 Spectral properties of bags of gold spacetimes: Contradictions and Resolution

Till now we have discussed the paradox of the coarse-grained entropy of bags of gold spacetimes. Let us now understand the spectral features of these spacetimes in the context of effective field theory. Firstly we will work with the semiclassical Hilbert space of the bags of gold spacetime spanned by the excitations placed far apart from each other. We will argue that such an effective field-theoretic description of the Hilbert space potentially contradicts with black holes' spectral observables' predicted behaviour.

Consider the phase space of a classical system exhibiting chaos. It was conjectured in [41] that the quantum counterpart of such a system should have an energy level spacing distribution which matches one of the three standard random matrix ensembles - Gaussian orthogonal (GOE), unitary (GUE) or symplectic (GSE), depending on the inherent symmetries of the system. Since black holes display scrambling properties, we expect that their level spacing distribution matches the one given by the Gaussian unitary dis-

⁷The derivation of $O(e^{-\frac{S}{2}})$ is straightforward, it is the same as calculating the expected displacement in a random walk problem after n steps.

tribution. Therefore a convenient way to model black holes is by using random matrices constructed using Gaussian unitary ensemble. For GUE, the expression for the level spacing distribution is [42]:

$$P(s) = \frac{32 s^2}{\pi^2} \exp\left[-\frac{4s^2}{\pi}\right], \quad (3.47)$$

where s is the distance between two consecutive eigenvalues, which we expect to be the energy level spacing distribution of black holes as well. The conjecture [41] proposes that the quantum counterpart of a classical system exhibiting chaos possesses either above level spacing distribution or that of its two cousins, the GOE or GSE. This is in contrast to the level spacing distribution obeyed by non-chaotic systems. As a drastically different example, for integrable systems, the Berry-Tabor conjecture states that the level spacing distribution should be Poissonian $P(s) = \exp\{-s\}$ [119].

Using the formalism of random matrix theory, we will argue that the violations of spectral observables can be classified into two types. The nature of the first violation is characterized by qualitative deviation from the expected GUE energy level spacing distribution. To overcome this violation, we will demand that the only bags of gold configurations which are allowed are strictly consistent with a GUE description. Such an imposition drastically constrains the space of allowed bags of gold configurations. We will observe that even after enforcing this condition, bags of gold configurations can still be captured using the spectral form factor; a spectral observable which quantifies the discrete nature of the system. Thus the effective field-theoretic description of the bags of gold's Hilbert space suffers from serious contradictions as compared to observed characteristics of black holes. Towards the end of this section, we demonstrate how our overcounting hypothesis resolves these contradictions in the spectral form factor.

3.6.1 Spectral observables in random matrix theory and discrete systems

Random matrix theory observables

In this section, we briefly review the spectral observables of random matrices belonging to the Gaussian unitary ensemble. The Gaussian unitary ensemble of Hermitian matrices H of dimension $N \times N$ is defined as follows

$$Z_G = \int [dH] \exp\left(-\frac{H^2}{4v^2}\right). \quad (3.48)$$

Here v^2 is a real number which is $O(1)$ and does not scale with N . A convenient way to solve this integral is by decomposing these matrices in terms of their eigenvalues.

From (3.48) the joint probability distribution of the eigenvalues $\lambda_i, i \in [1, N]$ belonging to Gaussian unitary ensemble is given by

$$p(\lambda_1, \dots, \lambda_N) = C \exp \left[2 \sum_{j < k} \log |\lambda_j - \lambda_k| - \sum_j \frac{\lambda_j^2}{2v^2} \right]. \quad (3.49)$$

The first term in the exponential of (3.49) arises from the Van der Monde determinant, which comes from the Jacobian of the transformation in the measure, while the second term arises due to the Gaussian potential from (3.48). The average density of eigenvalues $\bar{\rho}(\lambda)$, where $\bar{\rho}(\lambda) = \int d\lambda_2 \dots d\lambda_N p(\lambda, \lambda_2 \dots \lambda_N)$; is given by

$$\bar{\rho}(\lambda) = \frac{2\sqrt{R^2 - \lambda^2}}{\pi R^2}; \quad R^2 = 8v^2 N, \quad -R \leq \lambda \leq R. \quad (3.50)$$

Given this setup, we focus on the fluctuations of the eigenvalues, which are independent of the potential $V(\lambda)$ in the large- N limit. Regarding fluctuations, the vital quantity of interest related to quantum chaos is the level spacing distribution given by $P(s)$ as given in (3.47).

Measure of discreteness: Spectral form factor

As we discussed, apart from the chaotic signatures, since the systems we are studying are black holes which have discrete spectra, it is useful to look at physical observables which can capture discreteness. In this regard, it is useful to understand the typical size of the fluctuations at late times, which in turn characterizes the discreteness of the energy spectrum. In order to define such a quantity, let us first generalize the partition function of a system to include Lorentzian time along with the temperature:

$$Z(\beta + iT) = \text{Tr} \left[e^{-\beta H - iT H} \right]. \quad (3.51)$$

At late times, this generalized partition function oscillates, and the time average of this quantity is zero. Using this partition function, we will now define the spectral form factor which captures the magnitude of such oscillations:

$$S(\beta, T) = \frac{Z(\beta + iT) Z(\beta - iT)}{Z(\beta)^2} = \frac{1}{Z(\beta)^2} \sum_{i,j=1}^N e^{-\beta(E_i + E_j)} e^{i(E_i - E_j)t} \quad (3.52)$$

Since the systems in consideration are chaotic, we now demand that the Hamiltonian in consideration is described by a random matrix obeying GUE statistics. Therefore we write the expression for the generalized partition function in Gaussian unitary ensemble:

$$\langle Z(\beta + iT) \rangle_G = \frac{1}{Z_G} \int [dH] e^{-\frac{H^2}{4v^2}} \text{Tr} \left[e^{-\beta H - iT H} \right], \quad (3.53)$$

where Z_G is given by (3.48). Equation (3.53) can now be used to calculate the spectral form factor in (3.52). It was shown in [49] that the curve describing the logarithm of spectral form factor versus the logarithm of T obeys the following features:

1. The curve starts from 1 and starts decaying with a constant slope at early times. This behaviour can be understood by plugging in the level density in (3.50) into (3.53), and then using it to evaluate the spectral form factor in (3.52). The late-time decay of the spectral form factor at high temperature is captured by $S(\beta \approx 0, T) \sim T^{-3}$.
2. The decaying behaviour continues until the dip time, after which the curve rises with a constant slope. The physical reason behind this is as follows: $S(T, \beta)$ is roughly a sum of connected and disconnected parts. The disconnected part contributes to the decay which dominates until the "dip time". Equating the late time decay of the spectral form factor and the ramp growth gives the value for the dip time, which is $t_d \sim e^{N/2}$. After dip time the connected part dominates giving rise to the increasing ramp, which at high temperature is given by $S(\beta \approx 0, T) \sim \frac{T}{2\pi \exp\{2N\}}$.
3. At a certain time called the plateau time, the ramp stops increasing and gives rise to a constant plateau. Physically the plateau appears because oscillations in the generalized partition function are random and out of phase at very late times, contributing to a small but non-zero number. After the plateau time $t_p > 2e^N$, the constant plateau of the spectral form factor is given by $S(\beta \approx 0, T) \sim \frac{1}{\pi e^N}$.

This behaviour of the spectral form factor captures the discrete features of black holes, which can be seen from the red curve in Fig. 3.5 for $\beta = 1$. We will now see that treating the bulk effective degrees of freedom as independent degrees of freedom violates the delicate structure expected from the above description.

3.6.2 Spectral properties of bags of gold excitations

As before, we construct several unitary excitations behind the horizon creating a bags of gold configuration. These excitations are of the form given in (3.6), which we restate in the frequency basis:

$$|\psi_i\rangle = K_i e^{-\frac{\beta H}{2}} V_i(O_{\omega_i}, T_i) e^{\frac{\beta H}{2}} |\psi_{BH}\rangle, \quad i \in (1, m) \quad (3.54)$$

Here as previously, $V_j(O_{\omega_j}, T_j)$ includes the e^{iHt} insertions, i.e.

$$V_j(O_{\omega_j}, T_j) = e^{iHt} U_j(O_{\omega_j}) e^{-iHt}.$$

Here $O_i \in \mathcal{A}$, where \mathcal{A} is the algebra of simple operators. As argued before, these operator insertions have small energies $\omega_i \ll \mathcal{N} = N^2$. Consequently the energies of these excitations belong to a small interval $(E - \Delta E, E + \Delta E)$, where $E \sim O(\mathcal{N})$, and $\Delta E \sim O(1)$.

In the semiclassical description since states of the form (3.54) are spread wide apart spatially, we naively think that such distinct configurations have zero inner product. Let us represent the Hilbert space of the effective field theory of the bags of gold spacetime by $\mathcal{H}_{BOG} := \{|\psi_i\rangle\}$, $i \in (1, m)$, which is m -dimensional. Following this semiclassical logic, we saw previously that the m -dimensional space is very large as compared to the n -dimensional black hole's Hilbert space. Since the excitations are placed far apart, this naive reasoning leads us to conclude that the vectors denoting the bags of gold excitations in \mathcal{H}_{BOG} are orthonormal:

$$\langle \psi_j | \psi_i \rangle = 0, \quad \langle \psi_j | e^{iHt} | \psi_i \rangle. \quad (3.55)$$

Violations of Type 1

We will now see how this naive EFT description violates the spectral properties expected from §3.6.1. It is straightforward to construct bags of gold Hilbert spaces spanned by vectors $\mathcal{H}_{BOG} := \{|\psi_i\rangle\}$ such that the difference in the energy levels of these vectors do not obey the expected level spacing distribution given by GUE, which is given in (3.47).

A trivial example of such an EFT Hilbert space can be constructed by using vectors of the form (3.54) such that $V_i(O_i(\omega_i), T_i)$ has energies ω_i in integer multiples of a constant $\omega_i = k_i c$, $k_i \in \mathbb{R}$. The above example is an allowed bags of gold configuration because the only physical condition we have enforced is $\sum_i \omega_i \ll O(\mathcal{N})$, with no condition on the individual energies of the excitations. As before, we have denoted the number of black hole states as n and the bags of gold configuration as m with $m \gg n$. Thus the Hilbert space is spanned almost exclusively by the bags of gold states, since $m \gg n$. Therefore in this scenario, the level spacing distribution is that of a bunch of simple harmonic oscillators, which is an integrable system and thus is drastically different from the expected distribution in (3.47). In addition, we can see from Fig. 3.4 that the spectral form factor does not qualitatively match with the curve expected of black holes. Thus this bags of gold configuration contradicts with spectral features expected from a black hole.

In general, we can construct various bags of gold spacetimes by spanning the Hilbert space of the EFT using appropriate vectors such that the energy level spacing distribution and the spectral form factor deviates from the spacing distribution and spectral form factor predicted by GUE. We will call these examples where the energy level spacing distribution and spectral form factor do not qualitatively follow the GUE distribution as **violations of type 1**.

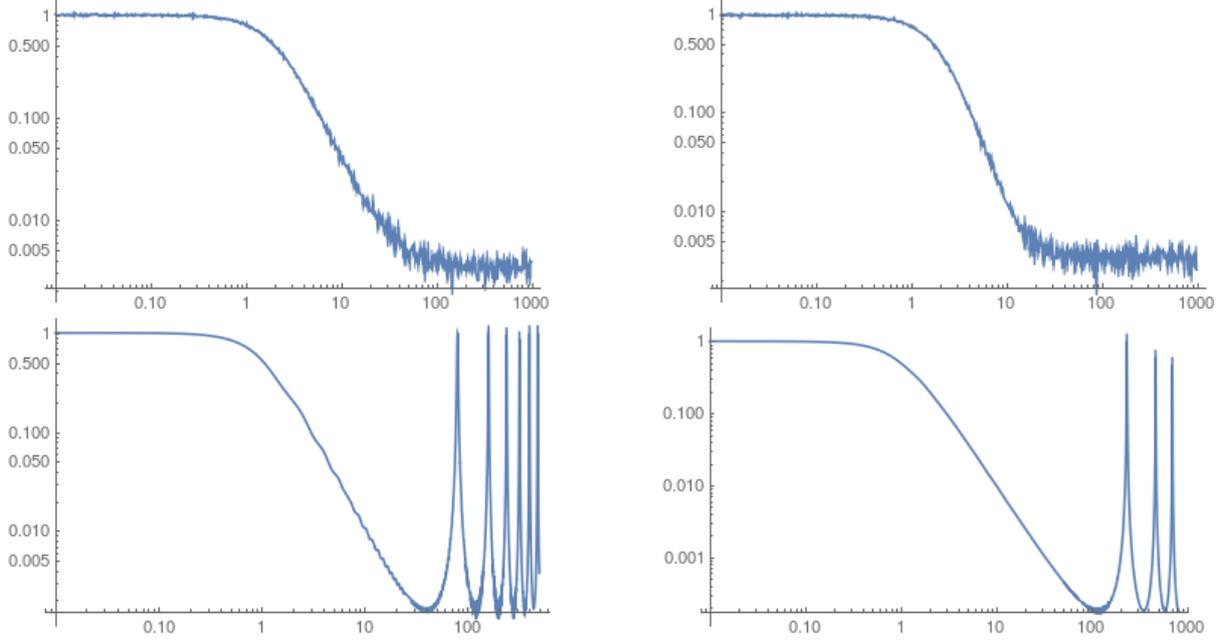


Figure 3.4: The spectral form factors for different distributions obeying violations of type 1 (See Table 3.1). The figure on the top left consists of random energy levels taken from a uniform probability distribution. The figure on the top right has energy levels picked from a near-uniform probability distribution. Both plots are for 1000 energy levels at $\beta = 2$ over 100 iterations. The bottom figures have uniformly spaced energy levels, and as a consequence, are integrable. The bottom left figure is plotted with 50 energy levels, with $\beta = 1$ over 50 iterations, while the bottom right figure is plotted with 250 energy levels, with $\beta = 1$ over 50 iterations.

Type	Energy level spacing distribution $P(s)$	Spectral form factor $S(\beta, T)$
Violation 1	No specified distribution	No correlation with black hole's curve
Violation 2	Follows GUE distribution	Different t_d , t_p and plateau height

Table 3.1: The two types of violations in spectral properties between EFT treatment of bags of gold excitations as independent states and black holes.

Configuration	Dip Time (t_d) \sim	Plateau time (t_p) \sim	Plateau Height \sim	2-pt function $\langle O(t)O(0) \rangle - G_p \sim$
Black hole	\sqrt{n}	n	n^{-1}	$t \exp\{-2S_{BH}\} - \exp\{-S_{BH}\}$
Bags of gold	\sqrt{m}	m	m^{-1}	$t \exp\{-2S_{BOG}\} - \exp\{-S_{BOG}\}$

Table 3.2: Violation 2 - The dip time, plateau time and plateau height for a black hole and a bags of gold configuration in EFT description obeying random matrix statistics in terms of the dimensionality of their Hilbert spaces at $\beta \approx 0$.

Violations of Type 2

As seen from the type 1 violations, the effective field-theoretic treatment of bags of gold scenarios can not only lose important features like scrambling etc. but may also result

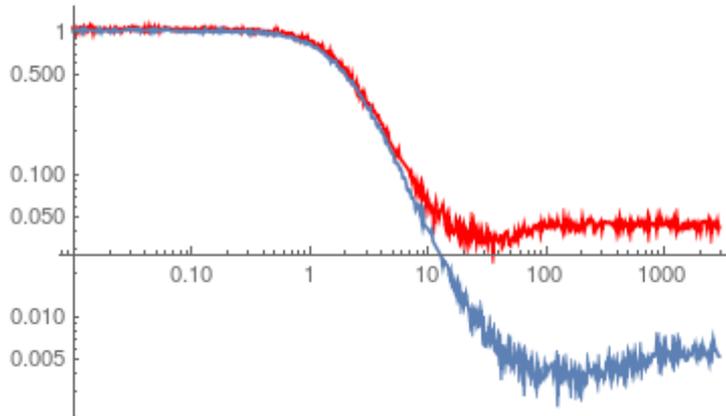


Figure 3.5: The spectral form factors for the "black hole" of 100 states in red and a "bags of gold distribution" with 1000 states in blue plotted using GUE ensemble, over 50 iterations with $\beta = 2$. The dip height, dip time, plateau time and plateau height are visible here which are different for both these configurations, which exemplifies violation 2.

in a completely different description which is integrable. In order to overcome these contradictions, one can demand to consider only those bags of gold spacetimes in which the energy level spacing distribution matches with the GUE level spacing distribution. Such a demand substantially reduces the space of allowed bags of gold spacetimes. Consequently, we have a more refined version of the paradox formulated in the effective field theory Hilbert space which is seemingly consistent with a few basic spectral properties of quantum chaotic systems.

However, we will show that even this restricted space of bags of gold spacetimes which obeys naive GUE level spacing statistics is inconsistent with quantitative features of the spectral form factor involving the height and time of the plateau, dip time and the slope of the ramp, which is due to the fact that $m \gg n$. We will call these examples where the level spacing distribution follows GUE statistics along with a quantitative deviation from the black hole's spectral form factor as **violations of type 2**. For convenience, we mention the properties characterizing these two classes of violations in Table 3.1. In order to evaluate the plateau height, we need to look at the long term average of the spectral form factor. The only terms which survive over large times are those with $E_i = E_j$, as the rest of the terms cancel out due to dephasing and thus die off. The long time average of the spectral form factor is thus given by:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt S(\beta, t) = \frac{1}{Z(\beta)^2} \sum_{i=1}^N g^2(E_i) e^{-2\beta E_i} = \frac{1}{Z(\beta)^2} \sum_{k=1}^n e^{-2\beta E_k}, \quad (3.56)$$

where $g(E_i)$ denotes the degeneracy of states at energy E_i . For the EFT description of the bags of gold spacetime, the plateau height is given by:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt S(\beta, t) = \frac{1}{Z(\beta)^2} \sum_{i=1}^{N'} g^2(E_i) e^{-2\beta E_i} = \frac{1}{Z(\beta)^2} \sum_{k=1}^m e^{-2\beta E_k}, \quad (3.57)$$

Here m is the number of the bags of gold states, such that $m \gg n$. Since $m \gg n$ there is a quantitative disagreement between the plateau height of the original black hole and the bags of gold spacetime. For the high temperature case with $\beta \approx 0$ as described in the §3.6.1, we can conclude that the plateau height is $e^{-N} = \frac{1}{n}$ for the original black hole and $e^{-N'} = \frac{1}{m}$ for the bags of gold spacetime. In addition the dip time is $t_d \sim e^{\frac{N}{2}}$ for the black hole and $t_d \sim e^{\frac{N'}{2}}$ for the bags of gold respectively, while the plateau time is $t_p \sim e^N$ for the original black hole, while $t_p \sim e^{N'}$ for the bags of gold respectively. These values are collectively summarized in Table 3.2. Thus even if we choose the bags of gold configurations in such a way that they obey naively obey qualitative spectral properties, there are quantitative differences which are captured using the spectral form factor.

[49] also pointed out the behaviour of the two-point function with the assumption that the system obeys the eigenstate thermalization hypothesis, and has a ramp at late times. They predicted that the two-point function should be of the following form:

$$G(t) = \langle O(t)O(0) \rangle \sim G_p + \frac{t}{L^2} - \frac{1}{L}, \quad (3.58)$$

where $L \sim \exp\{S\}$ of the system. Thus the two-point function for the effective field theory of bags of gold and the black hole has different behaviour, as mentioned in Table 3.2.

3.6.3 Resolution of spectral puzzles using overcounting

We now ask whether our earlier proposed resolution to the paradox reconciles these disagreements. We study the spectral properties in the context of pure state black holes for convenience, and we are interested in the order of magnitude of the partition function. Our conclusions can be extrapolated to general black holes as well. As before, we consider typical states defined on the interval $(E - \Delta E, E + \Delta E)$ to represent a pure state black hole. The partition function of the dual CFT describing the original black hole over this interval has the order of magnitude:

$$Z(\beta) = \text{Tr} \left(e^{-\beta H} \right) \sim O(ne^{-\beta E}). \quad (3.59)$$

Here we have considered $\Delta E \ll E$, which gives us the above order of magnitude of the partition function. We now evaluate the partition function of the bags of gold case where we assume that the m states spanning the EFT Hilbert space are orthogonal. Therefore the order of magnitude of the partition function is given by:

$$Z_{BOG}(\beta) = \text{Tr} \left(e^{-\beta H} \right) = \sum_{l=1}^m \sum_{i,j=1}^n \langle l|i \rangle \langle i|e^{-\beta H}|j \rangle \langle j|l \rangle = O(me^{-\beta E}). \quad (3.60)$$

This overcounting in the partition function manifests itself in wrong quantitative

values for the entropy of the black hole, spectral form factor and the two-point function at late times. As earlier, we will argue that the bags of gold states in quantum gravity are not independent but have small inner products with each other. Therefore the actual Hilbert space is spanned by n vectors, with m embedded Bags of gold vectors which have tiny but non-zero inner products between each other. Thus we can arrive at the correct conclusion that $Z_{BOG} \sim O(ne^{-\beta E})$ by working with bulk states such that they have small but finite inner products. The above conclusion holds as the correct sum over l in (3.60) is really up to n instead of up to m . The conclusion that $Z_{BOG} \sim O(ne^{-\beta E})$ is also consistent with the entropy of the black hole as seen before. Similarly we repeat this analysis for $Z(\beta, T)$ as well, and hence we argue that the correct spectral form factor for the bags of gold spacetime should match the black hole's spectral form factor by thinking about the bulk interior states as embedded in the n -dimensional Hilbert space with small inner products.

Another way to verify that overcounting resolves discrepancies is by observing that the spectral form factor for the bags of gold configuration quantitatively matches with the black hole's spectral form factor in Table 3.2 if overcounting is taken into account. Given that the actual dimensionality of the m dimensional overcounted Hilbert space is n , we see that the dip time, the plateau time and the height of the plateau for the bags of gold configurations match with the original black hole's curve's features. Similarly, overcounting resolves the discrepancy between the 2-pt function in the bags of gold spacetime and the black hole as well, in accordance with our earlier argument that Bekeknstein-Hawking entropy gives the correct entropy of bags of gold configurations.

3.7 Study of the paradox using toy matrix models

In this section we explicitly demonstrate how overcounting allows us to construct an immense number of bulk excitations in the context of toy matrix models. Even though the bulk states arise from matrix models in the large N limit, we can understand aspects of overcounting by performing computations even in small N toy matrix models. Such matrix models have a small dimension of the Hilbert space, and it is possible to list out the state-space explicitly.

By calculating the partition function of a matrix model at temperature T using the canonical ensemble, we can extract out the average energy and entropy of the system. Exponentiating the entropy gives us the dimension of the Hilbert space. We are interested in the regime of small N and temperature such that the dimension of the Hilbert space is less than 1000.

We will demonstrate overcounting in two different toy matrix models. The first example is of a $(0 + 1)$ dimensional two matrix model which has a $U(N)$ global symmetry group. We construct a typical state using the microcanonical ensemble. Afterwards, we

will write down the small Hilbert space and demonstrate that we can embed a larger number of vectors compared to the dimension of the small Hilbert space. The second example deals with a CFT consisting of 2 matrices defined on $S_3 \times \mathbb{R}$. Here we will calculate the Hagedorn temperature and construct the typical state above the Hagedorn temperature. Again we will construct the small Hilbert space and demonstrate overcounting. These toy examples show that overcounting with small inner products is natural in the small Hilbert space.

A similar construction of states follows for bulk states at large N . Apart from computational problems with enumerating the states explicitly, there is no further restriction to doing the same for CFTs with holographic duals in the large N limit. Some results obtained in this section are in agreement with recent related work [120].

3.7.1 Toy Model I: A (0 + 1)-d two matrix model

We will work with the two matrix model given by

$$L = \frac{1}{2} \left[\left(\partial_t A'_{ij} \right)^2 - \omega_A^2 \left(A'_{ij} \right)^2 + \left(\partial_t B'_{ij} \right)^2 - \omega_B^2 \left(B'_{ij} \right)^2 + \lambda A'_{ij} B'_{ij} \right]. \quad (3.61)$$

Here we have put the interaction term with coupling λ so that A' and B' are not independently diagonalized. We also demand that $\lambda \approx 0$, so that this coupling term does not have a significant contribution to the energy and the low energy states are the same as the states in the free field theory to a very good approximation. The ω 's here enforce an IR cutoff, and as a result, we do not have any soft modes in the problem.

We will also impose that our physical observables are singlets of the global group⁸. Therefore we diagonalize B_{ij} by using $U^{-1}B'U = B$, with B being a diagonal matrix comprising of the eigenvalues of B . Under the same transformation, $U^{-1}A'U = A$ which is a non-diagonal matrix. Note that this transformation is akin to gauge fixing and a similar transformation in gauged matrix models removes most of the gauge freedom. In the limit $\lambda \approx 0$, the equations of motion of A and B are given by

$$\left[\frac{\partial^2}{\partial t^2} + \omega_A^2 \right] A_{ij}(\vec{x}, t) = 0, \quad \left[\frac{\partial^2}{\partial t^2} + \omega_B^2 \right] B_{ii}(\vec{x}, t) = 0. \quad (3.62)$$

In this case, we will have $N^2 + N$ number of independent oscillators, with N^2 coming from A and N coming from B . In the large N limit, the N^2 oscillators are responsible for the Hagedorn growth of states. We will quantize the system by imposing the commutation relations

$$\left[a_{ij}(k), a_{i'j'}^\dagger(k') \right] = \delta_{ii'} \delta_{jj'} \quad \left[b_{ii}(k), b_{i'i'}^\dagger(k') \right] = \delta_{ii'} \delta_{ii'}. \quad (3.63)$$

⁸In a sense this replicates features of matrix models in which the matrices transform under a gauge group, where the relevant observables are gauge singlets.

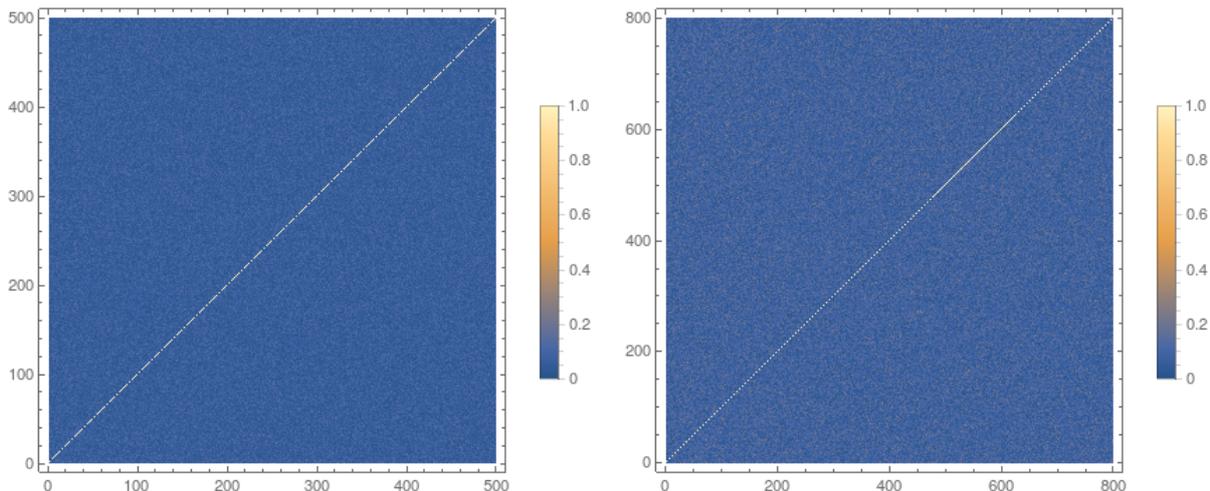


Figure 3.6: We display the inner products which arise from embedding $m = 500$ approximately equidistant vectors in the $n = 104$ -dimensional small Hilbert space of Toy Model I in the left figure and $m = 800$ vectors in the same Hilbert space in the right figure. A point on these plots corresponds to the absolute value of the inner product between vectors lying on the x -axis and the y -axis and hence $x = y$ line has inner product equal to 1.

The vacuum of this system is given in the following equation. We generate the state space by the repeated application of these oscillators on the vacuum.

$$a_{ij} |0\rangle = 0, \quad b_{ii} |0\rangle = 0. \quad (3.64)$$

Typical states, the small algebra and the small Hilbert space

We work with $N = 4$ for Toy Model I, for which we have $N^2 + N = 20$ creation and annihilation operators. We will set the zero-point energy of the matrix model to zero for our case by subtracting it off from the energy and thus redefining it, and set ω_A and ω_B both to 1 while setting $\lambda = 0.01$.

The first thing to construct here is the typical state. To do this we first select energy eigenstates in a range ΔE about average energy E such that the energies lie in the interval $E \pm \frac{\Delta E}{2}$. The typical state is now created using a random superposition of these energy eigenstates. We take the $E = 16$ with an interval $\frac{\Delta E}{2} = 3$. We now construct a typical state with random c_i 's weighing energy eigenstates in the interval ΔE such that $\sum_i |c_i|^2 = 1$. The inverse temperature β of this system is calculated using the first law $\beta = \frac{\Delta S}{\Delta E} = 0.92$.

We will now construct the small algebra and subsequently, create the small Hilbert space. We will demand the following three conditions on the small algebra:

- None of the operators in the small algebra annihilates the typical state.
- The maximum number of operator insertions on the state is less than 20, i.e. should be lesser than $O(N^2 + N)$. We take the maximum number 4.

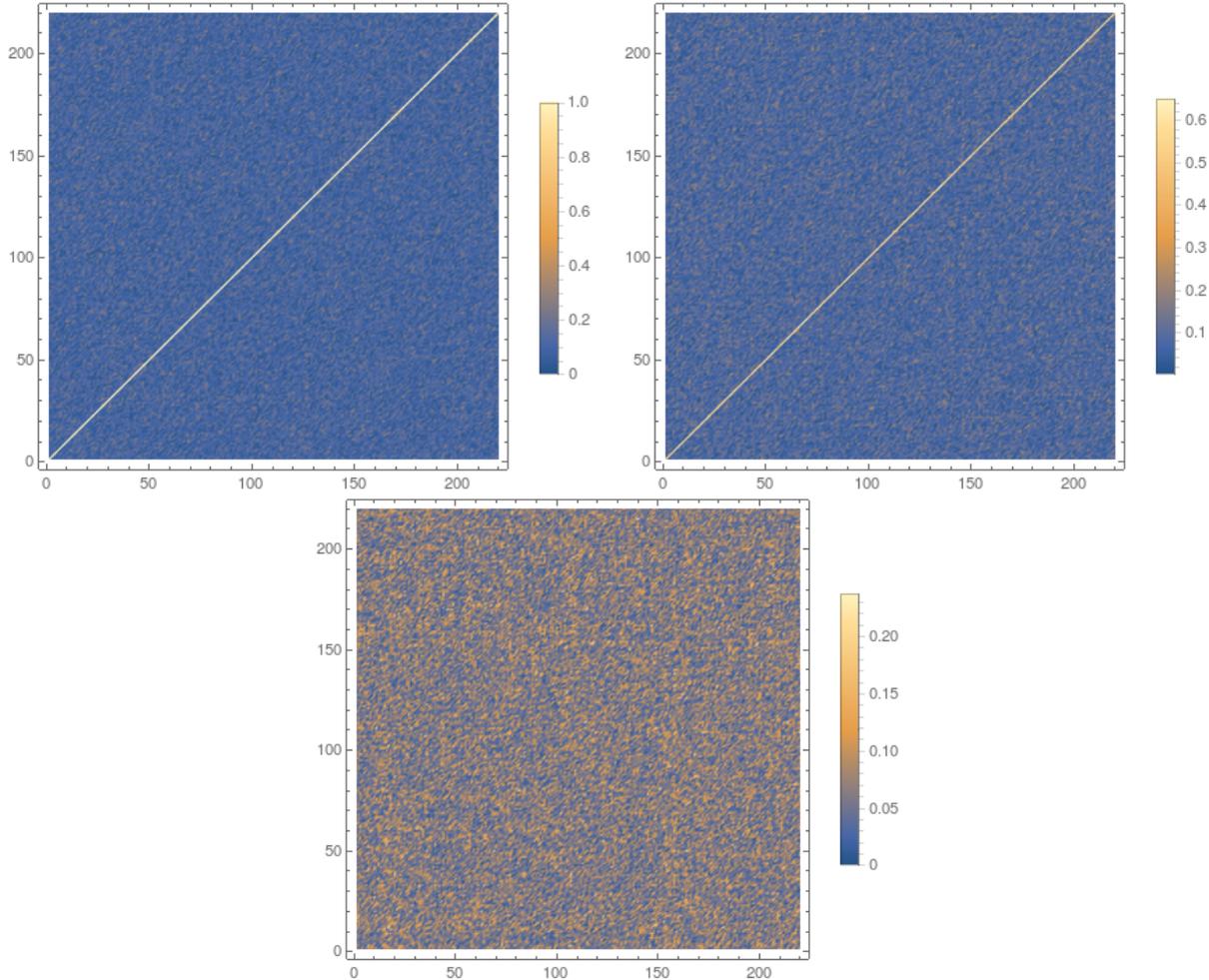


Figure 3.7: We create $m = 220$ excitations in the small Hilbert space of Toy model I. Each excitation is separated from the previous one by time $T = 0$ in the figure on the top left, $T = 1000$ in the top right and $T = 10^5$ in the bottom. We see that increasing time separation gradually washes out the inner products, especially the correlations on the line $x = y$. This behaviour of the inner products indicates that spatially separated excitations on the maximal volume slice have small inner products and hence a "fat tail" in quantum gravity and deviates from the semiclassical zero overlap prediction.

- The maximum energy of the operator insertions is 3, i.e. much less than average energy E which in our case is 16. The energy of operator insertions should not take us outside ΔE about E in order to ensure that the backreaction is small.

Using the above conditions we can identify all 104 possible operators and act them on the typical state to generate the 104-dimensional small Hilbert space.

$$\mathcal{H}_{|\text{TYP}\rangle} := \mathcal{A} |\text{TYP}\rangle, \quad (3.65)$$

Although not orthogonal these vectors are all linearly independent. As a cross-check, we computed the rank of the matrix constructed with all these vectors, which was found to be 104.

Kinematical demonstration of overcounting in toy model I

We will now use the small Hilbert space to create the interior states as given in equation (3.66) where $O_i(\omega) \in \mathcal{A}$.

$$|\psi_i\rangle = K_i e^{-\frac{\beta H}{2}} O_i(\omega) e^{\frac{\beta H}{2}} |\text{TYP}\rangle \quad (3.66)$$

We now construct interior "bulk-like states" by taking combinations of singlet states in (3.66). Each of these states corresponds to the action of an "interior bulk" operator on the typical state. We generate $m = 220$ vectors $|v^j\rangle$ spaced apart from each other in the Hilbert space by defining an energy cost between them, which minimizes their inner products. We implement this energy cost numerically by pushing the vectors around in the small Hilbert space (the sphere discussed in §3.2.1) such that they roughly become equidistant. We discuss this technique in detail in Appendix 6.4. The resulting "interior bulk states" are given below where each of them depends on the choice of coefficients Z_i^j

$$|v^j\rangle = \sum_i Z_i^j |\psi_i\rangle. \quad (3.67)$$

Here the choice of Z_i^j is determined by the energy cost which we can manually select. We plot the vectors' inner products in Figure 3.6, where each point denotes the absolute value of the inner product between a vector on the x -axis and a vector on the y -axis. The $x = y$ line has inner product 1, which indicates that these vectors are normalized. As a consistency check, the 220×220 matrix generated by these "bulk states" has rank 104. It can be seen from Figure 3.6 that there is a finitely non-zero inner product between these bulk vectors. As we increase the dimension of the Hilbert space, these inner products can be made quite small yet finite.

Excitations separated far apart in the "interior"

Till now, we have given a kinematical description of the "bulk excitations", i.e. we took the Hilbert space and showed that there exist vectors which have small inner products. In order to model the placement of the "bulk excitations" far apart on the maximum volume slices of the black hole, we need to send in each excitation long after the previous one. The static description corresponds to the excitations all sent in at the same time, which means that independent excitations are lying nearby close enough on the Cauchy slice and are not separated far apart. We now plot the dynamical case in Figure 3.7 where we send subsequent excitations with a time T separated between them. We now model the bags of gold paradox as given in §3.1 by placing the excitations far apart from each other on the Cauchy slice which corresponds to large numerical values of T .

We note a few interesting observations regarding the dynamical plot. The diagonal line here is the inner product of a vector on the x -axis with time evolution acting on the

same vector on the y -axis. At $T = 1000$, we see that the diagonal line fades away a bit and the larger inner products get slowly washed out. At a very late time, $T = 10^5$ the diagonal line completely vanishes. This disappearance corresponds to the case when the excitations on the bulk are placed quite far apart on the maximal volume slices. As we can see, the time evolution washes out correlations between the vectors, and the larger inner products cease to exist. Such a washing-out behaviour verifies the "fat tail" of inner products, which means that at late times the CFT excitations have a small overlap and is consistent with our derivation in §3.5.2. This numerical overlap becomes lesser and lesser if the dimension of the Hilbert space increases because there is much more space in the Hilbert space to accommodate all the vectors.

3.7.2 Toy Model II: A (3 + 1)-d CFT on $S_3 \times \mathbb{R}$

Toy matrix model I illustrates basic overcounting features for a thermal state constructed out of a matrix model. We will now proceed onto another example which is given by a CFT toy model. Here we first write down the CFT partition function and use it to calculate the Hagedorn temperature which allows us to work in the regime of big AdS black holes. We will construct a typical state at a temperature just above the Hagedorn temperature and demonstrate overcounting of bulk excitations. The metric on $S_3 \times \mathbb{R}$ is given by:

$$ds^2 = -dt^2 + a^2 d\psi^2 + a^2 \sin^2 \psi \left(d\theta^2 + \sin^2 \theta d\phi^2 \right), \quad (3.68)$$

where a is the radius of the S_3 ; ψ and θ go from $(0, \pi)$ and ϕ goes from $(0, 2\pi)$. On this manifold, we write down a CFT action of two matrix-valued bosonic oscillators A and B transforming under the adjoint representation of $U(N)$ global group in (3.69).

$$S_{\text{CFT}} = -\frac{1}{2} \int d^4x \sqrt{-g} \left(g^{\mu\nu} \partial_\mu A'_{ij} \partial_\nu A'_{ij} + g^{\mu\nu} \partial_\mu B'_{ij} \partial_\nu B'_{ij} + \frac{R}{6} \left[(A'_{ij})^2 + (B'_{ij})^2 \right] \right) \quad (3.69)$$

For the metric given in (3.68) the Ricci scalar is given by $R = \frac{6}{a^2}$. As in the previous toy model, we will add a small interaction term with a coupling $\lambda \approx 0$. The small coupling ensures that matrices A' and B' cannot be diagonalized independently, and the energy eigenstates are approximately the same as that of free matrix models.

$$S = S_{\text{CFT}} - \frac{\lambda}{2} \int d^4x \sqrt{-g} A'_{ij} B'_{ij} \quad (3.70)$$

We will again demand that the physical observables are global group singlets. This time instead of fixing the $U(N)$ matrices using diagonalization, we will perform a precise counting of the number of global group singlets constituting a thermal ensemble. We are interested in the following physical observables: average energy, entropy and the dimensionality of the Hilbert space. We will derive these quantities by evaluating the

thermal partition function of the matrix model. We outline this calculation in Appendix 6.2 where we count the number of group singlets using characters of $U(N)$ group and use it to write down the partition function in terms of a Coulomb gas problem with an attractive and a repulsive term. Counting only the group singlets allows us to model the confinement-deconfinement phase transition in the matrix model [121–123]. We calculate that the "Hagedorn temperature" of this system is given by $T_H = 0.63$. The thermodynamic observables at a temperature slightly above Hagedorn temperature $T = 0.64$ are listed in Table 3.3.

N	Entropy (S_N)	Average energy (E_N)	Dimension of Hilbert space ($D \approx e^{S_N}$)
2	3.43	0.54	31
3	4.76	1.64	116
4	5.68	3.25	293
5	6.59	5.63	725

Table 3.3: Entropy, Average energy and dimension of Hilbert space for small N Toy Matrix model II at $T = 0.64$

Typical states, the small algebra and the small Hilbert space

We will work with the $N = 5$ case, which gives us $N^2 + N = 30$ independent oscillators. We set the following parameters: the radius of S_3 is given by $a = 1.55$, $\lambda = 0.01$ and $T = 0.64$. As in the previous model we now construct a typical state with $T = 0.64$, which we accomplish by taking states in an interval ΔE such that $\frac{1}{T} = \frac{\Delta S}{\Delta E}$. We take energy eigenstates spreaded within $\Delta E = 3$ about $E = 5.63$ and create the typical state by random superposition of these vectors. This gives us a microcanonical description of the matrix model for $N = 5$ at $T = 0.64$, the canonical description of which is given in Table 3.3.

We again construct the small Hilbert space by the action of the small algebra on this typical state, where the small algebra satisfies the following conditions:

- The number of operator insertions on the state is much lesser than 30, i.e. should be lesser than $O(N^2 + N)$. We choose that the maximum number of operator insertions on the typical state is 1.
- The maximum energy of the operator insertions is 1.5, i.e. much less than average energy E which in our case is 16. The energy of operator insertions should not take us outside ΔE about E in order to ensure that the backreaction is small.
- None of the operators in the small algebra annihilates the typical state.

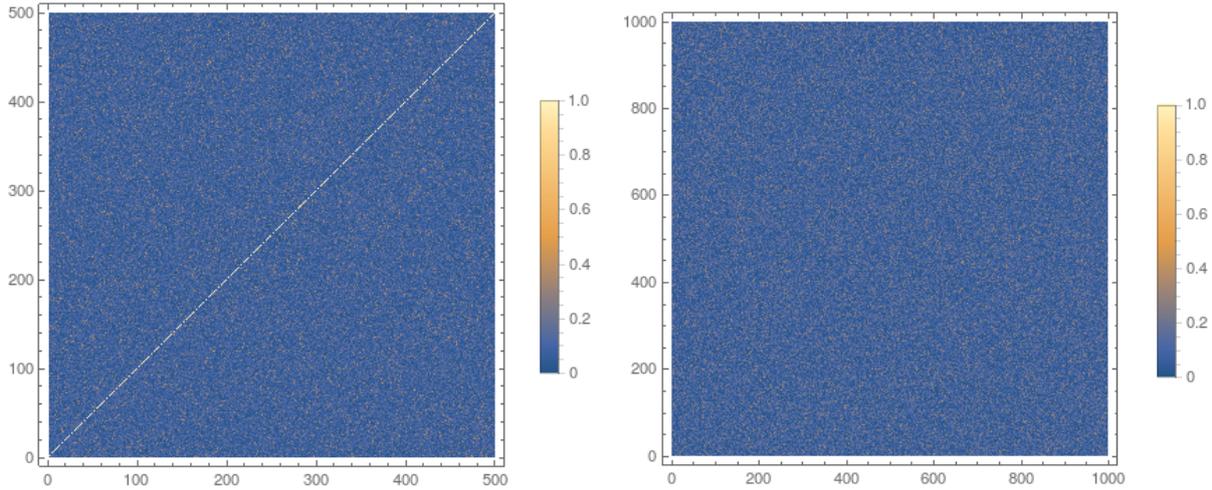


Figure 3.8: In the left figure, we have $m = 500$ excitations embedded in the $n = 61$ -dimensional small Hilbert space of Toy model II. This embedding corresponds to the static case where the time difference between consequent excitations is $T = 0$. The right figure deals with $m = 1000$ excitations created from the $n = 61$ -dimensional Hilbert space. Here we have a time difference of $T = 10^5$ between consecutive excitations. This case corresponds to inner products between excitations placed far apart on the maximal volume slice. The inner products saturate at late times which we can see from the washing out of inner products on the $x = y$ line.

Since the maximum number of operator insertions is 1, we have 61 states generated by the creation and annihilation operators, and hence the dimension of the small Hilbert space is 61. We will now construct the interior states using operators $O_i(\omega) \in \mathcal{A}$:

$$|\psi_i\rangle = K_i e^{-\frac{\beta H}{2}} O_i(\omega) e^{\frac{\beta H}{2}} |\text{TYP}\rangle \quad (3.71)$$

Kinematical demonstration of overcounting in toy model II

The vectors in (3.71) constitute the interior bulk excitations in the large N limit, where smeared semiclassical states correspond to combinations of these excitations living in the small Hilbert space. These bulk excitations have the form given in equation 3.67. Using the energy cost defined in Appendix 6.4, we now construct $m = 500$ bulk excitations as in Figure 3.8, which are approximately equidistant from each other. Each point in Figure corresponds to the inner product's absolute value between a vector on the x -axis and the y -axis. As expected, the $x = y$ line has an inner product of 1 along it since the states are normalized.

Excitations separated far apart in the "interior"

Till now, we have analyzed overcounting for the static case where the excitations are all situated close to each other on the maximal volume slices in bulk. We now proceed to the dynamical case where we separate the excitations in time, and the corresponding bulk states are spatially separated far apart from each other on the maximal volume nice slices. We expect from the previous toy model that the inner products between these

excitations get washed out at huge time separations. We now model $m = 1000$ vectors embedded in the 61 dimensional Hilbert space and confirm this in Figure 3.8. Here we see that the inner products saturate at a minuscule value if the time difference between two successive excitations is $T = 10^5$. Thus we obtain the predicted "fat tail" for this CFT as well.

We have thus shown using two toy matrix models that the inner products between excitations spaced far apart on the maximum volume slices deviate away from the semi-classically predicted inner product. This is consistent with our finding in §3.5.2 that the inner products in CFTs get saturated at a small number. These serve as examples demonstrating our resolution in §3.2 that the inner products are essential to resolve the bags of gold paradox.

3.8 General properties of systems with overcounted Hilbert spaces

We pose the following important question in this section: Since our proposed resolution says that the Hilbert space is overcounted due to small inner products between vectors, what are the physical consequences of such a resolution? In other words, can physical systems in our real-world also have a similar overcounting situation, thereby leading to a much smaller Hilbert space than what we think they have? We see that there exist some significant obstructions to such a situation.

Simulating a quantum system's Hilbert space using a smaller Hilbert space

Consider an m -dimensional "original" Hilbert space, which can be spanned by m orthonormal vectors. We will now simulate the m -dimensional Hilbert space using a smaller n -dimensional Hilbert space, such that $m > n$ and see whether it leads to any inconsistency in physical observables. We construct a nearly orthogonal basis of m vectors, which then "spans" the larger space with the following inner products:

$$\langle V_1 | V_m \rangle = 0, \quad \langle V_1 | V_i \rangle = \langle V_m | V_i \rangle = \epsilon \quad \forall i = 2, \dots, m-1; \quad \langle V_i | V_j \rangle = 0 \quad \forall i, j = 2, \dots, m-2 \quad (3.72)$$

These vectors are simulating orthogonal states in the larger Hilbert space. We will now consider the Hamiltonian acting on the m -dimensional space given by:

$$H = \sum_{i=2}^{m-1} |V_i\rangle\langle V_i|. \quad (3.73)$$

This Hamiltonian time evolves the state $|V_1\rangle$ to $e^{-iHt}|V_1\rangle$. Therefore starting from $|V_1\rangle$, time evolution will never lead to $|V_m\rangle$ in the original Hilbert space. We will keep the form of the Hamiltonian same in the smaller Hilbert space in order to not tamper with the energy spectrum. This time evolution takes place within the space spanned by the vectors:

$$|V_1\rangle, |\Psi\rangle = \frac{1}{\sqrt{m-2}} P \sum_{i=2}^{m-2} |V_i\rangle, |V_m\rangle,$$

where P denotes a projector which projects a vector onto the subspace orthogonal to $|V_1\rangle$ and $|V_m\rangle$. In this basis, the Hamiltonian in the above subspace takes the form:

$$\approx \begin{pmatrix} m\epsilon^2 & \sqrt{m}\epsilon & m\epsilon^2 \\ \sqrt{m}\epsilon & 1 & \sqrt{m}\epsilon \\ m\epsilon^2 & \sqrt{m}\epsilon & m\epsilon^2 \end{pmatrix}. \quad (3.74)$$

The original Hamiltonian acting on the m -dimensional Hilbert space expressed in the orthonormal basis is given by the following matrix,

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.75)$$

These matrices are not the same and their physical properties are very different for large m . The matrix given in equation (3.74) can almost perfectly transfer the state $|V_1\rangle$ to $|V_m\rangle$ in a time $t = \pi/(1 + 2m\epsilon^2)$ [124–126]. Therefore we arrive at a contradiction here. If we try to simulate a system without compromising upon the Hamiltonian's form, then they can behave erratically under time evolution. Conventional quantum systems thus cannot be described using a smaller Hilbert space as they can demonstrate forbidden quantum state transfers. Such quantum state transfers are a generic feature of simulated larger Hilbert spaces. Earlier, we argued that the semiclassical Hilbert space of gravity is a simulated Hilbert space with small inner products. It will be interesting to understand precisely what kind of such quantum state transfers occur in semiclassical gravity, and what novel physical features do they display.

Simulating a thermal system using a smaller Hilbert space

Consider vectors in an n dimensional Hilbert space simulating a larger m dimensional Hilbert space with $m \gg n$. Here we consider that the physical system is thermal. The vectors in the Hilbert space satisfy the following conditions:

$$\langle v_i | v_i \rangle = 1 \quad \& \quad |\langle v_i | v_j \rangle| \sim \epsilon, \quad i \neq j. \quad (3.76)$$

The thermal system under consideration is specified by energy levels spreaded over $(E \pm \Delta E)$ such that $\Delta E \ll E$. We are interested in the order of magnitude of the partition function, which is given by:

$$Z(\beta) = \text{Tr} \left(e^{-\beta H} \right) = O \left(m e^{-\beta |E|} \right). \quad (3.77)$$

We will now see that the simulated thermal system's partition function is has a significant correction.

$$Z_{\text{sim}}(\beta) = \text{Tr} \left(e^{-\beta H} \right) \sim \sum_i e^{-\beta E_i} + \sum_{i \neq j} \left[e^{-\beta E} \epsilon^2 \right]_{ij} = O \left[m e^{-\beta |E|} (1 + m \epsilon^2) \right]. \quad (3.78)$$

We can see that even with tiny corrections to the inner product of the order of $|\epsilon| \sim \frac{1}{\sqrt{m}}$ we will end up with an immense contribution to the partition function. Thermodynamic observables in a system are functions of the partition function and its derivatives. It is safe to say that such a significant contribution to the partition function messes up details of the thermodynamic observables in the system.

As long as $|\epsilon| \ll \frac{1}{\sqrt{m}}$ we don't have a problem with the thermodynamic observables. This is consistent with our observation from (3.28) that such a situation does not lead to the possibility of a big overcounting.

Chapter 4

Monogamy paradox in flat space

We now discuss a toy model of the monogamy paradox in flat space. An essential feature of the paradox is that such a paradox cannot be resolved using small corrections of $O(G_N)$ as proposed in our previous analysis of the bags of gold paradox, but is an $O(1)$ paradox.

4.1 Outline of our work

In §4.2 we review the construction of CHSH operators within quantum mechanics. We generalize this construction in §4.3 and calculate the CHSH correlation between regions A and B within a local quantum field theoretic framework for asymptotically flat space.

We now list the main non-trivial constructions and results of our work. In §4.4 we outline the construction of the operators C_i living on region C , which mimic the action of operators B_i on the global vacuum and write their CHSH correlation with A_i 's. We then use the CHSH correlations between AC and AB to set up the paradox in monogamy. In §4.4.4 we argue the resolution of the monogamy paradox in detail. In §5.2 we summarize our work and discuss related perspectives. In Appendix 6.10, we give a Fourier analytic proof for the existence of near boundary modes C_i subject to the constraints in our construction.

4.2 CHSH inequalities in quantum mechanics

This section reviews the CHSH inequality for quantum mechanical systems and uses them to provide a factual statement about the monogamy of entanglement.

4.2.1 CHSH operator and monogamy of entanglement

Consider a tripartite system composed of independent subsystems A, B , and C . We label operators belonging to the algebra of A as A_i and so on for the other subsystems. Let us look at two pairs of operators A_i and B_i where $i \in (1, 2)$ which satisfy the commutation

relations $[A_i, B_j] = 0$. These operators are constructed such that their eigenvalues lie in the interval $[-1, 1]$, or in other words $\|A\|, \|B\| \leq 1$. The CHSH operator is given by

$$C_{AB} = A_1B_1 + A_1B_2 + A_2B_1 - A_2B_2. \quad (4.1)$$

Classically, the maximum value of the CHSH operator is given by 2, which is the case when A_1 and A_2 are independent while $B_1 = B_2$ or $B_1 = -B_2$. However, this bound no longer holds in quantum mechanics if we evaluate the expectation value of the CHSH operator over a general state $|\psi\rangle$.

In order to estimate the quantum bound on the CHSH operator, let us square the same, which gives us $C_{AB}^2 = 4 - [A_1, A_2][B_1, B_2]$. Since the norm of the commutator is given by $|[A_1, A_2]| \leq 2$, we arrive at $|\langle C_{AB} \rangle| \leq 2\sqrt{2}$.

Now if we consider the square of the expectation value of the CHSH operators defined over AB and AC , then the statement of the monogamy of entanglement is as follows [87]:

$$\langle C_{AB} \rangle^2 + \langle C_{AC} \rangle^2 \leq 8. \quad (4.2)$$

The above relation statement quantifies the maximum entanglement that subsystem AC can possess provided there is a given entanglement among the subsystem AB . An interesting conclusion which follows is that there cannot be a scenario where the correlations between AB and AC both possess a non-classical description, i.e. both $\langle C_{AB} \rangle, \langle C_{AC} \rangle > 2$. Another outcome is that if the system AB is maximally entangled, i.e. $\langle C_{AB} \rangle = 2\sqrt{2}$, then AB cannot be entangled. Thus we have a precise statement regarding the violation of monogamy of entanglement, which violates the inequality given in (4.2).

4.2.2 Baby example: Bell operators using simple harmonic oscillators

Consider a pair of commuting simple harmonic oscillators living in separate regions A and B . We denote their corresponding annihilation operators as α_s , and their respective vacua as $|0\rangle_s$, where $s = A/B$. We want to evaluate the expectation value of the CHSH operator on the thermofield double state where $x^2 < 1$,

$$|\text{TFD}\rangle = \sqrt{1-x^2} e^{x\alpha_A^\dagger \alpha_B^\dagger} |0\rangle_A |0\rangle_B. \quad (4.3)$$

The above state reduces to the standard thermofield double case if we set $x^2 = e^{-\beta}$. Denoting projectors onto the s -th vacuum as P_s , we now choose Bell operators as follows:

$$\begin{aligned}
\text{A operators: } \quad & A_1 = P_A - \alpha_A^\dagger P_A \alpha_A \\
& A_2 = \alpha_A^\dagger P_A + P_A \alpha_A \\
\text{B operators: } \quad & B_1 = \frac{1}{\sqrt{2}} (P_B - \alpha_B^\dagger P_B \alpha_B + \alpha_B^\dagger P_B + P_B \alpha_B) \\
& B_2 = \frac{1}{\sqrt{2}} (P_B - \alpha_B^\dagger P_B \alpha_B - \alpha_B^\dagger P_B - P_B \alpha_B)
\end{aligned} \tag{4.4}$$

These operators are inspired by the Bell operators for spin- $\frac{1}{2}$ systems, and might look confusing at first glance. However, expanding the operators in the number basis gives us a much simpler looking form for the same.

$$\begin{aligned}
\text{A operators: } \quad & A_1 = |0\rangle_A \langle 0|_A - |1\rangle_A \langle 1|_A \\
& A_2 = |0\rangle_A \langle 1|_A + |1\rangle_A \langle 0|_A \\
\text{B operators: } \quad & B_1 = \frac{1}{\sqrt{2}} (|0\rangle_B \langle 0|_B - |1\rangle_B \langle 1|_B + |0\rangle_B \langle 1|_B + |1\rangle_B \langle 0|_B) \\
& B_2 = \frac{1}{\sqrt{2}} (|0\rangle_B \langle 0|_B - |1\rangle_B \langle 1|_B - |0\rangle_B \langle 1|_B - |1\rangle_B \langle 0|_B)
\end{aligned} \tag{4.5}$$

These are precisely the operators used in the spin- $\frac{1}{2}$ problem, with $|0\rangle / |1\rangle$ denoting the two states and the operators resembling combinations of Pauli matrices. We now evaluate the expectation value of the CHSH operator on the thermofield double state, which gives us

$$\langle C_{AB} \rangle = \sqrt{2} (1+x)^3 (1-x). \tag{4.6}$$

This takes a maximum value at $x = \frac{1}{2}$ with the maximum value being $\langle C_{AB} \rangle = \frac{27\sqrt{2}}{16} \approx 2.39 > 2$. Therefore using the above construction we see that the thermofield double state is entangled for $x = \frac{1}{2}$, though not maximally entangled.

4.3 CHSH inequalities in local quantum field theory

In this section, we will extend the above construction of the CHSH correlator for simple harmonic oscillators to analogously construct the CHSH correlator in a local quantum field theory [60, 127–129]. We will then utilize this formalism to calculate $\langle C_{AB} \rangle$ for smeared modes within a small interval on either side of an outgoing light cone in an empty flat space. This section is computationally intensive, and readers not interested

in the details of the computation can skip directly to §4.3.4, where we summarize the contents of this section.

4.3.1 Basic conventions and choice of operators

We define Hermitian operators (X_s, Π_s) on the spatially compact regions A and B , such that they satisfy canonical commutation relations. Consequently we can also define annihilation operators given by $\alpha_s = \frac{1}{\sqrt{2}}(X_s + i\Pi_s)$. These operators obey the simple harmonic commutation relations

$$[\alpha_s, \alpha_{s'}^\dagger] = \delta_{ss'}. \quad (4.7)$$

In addition to these modes, there also exist global modes for flat spacetime. These global modes in flat space obey the canonical commutators

$$[a_{\omega l}, a_{\omega' l'}^\dagger] = \delta_{l,l'} \delta(\omega - \omega'). \quad (4.8)$$

The global modes are related to α_s by Bogoliubov coefficients

$$\alpha_s = \sum_l \int d\omega \left(h_s(\omega, l) a_{\omega, l} + g_s^*(\omega, l) a_{\omega, l}^\dagger \right), \quad (4.9)$$

where the functions $h_s(\omega, l)$ and $g_s^*(\omega, l)$ are related by

$$\sum_l \int d\omega [h_s(\omega, l) h_{s'}^*(\omega, l) - g_s^*(\omega, l) g_{s'}(\omega, l)] = \delta_{s,s'}. \quad (4.10)$$

We rewrite (4.10) in the following fashion for convenience

$$h_s \cdot h_{s'}^* - g_s^* \cdot g_{s'} = \delta_{s,s'}, \quad (4.11)$$

where we have defined $h_s \cdot h_{s'}^* = \sum_l \int d\omega h_s(\omega, l) h_{s'}^*(\omega, l)$. Let us now consider the scenario where the CHSH correlators are evaluated on the global vacuum state, while the CHSH operators are following combinations of $\alpha_s/\alpha_s^\dagger$, which are precisely the same operators in (4.4).

$$\begin{aligned} \text{A operators:} \quad & A_1 = P_A - \alpha_A^\dagger P_A \alpha_A \\ & A_2 = \alpha_A^\dagger P_A + P_A \alpha_A \\ \\ \text{B operators:} \quad & B_1 = \frac{1}{\sqrt{2}} \left(P_B - \alpha_B^\dagger P_B \alpha_B + \alpha_B^\dagger P_B + P_B \alpha_B \right) \\ & B_2 = \frac{1}{\sqrt{2}} \left(P_B - \alpha_B^\dagger P_B \alpha_B - \alpha_B^\dagger P_B - P_B \alpha_B \right) \end{aligned} \quad (4.12)$$

We will proceed using general α_s in §4.3.2. Our physical case of interest is described in §4.3.3, where we will take α_s to be Rindler annihilation modes. Consequently, we have an analogous interpretation of the global state as the thermofield double state as defined in (4.3).

4.3.2 Vacuum projector and the most general two-point correlator

In order to define Bell operators as given in (4.12), we need to construct projectors onto the ground states of each oscillator, which are given by

$$P_s = -\frac{1}{\pi^2} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \int_0^{2\pi} d\theta_s \frac{e^{-(t_1^2+t_2^2)+\kappa(\theta_s)(t_1 X_s - t_2 \Pi_s)}}{e^{i\theta_s} - 1 - \epsilon} \quad (4.13)$$

where $\kappa(\theta) \equiv 2\sqrt{i \tan \theta}$ and ϵ is a small positive constant. The detailed construction of this projector is given in Appendix 6.5.

We can conveniently extract the CHSH correlator from the expression for the *most general two-point correlator*. Using the definition of the projector in (4.13), the most general two-point function is given by:

$$Q[\{v_i, \zeta_i\}] = \frac{1}{\pi^4} \int d^2\vec{t} d^2\vec{y} \int_0^{2\pi} d\theta_A d\theta_B \frac{e^{-(\vec{t}^2 + \vec{y}^2)}}{(e^{i\theta_A} - 1 - \epsilon)(e^{i\theta_B} - 1 - \epsilon)} \\ \times \langle e^{v_2 \alpha_B^\dagger} e^{\kappa(\theta_B)(y_1 X_B - y_2 \Pi_B)} e^{\zeta_2 \alpha_B} e^{v_1 \alpha_A^\dagger} e^{\kappa(\theta_A)(t_1 X_A - t_2 \Pi_A)} e^{\zeta_1 \alpha_A} \rangle \quad (4.14)$$

where $\vec{t} = (t_1, t_2)$. Let us define $\tilde{y}_i = \kappa(\theta_B)y_i$ and $\tilde{t}_i = \kappa(\theta_A)t_i$ so as to write the expectation value in the above integral as

$$\langle \mathcal{G} \rangle \equiv \langle e^{v_2 \alpha_B^\dagger} e^{(\tilde{y}_1 X_B - \tilde{y}_2 \Pi_B)} e^{\zeta_2 \alpha_B} e^{v_1 \alpha_A^\dagger} e^{(\tilde{t}_1 X_A - \tilde{t}_2 \Pi_A)} e^{\zeta_1 \alpha_A} \rangle \quad (4.15)$$

The above two-point correlator and its derivatives at $v_i = 0, \zeta_i = 0$ can be used to obtain the correlators of all relevant CHSH operators as defined in (4.12). As a demonstration, the derivatives of $Q[\{v_i, \zeta_i\}]$ can be easily used to generate correlators of the following form:

$$\left. \partial_{v_2}^{m_2} \partial_{v_1}^{m_1} \partial_{\zeta_2}^{n_2} \partial_{\zeta_1}^{n_1} Q[\{v_i, \zeta_i\}] \right|_{v_i=\zeta_i=0} = \langle \alpha_B^{\dagger m_2} P_B \alpha_B^{n_2} \alpha_A^{\dagger m_1} P_A \alpha_A^{n_1} \rangle. \quad (4.16)$$

We will write the expression for $\langle \mathcal{G} \rangle$ in terms of the global modes. This is performed by expressing α_s in terms of global modes using (4.10). This computation requires repeated application of the BCH lemma while working in a coherent state basis. The detailed calculation is given in § 6.7, and we state the final result here.

$$\langle \mathcal{G} \rangle = \exp \left(\frac{1}{8} \sum_{p,q=1}^4 (f_p \cdot f_q^* + f_p^* \cdot f_q) m_p m_q - \frac{\mathcal{R}}{2} \right) + \mathcal{O} \left(\sqrt{G_N} \right). \quad (4.17)$$

Here we have added corrections of $O(\sqrt{G_N})$ to include the effects of interactions in an interacting theory of the scalar field coupled to gravity, since the interacting vacuum is different from the global vacuum upto $O(\sqrt{G_N})$. Defining $\zeta_i^\pm = \frac{(\zeta_i \pm v_i)}{\sqrt{2}}$, the expression for \mathcal{R} is given by

$$\mathcal{R} = \left(m_1 \zeta_1^+ + i m_2 \zeta_1^- + m_3 \zeta_2^+ + i m_4 \zeta_2^- \right) - v_1 \zeta_1 - v_2 \zeta_2, \quad (4.18)$$

where the quantities f_i, m_i are defined as:

$$f_1 = (h_A + g_A) ; f_2 = -i(h_A - g_A) ; f_3 = (h_B + g_B) ; f_4 = -i(h_B - g_B) \quad (4.19a)$$

$$m_1 = (\tilde{t}_1 + \zeta_1^+) ; m_2 = (-\tilde{t}_2 + i\zeta_1^-) ; m_3 = (\tilde{y}_1 + \zeta_2^+) ; m_4 = (-\tilde{y}_2 + i\zeta_2^-) \quad (4.19b)$$

We can use (4.17) to obtain an expression for $Q[\{v_i, \zeta_i\}]$ in (4.14), since the integrals over \vec{t} and \vec{h} are Gaussian. The θ integration involves a trivial calculation of the residue in the complex plane. We will not write the expression for general h_s and g_s but will calculate the same for the Rindler to Minkowski Bogoliubov coefficients in the following subsection.

4.3.3 CHSH correlation between regions A and B in field theory

Note that our specific case of interest involves smearing operators on bounded regions A and B close to the light cone (See Fig. 1.2). Our smearing choice is such that the operators α_s denote the Rindler oscillators, and h_s, g_s denote the corresponding Rindler to Minkowski Bogoliubov coefficients. While we express the CHSH operators in terms of Rindler operators as given in (4.12), we take the expectation value in the CHSH correlator over the global vacuum, which is a thermofield double state in terms of the Rindler oscillators.

Massless modes in flat space

Consider a massless scalar coupled to gravity in d dimensional Minkowski space. The modes of the massless scalar end up at future null infinity, a fact that will be important in our posing of the monogamy paradox. The equation for the scalar field is given by:

$$\frac{\partial}{\partial x^\mu} \left[\sqrt{-g} g^{\mu\nu} \frac{\partial \Phi}{\partial x^\nu} \right] = 0. \quad (4.20)$$

We solve the above equation in global spherical coordinates (valid for $d \geq 3$):

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega_{d-2}^2 \quad (4.21)$$

where Ω_{d-2} denotes the angles of the $d - 2$ dimensional sphere. The equation of motion can be solved by putting in the ansatz $\Phi(r, t, \Omega) = T(t)\chi(r)Y_l(\Omega)$, where $Y_l(\Omega)$ denotes spherical harmonics of a $d - 2$ dimensional sphere. Here $\chi(r)$ satisfies:

$$\frac{d^2\chi(r)}{dr^2} + \frac{d-2}{r} \frac{d\chi(r)}{dr} + \left(\omega^2 - \frac{l(l+(d-3))}{r^2} \right) \chi(r) = 0, \quad (4.22)$$

where ω is the frequency given by:

$$\frac{d^2T(t)}{dt^2} = -\omega^2 T(t). \quad (4.23)$$

The solution for $\chi(r)$ is given by:

$$\chi(r) = \frac{C_1}{r^{\frac{d-3}{2}}} J_m(r\omega) + \frac{C_2}{r^{\frac{d-3}{2}}} Y_m(r\omega) \quad (4.24)$$

where J and Y denote the standard Bessel functions and $m = l + \frac{d-3}{2}$. We discard the Y term since it blows up at the origin. Thus the complete solution is given by:

$$\Phi(r, t, \Omega) = K \sum_l \int d\omega a_{\omega,l} \frac{J_m(r\omega)}{r^{\frac{d-3}{2}}} e^{-i\omega t} Y_l(\Omega) + \text{h.c.} \quad (4.25)$$

Here K is a normalization constant used to impose the normalization of the canonical commutator $[a_{\omega,l}, a_{\omega',l'}^\dagger] = \delta(\omega - \omega') \delta_{l,l'}$. Computing the momenta from the action of the massless scalar and using the equal time canonical commutation relation:

$$[\Phi(r, t, \Omega), \Pi(r', t, \Omega')] = i\delta(r - r')\delta(\Omega_1 - \Omega_2), \quad (4.26)$$

we obtain $K = \frac{1}{\sqrt{2}}$. Therefore the scalar field is expressed as

$$\Phi(r, t, \Omega) = \frac{1}{\sqrt{2}} \sum_l \int d\omega a_{\omega,l} \frac{J_m(r\omega)}{r^{\frac{d-3}{2}}} e^{-i\omega t} Y_l(\Omega) + \text{h.c.} \quad (4.27)$$

Note that in the preceding discussion we have suppressed the extra indices of the spherical harmonics. As an example, we can explicitly write them for $d = 4$, which gives us

$$\Phi(r, t, \theta, \phi) = \frac{1}{\sqrt{2}} \sum_{l,\bar{m}} \int d\omega a_{\omega,l} \frac{J_{l+\frac{1}{2}}(r\omega)}{r^{\frac{1}{2}}} e^{-i\omega t} Y_l^{\bar{m}}(\theta, \phi) + \text{h.c.}, \quad (4.28)$$

where we have used $Y_l^{\bar{m}}$ to denote the standard spherical harmonics to avoid confusion with m from (4.25).

Smeared operators on A and B

We now outline our construction of *approximately local operators* by smearing the scalar field over the bounded interval in such a way that the Rindler modes are extracted out. To perform this, we introduce a tuning function such that it is supported only on the small bounded regions and smoothly dies off. Recall that the regions A and B are situated just inside and outside an outgoing light cone at r_0 respectively. Thus we define the smeared operators on the regions A and B by

$$\begin{aligned}
\alpha_A &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) \Phi(t_A(U), r_A(U), \Omega) \\
\alpha_B &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_B^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}(U) \Phi(t_B(U), r_B(U), \Omega) \\
\alpha_A^\dagger &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}^*(U) \Phi(t_A(U), r_A(U), \Omega) \\
\alpha_B^\dagger &= \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_B^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}^*(U) \Phi(t_B(U), r_B(U), \Omega)
\end{aligned} \tag{4.29}$$

Here r_s and t_s , $s = A, B$ denote the global spherical coordinates on the regions A and B , as given in (4.21). The smearing function oscillates increasingly as U tends to go near $U = 0$, and thus even a small interval near $U = 0$ is useful to extract out the Rindler modes. Consequently U is integrated from U_l to U_h such that the tuning function $\mathcal{T}(U)$ vanishes smoothly as it approaches U_l and U_h . We work in the limit $U_0 \rightarrow 0$, such that

$$\log \frac{U_l}{U_0} \rightarrow -\infty \quad \text{and} \quad \log \frac{U_h}{U_0} \rightarrow \infty. \tag{4.30}$$

Note that in our convention, we have included the sphere metric determinant $\sqrt{\gamma}$ inside the angular integral in (4.29), such that

$$V_\Omega = \int d^{d-2}\Omega \equiv \frac{2\pi^{\frac{d-1}{2}}}{\Gamma\left(\frac{d-1}{2}\right)}.$$

We assume that the errors due to these length scales are of $O(\epsilon)$ such that $O(\epsilon) \gg O(\sqrt{G_N})$. In order to impose that the regions A and B remain causally disconnected, we assume the following conditions:

$$\begin{aligned}
t_A(U) &= \frac{U}{2} - v_0 & r_A(U) &= r_0 - v_0 - \frac{U}{2} \\
t_B(U) &= -\frac{U}{2} + v_0 & r_B(U) &= r_0 + v_0 + \frac{U}{2}
\end{aligned} \tag{4.31}$$

We also impose the following conditions on the tuning function, so that it is sharply centred about a particular frequency ω_0

$$\mathcal{T}(U) \left[\frac{U}{U_0} \right]^{i\omega_0} = \int \tilde{\mathcal{T}}(\nu) \left[\frac{U}{U_0} \right]^{i\nu} d\nu, \quad \int \frac{d\nu}{\nu} |\tilde{\mathcal{T}}(\nu)|^2 = \frac{1}{\pi}. \quad (4.32)$$

using which we can recover the standard expressions for the commutator of the above defined modes $[\alpha_s, \alpha_{s'}^\dagger] = \delta_{ss'}$ (See Appendix 6.6 for the detailed calculation). Another relation which will be useful in the computation of $\langle C_{AB} \rangle$ is

$$\lim_{\nu \rightarrow 0} \frac{\tilde{\mathcal{T}}(\nu)}{\nu} = 0 \quad (4.33)$$

Bogoliubov coefficients for Rindler modes

In order to calculate $\langle C_{AB} \rangle$, we first need to determine the Bogoliubov coefficients so as to calculate the most general two-point correlator, whose simplification has been derived in (4.17). Since we have smeared the field over the entire sphere in (4.29) on either side of the light cone at $r = r_0$, therefore we only need to look at the $l = 0$ mode. This is because the modes $l \neq 0$ vanish due to the angular integral. The radial part of the $l = 0$ mode takes a very simple form in d -dimensions:

$$\chi(r) \sim \frac{J_{\frac{d-3}{2}}(\omega r)}{r^{\frac{d-3}{2}}} \quad (4.34)$$

We also note that since we have smeared our operators on very small spatial regions A and B , the smearing functions remain almost constant over the region. However using (4.33) our tuning function vanishes for small frequencies. Consequently the Bogoliubov coefficients in (4.9) have support only for large frequencies ω , which we denote by $\omega > \omega'$, where ω' is a large enough frequency above which the Bogoliubov coefficients are non-zero. In the large frequency limit, the above radial function simplifies to

$$\chi(r) \sim \sqrt{\frac{2}{\pi\omega}} \frac{1}{r^{\frac{d-2}{2}}} \cos\left(\omega r - \frac{(d-2)\pi}{4}\right) \quad (4.35)$$

Using the large frequency limit, we evaluate the Bogoliubov coefficients. We refer to Appendix 6.8 for the detailed calculation, and state the main result here.

$$\begin{aligned}
h_A(\omega, 0) &= \frac{e^{-i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{\pi\nu/2} (\omega U_0)^{-i\nu} \Gamma(i\nu) \tilde{\mathcal{T}}(\nu), \\
g_A^*(\omega, 0) &= \frac{e^{i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{-\pi\nu/2} (\omega U_0)^{-i\nu} \Gamma(i\nu) \tilde{\mathcal{T}}(\nu), \\
h_B(\omega, 0) &= \frac{e^{-i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{\pi\nu/2} (\omega U_0)^{i\nu} \Gamma(-i\nu) \tilde{\mathcal{T}}^*(\nu), \\
g_B^*(\omega, 0) &= \frac{e^{i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{-\pi\nu/2} (\omega U_0)^{i\nu} \Gamma(-i\nu) \tilde{\mathcal{T}}^*(\nu).
\end{aligned} \tag{4.36}$$

$\langle C_{AB} \rangle > 2$ for entangled Rindler modes in flat space

We will now use the Bogoliubov coefficients given in (4.36) to evaluate $\langle C_{AB} \rangle$, using (4.17). In order to do this, we need to calculate the 4×4 matrix $f_p \cdot f_q^* + f_p^* \cdot f_q$. The detailed calculation of this matrix is given in Appendix 6.8, and we state the final result.

$$f_p \cdot f_q^* + f_q \cdot f_p^* = \frac{2}{1-x^2} \begin{pmatrix} 1+x^2 & 0 & 2x & 0 \\ 0 & 1+x^2 & 0 & -2x \\ 2x & 0 & 1+x^2 & 0 \\ 0 & -2x & 0 & 1+x^2 \end{pmatrix}. \tag{4.37}$$

Note that the matrix in (4.37) turns out to be the same as obtained for the Rindler-to-global AdS case in [60]. Although solutions to the massless scalar field in AdS and flat space are quite different, it is not surprising that the matrix turns out to be the same. This is because the near-horizon local Rindler modes possess universal features as explained in [80].

We now substitute (4.37) in (4.17) to derive the expression for $\langle \mathcal{G} \rangle$. In order to do so, we perform the Gaussian integrals over \vec{t} and \vec{h} and evaluate the θ integral by calculating the residue about the pole. Using the expression for $\langle \mathcal{G} \rangle$, we derive the result for $\langle C_{AB} \rangle$, which is again given by

$$\langle C_{AB} \rangle = \sqrt{2} (1+x)^3 (1-x). \tag{4.38}$$

The reader might ask why our expression for the CHSH operator's expectation value in QFT is precisely the same as the expression we had derived for the quantum mechanical case. This is simply because our chosen operators and states were essentially the same in both cases. Again the expectation value is maximized at $x = \frac{1}{2}$, where $\langle C_{AB} \rangle$ takes the value

$$\langle C_{AB} \rangle = \frac{27\sqrt{2}}{16} + \mathcal{O}\left(\sqrt{G_N}\right) + \mathcal{O}(\epsilon) \tag{4.39}$$

Here we have included corrections since the interacting vacuum of the scalar-gravity theory is different from the free field vacuum using $O(\sqrt{G_N})$. As defined before, we denote small errors in length scales by $O(\epsilon)$.

4.3.4 Summary of this section

The main goal of this section was to show that within a local quantum field theoretic framework, using a careful choice of operators, we can violate the classical bound. To do this, we first developed the formalism for looking at CHSH correlators in terms of the most general two-point correlator acting on the global vacuum. The key here is to write down the CHSH correlator in terms of Bogoliubov coefficients between the spatially compact regions' modes and the global modes. We then wrote down creation and annihilation operators by smearing the massless scalar field with Rindler smearing functions on small bounded regions A and B situated just inside and outside an outgoing light cone at r_0 and calculated the corresponding Bogoliubov coefficients between these operators and the global Minkowski operators. We used these Bogoliubov coefficients to obtain $\langle C_{AB} \rangle$, where we take the expectation value over the global Minkowski vacuum, which looks like a thermofield double in terms of the Rindler oscillators. In particular, our construction of operators in the local QFT is the same as done for the quantum mechanical case described earlier in §4.2. Consequently the CHSH correlator is given by (4.38), whose maximum value is $\langle C_{AB} \rangle \approx 2.39$, which violates the classical bound.

4.4 The monogamy paradox in flat space

We will now outline the paradox in the monogamy of entanglement. Previously in (4.39) we have derived that up to small corrections, we can obtain $\langle C_{AB} \rangle = 2.39 > 2$, which indicates non-classicality. We will now consider another region C situated far away from our system AB at \mathcal{I}_-^+ , and consider operators C_i supported on the same (See Fig. 1.2). Applying (4.2) to a local QFT, we have the following upper bound on the CHSH correlators between systems AB and AC

$$\langle C_{AB} \rangle^2 + \langle C_{AC} \rangle^2 \leq 8. \quad (4.40)$$

In this section, we will show that using the Reeh-Schlieder theorem and the fact that in a theory of gravity, the Hamiltonian is a boundary term [89], we can create operators C_i such that their action on the vacuum is the same as the action of operators B_i . Consequently, the expectation in (4.40) based on local quantum field theory is violated up to an $O(1)$ extent.

Unless indicated otherwise, from here on, we will restrict ourselves to describing the

effects of gravity in four dimensions. Firstly we describe the Hilbert space of the theory and construct operators relevant to our calculation. Then we calculate the $\langle C_{AC} \rangle$ correlator and pose the paradox. We will further discuss conditions on the vacuum structure under which we can similarly pose the paradox in general dimensions. Towards the end of this section, we discuss the resolution of the paradox.

4.4.1 Gravity in asymptotically flat spacetime

In this subsection, we will describe the Hilbert space of the four-dimensional flat space theory and construct a boundary projector onto low energy states. This projector will be essential to construct bounded operators C_i within a small region at the past of future null infinity (\mathcal{I}_-^+). Readers familiar with the details of this section can directly proceed onto §4.4.2.

The Hilbert space

A good coordinate system which encapsulates the asymptotic large- r structure near the future null infinity is the retarded Bondi coordinates [130].

$$ds^2 = -du^2 - 2dudr + r^2 \gamma_{AB} d\Omega^A d\Omega^B + r C_{AB} d\Omega^A d\Omega^B + \frac{2m_B}{r} du^2 + \gamma^{DA} D_D C_{AB} du d\Omega^B + \dots \quad (4.41)$$

There is an infinite-dimensional symmetry group in the asymptotic region consistent with the leading falloff given above [130–135]. These symmetries are called supertranslations which are generated by the following charges:

$$Q_{lm} = \frac{1}{4\pi G_N} \int \sqrt{\gamma} d^2\Omega m_B(u = -\infty, \Omega) Y_{l,m}(\Omega) \quad (4.42)$$

The Bondi news is given by the u -derivative of the shear, $N_{AB} = \partial_u C_{AB}$. This tensor has a zero mode, which is used to split the supertranslation charges into two parts, a soft part and a hard part. Technically it is the soft part that leads to the asymptotic symmetries, while the hard part contains stress-energy contributions.

We will briefly talk about the Fock space of this asymptotic theory [90–92] which is elaborated in more detail in [79, 93]. Since the news tensor contains a zero mode, the vacuum must be specified not only by the annihilation of the positive frequency modes of the news tensor and the scalar field but should also be labelled by the eigenvalue under the supertranslation sector

$$Q_{lm} |0, \{s\}\rangle = s_{l,m} |0, \{s\}\rangle \quad (4.43)$$

Here 0 denotes that the positive frequency modes of the field, i.e. the hard part of the supertranslation charges annihilate the vacuum. By smearing over the energies using

suitable tuning functions¹, where the smearing scale can be taken to be arbitrarily small, the inner product between two states is given by

$$\langle \{n_\omega\}, \{s\} | \{n'_\omega\}, \{s'\} \rangle = \prod_{l,m} \delta_{\{n_\omega\}, \{n'_\omega\}} \delta(s_{l,m} - s'_{l,m}). \quad (4.44)$$

where the Dirac delta function goes over the space of all l, m . Consequently, we can build up the Hilbert space by acting with the massless scalar and the news field on the top of each vacuum labelled by $|0, \{s\}\rangle$. Thus the Hilbert space is fragmented into different sectors, with an element from one sector orthogonal to another from a different sector. Thus Hilbert space of canonical gravity is given by

$$\mathcal{H} = \bigoplus_{\{s\}} \mathcal{H}_{\{s\}} \quad (4.45)$$

We will pause here to clarify some important aspects while working with the Fock space as described in (4.43), (4.44) and (4.45). Physically in order to compute meaningful quantities, we write the state of our scalar field as follows:

$$|\{n_\omega\}, \mathcal{S}\rangle \equiv \int \left(\prod_{l,m} ds_{l,m} \right) \mathcal{S}(\{s\}) |\{n_\omega\}, \{s\}\rangle, \quad (4.46)$$

where we have smeared the supertranslation of the vacuum, with the peak of the smearing function $\mathcal{S}(\{s\})$ centred about a particular $s_{l,m}$ to ensure normalizability of states. The smearing function $\mathcal{S}(\{s\})$ is chosen such that our states have unit norm. Therefore using (4.46) and (4.44), the inner product between smeared states is given by

$$\langle \{n_\omega\}, \mathcal{S} | \{n'_\omega\}, \mathcal{S}' \rangle = \delta_{\mathcal{S}, \mathcal{S}'} \delta_{\{n_\omega\}, \{n'_\omega\}}. \quad (4.47)$$

We note a critical assumption in our discussion: we have ignored UV corrections, e.g., stringy effects, and assumed that the low energy effective physics correctly describes the low energy structure of quantum gravity. This assumption seems quite reasonable since gravity is an excellent effective field theory up to the Planck scale. In our work, we pose the paradox within a low energy framework where we perform only tree-level calculations, and hence we are not bothered by any possible modification to the Hilbert space introduced by a UV completion of gravity such as string theory.

Finally, we also note that our construction manifestly ensures that our Fock space is separable. This statement can also be motivated using constructive QFT [136–139].

¹To see why smearing over energy kets is convenient, note that in QFT without the inclusion of gravity, the vacuum is normalizable but the excited states are not. In order to work with states satisfying nice properties, we smear over energy kets there as well. Also note that the Fock space in both cases, i.e., QFT with or without gravity is separable.

Boundary projector

We will now use the gravity Hamiltonian to write down a projector in asymptotically flat spacetime [79]. We first write the Bondi mass, which is the integration of the Bondi mass aspect over the sphere at infinity.

$$M(u) = \int d^2\Omega \sqrt{\gamma} m_B(u, \Omega) \quad (4.48)$$

Note that the Bondi mass at $u \rightarrow -\infty$ is the $m = 0, l = 0$ component of supertranslation charges Q_{lm} . The Bondi mass reduces to the canonical ADM Hamiltonian in the limit $u \rightarrow -\infty$ [89, 140, 141]:

$$\lim_{u \rightarrow -\infty} \frac{M(u)}{4\pi G_N} = H. \quad (4.49)$$

The ADM Hamiltonian can be expressed in terms of the boundary metric, which is given by

$$H = \lim_{u \rightarrow -\infty} \frac{M(u)}{4\pi G_N} = \lim_{r \rightarrow \infty} \frac{1}{4\pi G_N} \int d^2\Omega \sqrt{\gamma} (r h_{00}(r, \Omega)). \quad (4.50)$$

Using this boundary Hamiltonian, we can write down a projector residing at \mathcal{I}_-^+ . The projector onto the subspace of vacuum states labelled by supertranslations is constructed by taking the following limit [79, 84]:

$$\mathcal{P}_0 = \lim_{a \rightarrow \infty} \exp(-aH). \quad (4.51)$$

where the subscript 0 in the projector denotes that we are projecting onto the degenerate subspace of zero energy states spanned by supertranslations. We can express this projector as an operator on the Fock space as follows²

$$\mathcal{P}_0 = \int \left(\prod_{l,m} ds_{l,m} \right) |0, \{s\}\rangle \langle 0, \{s\}| + \mathcal{O}(\sqrt{G_N}). \quad (4.52)$$

However, in practice, the projector written in (4.51) is defined only in an abstract sense. A more physically motivated projector in flat space should project only up to energies below an IR scale δ , such that $\mathcal{O}(\delta) \gg \mathcal{O}(G_N)$. We should be able to set the IR cutoff δ arbitrarily small, i.e.; it should not appear in answers to a well-defined physical problem. The expression for the projector onto low energy states in the Fock space is given by

$$\mathcal{P}_\delta = \Theta(\delta - H). \quad (4.53)$$

²Throughout this work, we will use the basis ($\{s\}$) to denote the supertranslation elements within projectors rather than the basis of smeared supertranslations (\mathcal{S}) to do so. The smeared basis is utilized while labelling the vacuum state denoting our system.

Since our projector is a function of the Hamiltonian, it is given by a boundary term as well. The representation of this operator over states labelled by the energy and super-translations is given by

$$\mathcal{P}_\delta = \int \left(\prod_{l,m} ds_{l,m} \right) \sum_i \Theta(\delta - E_i) |E_i, \{s\}\rangle \langle E_i, \{s\}| + \mathcal{O}\left(\sqrt{G_N}\right), \quad (4.54)$$

where for notational convenience we have relabelled the states as $|E_i, \{s\}\rangle$. However it should be kept in mind that states satisfy the inner product in (4.47). In particular states with different energy distributions but with same total energy should be thought of as labelled by different values of the index i .

4.4.2 CHSH correlation between regions A and C

We note that our calculation of $\langle C_{AB} \rangle$ in the absence of gravity remains unmodified when we turn on gravity (up to $\mathcal{O}(\sqrt{G_N})$) since we have simply fixed $s_{l,m}$ in (4.46). Physically, our operator insertions within C_{AB} are hard, and such operator insertions do not change the soft quantum numbers. As a result, the calculation of C_{AB} goes through in gravity.

Now we can construct a spacelike nice slice containing the regions A, B and C . On this slice, using the Reeh-Schlieder theorem [142, 143] we can construct local operators Q_i living on the region C which replicate the action of hard operators living on region B , such that

$$Q_i |0, \{s\}\rangle = B_i |0, \{s\}\rangle + \mathcal{O}\left(\sqrt{G_N}\right) \quad (4.55)$$

where as usual we have added contributions due to the interacting vacuum. Apart from the theorem guaranteeing their existence, in d -dimensions the operators Q_i can be explicitly constructed as follows.

Consider region C denoted by the Rindler wedge covered by the chart $z = Z + \zeta \cosh \tau$, $t = \zeta \sinh \tau$, in the domain $0 < \zeta < \infty$, $-\infty < \tau < \infty$ so as to have $z > Z + |t|$. To make this wedge spacelike separated from the region AB we keep $Z \gg r_0$. The metric is

$$ds^2 = -\zeta^2 d\tau^2 + d\zeta^2 + \sum_{i=1}^{d-2} dx_i^2. \quad (4.56)$$

We take a separable solution of the form $\Phi(\tau, \zeta, \mathbf{x}) = e^{-i(\omega\tau - \mathbf{k}\cdot\mathbf{x})} \chi(\zeta)$ in order to solve $\square\Phi = 0$. The ζ -equation is given by

$$\zeta^2 \frac{d^2\chi}{d\zeta^2} + \zeta \frac{d\chi}{d\zeta} + (\omega^2 - k^2 \zeta^2) \chi = 0, \quad (4.57)$$

where $k \equiv |\mathbf{k}|$. Imposing boundedness of the solution in the limit $\zeta \rightarrow \infty$ at fixed τ and

\mathbf{x} , the the field can be expressed as

$$\Phi(\tau, \zeta, \mathbf{x}) = \int_{\omega>0} \frac{d\omega d\mathbf{k}}{(2\pi)^{\frac{d-1}{2}}} \sqrt{\frac{2}{\omega}} b_{\omega, \mathbf{k}} e^{-i(\omega\tau - \mathbf{k}\cdot\mathbf{x})} \frac{K_{i\omega}(k\zeta)}{|\Gamma(i\omega)|} + \text{h.c.} \quad (4.58)$$

where $K_{i\omega}(k\zeta)$ is the modified Bessel function of the second kind. The ω dependent factors inside the integral ensure the canonical commutation relations³: $[b_{\omega, \mathbf{k}}, b_{\omega', \mathbf{k}'}^\dagger] = \delta(\omega - \omega')\delta(\mathbf{k} - \mathbf{k}')$ and $[b_{\omega, \mathbf{k}}, b_{\omega', \mathbf{k}'}] = 0$ [29].

On the complement of this Rindler wedge we can again write down Rindler-like coordinates, where in addition to the crossed over modes b and b^\dagger , there also exist a set of modes with support on $z < Z$ at $t = 0$ denoted by \tilde{b} and \tilde{b}^\dagger . Since these tilde operators are spacelike to Rindler wedge operators, they commute. Thus within the complement of the Rindler wedge, where the coordinates are $t = -\zeta \sinh \tau$, $z = Z - \zeta \cosh \tau$, the field operator can be written as

$$\Phi(\tau, \zeta, \mathbf{x}) = \int_{\omega>0} \frac{d\omega d\mathbf{k}}{(2\pi)^{\frac{d-1}{2}}} \sqrt{\frac{2}{\omega}} \tilde{b}_{\omega, \mathbf{k}} e^{i(\omega\tau - \mathbf{k}\cdot\mathbf{x})} \frac{K_{i\omega}(k\zeta)}{|\Gamma(i\omega)|} + \text{h.c.} \quad (4.59)$$

This is precisely how the smeared operators A_i and B_i in the previous section can be constructed from wedge operators and its complement. From the Bisognano-Wichmann construction [143], the complement operators are related to the wedge operators as:

$$\tilde{b}_{\omega, l} |0, \{s\}\rangle = e^{-\pi\omega} b_{\omega, l}^\dagger |0, \{s\}\rangle; \quad \tilde{b}_{\omega, l}^\dagger |0, \{s\}\rangle = e^{\pi\omega} b_{\omega, l} |0, \{s\}\rangle \quad (4.60)$$

where $|0, \{s\}\rangle$ denotes the global vacuum. Thus we systematically obtain (4.55). Using this construction, the action of the complement operators B_i on the vacuum can be written in terms of the action of the wedge operators Q_i on the vacuum.

The operators Q_i constructed above are in general unbounded, whereas in order to calculate CHSH correlations we require bounded operators. We will now construct operators C_i such that they satisfy

$$\|C_i\| = \langle B_i^2 \rangle + \mathcal{O}\left(\sqrt{G_N}\right) \quad \langle A_j C_i \rangle = \langle A_j B_i \rangle + \mathcal{O}\left(\sqrt{G_N}\right) \quad (4.61)$$

firstly using the projector \mathcal{P}_0 onto the flat vacua subspace. We will then use the physical projector \mathcal{P}_δ , and show that there exist our required operators C_i , and construct them explicitly.

³We make use of

$$\int_0^\infty \frac{dx}{x} K_{i\omega}(x) K_{i\omega'}(x) = \frac{\pi}{2} |\Gamma(i\omega)|^2 \delta(\omega - \omega')$$

$$\int_0^\infty \frac{d\omega}{|\Gamma(i\omega)|^2} K_{i\omega}(x) K_{i\omega}(x') = \frac{\pi}{2} x \delta(x - x')$$

Construction of C_i using \mathcal{P}_0

In this part, we outline the construction of operators C_i using the exact projector onto the vacuum. For notational simplicity, we will suppress factors of $O(\sqrt{G_N})$ within this subsection, and will reinstate the same in §4.4.3.

In order to construct bounded operators from Q_i , we take combinations of products of Q_i with the projector \mathcal{P}_0 . Consequently we recover the action of B_i on the vacuum, and therefore the resulting operator can be bounded. We define the operators C_i by the following expression

$$C_i \equiv \frac{\langle B_i^2 \rangle (Q_i \mathcal{P}_0 + \mathcal{P}_0 Q_i^\dagger - \langle B_i \rangle \mathcal{P}_0) - \langle B_i \rangle Q_i \mathcal{P}_0 Q_i^\dagger}{\langle B_i^2 \rangle - \langle B_i \rangle^2}, \quad (4.62)$$

where the cumulants are defined with respect to the smeared state $|0, \mathcal{S}\rangle$. The operators constructed in (4.62) might appear out of the blue, however they are systematically constructed by considering the subspace spanned by $\{|0, \{s\}\rangle, B_i |0, \{s\}\rangle\}$. For notational convenience, we also define

$$|B_i, \{s\}\rangle \equiv B_i |0, \{s\}\rangle \quad \text{and} \quad \beta_i \equiv \sqrt{\langle B_i^2 \rangle - \langle B_i \rangle^2}.$$

Then the construction of C_i is as follows. We start with a candidate C_i with linear combination of all possible outer products which do not involve cross terms from different superselection sectors, i.e.:

$$\{|0, \{s\}\rangle \langle 0, \{s\}|; |0, \{s\}\rangle \langle B_i, \{s\}|; |B_i, \{s\}\rangle \langle 0, \{s\}|; \& |B_i, \{s\}\rangle \langle B_i, \{s\}|\}$$

multiplied by undetermined coefficients. These coefficients can be systematically determined such that they satisfy the bounds in (4.61), which gives us (4.62). To demonstrate this, we rewrite the expression for C_i in (4.62) as a linearized sum of outer products with determined coefficients

$$\begin{aligned} C_i = & \int \left(\prod_{l,m} ds_{l,m} \right) \frac{\langle B_i^2 \rangle}{\beta_i^2} \left(|B_i, \{s\}\rangle \langle 0, \{s\}| \left[1 - \int \left(\prod_{l,m} ds'_{l,m} \right) \frac{|B_i, \{s'\}\rangle \langle B_i, \{s'\}|}{\langle B_i^2 \rangle} \right] \right) \\ & + \int \left(\prod_{l,m} ds_{l,m} \right) \frac{\langle B_i^2 \rangle}{\beta_i^2} \left(|0, \{s\}\rangle \langle B_i, \{s\}| \left[1 - \int \left(\prod_{l,m} ds'_{l,m} \right) |0, \{s'\}\rangle \langle 0, \{s'\}| \right] \right). \end{aligned} \quad (4.63)$$

The proof of boundedness of C_i as defined in (4.62) and $\langle A_j C_i \rangle = \langle A_j B_i \rangle + O(\sqrt{G_N})$ is given in Appendix 6.9.

Construction of C_i using \mathcal{P}_δ

We will now proceed with the construction of C_i using the more physical projector \mathcal{P}_δ . Motivated by (4.62), we can write a similar expression for C_i , which is valid up to an $O(\epsilon)$ correction.

$$C_i \equiv \frac{\langle B_i^2 \rangle (Q_i \mathcal{P}_\delta + \mathcal{P}_\delta Q_i^\dagger - \langle B_i \rangle \mathcal{P}_\delta) - \langle B_i \rangle Q_i \mathcal{P}_\delta Q_i^\dagger}{\langle B_i^2 \rangle - \langle B_i \rangle^2}. \quad (4.64)$$

To see why operators in (4.64) are valid operators upto $O(\epsilon)$, we firstly decompose the projector \mathcal{P}_δ as

$$\mathcal{P}_\delta = \mathcal{P}_0 + \delta\mathcal{P}. \quad (4.65)$$

The claim holds provided the contribution to C_i arises solely due to \mathcal{P}_0 , with $\delta\mathcal{P}$ not contributing to C_i . By acting operators C_i on the vacuum $|0, \mathcal{S}\rangle$, we can ensure that the chief contribution to C_i comes from \mathcal{P}_0 by demanding

$$|\langle 0, \mathcal{S} | Q_i | E_j, \mathcal{S} \rangle| \sim O(\epsilon) \quad \text{and} \quad |\langle 0, \mathcal{S} | A_i Q_k | E_j, \mathcal{S} \rangle| \sim O(\epsilon). \quad (4.66)$$

where the net energy E_j of the state satisfies $0 < E_j < \delta$. This renders $\delta\mathcal{P}$'s contribution within C_i very small, and consequently the C_i 's defined in (4.64) satisfy the constraints in (4.61). Note that these extra contributions arise due to the last term in (4.64).

Note here that $O(\epsilon)$ denotes minor errors introduced due to smearing scales, i.e., the operator smearing and the wedge smearing scales. We group all such scales as $O(\epsilon)$ since relatively these errors are of the same magnitude, in contrast to much more minor errors of $O(\sqrt{G_N})$.

Under what condition can we ensure (4.66)? To begin, consider a single-particle state $|j_\Omega\rangle = a_{E_j}^\dagger |0, \mathcal{S}\rangle$, such that $E_j < \delta$. To ensure (4.66), we first evaluate the expression $Q_1 |j_\Omega\rangle$.

$$Q_1 |j_\Omega\rangle = [Q_1, a_{E_j}^\dagger] |0, \mathcal{S}\rangle + a_{E_j}^\dagger Q_1 |0, \mathcal{S}\rangle \quad (4.67)$$

We will now argue that both the terms in (4.67) can be set small enough, thereby satisfying the conditions in (4.66). To see why the first term is small, let us discuss the energy scales in the problem. Apart from the Planck scale, there are two other energy scales in the problem: the energy ω' as defined in §4.3.3 (below which the Bogoliubov coefficients were close to zero); and δ , which denotes the IR cutoff. Now recall that B_1 is given by

$$B_1 = \frac{1}{\sqrt{2}} (P_B - \alpha_B^\dagger P_B \alpha_B + \alpha_B^\dagger P_B + P_B \alpha_B) \quad (4.68)$$

where P_B denotes the projector onto the B -vacuum, i.e., $P_B = |0_B\rangle \langle 0_B|$, and where we

have suppressed the supertranslation labels for convenience. Note that the modes α_B in (4.68) are related to the global modes as given in (4.9), and consequently the vacuum $|0_B\rangle$ is related to the global vacuum $|0\rangle$ as follows:

$$|0_B\rangle = \exp\left(\sum_{jk} \frac{1}{2} a_j^\dagger C_{jk} a_k^\dagger\right) |0\rangle. \quad (4.69)$$

where C_{jk} is the matrix outlined in the footnote⁴. Thus the operators B_i can be expressed in terms of the global modes as outlined above. Note that the global operators a_i^\dagger can be constructed only if we have access to the entire spacelike slice Σ , i.e.:

$$a_k^\dagger = \int_{\Sigma} \phi(x) e^{+ikx} \frac{d^{d-1}x}{(2\pi)^{d-1}} \quad (4.73)$$

and consequently B_i can only be written down provided we have access to the whole entire spacelike slice. However, since we have access only to the wedge and not the entire slice, an exact wedge reconstruction of the operator B_i is impossible. In particular, any attempt to reconstruct a_k^\dagger will also necessarily include other creation and annihilation operators.

$$\int_{\Sigma'} \phi(x) f_k(x) \frac{d^{d-1}x}{(2\pi)^{d-1}} = a_k^\dagger + \sum_j c_j a_j + \sum_{j \neq k} d_j a_j^\dagger \quad (4.74)$$

where $\Sigma' \in \Sigma$ denotes the spacelike part of the wedge and where $f_k(x)$ is a smearing function with support on Σ' . In spite of this obstruction, the Reeh-Schlieder theorem, and in particular our wedge reconstruction analysis in §4.4.2 gives us (4.55), i.e.:

$$Q_1 |0, \{s\}\rangle = B_1 |0, \{s\}\rangle$$

The critical point here is that there exist smearing functions $f_k(x)$, with support on the wedge, which convolves with field operator $\phi(x)$ using which we can construct such an operator Q_1 from the wedge. Then the practical way to construct Q_1 is as follows: we at-

⁴In general, for modes related by Bogoliubov transformations

$$a_i = \sum_{j} \alpha_{ij} b_j + \beta_{ij} b_j^\dagger \quad (4.70)$$

with $a_i |\Omega\rangle = 0$ and $b_j |X\rangle = 0$, the vacua $|\Omega\rangle$ and $|X\rangle$ are related by:

$$|\Omega\rangle = \exp\left(\frac{1}{2} b_j^\dagger C_{jk} b_k^\dagger\right) |X\rangle \quad (4.71)$$

where the matrix C_{jk} is given by

$$C_{mj} = -\sum_i \beta_{mi}^* \gamma_{ij} \quad \text{with} \quad \sum_i \alpha_{ji} \gamma_{ik} = \delta_{ik} \quad (4.72)$$

where δ_{ik} denotes the Kronecker delta function.

tempt to *closely simulate* B_i by wedge reconstructing the global creation and annihilation operators as in (4.74). We perform this attempt by choosing wedge smearing functions appropriately and substituting the closely simulated operators in (4.68) (which is essentially an infinite string of creation and annihilation operators from (4.9) and (4.69)). Consequently, we have a vast choice in choosing the smearing functions since each global operator insertion in (4.69) can be simulated using a reconstructed wedge operator. This method gives us the action of B_1 on the vacuum using Q_1 . As a result, Q_1 has additional terms than B_1 since we cannot precisely reconstruct the operator B_1 .

Upon normally ordering, Q_1 takes the following form:

$$Q_1 = B_1 + \sum_j p_{1j}(a_k^\dagger) + \sum_j q_{1j}(a_k) \quad (4.75)$$

where q_{1j} contains at least one annihilation operator, while p_{1j} contains the remaining terms with zero or more creation oscillators (Note that the operator $|0\rangle\langle 0|$ inside (4.68) cancels the remaining terms). In order to demand (4.55), the complex coefficients multiplying operator distributions inside p_{1j} in (4.75) are conveniently adjusted using smearing functions such that the following inner product is ensured:

$$\langle 0, \mathcal{S} | Q_1^\dagger Q_1 | 0, \mathcal{S} \rangle \approx \langle 0, \mathcal{S} | B_1^\dagger B_1 | 0, \mathcal{S} \rangle. \quad (4.76)$$

Now using (4.9), (4.68) and (4.69), we will argue that the wedge reconstructed operators Q_1 in (4.75) satisfy (4.55) along with ensuring the first term in (4.67) is small enough, provided that B_i and the smearing functions $f_k(x)$ satisfy the following conditions:

1. B_1 has a small overlap with $a_{E_j}^\dagger$. In other words, the modes constituting B_1 are sufficiently high energy modes which are constructed such that $\omega' \gg \delta$. This ensures that

$$[B_1, a_{E_j}^\dagger] \sim \mathcal{O}(\epsilon) \quad (4.77)$$

Here $\mathcal{O}(\epsilon)$ denotes the order of overlap between the high energy and the low energy modes due to smearing scales. A physically intuitive way to understand why we require $\omega' \gg \delta$ in flat space is to view the projection onto states below δ as noise in our description over the ground state subspace. Naturally, we do not want operator insertions inside the correlators characterized by frequencies within the noisy regime, rendering measurements meaningless. Therefore the noise δ needs to be set sufficiently low enough for the construction to work⁵.

We also note that if $\omega' < \delta$, the first term in (4.77) cannot be $\mathcal{O}(\epsilon)$, but constitutes an $\mathcal{O}(1)$ contribution.

⁵Note that in *AdS*, the *AdS* radius sets a natural length scale restricting δ once and for all, which is not the case here since the cosmological constant is zero.

2. This leaves us with the third term (note that the second term commutes with $a_{E_i}^\dagger$, and also has a very small magnitude) i.e., an infinite number of annihilator strings q_{1j} . These can provide a large contribution to the commutator in (4.67). To circumvent this, we require that our smearing functions is chosen such that the following contribution is ensured:

$$\sum_j [q_{1j}(a_k), a_{E_i}^\dagger] \sim O(\epsilon). \quad (4.78)$$

In particular, the constraint in (4.78) implies that the third term of the operator Q_1 given in (4.75) has a minimal contribution from annihilators below δ , and hence a slight overlap.

Thus using the conditions (4.76), (4.77) and (4.78) on Q_1 , we can ensure that the modes constituting Q_1 are engineered such that the following commutator in (4.67) is ensured:

$$[Q_1, a_{E_j}^\dagger] \sim O(\epsilon). \quad (4.79)$$

A more rigorous approach to showing the existence of a suitable boundary observable can be found in Appendix 6.10.

Let us now look at the second term in (4.67). This renders $\langle 0, \mathcal{S} | a_{E_j}^\dagger Q_1 | 0, \mathcal{S} \rangle = 0$, thereby satisfying the first condition in (4.66). Regarding the other condition in (4.66), using (4.67), the second term gives us $\langle 0, \mathcal{S} | A_i a_{E_j}^\dagger Q_1 | 0, \mathcal{S} \rangle$. Now consider the commutator

$$A_i a_{E_j}^\dagger = [A_i, a_{E_j}^\dagger] + a_{E_j}^\dagger A_i. \quad (4.80)$$

Since A_i is again an operator with energy much higher than δ , the first term in (4.80) is $O(\epsilon)$, and the second term is zero i.e. $\langle 0, \mathcal{S} | a_{E_j}^\dagger A_i B_1 | 0, \mathcal{S} \rangle = 0$. We can repeat the analysis for multi-particle states as well as straightforwardly generalize the result from B_1 to B_i . We obtain similar conclusions for multi-particle states, with products of creation operators replacing the single creation operator in the analog of (4.67) and (4.80). Given the above discussed smearing conditions, operators in (4.64) represent valid operators C_i satisfying constraints in (4.61), with errors from smearing again giving rise to an $O(\epsilon)$ correction.

Existence of C_i using \mathcal{P}_δ and the boundary algebra

In this subsection, we argue that we can always construct C_i using \mathcal{P}_δ and other elements of the boundary algebra which satisfy constraints in (4.61). This differs from our analysis in §4.4.2 since our expressions for C_i satisfying (4.61) are exact here, without any factors of $O(\epsilon)$.

In general, to exactly construct C_i satisfying the constraints (4.61), we require other boundary operators along with the projector \mathcal{P}_δ . We define

$$|B_i^\perp, \{s\}\rangle \equiv \frac{(1 - \mathcal{P}_\delta)}{\mathcal{N}_i} |B_i, \{s\}\rangle$$

where \mathcal{N}_i is a normalization constant. As an example, we can read off $\mathcal{N}_i = \beta_i$ from (4.63), when we work with the exact vacuum projector \mathcal{P}_0 . On the lines of the construction of C_i using \mathcal{P}_0 , we write a candidate C_i , which is a sum of all possible outer products multiplied by undetermined coefficients

$$\begin{aligned} C_i = & \int \left(\prod_{l,m} ds_{l,m} \right) \sum_{E_j, E_k} \Theta(\delta - E_j) \Theta(\delta - E_k) x_{j,k}^i |E_j, \{s\}\rangle \langle E_k, \{s\}| \\ & + \int \left(\prod_{l,m} ds_{l,m} \right) \sum_{E_j} \Theta(\delta - E_j) \left(y_j^i |B_i^\perp, \{s\}\rangle \langle E_j, \{s\}| + \text{h.c.} \right) \\ & + \int \left(\prod_{l,m} ds_{l,m} \right) z_i |B_i^\perp, \{s\}\rangle \langle B_i^\perp, \{s\}|. \end{aligned} \quad (4.81)$$

Note that here the elements $|E_j, \{s\}\rangle \langle E_k, \{s\}|$ belong to the boundary algebra as argued in [79]. We will systematically fix some of the coefficients in (4.81) as follows, where we will suppress corrections of $\mathcal{O}(\sqrt{G_N})$ for presentation. The Hermiticity of C_i implies $x_{j,k}^i = (x_{k,j}^i)^*$ and $z_i \in \mathbb{R}$. Imposing $C_i |0, \{s\}\rangle = B_i |0, \{s\}\rangle$, we fix the coefficients

$$x_{m,0}^i = \int \left(\prod_{l,m} ds_{l,m} \right) \langle E_m, \{s\} | B_i, \{s\} \rangle, \quad y_0^i = \mathcal{N}_i. \quad (4.82)$$

which ensures $\langle A_j C_i \rangle = \langle A_j B_i \rangle$. Given that we still have undetermined coefficients in C_i , we can always choose them in such a way that the absolute value of the largest eigenvalue is given by $\sqrt{\langle B_i^2 \rangle}$, thereby giving us $\|C_i\| = \langle B_i^2 \rangle$.

As an example, we will demonstrate this for the case of the exact projector \mathcal{P}_0 . Here we have $\mathcal{N}_i = \beta_i$. Consequently we obtain

$$\|C_i\| = \frac{1}{2} \left(\langle B_i \rangle + z_i + \sqrt{4\beta_i^2 + (\langle B_i \rangle - z_i)^2} \right).$$

Now requiring that the bound is satisfied, i.e. $\|C_i\|^2 = \langle B_i^2 \rangle$ gives us $z_i = -\langle B_i \rangle$, which again leads to the seemingly serendipitously constructed C_i in (4.62).

4.4.3 The paradox and generalization to higher dimensions

In §4.4.2, using the wedge reconstruction, the boundary algebra, and the fact that the Hamiltonian in gravity is a boundary term, we have constructed operators living in the

exterior region C which essentially replicate the action of operators B_i on the vacuum state, in three different fashions. Subsequently, we arrive at the following conclusion:

$$\langle C_{AC} \rangle = \langle C_{AB} \rangle + \mathcal{O}\left(\sqrt{G_N}\right). \quad (4.83)$$

Consequently at $x = \frac{1}{2}$, the correlator $\langle C_{AC} \rangle$ takes a maximum value $\langle C_{AC} \rangle = \frac{27\sqrt{2}}{16} + \mathcal{O}\left(\sqrt{G_N}\right) + \mathcal{O}(\epsilon)$.

After getting all our ingredients in place we will now pose the paradox in monogamy of entanglement. For the maximum violation at $x = \frac{1}{2}$, we obtain

$$\langle C_{AB} \rangle^2 + \langle C_{AC} \rangle^2 = 11.4 + \mathcal{O}\left(\sqrt{G_N}\right) + \mathcal{O}(\epsilon) > 8. \quad (4.84)$$

Equation (4.84) contradicts the upper bound in (4.40) and gives rise to the paradox in monogamy. As mentioned earlier, this is an $\mathcal{O}(1)$ violation. The violation does not have a leading dependence on the IR cutoff δ , an expected feature of a well-defined physical observable.

An immediate generalization of the paradox in four dimensions is extending the same to general dimensions. In $d \neq 4$, the low energy vacuum structure of gravity is not concretely established (See [144–150] for recent discussions on the subject). Provided that the vacuum structure of gravity in higher dimensions has a similar form, i.e., there is a unique vacuum or degenerate vacua labelled by supertranslations, we can pose the paradox in precisely the same fashion we have done presently. Regarding additional symmetries, we can again treat them in a fashion similar to our treatment of supertranslations.

We will point out why analogs of supertranslations in general dimensions are not in conflict with our calculation. The calculation of $\langle C_{AB} \rangle$ does not require us to go to the asymptotics since the operator insertions are deep inside the bulk, and hence our operator insertions do not change the supertranslation of the state on which they act. More precisely, these operator insertions are hard. The case of $\langle C_{AC} \rangle$ is a bit more subtle since it involves the construction of operators Q_i and the projector \mathcal{P}_δ both of which have support near \mathcal{I}^+ . However, from (4.55), the action of Q_i on the supertranslation fixed vacuum is precisely the action of the hard operators B_i on the vacuum. In addition, the projector \mathcal{P}_δ as defined in (4.54) is diagonal in supertranslation labelled vacua. Consequently, the insertion of the operator C_i within the vacuum to vacuum correlators does not introduce any new complications because of our construction and the very nature of the vacuum structure.

4.4.4 Resolution of the paradox

In our calculation, we have explicitly pointed out small corrections of $\mathcal{O}\left(\sqrt{G_N}\right)$ (See (4.84)). Hence the paradox cannot be resolved by introducing small corrections, as is the

case for Hawking’s original paradox and the bags of gold paradox [1, 28, 40, 58, 94–96]. Here the $O(1)$ violation indicates the existence of a severe flaw in our basic assumption, i.e., we have assumed that in the presence of gravity, our system admits a description in terms of a local quantum field theory. Building upon this assumption, we have factorized our Hilbert space into three different parts into three spatially disconnected and separated regions A , B , and C .

However, it is a well-known fact that in gravity, the Hilbert space cannot be factorized due to the Gauss constraint. Consequently, our factorization into a tripartite system each described by a local QFT is incorrect, which resolves the paradox posed above. With gravity turned on, degrees of freedom in the region B are secretly the same as degrees of freedom in the region C . Therefore it is incorrect to describe operators probing the underlying degrees of freedom using local quantum field theory, and we explicitly see an $O(1)$ violation if we assume a local quantum field theory setup in our case of empty flat space.

In a certain sense, we can observe this non-factorization of Hilbert space of the effective field theory based on spatial partitioning at the level of commutators itself ⁶. Note that since our operator insertions A_i/B_i introduce energy into the bulk, the commutator $[H, B_i] \neq 0$. Since C_i is a function of boundary projectors, in general the commutator $[C_j, B_i] \neq 0$. Following the Gauss constraint, this is a complementary objection to why we should not expect factorization based on spatial partitioning within a theory of gravity, even though effective field theory reasoning naively indicates otherwise.

Since our calculation is performed in a general fashion, we can equivalently interchange the operators A and B describing the interior and the exterior respectively in (4.84) to set up an information-theoretic inequality again. The resolution for this situation is precisely what black hole complementarity states, i.e., the interior operators are complicated polynomials of the exterior operators. In principle, our setup provides an explicit demonstration of the complementarity principle in flat space. Our boundary projector is a vital ingredient in this construction, allowing us to write down operators far away from the light cone that can probe the interior of the light cone.

⁶We thank Simon Caron-Huot for pointing out this issue.

Chapter 5

Summary and discussion

The underlying theme of the thesis was to pose these paradoxes using bulk effective field theory and precisely understand what goes wrong while using such a description. Another theme was understanding how quantum information is localized in a theory of quantum gravity. In this light, we conveyed the basic idea that studying puzzles regarding the black hole interior allows us to understand such issues better.

As a brief recap, we started by motivating our work from various angles in our introductory Chapter. In §2, we provided the basic background and technical tools necessary to study various aspects of the rest of the thesis. In this chapter, we will be discussing our work in §3 and §4.

We have hereby looked into two paradoxes from the bulk perspective: the bags of gold, which can be resolved using small gravitational corrections, while the monogamy paradox requires $O(1)$ corrections for its resolution. Both approaches demonstrate the failure of local effective field theory and require us to go beyond it. The case of bags of gold requires quantum gravity corrections, i.e., the fat tail effects for its resolution. In contrast, the monogamy paradox arises due to the assumption that the Hilbert space does not factorize upon spatial partitioning.

We will now summarize different aspects of our work, outline how our work connects to existing literature, and conclude by mentioning future possibilities.

5.1 The bags of gold paradox

Regarding our work on the bags of gold, we have demonstrated a possible resolution to understand the case of several excitations living in the black hole interior. We have proposed that these numerous excitations living on large volume Cauchy slices in the interior are not inconsistent with the Bekenstein-Hawking entropy, as they have small inner products and thus are not independent excitations. We advocate that such a situation is not a problematic feature of effective field theory but an essential aspect of

quantum gravity. This overcounting naturally arises in the context of boundary theories, as shown using toy matrix models. We also showed that spectral observables like the form factor and the level spacing distribution are violated in the semiclassical treatment of the interior excitations, and our proposal also resolves these contradictions.

The reader, at this point, might wonder what exactly our proposal accomplishes since the bags of gold can also be conveniently settled using the fine-grained entropy, as we outlined in §3.3. In a certain sense, as outlined there, the fine-grained entropy quite cleanly settles this question by giving us the correct quantitative answer. However, while these calculations potentially give us the correct answer, quite a few aspects of bulk physics remain unknown. An important aspect involves a precise understanding of the physics of the interior. In this regard, one may ask that even though the fine-grained entropy predicts that the Bekenstein-Hawking formula gives the actual entropy, what exactly are the bulk criteria for discarding most of the interior excitations in the first place. More precisely, given some semiclassical excitations, what are the criteria for discarding them as not contributing to the fine-grained entropy. In other words, since the entanglement wedge goes only up to the neck, by performing entanglement wedge reconstruction, one can only reconstruct operators within the wedge without a precise bulk understanding of what exactly happens to the interior operators.

In this regard, the quantum extremal surface, which calculates the fine-grained entropy, correctly counts the *necessary degrees of freedom*. We can use this correct counting to construct a suitable basis spanning the actual Hilbert space. The rest of the interior region, with numerous semiclassical excitations giving rise to the paradox, can be described using overcounting within this Hilbert space (i.e., embedded into the actual Hilbert space via small inner products). This provides a bulk understanding of what gives rise to the paradox.

The notion of small inner products is consistent with the breakdown of locality in quantum gravity. We expect locality to hold in ordinary effective field theories describing nature. In contrast, we expect locality to hold only in an approximate sense in quantum gravity. Here locality can break down in various situations, such as the case where we act with too many probes on the spacetime [151]. Small inner products between spatially separated excitations can be understood as another such situation that demonstrates the breakdown of locality in quantum gravity.

We saw some examples of the grave problems associated with simulating a Hilbert space with a much smaller one. So we naturally ask: why are black holes unique? Quantum and thermal systems have macroscopic observables which can be measured experimentally, and in a certain sense, we can find out the Hilbert space's correct dimension. Thus the possible kinematic overcounting of Hilbert space is not realized in these systems. For the case of black holes, thermodynamic observables do point out that the Hilbert space is far smaller than what bulk semiclassical quantization indicates. As an

example, we know that the thermodynamic entropy of a black hole should go as area. Further, from the CFT side, we showed using various examples that spatially separated bulk interior states indeed have small inner products. These small overlaps between the dual bulk semiclassical excitations hint at how the Hilbert space of gravity embeds bulk states.

We now ask why we do not see an overcounting using effective field theory in empty AdS. Given a holographic CFT, the HKLL prescription wholly reconstructs the empty AdS bulk. As a result, the CFT description captures all the bulk excitations, and consequently, there is no question of any overcounting. The HKLL prescription also reconstructs the exterior regions of eternal black holes. The only places where overcounting using effective field theory can arise are causally inaccessible regions from the boundary. The black hole interior is an example of such an inaccessible region, and reconstruction using state-dependent operators allows us to resolve apparent paradoxical situations like the one we have treated here.

An important aspect that we have briefly touched upon here is the subject of quantum state transfers in semiclassical gravity. We showed via an example in §3.8 that overcounted Hilbert spaces can lead to quantum state transfers which are forbidden if orthogonal vectors span the Hilbert space. Since we have argued that effective field theory in the black hole interior leads to an overcounting of the Hilbert space, it will be helpful to understand which forbidden semiclassical quantum state transfers are allowed in quantum gravity. We think this holds important implications for black holes in AdS and possibly in flat space as well.

For the general reader not interested in details, we have thus answered a seemingly interesting puzzle: Can there exist a giant universe inside an AdS black hole, with the universe's coarse-grained entropy exceeding the Bekenstein-Hawking entropy of the AdS black hole, while an external observer in the black hole spacetime is utterly oblivious to the universe's existence? The answer to this question is yes, provided that the Hilbert space of this universe is constructed from the small Hilbert space of the AdS black hole using small inner products. A significant number of excitations can be described using the overcounted Hilbert space, provided that their backreaction on the black hole is extremely small (or, in other words, their states belong to the small Hilbert space). Careful measurements of thermodynamic observables and state transfers in semiclassical gravity can conclude that the Hilbert space of this universe is constructed from the small Hilbert space of the black hole itself. The initial state giving rise to this universe is a Euclidean state as mentioned in §3.4, while the future of this universe is doomed anyway since it will eventually fall into the black hole.

5.2 The monogamy paradox

Regarding our work on the monogamy paradox, firstly, we introduced a formalism for treating Bell inequalities in a local QFT in an asymptotically flat space. We utilized the fact that monogamy of Bell correlations is a concrete measure for monogamy of entanglement and consequently used our formalism to compute Bell correlation between regions A and B . Then bounded operators C_i were constructed in a spacelike separated region C close to the boundary using the Reeh Schlieder theorem and the boundary projector, replicating the operators B_i 's action on the vacuum. Using this, a concrete paradox was posed in the monogamy of entanglement between the regions A , B and C . We argued that the resolution to the paradox is as follows: in a theory of gravity, one cannot factorize the Hilbert space into subspaces describing spatially separated regions, which is necessary to set up a paradox in the monogamy of entanglement.

As discussed in our work, in canonical gravity, the Hamiltonian is a boundary term that plays a crucial role in constructing bounded operators in the region C that replicate operators B_i . The fact that the Hamiltonian is a boundary term is an essential feature of gravity, which strongly hints at non-local aspects inbuilt within theories of gravity¹. Note that this feature is unique to gravity and is not true of other theories, say theories with Gauss constraints. Case in point, operator insertions with zero charge in gauge theories do not affect the field strength residing on the Gaussian surface. In gravity, an operator insertion necessarily changes the stress-energy tensor, and consequently, one cannot introduce invisible operator insertions².

We now discuss the relation of our model to the monogamy paradox for old black holes in flat space. Since non-local effects of gravity play a primary role in our problem, it is only natural to assume that such effects play a similar role in the black hole problem³. The operators C_i in our problem are in a spirit similar to complicated operators situated far away from a black hole used to extract information from Hawking radiation. Our construction also emphasizes the usage of CHSH correlations in studying the monogamy of entanglement paradox, primarily how CHSH correlations can be used to quantify entanglement. The study of these correlators is necessary since standard measures of entanglement like Von Neumann entropy are not well defined in gravity.

Note that the monogamy paradox is conventionally posed within the context of old black holes. However, our discussion only relies upon the entanglement of modes across the horizon and the boundary. Consequently, we do not require an old black hole to pose

¹In a certain sense, this is a generalization of the Gauss law in Newtonian gravity, where the effect of any massive insertion at a point within a spacelike separated Gaussian surface is necessarily manifested on the surface.

²This argument is also tied to why we cannot write diffeomorphism invariant local observables in gravity.

³In fact, such non-local effects are a generic feature of quantum gravity, and manifestly reveal them under extreme situations [151].

the paradox, which is reflected in the fact that the validity of operators C_i does not involve any particular time scale. In line with the principle of holography of information [79–81], this is because the information about the non-boundary regions is *always contained* within \mathcal{I}_-^+ .

The issue with writing down a similar construction for evaporating black holes is that we require a projector onto the space of black hole microstates, which we presently do not understand how to construct. Consequently, it is not easy to write bounded operators in a region far away from the black hole, which can be used to write down CHSH correlations. However, there is no problem with calculating C_{AB} correlator between modes just inside and outside the horizon. Formulating the paradox in our toy model’s fashion also shows that we do not need any modified structure in the black hole interior, as with firewall and fuzzball constructions. Instead, such a paradox in monogamy is a natural consequence of wrongly treating gravity as a local quantum field theory.

Before we conclude, we list out some related open questions. In our case, we need to go very close to future null infinity to construct a projector onto our ground state. In line with our holographic intuition that gravity knows about quantum information inside a given region, is it possible to construct a similarly approximate projector onto the vacuum at a finite radius? Finding such a projector will be pretty valuable not only as an independent problem for our flat space toy model but also to pose a similar resolution of the monogamy paradox for dS black holes. Besides, such a projector will be pretty valuable for understanding aspects of the principle of holography of information for compact spacetimes, where naively, a projector will project onto all physical states in the Hilbert space since there is no boundary. Consequently, we need a projector at a finite radius.

Another problem is to write down the projector onto the space of all black hole microstates in flat space and AdS, which will allow us to write down a more accurate toy model. A distant direction is to understand the asymptotic vacuum structure in general dimensions, which will help pose the toy model concretely in such dimensions. We envisage our present work as a starting point to address some of these issues in the near future.

While at the risk of trespassing into areas wherein the author’s expertise is almost non-existent, we wish to state a couple of lines about the possible relevance of our work to low-energy experiments. The work discussed here was related to potentially resolving inconsistencies arising in quantum gravity from a theoretical standpoint. However, in general, the analysis of CHSH correlators in flat space could be used in the future to test the localization of quantum information in quantum field theories. Another important direction is to analogously extend this formalism to understand information-theoretic statements in cosmology, especially in the context of CMB experiments. We believe these issues deserve further attention.

Chapter 6

Appendices

6.1 Explicit examples of overcounting in small vector spaces

In this appendix we explicitly demonstrate that there can be many more vectors than n , where n is the dimensionality of the vector space if $\epsilon \neq 0$. We will demonstrate this using regular polyhedra. The study of overcounting using regular polyhedra serves as an easy way to develop our intuition for understanding overcounting by starting from small dimensions and gradually building up to higher dimensional examples. Note that our vector space defined over reals in contrast to the Hilbert spaces in quantum mechanics. We will denote the maximum number of vectors as a function of ϵ to be $\hat{m}_n(\epsilon)$ in the previous case. Now regular polyhedra are classified into three classes:

- *Simplex*: This polyhedra is defined by the condition that the distance between any two vertices is the same. Examples are : equilateral triangle with $n = 2$ and tetrahedron with $n = 3$. The vectors corresponding to neighbouring vertices have an inner product given by $\vec{p}_i \cdot \vec{p}_j = -\frac{1}{n}$. Therefore we have $\hat{m}_n\left(-\frac{1}{n}\right) = n + 1$ number of vectors for the simplex.
- *Orthoplex*: These polyhedra are defined such that they have a vector each pointing towards each coordinate direction, such that the inner product between the neighbouring vertices is given by $\vec{p}_i \cdot \vec{p}_j = 0$. Consequently a simplex has $\hat{m}_n(0) = 2n$ number of vectors.
- *Hypercube*: In the Cartesian coordinate system, these polyhedra have vertices situated at the coordinates $(\pm 1, \pm 1, \pm 1, \pm 1 \dots) / \sqrt{n}$. Examples are : square with $n = 2$ and cube with $n = 3$. The neighbouring vertices have inner products given by $\vec{p}_i \cdot \vec{p}_j = 1 - \frac{2}{n}$. Consequently a hypercube has $\hat{m}_n\left(1 - \frac{2}{n}\right) = 2^n$ number of vectors.

We will now compute the inner products for representatives of these above-mentioned classes of polyhedra. We will now give some examples in low dimensions below:

- *Two dimensions:* In $n = 2$ for a regular polygon, the scalar product between position vectors of m neighbouring vertices is given by $\vec{p}_i \cdot \vec{p}_j = \cos \left[\frac{2\pi}{m} \right]$, where $\hat{m}_2 \left(\frac{2\pi}{m} \right) = m$.
- *Three dimensions:* In $n = 3$, the icosahedron has $\hat{m}_3 \left(\frac{1}{\sqrt{5}} \right) = 12$ vertices while the dodecahedron has $\hat{m}_3 \left(\frac{\sqrt{5}}{3} \right) = 20$ vertices.
- *Four dimensions:* In $n = 4$, we consider the 24-cell which has total number of vertices given by $\hat{m}_4 \left(\frac{1}{2} \right) = 24$. Similarly the 120-cell has $\hat{m}_4 \left(\frac{1+\sqrt{5}}{4} \right) = 120$ vertices while the 600-cell has $\hat{m}_4 \left(\frac{1+3\sqrt{5}}{8} \right) = 600$ vertices.

Overcounting in the limit $m \rightarrow \infty$?

We will construct a situation where m vectors are approximately equidistant on the sphere S_{n-1} which has surface area of $\frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$. Each unit vector has an exclusion zone given by $\frac{2\pi^{n/2}}{m\Gamma(\frac{n}{2})}$, where we cannot have any other vector. These exclusion zones have a radius r , with the volume of these $n - 1$ dimensional zones given by $\frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n+1}{2})} r^{n-1}$. Therefore we can solve for the radius of the exclusion zone as done in the following equation:

$$r \approx \left(\frac{2\sqrt{\pi}\Gamma(\frac{n+1}{2})}{m\Gamma(\frac{n}{2})} \right)^{\frac{1}{n-1}}. \quad (6.1)$$

The distance between two neighbouring unit vectors is given by $d \approx 2r$. Therefore the inner product between two neighbouring unit vectors can be easily computed and is found to be:

$$\vec{p}_i \cdot \vec{p}_j \approx 1 - \frac{1}{2}d^2 = 1 - 2 \left(\frac{2\sqrt{\pi}\Gamma(\frac{n+1}{2})}{m\Gamma(\frac{n}{2})} \right)^{\frac{2}{n-1}}. \quad (6.2)$$

We can now solve the expression for inner products to obtain m . In the limit $\epsilon \rightarrow 1$ with $0 < 1 - \epsilon \ll 1$, we obtain the following expression for m :

$$\hat{m}_n(\epsilon) \approx 2^{n/2} \frac{\sqrt{2\pi}\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} (1 - \epsilon)^{-\frac{n-1}{2}}. \quad (6.3)$$

For $0 < \epsilon \ll 1$, m takes the values:

$$\hat{m}_n(\epsilon) \approx 2n \left(\frac{2n}{n+1} \right)^{n\epsilon} \quad (6.4)$$

This expression for m is in agreement with our derivation of inner products between neighbouring vectors of the simplex and orthoplex, which are given by $\hat{m}_n \left(-\frac{1}{n} \right) = 1 + n$

and $\hat{m}_n(0) = 2n$ respectively. Equation (6.4) is also in agreement with our derivation for the hypercube's case, up to a minor factor of $\frac{n}{2e}$.

$$\lim_{n \rightarrow \infty} \hat{m}_n \left(1 - \frac{2}{n}\right) \rightarrow 2^n \times \frac{n}{2e}. \quad (6.5)$$

6.2 Partition function of the $U(N)$ two-oscillator model

We outline the calculation of the partition function of two bosonic oscillators denoted by A and B transforming under the adjoint representation of $U(N)$ global group [121, 122]. We hereby define $x = e^{-\beta}$.

Since we are interested only in the global group singlets, we calculate the partition function by summing over all Boltzmann weights multiplied by the number of group singlets at each Boltzmann weight. With E_i being the energy of the bosonic modes, the partition function is given by

$$Z(x) = \sum_{n_1} \sum_{n_2} x^{n_1 E_1} x^{n_2 E_2} \left(\text{All singlets } \text{sym}_{n_1}[\text{adj}] \times \text{sym}_{n_2}[\text{adj}] \right) \quad (6.6)$$

Here the index n_1 goes over the A oscillators and n_2 goes over B oscillators. Note that here we have set the ground state energy of the oscillators to zero. We now return to the problem of counting the group singlets. A convenient way to count the number of group singlets is by using properties of characters, which are maps from the representation of the group to complex numbers defined as

$$\chi_{\mathbb{R}} : G \rightarrow \mathbb{C} \quad (6.7)$$

satisfying the property that for $g \in G$,

$$\chi_{\mathbb{R}}(g) = \text{Tr}_{\mathbb{R}}(g). \quad (6.8)$$

These satisfy the orthonormality relation, where $[dg]$ is the Haar measure

$$\int [dg] \chi_{\mathbb{R}_i}^*(g) \chi_{\mathbb{R}_j}(g) = \delta_{R_i R_j}, \quad (6.9)$$

chosen such that $\int [dg] = 1$. Since the character is the trace of the element, therefore $\chi_{\mathbb{R}_1 \times \mathbb{R}_2} = \chi_{\mathbb{R}_1} \chi_{\mathbb{R}_2}$. The number of irreps can thus be counted using the above relation as

$$n_{R_I} = \int [dg] \chi_{\mathbb{R}_I}^*(g) \prod_j \chi_{\mathbb{R}_j}(g). \quad (6.10)$$

We will now use this relation to count the number of group singlets. For a singlet representation, by definition we have $\chi_s = 1$. Therefore the number of singlets is given

by

$$\text{All singlets} = \int [dg] \prod_j \chi_{\mathbb{R}_j}(g) \quad (6.11)$$

We will now use (6.11) to rewrite the partition function in (6.6) in terms of the characters of $U(N)$:

$$Z(x) = \int [dU] \prod_{i=A,B} \sum_{n_i=0}^{\infty} x^{n_i E_i} \chi_{\text{sym}^{n_i}}(U). \quad (6.12)$$

We will utilize the following relation for the characters in order to simplify the partition function:

$$\sum_{n=0}^{\infty} t^n \chi_{\text{sym}^n}(g) = \exp \left\{ \sum_{l=1}^{\infty} \frac{t^l \chi(g^l)}{l} \right\} \quad (6.13)$$

We denote z_A and z_B as the single particle partition functions for the A and the B harmonic oscillators. Using (6.13), and using the bosonic partition function $z(x) = z_A + z_B$ and $\chi_{\text{adj}}(U) = \text{Tr}(U) \text{Tr}(U^\dagger)$ we get the partition function as

$$Z(x) = \int [dU] \exp \left[\sum_{k=1}^{\infty} \frac{z(x^k)}{k} \text{Tr} U^k \text{Tr}(U^\dagger)^k \right]. \quad (6.14)$$

Rewriting this unitary matrix model in terms of the eigenvalues of the unitary matrix easily solves the model [42, 152]. Denoting the eigenvalues of U by $e^{i\alpha_i}$, we write the measure of the model as:

$$\int [dU] = \prod_i \int_{-\pi}^{\pi} [d\alpha_i] \prod_{i < j} \sin^2 \frac{(\alpha_i - \alpha_j)}{2} \quad (6.15)$$

Now the partition function is a function of the eigenvalues α_i and is given by

$$Z(x) = \int [d\alpha_i] \exp \left\{ - \sum_{i \neq j} V(\alpha_i - \alpha_j) \right\} \quad (6.16)$$

where the potential is given by

$$V(\theta) = - \log \left| \sin \frac{\theta}{2} \right| - \sum_{k=1}^{\infty} \frac{z(x^k)}{k} \cos k\theta \quad (6.17)$$

Equation (6.17) is reminiscent of the Coulomb potential for charges on a sphere where the term coming from the measure is repulsive interaction between the like charges, and the other term is attractive interaction due to electric field. The dynamics are similar to the partition function of the Gross-Witten-Wadia model [153–155], and has a third-order phase transition in the $N \rightarrow \infty$ limit and has a free energy of $O(N^2)$. The average energy is $O(N^2)$, and as a consequence, the entropy is also of the same order. Therefore

we can see that the Hilbert space's dimensionality is $O(\exp N^2)$.

Entropy, average energy and dimensionality of the Hilbert space for oscillators on $S_3 \times \mathbb{R}$

We now proceed to calculate the Hagedorn temperature of the CFT, which is essential because we want to describe black holes, and going above the Hagedorn temperature is the regime where we have black holes in the bulk. We will then calculate the entropy, the average energy and the dimensionality of the Hilbert space.

The single particle partition functions

We evaluate the single-particle partition functions for the CFT living on $S^3 \times \mathbb{R}$. With $\lambda \approx 0$ our bosonic harmonic oscillators A and B obey the equation of motion given by

$$\left(-\partial^2 + a^{-2}\right) A_{ij} = 0 \quad \& \quad \left(-\partial^2 + a^{-2}\right) B_{ij} = 0 \quad (6.18)$$

In four dimensions we utilize the conformal map from $S^3 \times \mathbb{R} \rightarrow \mathbb{R}^4$ to write the partition function $Z = \sum_{E_i} e^{-\beta E_i}$ in the form $Z = \sum_{\Delta} e^{-\beta \Delta}$ where Δ is the scaling dimension. The scaling dimension goes over all the local operators in the theory which are generated by repeated applications of the derivatives ∂_μ on the fields, i.e. $(A, \partial_\nu A, \partial_\mu \partial_\nu A \dots)$ and similarly for the B field modulo the equation of motion. Since $[\partial] = 1$ a single derivative gives rise to a factor in:

$$x^0 + x^1 + x^2 + x^3 + \dots = \frac{1}{1-x} \quad (6.19)$$

Four such derivatives will give rise to $\frac{1}{(1-x)^4}$. The mass dimension of the matrix oscillators is $[A], [B] = 1$. Without incorporating the equation of motion, the naive partition function constructed using all local operators arising in the matrix model is given by

$$\frac{x}{(1-x)^4} \quad (6.20)$$

We now need to take the modulus by the equation of motion. Notice that the equation of motion imposes a condition on any local operator O of the theory:

$$\left(-\partial^2 + a^{-2}\right) O = 0. \quad (6.21)$$

The factor $(-\partial^2 + a^{-2})$ has a mass dimension x^2 which we need to subtract off. The CFT partition function for oscillators A and B upon this subtraction is therefore given by

$$z_A(x) = z_B(x) = \frac{x - x^3}{(1-x)^4} = \frac{x + x^2}{(1-x)^3}. \quad (6.22)$$

The Hagedorn transition

We calculate the entropy in this subsection and deduce the dimensionality of the Hilbert space from it. First of all we set $m_A = m_B = 1$ for our convenience. The entropy from the partition function with $k_B = 1$ is given by

$$S = -\frac{\partial F}{\partial T} = \frac{\partial (T \log Z)}{\partial T} = \log Z + \frac{T}{Z} \frac{\partial Z}{\partial T}. \quad (6.23)$$

We can write the potential in (6.17) as

$$V(\theta) = \log 2 + \sum_{k=1}^{k=\infty} \frac{1}{k} \left(1 - \frac{2(x+x^2)}{(1-x)^3} \right) \cos k\theta. \quad (6.24)$$

In the low-temperature phase, it follows from (6.17) that the attractive second term goes to zero as $T \rightarrow 0$. Therefore at low temperature, the repulsive interactions dominate. The way to solve this matrix model in (6.17) is to introduce the eigenvalue density and solve using the mean-field theory approximation. The level density of the eigenvalues is spread uniformly over the circle as $T \rightarrow 0$. As we increase the temperature, this distribution becomes more and more non-uniform, and the phase ceases to be stable when the terms in the potential turn negative. The condition for the stability of the potential then becomes

$$\frac{2(x+x^2)}{(1-x)^3} < 1. \quad (6.25)$$

Since $0 < x < 1$ and x monotonically increases with temperature in this regime, the leading $k = 1$ order well approximates the above condition since it gives the strongest contribution. The temperature at which this phase becomes unstable is the Hagedorn temperature and is given by

$$\frac{2(x+x^2)}{(1-x)^3} = 1. \quad (6.26)$$

Solving this the Hagedorn temperature is given by $k_b T_H = 0.634484$.

Evaluation of the partition function for small N

We use the single-particle bosonic partition functions to expand the complete partition function up to the first two powers in cosines, as the remaining terms fall off quite rapidly and therefore have negligible contributions. We will demonstrate this for the $N = 2$ case and will treat higher N similarly. The partition function for $N = 2$ is given by

$$Z(x) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d\alpha_1 d\alpha_2 \sin^2 \left(\frac{\alpha_2 - \alpha_1}{2} \right) \exp \left[\sum_{k=1}^{\infty} \frac{4 \cos k(\alpha_2 - \alpha_1)}{k} \frac{x + x^2}{(1-x)^3} \right]. \quad (6.27)$$

Upon the expansion to the first two orders, we obtain the partition function to be

$$Z(x) = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d\alpha_1 d\alpha_2 \sin^2 \left(\frac{\alpha_2 - \alpha_1}{2} \right) \exp \left[\frac{4(x + x^2) \cos(\alpha_2 - \alpha_1)}{(1 - x)^3} + \frac{2(x^2 + x^4) \cos[2(\alpha_2 - \alpha_1)]}{(1 - x^2)^3} \right]. \quad (6.28)$$

We evaluate the partition function numerically for $N = 2$ using equation (6.28) at temperature $T = 0.64$, which is just above the Hagedorn temperature. Similarly, we can explicitly write down the first two terms in the partition function, which are the leading contributions and numerically integrate them for $N = 3, 4, 5$ at $T = 0.64$. We use these to derive the numerical values of entropy, average energy and use the entropy to calculate the dimensionality of the Hilbert space. These values are given in Table 3.3.

6.3 Maximum volume slices for the AdS black hole

In this appendix we will maximize the volume of the nice slices of AdS_{d+1} black holes whose endpoints are at $(u_0, 0)$ on the left horizon and $(0, v_0)$ on the right horizon. Using the isometry of AdS , we will set $u_0 = v_0$. The following calculation holds for both the one-sided black hole and the eternal black hole. The metric in the Kruskal coordinates for the eternal black hole is given by

$$ds^2 = -\frac{4f(r)}{f'(r_h)^2} e^{-f'(r_h)r^*} du_k dv_k + r^2 d\Omega_{d-1}^2, \quad (6.29)$$

where $f(r) = r^2 + 1 - \frac{C}{r^{d-2}}$, r_h is the black hole horizon and r^* is the tortoise coordinate. The Kruskal coordinates are denoted with a subscript k in order to avoid potential confusion with Eddington-Finkelstein coordinate v which we will be using later on.

The variational problem

In this subsection, we utilize the method given in [106, 107] to compute the volumes of maximal slices. First, we will define a conserved quantity E and write the maximum volume in terms of it. Afterwards, we will fix E in terms of the Kruskal coordinate u_0 .

Note that the method used to calculate the volume-maximizing slices is not restricted to AdS black holes. As an example, we will calculate the maximum volume for slices in the interior of a $2 + 1$ dimensional AdS black brane.

Expression for the maximum volume in terms of the conserved quantity E

We write the black hole metric in infalling Eddington-Finkelstein coordinates where we will be using $v = t + r^*$. The method used to compute the volume does not depend on a

specific $f(r)$. In these coordinates, the metric takes the form:

$$ds^2 = -f(r) dv^2 + 2dr dv + r^2 d\Omega_{d-1}^2, \quad (6.30)$$

where again $f(r) = r^2 + 1 - \frac{C}{r^{d-2}}$. We assume that the nice slice of maximum volume has the same symmetry as that of the $(d-1)$ sphere. Extremizing (6.31) gives the maximum volume, where a dot denotes the derivative with respect to σ , which is a parameter characterizing the nice slices. Here V_{d-1} is the volume of the $(d-1)$ spherical ball.

$$V = V_{d-1} \int d\sigma r^{d-1} \left(-f(r) \dot{v}^2 + 2\dot{r} \dot{v} \right)^{\frac{1}{2}} \quad (6.31)$$

Extremizing the above equation follows the same procedure for extremizing action, with the integrand playing the role of a Lagrangian. Since the Lagrangian does not depend on v , therefore we have a conserved quantity $E = -\frac{\partial L}{\partial \dot{v}}$.

The volume here (6.31) is reparametrization invariant as it does not depend on the choice of σ . We will fix the parametrization as follows:

$$r^{d-1} \left(-f(r) \dot{v}^2 + 2\dot{r} \dot{v} \right)^{\frac{1}{2}} - 1 = 0. \quad (6.32)$$

We can now write down equations determining $r(\sigma)$ and $v(\sigma)$ using the fixed parametrization and the expression for E , which allows us to write down r and v as coupled differential equations in terms of E .

$$E - r^{2(d-1)} [f(r)\dot{v} - \dot{r}] = 0 \quad (6.33)$$

$$r^{2(d-1)} \dot{r}^2 - f(r) - r^{-2(d-1)} E^2 = 0. \quad (6.34)$$

Eliminating \dot{v} using the above equation, the expression for the maximum volume takes the following form:

$$V = 2V_{d-1} \int_{r_{min}}^{r_h} \frac{dr}{\dot{r}} = 2V_{d-1} \int_{r_{min}}^{r_h} dr \frac{r^{2(d-1)}}{\sqrt{f(r)r^{2(d-1)} + E^2}}, \quad (6.35)$$

where r in the integral goes to a minimum value of r_{min} which is determined by substituting $\dot{r} = 0$ in (6.34).

$$f(r_{min})r_{min}^{2(d-1)} + E^2 = 0. \quad (6.36)$$

Fixing E in terms of Kruskal coordinate u_0

Here we fix E in terms of u_0 . We see that E is negative since at $r = r_{min}$ as we have $\dot{r}|_{r=r_{min}} = 0$ and $\dot{v}|_{r=r_{min}} > 0$. From the definition of the coordinate v , and using equations

(6.33) and (6.34) we get

$$t_r + r_h^* - r^*(r_{min}) = \int_{v_{min}}^{v_{r_h}} dv = \int_{r_{min}}^{r_h} dr \left[\frac{E}{f(r)\sqrt{f(r)r^{2(d-1)} + E^2}} + \frac{1}{f(r)} \right] \quad (6.37)$$

Inside the horizon, the Kruskal coordinate u_k is related to r^* and t as $u = e^{\frac{f'(r_h)}{2}(r^*-t)}$. Expressed in Kruskal coordinates (6.37) is given by:

$$\log(u_0) = \log u_{min} + \frac{f'(r_h)}{2} \int_{r_{min}}^{r_h} dr \left[\frac{E}{f(r)\sqrt{f(r)r^{2(d-1)} + E^2}} + \frac{1}{f(r)} \right] \quad (6.38)$$

Analytic expression for the volume growth of maximum volume surfaces in terms of Kruskal coordinate u_0

We derive an analytic expression relating the volume growth of the maximum volume surfaces in terms of the Kruskal coordinate u_0 . Notice that the integrands in (6.35) and (6.38) are regular and don't blow up at r_h . The integrands are denoted below by V' and I' respectively.

$$\lim_{r \rightarrow r_h} V' = \frac{V_{d-1} r_h^{2(d-1)}}{|E|} \quad (6.39)$$

$$\lim_{r \rightarrow r_h} I' = \lim_{r \rightarrow r_h} \left[\frac{E}{(r^2 - r_h^2)\sqrt{(r^2 - r_h^2)r^2 + E^2}} + \frac{1}{f(r)} \right] = \frac{1}{4} \frac{f'(r_h) r_h^{2(d-1)}}{E^2} \quad (6.40)$$

At r_{min} also both these integrands encounter a similar logarithmic blow-up.

$$\lim_{r \rightarrow r_{min}} V' = \frac{V_{d-1} r_{min}^{2(d-1)}}{\sqrt{E^2 + r^{2(d-1)} f(r)}} \quad (6.41)$$

$$\lim_{r \rightarrow r_{min}} I' = \lim_{r \rightarrow r_{min}} \left[\frac{E}{(r^2 - r_h^2)\sqrt{(r^2 - r_h^2)r^2 + E^2}} + \frac{1}{f(r)} \right] = -\frac{f'(r_h)|E|}{2f(r_{min})r_{min}^{2(d-1)}} \frac{r_{min}^{2(d-1)}}{\sqrt{E^2 + r^{2(d-1)} f(r)}} \quad (6.42)$$

Since these integrands have similar blow up, one can relate the volume with u_0 as follows using (6.36) and the definition of Hawking temperature β ,

$$V = \frac{\beta A(r_{min})}{2\pi} \log u_0 + O(1) \quad (6.43)$$

Here $O(1)$ is a subleading quantity which does not grow with u_0 .

An example: The 2+1 black brane

As an example of this above volume maximization we will take a look at 2+1 dimensional branes. This provides us an opportunity to study the late-time behaviour of E , thereby allowing us to understand how E characterizes the slice. The general $d + 1$ black brane metric is given by:

$$ds^2 = -f(r) dv^2 + 2dr dv + r^2 dx_{d-1}^2, \quad (6.44)$$

where $f(r) = r^2 \left(1 - \left(\frac{r_h}{r}\right)^d\right)$. The inverse temperature for the black brane is given by $\beta = \frac{4\pi}{f'(r_h)} = \frac{4\pi}{dr_h}$. The expression for u_{min} takes the form

$$u_{min} = e^{\frac{dr_h}{2} r^*(r_{min})}, \quad (6.45)$$

where r_{min} is calculated from (6.36) for $d = 2$ to be

$$r_{min} = \sqrt{\frac{r_h^2 + \sqrt{r_h^4 - 4E^2}}{2}}. \quad (6.46)$$

Note that (6.46) has 3 more roots, two of which are dropped because they are negative. The third root $r = \sqrt{\frac{r_h^2 - \sqrt{r_h^4 - 4E^2}}{2}}$ is dropped as the integrand becomes imaginary once the lower limit of integration goes below $r_{min} = \sqrt{\frac{r_h^2 + \sqrt{r_h^4 - 4E^2}}{2}}$. The volume and the relation of E with u_0 for $d = 2$ are respectively given by

$$V = \int_{r_{min}}^{r_h} dr \frac{r^2}{\sqrt{(r^2 - r_h^2)r^2 + E^2}} \quad (6.47)$$

and

$$\log u_0 = \log u_{min} + r_h \int_{r_{min}}^{r_h} dr \left[\frac{E}{(r^2 - r_h^2)\sqrt{(r^2 - r_h^2)r^2 + E^2}} + \frac{1}{f(r)} \right] \quad (6.48)$$

Late time behaviour of r_{min} and E

The indefinite integrals indicate that the volume tends to infinity as $E^2 \rightarrow \frac{r_h^4}{4}$. We took into account the largest root as r_{min} while solving the minimization equation $E^2 + r_{min}^2 f(r_{min}) = 0$. $|E|$ characterizes the nice slice and it increases monotonically with Kruskal time. Therefore at late times $r_m = \lim_{(u_0 \rightarrow \infty)} r_{min}$ is an extremum of $r^2 f(r)$, which translates to $[r^2 f(r)]' = 0$. By definition r_m is also a root of $E^2 + r^2 f(r) = 0$. Therefore we see at late times,

$$r_m = \frac{r_h}{\sqrt{2}}, \quad (6.49)$$

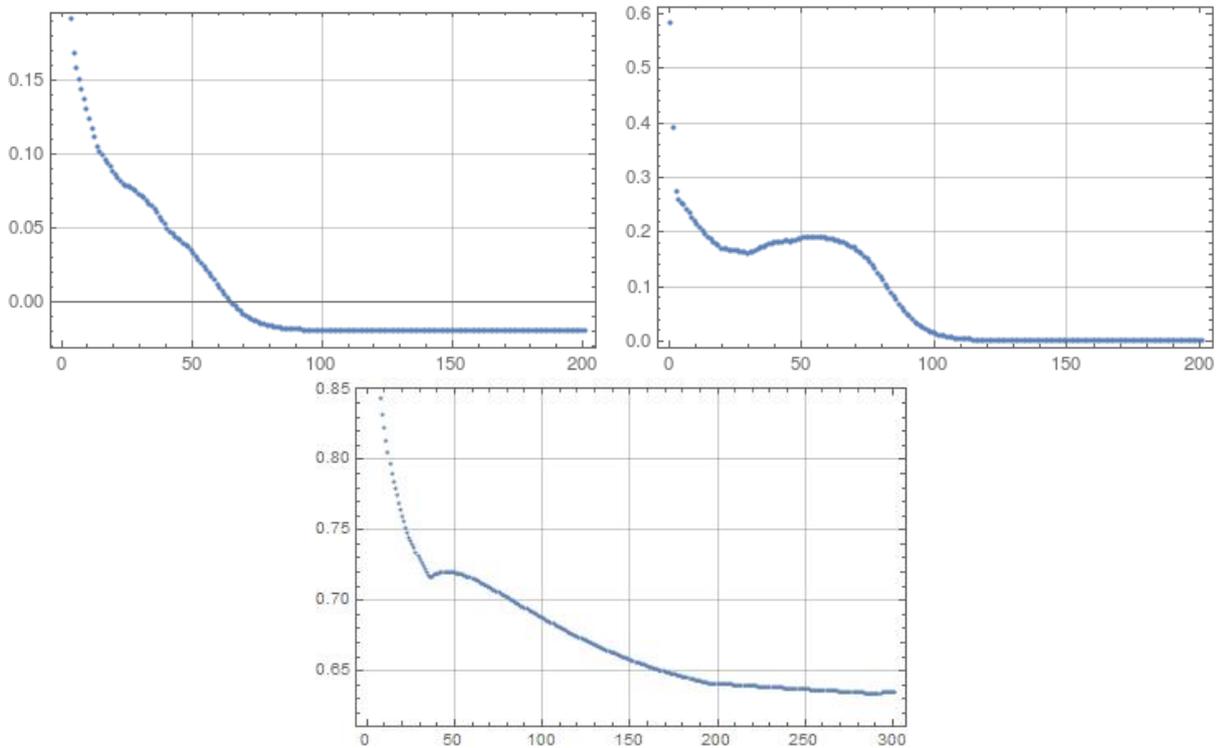


Figure 6.1: We test the pushing technique for simplex in the left figure, orthoplex on the right and hypercube in the bottom figure. We see that the inner products converge to their actual value after around 100 iterations for the left and the right figures where we set the parameters $\alpha = 100$, $\beta = 4$. In comparison, inner products converge to their actual value after 200-220 iterations for the bottom figure which we perform with $\alpha = 1$ and $\beta = 1$.

while late times E is related to r_h by

$$E^2 = \frac{r_h^4}{4}. \quad (6.50)$$

The volume is given by (6.43), where we substitute r_{min} from (6.46).

6.4 Technique used to accommodate vectors on the unit sphere in Hilbert space

In order to demonstrate overcounting, we construct a larger number of vectors on a sphere than its dimension such that they are almost equally separated from each other. This procedure is similar to the one we used to derive the worst-case overcounting formula.

In order to do this we define an energy function with a positive energy cost, i.e. a repelling force if vectors are too close to each other. The energy is minimized when the points are evenly distributed. We implement this numerically by pushing vectors away from each other. For vector \vec{v}_i where $i \neq j$,

$$\vec{v}_i \rightarrow \vec{v}_i + \alpha \frac{(\vec{v}_i - \vec{v}_j)}{|\vec{v}_i - \vec{v}_j|^{2\beta}} \quad (6.51)$$

This map drives a vector away from the nearest vectors and brings it closer to other distant vectors. The action pushes the vector out of the sphere, and we compensate this by normalizing the vector to bring them back on the sphere. α and β are parameters which make the pushing action more or less. As a result, the pushing action separates the vectors until they come close to equilibrium and are almost equally separated from each other.

As a test, we check our program for the inner products of simplex, orthoplex and hypercube whose inner product we have already computed in Appendix 6.1. We push these vectors 200, 200, 300 times respectively with $\alpha = 100, 100, 1$ for these three cases and $\beta = 4, 4, 1$, and find that there is convergence to the theoretically estimated dot product after approximately 120 iterations in each case. Figure 6.1 demonstrates the convergence of the maximum inner products to their theoretical values.

Figure 6.1 gives a nice description of what the method does. Stronger values of α and β means a stronger repelling force from nearby vectors and hence more energy cost. Hence for $\alpha = 100, \beta = 4$ we require about 120 iterations for the 300 points to converge. However, for a much lower value of $\alpha = 1, \beta = 1$ the repelling force is not that strong, and hence the energy cost is not great. Hence it takes many iterations for the inner products to converge to their theoretical value.

6.5 Projectors onto smeared modes' vacua

In this section we shall verify the expression for projector onto vacuum (4.13). We first take a variable transformation, $z = t_1 + it_2$ and $z^* = t_1 - it_2$. With $\alpha_s = \frac{1}{\sqrt{2}}(X_s + i\Pi_s)$ and using Baker–Campbell–Hausdorff lemma, we can write the projector as,

$$P_s = -\frac{1}{\pi^2} \int d^2z \int_0^{2\pi} d\theta_s \frac{e^{-z\bar{z}(1-i\tan\theta_s)}}{e^{i\theta_s} - 1 - \epsilon} e^{-\beta(\theta_s)\bar{z}\alpha_s^\dagger} e^{-\beta(\theta_s)z\alpha_s}$$

where $\beta(\theta) = \sqrt{2i\tan\theta}$. Let us calculate $\langle i_s | P_s | j_s \rangle$, where $|i_s\rangle, |j_s\rangle$ are number states corresponding to oscillator labelled by s . We get,

$$\langle i_s | e^{-\beta(\theta)\bar{z}\alpha_s^\dagger} e^{-\beta(\theta)z\alpha_s} | j_s \rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (-\beta(\theta_s))^{m+n} \frac{\bar{z}^m z^n}{m! n!} \langle i_s | \alpha_s^\dagger{}^m \alpha_s^n | j_s \rangle \quad (6.52)$$

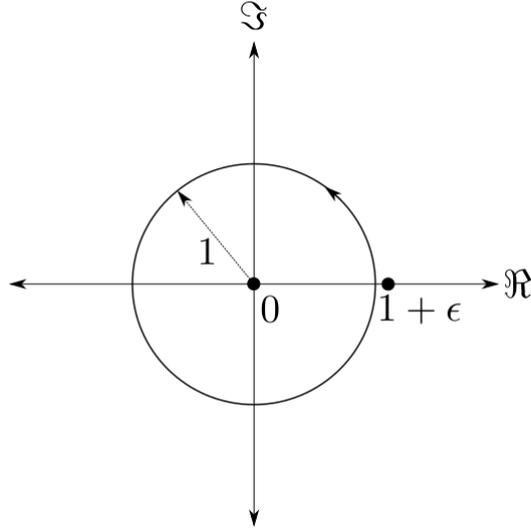


Figure 6.2: Contour used in the smeared vacuum projector calculation. There exist two poles at 0 and $1 + \epsilon$, amongst which only the former contributes to the unit circle contour integral.

This is only non-zero if $n + i_s = m + j_s$. If we also perform the z, \bar{z} integral with $z = re^{i\phi}$ and $\bar{z} = re^{-i\phi}$, that further constrains us with a δ_{i_s, j_s} factor. Hence,

$$\langle i_s | P_s | j_s \rangle = -\frac{1}{\pi} \int_0^{2\pi} d\theta_s \sum_{n=0}^{\infty} \frac{(2i \tan \theta_s)^n}{(1 - i \tan \theta_s)^{n+1} (e^{i\theta_s} - 1 - \epsilon)} \left[\frac{\delta_{i_s, j_s}}{n!} \langle i_s | \alpha_s^{\dagger n} \alpha_s^n | j_s \rangle \right]_{m=n} \quad (6.53)$$

The term inside third braces is,

$$\delta_{i_s, j_s} \frac{1}{n!} \langle i_s | \alpha_s^{\dagger n} \alpha_s^n | j_s \rangle = \begin{cases} \delta_{i_s, j_s} \frac{i_s!}{n!(n - i_s)!} & \text{for } n \leq i_s, \\ 0 & \text{for } n > i_s. \end{cases} \quad (6.54)$$

Summing over n and further changing variable $\omega = e^{i\theta_s}$, we have a contour integral

$$\langle i_s | P_s | j_s \rangle = -\frac{\delta_{i_s, j_s}}{2\pi i} \oint_{|\omega|=1} d\omega \frac{\omega^{2i_s-1} (\omega^2 + 1)}{(\omega - 1 - \epsilon)}. \quad (6.55)$$

With $\epsilon > 0$ as shown in Fig. 6.2 the contour evaluates to

$$\langle i_s | P_s | j_s \rangle = \delta_{i_s, j_s} \delta_{i_s, 0} \quad (6.56)$$

Hence,

$$P_s = -\frac{1}{\pi^2} \int dt_1 dt_2 \int_0^{2\pi} d\theta_s \frac{1}{e^{i\theta_s} - 1 - \epsilon} e^{-(t_1^2 + t_2^2) - \kappa(\theta_s)(t_1 X_s - t_2 \Pi_s)} = |0_s\rangle \langle 0_s| \quad (6.57)$$

6.6 Explicit commutator of smeared Rindler modes

Consider the commutator of the modes on region A first. The modes are given by:

$$\alpha_A = \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) \phi(t_A(U), r_A(U), \Omega), \quad (6.58)$$

$$\alpha_A^\dagger = \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_A^{\frac{(d-2)}{2}} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}^*(U) \phi(t_A(U), r_A(U), \Omega) \quad (6.59)$$

We integrate α_A^\dagger by parts, and the only part of α_A^\dagger that contributes to the commutator $[\alpha_A, \alpha_A^\dagger]$ is

$$\frac{1}{\sqrt{V_\Omega}} \int d^{d-2}\Omega \int dU r_A^{\frac{(d-2)}{2}} \partial_U \phi(U, -2v_0, \Omega) \int \frac{d\nu}{i\nu} \tilde{\mathcal{T}}^*(\nu) \left(\frac{U}{U_0}\right)^{-i\nu} \quad (6.60)$$

Note that here we have used $\mathcal{T}(U) \left(\frac{U}{U_0}\right)^{i\omega_0} = \int d\nu \tilde{\mathcal{T}}(\nu) \left(\frac{U}{U_0}\right)^{i\nu}$ to perform the following replacement

$$\int_0^U \frac{dU'}{U'} \mathcal{T}(U') \left(\frac{U'}{U_0}\right)^{i\omega_0} = \int \frac{d\nu}{i\nu} \tilde{\mathcal{T}}(\nu) \left(\frac{U}{U_0}\right)^{i\nu}. \quad (6.61)$$

In terms of the light cone coordinates, the annihilation operator is given by

$$\alpha_A = \frac{1}{\sqrt{V_\Omega}} \int \frac{dU}{U} \int d^{d-2}\Omega r_A^{\frac{(d-2)}{2}} \phi(U, -2v_0, \Omega) \int d\nu \tilde{\mathcal{T}}(\nu) \left(\frac{U}{U_0}\right)^{i\nu} \quad (6.62)$$

Using the null surface canonical commutation [156] relation

$$[\phi(U_1, V, \Omega_1), \partial_U \phi(U_2, V, \Omega_2)] = \frac{i\delta(U_1 - U_2)\delta(\Omega_1, \Omega_2)}{2r_1^{d-2}} \quad (6.63)$$

we get

$$[\alpha_A, \alpha_A^\dagger] = \frac{1}{2} \int \frac{d\Omega}{V_\Omega} \int \frac{d\nu_1 d\nu_2}{\nu_2} \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i(\nu_1 - \nu_2)} \quad (6.64)$$

$$= \frac{1}{2} \int \frac{d\Omega}{V_\Omega} \int \frac{d\nu_1 d\nu_2}{\nu_2} \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) 2\pi\delta(\nu_1 - \nu_2) \quad (6.65)$$

$$= 1 \quad (6.66)$$

where we have used the normalization $\int \frac{d\nu}{\nu} |\tilde{\mathcal{T}}(\nu)|^2 = \frac{1}{\pi}$.

6.7 Computation of $\langle \mathcal{G} \rangle$

We will discretize the frequency space with a Δ gap for ease of calculation. Towards the end of this appendix, we will go back to the continuous limit by taking $\Delta \rightarrow 0$.

Performing the discretization, the global mode commutation relation (4.8) becomes

$$[a_{n,l}, a_{n',l'}^\dagger] = \frac{\delta_{n,n'} \delta_{l,l'}}{\Delta}, \quad (6.67)$$

where we have labelled the frequency ω with integer n . The global mode decomposition (4.9) now looks like $\alpha_s = \Delta \sum_{n,l} h_s(n,l) a_{n,l} + g_s^*(n,l) a_{n,l}^\dagger$. Using the BCH lemma, we can decompose the B piece in $\langle \mathcal{G} \rangle$ in terms of creation and annihilation operators as

$$e^{v_2 \alpha_B^\dagger} e^{(\tilde{y}_1 X_B - \tilde{y}_2 \Pi_B)} e^{\zeta_2 \alpha_B} = e^{\left(v_2 + \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}}\right) \alpha_B^\dagger + \left(\zeta_2 + \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}}\right) \alpha_B} e^{-\frac{1}{2} \left(v_2 \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}} + \zeta_2 \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}} + v_2 \zeta_2\right)}. \quad (6.68)$$

Similarly decomposing the A piece and then writing both in terms of global modes, we get

$$\langle \mathcal{G} \rangle = \frac{\left\langle \exp \left(\Delta \sum_{n,l} \left(u_B(n,l) + u_A(n,l) \right) a_{n,l} + \left(u'_B(n,l) + u'_A(n,l) \right) a_{n,l}^\dagger \right) \right\rangle}{e^{\frac{1}{2} \left(v_2 \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}} + \zeta_2 \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}} + v_2 \zeta_2 \right)} e^{\frac{1}{2} \left(v_1 \frac{(\tilde{t}_1 + i\tilde{t}_2)}{\sqrt{2}} + \zeta_1 \frac{(\tilde{t}_1 - i\tilde{t}_2)}{\sqrt{2}} + v_1 \zeta_1 \right)}}, \quad (6.69)$$

where

$$u_A(n,l) = \left(v_1 + \frac{(\tilde{t}_1 - i\tilde{t}_2)}{\sqrt{2}} \right) g_A(n,l) + \left(\zeta_1 + \frac{(\tilde{t}_1 + i\tilde{t}_2)}{\sqrt{2}} \right) h_A(n,l), \quad (6.70)$$

$$u'_A(n,l) = \left(v_1 + \frac{(\tilde{t}_1 - i\tilde{t}_2)}{\sqrt{2}} \right) h_A^*(n,l) + \left(\zeta_1 + \frac{(\tilde{t}_1 + i\tilde{t}_2)}{\sqrt{2}} \right) g_A^*(n,l), \quad (6.71)$$

$$u_B(n,l) = \left(v_2 + \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}} \right) g_B(n,l) + \left(\zeta_2 + \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}} \right) h_B(n,l), \quad (6.72)$$

$$u'_B(n,l) = \left(v_2 + \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}} \right) h_B^*(n,l) + \left(\zeta_2 + \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}} \right) g_B^*(n,l). \quad (6.73)$$

Next we use a simple result involving coherent states of harmonic oscillators to simplify our expressions further. Consider a system of oscillators with ground states $|0_i\rangle$, where ($i = 1, 2, \dots, \infty$), with commutation relations $[\hat{\alpha}_i, \hat{\alpha}_j^\dagger] = \delta_{ij}$. Define the combined ground state of the system as $|0\rangle \equiv \otimes_i |0_i\rangle$. This setup is intended to mimic the global modes $a_{n,l}$, as in our theory the global vacuum is indeed the tensor product of all the different global mode vacua. A coherent state in the j 'th oscillator is given by $|z_j\rangle \equiv e^{z_j \hat{\alpha}_j^\dagger} |0_j\rangle$, and the inner product between two such states is $\langle z_j | z'_j \rangle = e^{z_j^* z'_j}$ [157]. Then we have

$$\langle 0 | e^{\sum_i z_i^* \hat{\alpha}_i + z'_i \hat{\alpha}_i^\dagger} | 0 \rangle = e^{-\frac{1}{2} \sum_i z_i^* z'_i} \langle 0 | e^{\sum_i z_i^* \hat{\alpha}_i} e^{\sum_j z'_j \hat{\alpha}_j^\dagger} | 0 \rangle \quad (6.74)$$

$$= e^{-\frac{1}{2} \sum_i z_i^* z'_i} \prod_{ij} \langle 0_i | e^{z_i^* \hat{\alpha}_i} e^{z'_j \hat{\alpha}_j^\dagger} | 0_j \rangle \quad (6.75)$$

$$= e^{-\frac{1}{2} \sum_i z_i^* z'_i} \prod_j \langle z_j | z'_j \rangle \quad (6.76)$$

$$= e^{\frac{1}{2} \sum_i z_i^* z'_i}. \quad (6.77)$$

To make use of this in simplifying (6.69), we identify $\sqrt{\Delta}a_{n,l}$ with \hat{a}_j . This gives us terms like $\sum_n \Delta u_A(n,l)u'_B(n,l)$ on top of the exponential. Taking the limit $\Delta \rightarrow 0$, the sum \sum_j goes to an integral and the whole expression simplifies to

$$\langle \mathcal{G} \rangle = \frac{\exp\left(\frac{1}{2}(u_B + u_A) \cdot (u'_B + u'_A)\right)}{e^{\frac{1}{2}\left(v_2 \frac{(\tilde{y}_1 + i\tilde{y}_2)}{\sqrt{2}} + \zeta_2 \frac{(\tilde{y}_1 - i\tilde{y}_2)}{\sqrt{2}} + v_2 \zeta_2\right)} e^{\frac{1}{2}\left(v_1 \frac{(\tilde{t}_1 + i\tilde{t}_2)}{\sqrt{2}} + \zeta_1 \frac{(\tilde{t}_1 - i\tilde{t}_2)}{\sqrt{2}} + v_1 \zeta_1\right)}}, \quad (6.78)$$

where $u_A \cdot u_B \equiv \sum_l \int d\omega u_A(\omega, l)u_B(\omega, l)$. Using the f_p defined in (4.19), we obtain

$$u_A + u_B = \frac{1}{\sqrt{2}}[f_1(\tilde{t}_1 + \zeta_1^+) + f_2(-\tilde{t}_2 + i\zeta_1^-) + f_3(\tilde{y}_1 + \zeta_2^+) + f_4(-\tilde{y}_2 + i\zeta_2^-)] \quad (6.79)$$

$$u'_A + u'_B = \frac{1}{\sqrt{2}}[f_1^*(\tilde{t}_1 + \zeta_1^+) + f_2^*(-\tilde{t}_2 + i\zeta_1^-) + f_3^*(\tilde{y}_1 + \zeta_2^+) + f_4^*(-\tilde{y}_2 + i\zeta_2^-)]. \quad (6.80)$$

Re-arranging the terms to gather the f_p 's together and using the m_q defined in (4.19), we finally obtain

$$\langle \mathcal{G} \rangle = \exp\left(\frac{1}{8} \sum_{p,q=1}^4 (f_p \cdot f_q^* + f_q \cdot f_p^*)m_p m_q - \frac{\mathcal{R}}{2}\right). \quad (6.81)$$

6.8 Bogoliubov coefficients and $\langle C_{AB} \rangle \geq 2$

In this appendix, we demonstrate the calculation of the Bogoliubov coefficients and show that $\langle C_{AB} \rangle \geq 2$.

Bogoliubov coefficients of local Rindler-to-global modes

From (4.35) we can read off the Bogoliubov coefficients using the large frequency limit, which are given by

$$\begin{aligned} h_A(\omega, 0) &= \frac{1}{\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) e^{-i\omega t_A} \cos\left(\omega r_A - \frac{(d-2)\pi}{4}\right), \\ g_A^*(\omega, 0) &= \frac{1}{\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) e^{i\omega t_A} \cos\left(\omega r_A - \frac{(d-2)\pi}{4}\right), \\ h_B(\omega, 0) &= \frac{1}{\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}(U) e^{-i\omega t_B} \cos\left(\omega r_B - \frac{(d-2)\pi}{4}\right), \\ g_B^*(\omega, 0) &= \frac{1}{\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{-i\omega_0} \mathcal{T}(U) e^{i\omega t_B} \cos\left(\omega r_B - \frac{(d-2)\pi}{4}\right). \end{aligned} \quad (6.82)$$

The above Bogoliubov coefficients are written in terms of integrals over U . We can perform these integrals using our conditions on the tuning function in (4.32) and (4.33). Since the form of the integrals is similar, we will demonstrate this by evaluating the

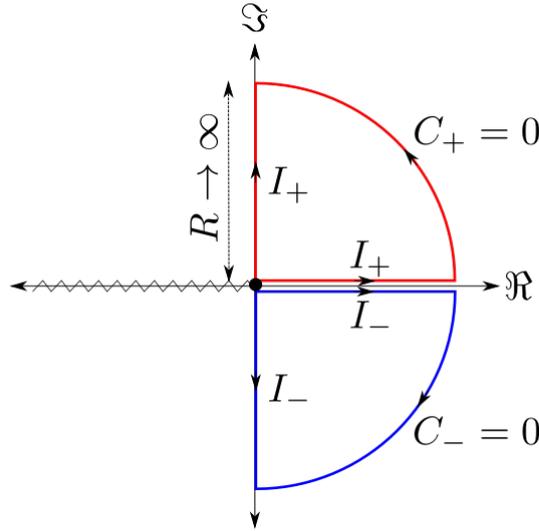


Figure 6.3: The red and the blue contours are respectively used for the integrals $I_{\pm} = \int dx e^{\pm ix} x^{i\nu-1}$. Both the curved contours C_{\pm} give 0 in the limit $R \rightarrow \infty$. Since there are no poles inside either contour (they are slightly separated from the real axis), the total contours also give 0. This lets us equate the integrals over the real and the imaginary axes for both the e^{ix} and e^{-ix} cases. We keep the branch cut (occurring due to $x^{i\nu}$) on the negative real axis so it doesn't interfere with the calculation.

$h_A(\omega, 0)$ integral:

$$\begin{aligned}
 h_A(\omega, 0) &= \frac{1}{\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) e^{-i\omega t_A} \cos\left(\omega r_A - \frac{(d-2)\pi}{4}\right) \\
 &= \frac{1}{2\sqrt{\pi\omega}} \int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i\omega_0} \mathcal{T}(U) [e^{-i\xi_1} e^{-i\omega U} + e^{i\xi_2}] \\
 &= \frac{1}{2\sqrt{\pi\omega}} \int d\nu \frac{\tilde{\mathcal{T}}(\nu)}{(U_0)^{i\nu}} \left[e^{-i\xi_1} \int dU U^{i\nu-1} e^{-i\omega U} + e^{i\xi_2} \int \frac{dU}{U} U^{i\nu} \right] \\
 &= \frac{e^{-i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu \frac{\tilde{\mathcal{T}}(\nu)}{(\omega U_0)^{i\nu}} \int dx x^{i\nu-1} e^{-ix} + 0
 \end{aligned}$$

where $\xi_1 = \frac{(d-2)\pi}{4} - \omega r_0$, $\xi_2 = \frac{(d-2)\pi}{4} - \omega(r_0 - 2v_0)$. The second term in the third line vanishes because $\int \frac{dU}{U} \left(\frac{U}{U_0}\right)^{i\nu} = 2\pi\delta(\nu)$, and using (4.33) $\tilde{\mathcal{T}}(\nu)$ vanishes at $\nu = 0$. The x integral can be evaluated by choosing a contour shown in Fig. 6.3. We encounter such x integrals in the expressions for the other Bogoliubov coefficients as well, where we similarly choose appropriate contours and obtain the following values for the integrals

$$\int_0^{\infty} dx x^{i\nu-1} e^{\pm ix} = e^{\mp\pi\nu/2} \Gamma(i\nu). \quad (6.83)$$

Thus the Bogoliubov coefficients can be conveniently summarized as

$$\begin{aligned}
h_A(\omega, 0) &= \frac{e^{-i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{\pi\nu/2} (\omega U_0)^{-i\nu} \Gamma(i\nu) \tilde{\mathcal{T}}(\nu), \\
g_A^*(\omega, 0) &= \frac{e^{i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{-\pi\nu/2} (\omega U_0)^{-i\nu} \Gamma(i\nu) \tilde{\mathcal{T}}(\nu), \\
h_B(\omega, 0) &= \frac{e^{-i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{\pi\nu/2} (\omega U_0)^{i\nu} \Gamma(-i\nu) \tilde{\mathcal{T}}^*(\nu), \\
g_B^*(\omega, 0) &= \frac{e^{i\xi_1}}{2\sqrt{\pi\omega}} \int d\nu e^{-\pi\nu/2} (\omega U_0)^{i\nu} \Gamma(-i\nu) \tilde{\mathcal{T}}^*(\nu).
\end{aligned} \tag{6.84}$$

Calculation of $\langle C_{AB} \rangle \geq 2$ for entangled Rindler modes

Here we will demonstrate the calculation of the element $f_1 \cdot f_1^* = h_A^* \cdot h_A + g_A^* \cdot g_A + h_A^* \cdot g_A + g_A^* \cdot h_A$. The typical integral encountered here is of the form

$$\begin{aligned}
h_A \cdot h_A^* &= \int \frac{d\omega}{4\pi\omega} \int d\nu_1 d\nu_2 e^{\pi(\nu_1+\nu_2)/2} (\omega U_0)^{i(\nu_2-\nu_1)} \Gamma(i\nu_1) \Gamma^*(i\nu_2) \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \\
&= \frac{1}{2} \int d\nu_1 d\nu_2 e^{\pi(\nu_1+\nu_2)/2} \Gamma(i\nu_1) \Gamma^*(i\nu_2) \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \int \frac{d\omega}{2\pi\omega} (\omega U_0)^{i(\nu_2-\nu_1)} \\
&= \frac{1}{2} \int d\nu_1 d\nu_2 e^{\pi(\nu_1+\nu_2)/2} \Gamma(i\nu_1) \Gamma^*(i\nu_2) \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \delta(\nu_1 - \nu_2) \\
&= \frac{1}{2} \int d\nu e^{\pi\nu} |\Gamma(i\nu)|^2 |\tilde{\mathcal{T}}(\nu)|^2 \\
&= \frac{\pi}{2} \int d\nu \frac{e^{\pi\nu}}{\nu \sinh(\pi\nu)} |\tilde{\mathcal{T}}(\nu)|^2 \\
&= \frac{e^{\pi\omega_0}}{2 \sinh(\pi\omega_0)} \equiv \frac{1}{1-x^2}.
\end{aligned}$$

where $x \equiv e^{-\pi\omega_0}$. Here in the fifth step, we have used the identity $|\Gamma(i\nu)|^2 = \frac{\pi}{\nu \sinh(\pi\nu)}$ and used the fact that $|\tilde{\mathcal{T}}(\nu)|^2/\nu$ is sharply peaked around ω_0 to go the sixth step. We will now show that $h_A \cdot g_A^*$ is zero.

$$\begin{aligned}
h_A \cdot g_A^* &= \int \frac{d\omega}{4\pi\omega} \int d\nu_1 d\nu_2 e^{\pi(\nu_1-\nu_2)/2} (\omega U_0)^{-i(\nu_2+\nu_1)} \Gamma(i\nu_1) \Gamma^*(i\nu_2) \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \\
&= \frac{1}{2} \int d\nu_1 d\nu_2 e^{\pi(\nu_1-\nu_2)/2} \Gamma(i\nu_1) \Gamma^*(i\nu_2) \tilde{\mathcal{T}}(\nu_1) \tilde{\mathcal{T}}^*(\nu_2) \delta(\nu_1 + \nu_2) \\
&= \frac{1}{2} \int d\nu e^{\pi\nu} \Gamma(i\nu) \Gamma^*(-i\nu) \tilde{\mathcal{T}}(\nu) \tilde{\mathcal{T}}^*(-\nu) \\
&= 0
\end{aligned}$$

The final step follows due to the fact that within the integral $\int d\nu \tilde{\mathcal{T}}(\nu) \tilde{\mathcal{T}}^*(-\nu)$, when $\tilde{\mathcal{T}}(\nu)$ peaks at $\nu = \omega_0$, the other term goes to zero, i.e. $\tilde{\mathcal{T}}^*(-\nu) = \tilde{\mathcal{T}}^*(-\omega_0) \approx 0$. The rest of the terms are evaluated by straightforward replication of the above logic. We similarly

evaluate the following expressions in order to completely determine the ff matrix.

$$\begin{aligned}
 h_A \cdot h_A^* &= h_B \cdot h_B^* = \frac{1}{1-x^2}; \\
 g_A \cdot g_A^* &= g_B \cdot g_B^* = \frac{x^2}{1-x^2}; \\
 h_A \cdot g_B^* &= g_A \cdot h_B^* = \frac{x}{1-x^2}; \\
 h_A \cdot g_A^* &= h_A \cdot h_B^* = g_A \cdot g_B^* = h_B \cdot g_B^* = 0.
 \end{aligned} \tag{6.85}$$

Substituting the expressions in (6.85) in $f_p \cdot f_q^* + f_p^* \cdot f_q$, we obtain

$$f_p \cdot f_q^* + f_q \cdot f_p^* = \frac{2}{1-x^2} \begin{pmatrix} 1+x^2 & 0 & 2x & 0 \\ 0 & 1+x^2 & 0 & -2x \\ 2x & 0 & 1+x^2 & 0 \\ 0 & -2x & 0 & 1+x^2 \end{pmatrix}. \tag{6.86}$$

6.9 Proof of $\langle A_j C_i \rangle = \langle A_j B_i \rangle + O(\sqrt{G_N})$ and boundedness of C_i of §4.4.2

In this appendix, we will show that the operators C_i constructed as

$$C_i \equiv \frac{\langle B_i^2 \rangle (Q_i \mathcal{P}_0 + \mathcal{P}_0 Q_i^\dagger - \langle B_i \rangle \mathcal{P}_0) - \langle B_i \rangle Q_i \mathcal{P}_0 Q_i^\dagger}{\langle B_i^2 \rangle - \langle B_i \rangle^2}. \tag{6.87}$$

do indeed mimic the contribution of operators B_i in the two-point correlators and are bounded.

Proof of $\langle A_j C_i \rangle = \langle A_j B_i \rangle$

First we note that by construction we have $Q_i |0, \{s\}\rangle = B_i |0, \{s\}\rangle$ for all sectors $\{s\}$, which guarantees $\mathcal{P}_0 Q_i^\dagger = \mathcal{P}_0 B_i^\dagger = \mathcal{P}_0 B_i$. Further, the fact that B_i is block diagonal in and independent of supertranslation sectors, allows us to write $\langle 0, \{s\} | B_i | 0, \{s'\} \rangle = K \delta(\{s\} - \{s'\})$. Taking the expectation value w.r.t. the smeared vacuum $|0, \mathcal{S}\rangle$ gives us $K = \langle 0, \mathcal{S} | B_i | 0, \mathcal{S} \rangle$. Thus we have

$$\langle 0, \{s'\} | B_i | 0, \{s\} \rangle = \langle 0, \mathcal{S} | B_i | 0, \mathcal{S} \rangle \delta(\{s'\} - \{s\}). \tag{6.88}$$

So, we have

$$\begin{aligned}
\mathcal{P}_0 Q_i^\dagger |0, \mathcal{S}\rangle &= \mathcal{P}_0 B_i |0, \mathcal{S}\rangle \\
&= \int \left(\prod_{l,m} ds_{l,m} \right) \int \left(\prod_{l,m} ds'_{l,m} \right) |0, \{s'\}\rangle \langle 0, \{s'\} | B_i |0, \{s}\rangle \mathcal{S}(\{s\}) + \mathcal{O}(\sqrt{G_N}) \\
&= \int \left(\prod_{l,m} ds_{l,m} \right) \int \left(\prod_{l,m} ds'_{l,m} \right) |0, \{s'\}\rangle \langle B_i | \delta(\{s\} - \{s'\}) \mathcal{S}(\{s\}) + \mathcal{O}(\sqrt{G_N}) \\
&= \langle B_i | \int \left(\prod_{l,m} ds_{l,m} \right) \mathcal{S}(\{s\}) |0, \{s'\}\rangle + \mathcal{O}(\sqrt{G_N}) \\
&= \langle B_i |0, \mathcal{S}\rangle + \mathcal{O}(\sqrt{G_N}).
\end{aligned} \tag{6.89}$$

Now

$$\begin{aligned}
C_i |0, \mathcal{S}\rangle &= \frac{\langle B_i^2 \rangle (Q_i |0, \mathcal{S}\rangle + \mathcal{P}_0 Q_i^\dagger |0, \mathcal{S}\rangle - \langle B_i |0, \mathcal{S}\rangle) - \langle B_i | Q_i \mathcal{P}_0 Q_i^\dagger |0, \mathcal{S}\rangle}{\langle B_i^2 \rangle - \langle B_i \rangle^2} + \mathcal{O}(\sqrt{G_N}) \\
&= \frac{\langle B_i^2 \rangle (Q_i |0, \mathcal{S}\rangle + \langle B_i |0, \mathcal{S}\rangle - \langle B_i |0, \mathcal{S}\rangle) - \langle B_i \rangle \langle B_i \rangle Q_i |0, \mathcal{S}\rangle}{\langle B_i^2 \rangle - \langle B_i \rangle^2} + \mathcal{O}(\sqrt{G_N}) \\
&= \frac{(\langle B_i^2 \rangle - \langle B_i \rangle^2) Q_i |0, \mathcal{S}\rangle}{\langle B_i^2 \rangle - \langle B_i \rangle^2} + \mathcal{O}(\sqrt{G_N}) \\
&= B_i |0, \mathcal{S}\rangle + \mathcal{O}(\sqrt{G_N}).
\end{aligned} \tag{6.90}$$

Thus we can clearly see that $\langle A_j C_i \rangle = \langle A_j B_i \rangle + \mathcal{O}(\sqrt{G_N})$ from here.

Boundedness of C_i

Let us define the orthonormal states $|B_i^\perp, \{s\}\rangle \equiv \frac{1}{\beta_i} (1 - \mathcal{P}_0) B_i |0, \{s\}\rangle$, where $\beta_i \equiv \sqrt{\langle B_i^2 \rangle - \langle B_i \rangle^2}$. Then C_i can be written as

$$\begin{aligned}
C_i &= \int \left(\prod_{l,m} ds_{l,m} \right) \left(|0, \{s\}\rangle \langle 0, \{s\}| + \beta_i \left(|0, \{s\}\rangle \langle B_i^\perp, \{s\}| + |B_i^\perp, \{s\}\rangle \langle 0, \{s\}| \right) \right. \\
&\quad \left. - \langle B_i \rangle |B_i^\perp, \{s\}\rangle \langle B_i^\perp, \{s\}| \right).
\end{aligned} \tag{6.91}$$

In terms of the orthonormal basis $\{|0, \{s\}\rangle, |B_i^\perp, \{s\}\rangle\}$, it takes the form

$$C_i = \begin{pmatrix} \langle B_i \rangle & \beta_i \\ \beta_i & -\langle B_i \rangle \end{pmatrix} \otimes \mathbb{1} + \mathcal{O}(\sqrt{G_N}) \tag{6.92}$$

where the $\otimes \mathbb{1}$ stands for C_i 's identity action on supertranslation sectors. The eigenvalues of C_i are $\pm\sqrt{\langle B_i^2 \rangle} + O(\sqrt{G_N})$ and hence the norm is

$$\|C_i\|^2 = \langle B_i^2 \rangle + O(\sqrt{G_N}) < 1. \quad (6.93)$$

6.10 Proof of existence of C_i of §4.4.2

Let's split the boundary observable as $C = C^{(0)} + C^{(\delta)}$ where $C^{(0)}$ contains \mathcal{P}_0 and $C^{(\delta)}$ contains $\delta\mathcal{P}$ (CHSH label is suppressed). We know that the $C^{(0)}$ part gives the desired correlators and is bounded in a desired way by constructing a boundary Q_i such that $Q_i|0, \mathcal{S}\rangle = B_i|0, \mathcal{S}\rangle$, which Reeh-Schlieder guarantees can always be done. We need to show that $\langle A_j C_i^{(\delta)} \rangle$ can be made arbitrarily small. We have

$$\beta_i^2 \langle A_j C_i^{(\delta)} \rangle = \langle B_i^2 \rangle \underbrace{\langle \Omega | A_j \delta\mathcal{P} Q_i^\dagger | 0, \mathcal{S} \rangle}_{\text{energy} < \delta} - \langle B_i \rangle \underbrace{\langle \Omega | A_j Q_i \delta\mathcal{P} Q_i^\dagger | 0, \mathcal{S} \rangle}_{\text{energy} < \delta}. \quad (6.94)$$

Both of these terms can be interpreted as the inner product of $Q_i^\dagger|0, \mathcal{S}\rangle$ and a bra which contains excitations on $\langle \Omega |$ with energy less than δ . The latter is a linear combination of bras of the kind $\langle \Omega | a_{\omega_1} a_{\omega_2} \dots a_{\omega_n}$ such that $\sum_j \omega_j < \delta$. Here we have suppressed the l label of the global annihilators because they don't contribute to energy. Each of these terms is

$$\langle \Omega | a_{\omega_1} a_{\omega_2} \dots a_{\omega_n} Q_i^\dagger | \Omega \rangle = \langle \Omega | a_{\omega_1} \dots a_{\omega_{j-1}} a_{\omega_{j+1}} \dots a_{\omega_n} [a_{\omega_j}, Q_i^\dagger] | \Omega \rangle, \quad (6.95)$$

where ω_j is any of the n different energies. So these terms can be made arbitrarily small individually if we can guarantee

$$[Q_i, a_\omega^\dagger] \approx 0 \quad \forall 0 < \omega < \delta. \quad (6.96)$$

where \approx has been used to mean ‘‘arbitrarily close to’’. This condition requires that Q_i in addition to satisfying

$$Q_i|0, \mathcal{S}\rangle \approx B_i|0, \mathcal{S}\rangle, \quad (6.97)$$

needs to be constructed in a way such that it contains (arbitrarily) small contribution from $a_{\omega,l}$ for $\omega < \delta$. To make this condition more precise we inspect how smearing of the field operator translates into smearing of creation and annihilation operators in energy domain. Consider a smearing of the kind $\phi_f = \int dt f(t)\phi(t)$. We have suppressed the position argument of both the field ϕ and the smearing function f for simplicity. This

decomposes as,

$$\begin{aligned}
\phi_f &= \int dt f(t)\phi(t) \\
&= \int dt f(t) \int_0^\infty d\omega \left(e^{-i\omega t} a_\omega + e^{i\omega t} a_\omega^\dagger \right) \\
&= \int_0^\infty d\omega \left(\hat{f}(\omega) a_\omega^\dagger + \hat{f}(-\omega) a_\omega \right) \\
&= a_{\hat{f}^+}^\dagger + a_{\hat{f}^-},
\end{aligned} \tag{6.98}$$

where the hats represent time domain Fourier transforms, the subscripts on mode operators denote the frequency space smearing: $a_{\hat{f}^\pm}^\dagger \equiv \int_0^\infty d\omega \hat{f}^\pm(\omega) a_\omega^\dagger$ and $\hat{f}^\pm(\omega) = \hat{f}(\pm\omega)$. Again, we have suppressed the sum over spherical mode information l for the creation and annihilation operators for brevity. Evidently, creation and annihilation operators are weighted by the positive and negative Fourier modes of $f(t)$ respectively.

Now we make the condition set on Q_i more precise. Let \mathcal{A} be the algebra generated by all ϕ smearings in region C . Also, let $\mathcal{A}_{\delta,\theta}$ be the subset of \mathcal{A} containing operators of the kind $\phi_{f_1} + \phi_{f_2}\phi_{f_3} + \dots$ such that $\sum_i \int_0^\delta d\omega \left| \hat{f}_i^-(\omega) \right|^2 < \theta$. In simple terms, $\mathcal{A}_{\delta,\theta}$ is a subset of \mathcal{A} in which all elements obey (6.96) up to precision θ (this may not be the maximal subset with this property). We shall argue for the existence of a Q_i obeying both Eqs. (6.96) and (6.97) by showing that $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense in the entire Hilbert space \mathcal{H} which is generated by field operations on $|0, \mathcal{S}\rangle$ for any $\theta > 0$ however small. Notice that the denseness of $\mathcal{A} |0, \mathcal{S}\rangle$ in \mathcal{H} is just the statement of the Reeh-Schlieder theorem and hence the denseness of $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ in \mathcal{H} is not too surprising.

To simplify things a little, we take \mathcal{A} to be the algebra of all operator smearings with support in some time band $[0, \epsilon]$ (as a simplified model of region C). We also simplify the definition of $\mathcal{A}_{\delta,\theta}$ accordingly. Now, let us split the full Hilbert space into particle number sectors as $\mathcal{H} = \bigoplus_{n=0}^\infty \mathcal{H}_n$, where the n -particle sector \mathcal{H}_n contains states like $a_{\hat{f}_1}^\dagger \dots a_{\hat{f}_n}^\dagger |0, \mathcal{S}\rangle$. We shall show that $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense in both even and odd particle sectors by induction.

Even sector: Consider the hypothesis: $\mathcal{A}_{\delta,\theta}$ is dense in \mathcal{H}_n . By axiom, the global identity operator \mathbb{I} exists in $\mathcal{A}_{\delta,\theta}$ (\mathcal{A} is a von Neumann algebra), and hence $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ contains $|0, \mathcal{S}\rangle$. Therefore $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense in \mathcal{H}_0 , i.e. the hypothesis is true for $n = 0$. Consider a general $(2n + 2)$ particle term $a_{\hat{f}_1}^\dagger \dots a_{\hat{f}_{2n+2}}^\dagger |0, \mathcal{S}\rangle$. Because we have all the $\hat{f}_i(\omega)$ at our disposal, by the corollary stated and proved in §6.10.1, we can construct a $g_i(t)$ with support in $[0, \epsilon]$ such that $\hat{g}_i^+(\omega) \approx \hat{f}_i(\omega)$ and $\hat{g}_i^-(\omega) \approx 0$ for $0 < \omega < \delta$. This gives

$$\phi_{g_1} \phi_{g_2} \dots \phi_{g_{2n+2}} |0, \mathcal{S}\rangle \approx a_{\hat{f}_1}^\dagger a_{\hat{f}_2}^\dagger \dots a_{\hat{f}_{2n+2}}^\dagger |0, \mathcal{S}\rangle + (\mathcal{H}_{2n} \text{ term}) + \dots + (\mathcal{H}_2 \text{ term}) + (\mathcal{H}_0 \text{ term}). \tag{6.99}$$

The first term on the RHS is the one we need to approximate, but other lower particle number terms show up due to the non-commutativity of creation and annihilation oper-

ators. If the hypothesis is true for $n = 2, 4, \dots, 2n$, then these residual terms are also limit points of $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ and can be cancelled off to any precision by summoning a state from $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$. But this means the $(2n+2)$ particle term is also a limit point of $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$. Hence the hypothesis is true for all even n (including 0) and we have proved by strong induction that $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense in all even number particle sectors.

Odd sector: Consider a state $a_{\hat{f}}^\dagger |0, \mathcal{S}\rangle \in \mathcal{H}_1$. Just like in the even case, $\hat{f}_1(\omega)$ lets us construct $g_1(t)$ with support in $[0, \epsilon]$ such that $\hat{g}_1^+(\omega) \approx \hat{f}_1(\omega)$ and $\hat{g}_1^-(\omega) \approx 0$ for $0 < \omega < \delta$. Then we have

$$\phi_g |0, \mathcal{S}\rangle \approx a_{\hat{f}_1}^\dagger |0, \mathcal{S}\rangle. \quad (6.100)$$

Since $\hat{f}_1(\omega)$ was a general smearing function, we know $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense in \mathcal{H}_1 and the hypothesis is true for $n = 1$. Tracing the exact same inductive steps as the above case, we obtain that $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ is dense for all odd n .

This concludes the proof for denseness of $\mathcal{A}_{\delta,\theta} |0, \mathcal{S}\rangle$ in \mathcal{H} , and hence also the proof for the existence of a Q_i localised in region C and satisfying both Eqs. (6.96) and (6.97).

6.10.1 Positive Fourier mode reconstruction

Lemma: Given $\delta, \epsilon > 0$, the space of $L^2\mathbb{R}$ functions with support in $[0, \epsilon]$ is dense in $L^2\mathbb{R}$ under the norm defined by $\|f\|_\delta^2 = \int_{-\delta}^\infty d\omega |\hat{f}(\omega)|^2$, where $\hat{f}(\omega) = \int_{\mathbb{R}} \frac{dt}{2\pi} f(t) e^{i\omega t}$. More explicitly, given a function $f \in L^2\mathbb{R}$, and $\delta, \epsilon, w > 0$, there exists a function $g \in L^2[0, \epsilon]$ such that

$$r \equiv \int_{-\delta}^\infty d\omega |\hat{f}(\omega) - \hat{g}(\omega)|^2 < w. \quad (6.101)$$

Proof: Let P_ϵ be the projector onto the space of all functions supported in $[0, \epsilon]$, and $P_{-\delta}$ be the projector onto the space of all functions which contain no Fourier modes in the range $(-\infty, -\delta)$. The quantity in question, $r = \int_{-\delta}^\infty d\omega |\hat{f}(\omega) - \hat{g}(\omega)|^2$ is manifestly equal to $\frac{1}{2\pi} \|P_{-\delta}f - P_{-\delta}g\|^2$, where $\|\cdot\|$ is the standard L^2 norm. Since, we need to show the existence of a $g \in P_\epsilon L^2\mathbb{R}$ such that r can be made arbitrarily small, it is enough to show that the subspace $P_{-\delta}P_\epsilon L^2\mathbb{R}$ is dense in $P_{-\delta}L^2\mathbb{R}$. Let $C_C^\infty(\mathbb{R})$ be the subspace of all smooth functions with compact support in $L^2\mathbb{R}$. This subspace happens to be dense in $L^2\mathbb{R}$. Let us first show the denseness of $P_{-\delta}P_\epsilon L^2\mathbb{R}$ in the subspace $P_{-\delta}C_C^\infty(\mathbb{R})$. We shall show this by contradiction.

Let $P_{-\delta}P_\epsilon L^2\mathbb{R}$ not be dense in $P_{-\delta}C_C^\infty(\mathbb{R})$. Then there exists a non-zero function

$\chi \in P_{-\delta}C_C^\infty(\mathbb{R})$ such that $(\phi, \chi) = \int_{\mathbb{R}} dt \phi^*(t)\chi(t) = 0$ for all $\phi \in P_{-\delta}P_\epsilon L^2\mathbb{R}$. So,

$$\begin{aligned}
& (P_{-\delta}P_\epsilon\psi, \chi) = 0 \quad \forall \psi \in L^2\mathbb{R} \\
\Rightarrow & (\psi, P_\epsilon P_{-\delta}\chi) = 0 \quad \because \text{both projectors are hermitian} \\
\Rightarrow & (\psi, P_\epsilon\chi) = 0 \quad \because P_{-\delta}\chi = \chi \\
& \Rightarrow P_\epsilon\chi = 0 \quad \because \psi \text{ is arbitrary} \\
\Rightarrow & \chi(t) = 0 \quad \forall t \in [0, \epsilon].
\end{aligned} \tag{6.102}$$

Now, since $\chi(t)$ is a smooth function on \mathbb{R} , identically vanishing over the interval $[0, \epsilon]$ means while we Taylor expand it around some point in this interval, say 0, all the Taylor coefficients turn out to be 0. $\chi(t)$ therefore vanishes identically all throughout the real line. This is in contradiction to the hypothesis that $\chi(t)$ is non-zero. Hence we have shown that $P_{-\delta}P_\epsilon L^2\mathbb{R}$ is dense in $P_{-\delta}C_C^\infty(\mathbb{R})$. On the other hand, $P_{-\delta}C_C^\infty(\mathbb{R})$ is dense in $P_{-\delta}L^2\mathbb{R}$ because $C_C^\infty(\mathbb{R})$ is dense in $L^2\mathbb{R}$. Hence, by transitivity of denseness of topological spaces, we have shown $P_{-\delta}P_\epsilon L^2\mathbb{R}$ is dense in $P_{-\delta}L^2\mathbb{R}$.

Corollary: Given a function $f \in L^2\mathbb{R}$, and $\delta, \epsilon, w > 0$, there exists a function $g \in L^2[0, \epsilon]$ such that

$$\int_{-\delta}^0 d\omega |\hat{g}(\omega)|^2 + \int_0^\infty d\omega |\hat{f}(\omega) - \hat{g}(\omega)|^2 < w. \tag{6.103}$$

In other words, for any function $f(t)$, there exists a $g(t)$ supported in $[0, \epsilon]$ such that it approximates $f(t)$ in the positive Fourier modes with its modes in the $[-\delta, 0)$ range suppressed to arbitrary precision.

Proof: Given $f \in L^2\mathbb{R}$, construct $f_1 \in L^2\mathbb{R}$ by deleting its Fourier modes in the $[-\delta, 0)$ range. That is

$$\hat{f}_1(\omega) = \begin{cases} 0 & \text{if } \omega \in [-\delta, 0) \\ \hat{f}(\omega) & \text{if } \omega \notin [-\delta, 0) \end{cases}. \tag{6.104}$$

Now applying the above lemma to $f_1(t)$ instead of $f(t)$ proves the existence of the desired $g(t)$.

Bibliography

- [1] J. Chakravarty, *Overcounting of interior excitations: A resolution to the bags of gold paradox in AdS*, *JHEP* **02** (2021) 027, [[arXiv:2010.03575](#)].
- [2] T. Chakraborty, J. Chakravarty, and P. Paul, *Monogamy paradox: A toy model in flat space*, [arXiv:2107.06919](#).
- [3] B. Chakraborty, J. Chakravarty, S. Chaudhuri, C. Jana, R. Loganayagam, and A. Sivakumar, *Nonlinear Langevin dynamics via holography*, *JHEP* **01** (2020) 165, [[arXiv:1906.07762](#)].
- [4] J. Chakravarty and D. Jain, *Critical exponents for higher order phase transitions: Landau theory and RG flow*, *J. Stat. Mech.* **2109** (2021) 093204, [[arXiv:2102.08398](#)].
- [5] J. Chakravarty, *Small corrections to semiclassical gravity and their role in information paradoxes*, [arXiv:2105.09924](#).
- [6] J. Chakravarty and A. Sen, *Normalization of D instanton amplitudes in two dimensional type 0B string theory*, [arXiv:2207.07138](#).
- [7] J. Chakravarty, P. Maity, and A. Mishra, *On the positivity of Coon amplitude in $D=4$* , [arXiv:2208.02735](#).
- [8] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113–1133, [[hep-th/9711200](#)].
- [9] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253–291, [[hep-th/9802150](#)].
- [10] S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys. Lett. B* **428** (1998) 105–114, [[hep-th/9802109](#)].
- [11] S. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199–220. [Erratum: *Commun.Math.Phys.* 46, 206 (1976)].

- [12] S. D. Mathur, *The Information paradox: A Pedagogical introduction*, *Class. Quant. Grav.* **26** (2009) 224001, [[arXiv:0909.1038](#)].
- [13] S. D. Mathur, *The information paradox: conflicts and resolutions*, *Pramana* **79** (2012) 1059–1073, [[arXiv:1201.2079](#)].
- [14] L. Susskind, L. Thorlacius, and J. Uglum, *The Stretched horizon and black hole complementarity*, *Phys. Rev. D* **48** (1993) 3743–3761, [[hep-th/9306069](#)].
- [15] L. Susskind and L. Thorlacius, *Gedanken experiments involving black holes*, *Phys. Rev. D* **49** (1994) 966–974, [[hep-th/9308100](#)].
- [16] D. N. Page, *Information in black hole radiation*, *Phys. Rev. Lett.* **71** (1993) 3743–3746, [[hep-th/9306083](#)].
- [17] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys. Lett. B* **379** (1996) 99–104, [[hep-th/9601029](#)].
- [18] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323–2329, [[arXiv:1005.3035](#)].
- [19] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *JHEP* **09** (2007) 120, [[arXiv:0708.4025](#)].
- [20] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, *An Apologia for Firewalls*, *JHEP* **09** (2013) 018, [[arXiv:1304.6483](#)].
- [21] E. Verlinde and H. Verlinde, *Black Hole Entanglement and Quantum Error Correction*, *JHEP* **10** (2013) 107, [[arXiv:1211.6913](#)].
- [22] S. H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067, [[arXiv:1306.0622](#)].
- [23] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781–811, [[arXiv:1306.0533](#)].
- [24] D. L. Jafferis, *Bulk reconstruction and the Hartle-Hawking wavefunction*, [arXiv:1703.01519](#).
- [25] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, [arXiv:1905.08255](#).
- [26] A. Almheiri, R. Mahajan, J. Maldacena, and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149, [[arXiv:1908.10996](#)].

- [27] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013, [[arXiv:1911.12333](#)].
- [28] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, *Replica wormholes and the black hole interior*, [arXiv:1911.11977](#).
- [29] K. Papadodimas and S. Raju, *An Infalling Observer in AdS/CFT*, *JHEP* **10** (2013) 212, [[arXiv:1211.6767](#)].
- [30] K. Papadodimas and S. Raju, *State-Dependent Bulk-Boundary Maps and Black Hole Complementarity*, *Phys. Rev. D* **89** (2014), no. 8 086010, [[arXiv:1310.6335](#)].
- [31] K. Papadodimas and S. Raju, *Black Hole Interior in the Holographic Correspondence and the Information Paradox*, *Phys. Rev. Lett.* **112** (2014), no. 5 051301, [[arXiv:1310.6334](#)].
- [32] K. Papadodimas and S. Raju, *Local Operators in the Eternal Black Hole*, *Phys. Rev. Lett.* **115** (2015), no. 21 211601, [[arXiv:1502.06692](#)].
- [33] K. Papadodimas and S. Raju, *Remarks on the necessity and implications of state-dependence in the black hole interior*, *Phys. Rev. D* **93** (2016), no. 8 084049, [[arXiv:1503.08825](#)].
- [34] J. Wheeler, in *Relativity, Groups and Fields*, edited by B.S. DeWitt and C.M. DeWitt, Gordon and Breach, New York (1964).
- [35] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062, [[arXiv:1207.3123](#)].
- [36] J. Chakravarty, *Overcounting of interior excitations: A resolution to the bags of gold paradox in AdS*, *JHEP* **02** (2021) 027, [[arXiv:2010.03575](#)].
- [37] T. Chakraborty, J. Chakravarty, and P. Paul, *Monogamy paradox: A toy model in flat space*, [arXiv:2107.06919](#).
- [38] J. D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (Apr, 1973) 2333–2346.
- [39] S. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199–220.
- [40] J. M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021, [[hep-th/0106112](#)].

- [41] O. Bohigas, M. J. Giannoni, and C. Schmit, *Characterization of chaotic quantum spectra and universality of level fluctuation laws*, *Phys. Rev. Lett.* **52** (Jan, 1984) 1–4.
- [42] M. Mehta, *Random Matrices*. ISSN. Elsevier Science, 2004.
- [43] F. J. Dyson, *Statistical theory of the energy levels of complex systems. i*, *Journal of Mathematical Physics* **3** (1962), no. 1 140–156, [<https://doi.org/10.1063/1.1703773>].
- [44] F. J. Dyson, *Statistical theory of the energy levels of complex systems. iii*, *Journal of Mathematical Physics* **3** (1962), no. 1 166–175, [<https://doi.org/10.1063/1.1703775>].
- [45] M. Mehta, *On the statistical properties of the level-spacings in nuclear spectra*, *Nuclear Physics* **18** (1960) 395 – 419.
- [46] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065, [[arXiv:0808.2096](https://arxiv.org/abs/0808.2096)].
- [47] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, *Towards the Fast Scrambling Conjecture*, *JHEP* **04** (2013) 022, [[arXiv:1111.6580](https://arxiv.org/abs/1111.6580)].
- [48] J. Maldacena, S. H. Shenker, and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106, [[arXiv:1503.01409](https://arxiv.org/abs/1503.01409)].
- [49] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, *Black Holes and Random Matrices*, *JHEP* **05** (2017) 118, [[arXiv:1611.04650](https://arxiv.org/abs/1611.04650)]. [Erratum: *JHEP* 09, 002 (2018)].
- [50] A. V. Andreev and B. L. Altshuler, *Spectral statistics beyond random matrix theory*, *Phys. Rev. Lett.* **75** (Jul, 1995) 902–905.
- [51] E. Brézin and S. Hikami, *Spectral form factor in a random matrix theory*, *Phys. Rev. E* **55** (Apr, 1997) 4067–4083.
- [52] E. Brézin and S. Hikami, *Extension of level-spacing universality*, *Phys. Rev. E* **56** (Jul, 1997) 264–269.
- [53] D. Marolf, *Black Holes, AdS, and CFTs*, *Gen. Rel. Grav.* **41** (2009) 903–917, [[arXiv:0810.4886](https://arxiv.org/abs/0810.4886)].
- [54] S. D. Hsu and D. Reeb, *Monsters, black holes and the statistical mechanics of gravity*, *Mod. Phys. Lett. A* **24** (2009) 1875–1887, [[arXiv:0908.1265](https://arxiv.org/abs/0908.1265)].

- [55] B. Freivogel, V. E. Hubeny, A. Maloney, R. C. Myers, M. Rangamani, and S. Shenker, *Inflation in AdS/CFT*, *JHEP* **03** (2006) 007, [[hep-th/0510046](#)].
- [56] Z. Fu and D. Marolf, *Bag-of-gold spacetimes, Euclidean wormholes, and inflation from domain walls in AdS/CFT*, *JHEP* **11** (2019) 040, [[arXiv:1909.02505](#)].
- [57] Y. C. Ong and P. Chen, *The Fate of Monsters in Anti-de Sitter Spacetime*, *JHEP* **07** (2013) 147, [[arXiv:1304.3803](#)].
- [58] K. Langhoff and Y. Nomura, *Ensemble from Coarse Graining: Reconstructing the Interior of an Evaporating Black Hole*, *Phys. Rev. D* **102** (2020), no. 8 086021, [[arXiv:2008.04202](#)].
- [59] Y. Nomura, *From the Black Hole Conundrum to the Structure of Quantum Gravity*, [arXiv:2011.08707](#).
- [60] S. Raju, *A Toy Model of the Information Paradox in Empty Space*, *SciPost Phys.* **6** (2019), no. 6 073, [[arXiv:1809.10154](#)].
- [61] R. Bousso, *Complementarity Is Not Enough*, *Phys. Rev. D* **87** (2013), no. 12 124023, [[arXiv:1207.5192](#)].
- [62] L. Susskind, *Computational Complexity and Black Hole Horizons*, *Fortsch. Phys.* **64** (2016) 24–43, [[arXiv:1403.5695](#)]. [Addendum: *Fortsch.Phys.* 64, 44–48 (2016)].
- [63] S. D. Mathur and D. Turton, *Comments on black holes I: The possibility of complementarity*, *JHEP* **01** (2014) 034, [[arXiv:1208.2005](#)].
- [64] Y. Nomura, J. Varela, and S. J. Weinberg, *Complementarity Endures: No Firewall for an Infalling Observer*, *JHEP* **03** (2013) 059, [[arXiv:1207.6626](#)].
- [65] Y. Nomura, J. Varela, and S. J. Weinberg, *Black Holes, Information, and Hilbert Space for Quantum Gravity*, *Phys. Rev. D* **87** (2013) 084050, [[arXiv:1210.6348](#)].
- [66] E. Verlinde and H. Verlinde, *Passing through the Firewall*, [arXiv:1306.0515](#).
- [67] A. Giveon and N. Itzhaki, *String Theory Versus Black Hole Complementarity*, *JHEP* **12** (2012) 094, [[arXiv:1208.3930](#)].
- [68] M. Van Raamsdonk, *Evaporating Firewalls*, *JHEP* **11** (2014) 038, [[arXiv:1307.1796](#)].
- [69] J. Hutchinson and D. Stojkovic, *Icezones instead of firewalls: extended entanglement beyond the event horizon and unitary evaporation of a black hole*, *Class. Quant. Grav.* **33** (2016), no. 13 135006, [[arXiv:1307.5861](#)].

- [70] A. Karlsson, *A paradox regarding monogamy of entanglement*, [arXiv:1911.09226](#).
- [71] A. Karlsson, *Replica wormhole and island incompatibility with monogamy of entanglement*, [arXiv:2007.10523](#).
- [72] K. L. H. Bryan and A. J. M. Medved, *Black holes and information: A new take on an old paradox*, *Adv. High Energy Phys.* **2017** (2017) 7578462, [[arXiv:1603.07569](#)].
- [73] Y. Nomura, *Reanalyzing an Evaporating Black Hole*, *Phys. Rev. D* **99** (2019), no. 8 086004, [[arXiv:1810.09453](#)].
- [74] L. Susskind, *The Transfer of Entanglement: The Case for Firewalls*, [arXiv:1210.2098](#).
- [75] B. Yoshida, *Firewalls vs. Scrambling*, *JHEP* **10** (2019) 132, [[arXiv:1902.09763](#)].
- [76] J. Polchinski, *The Black Hole Information Problem*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, 9, 2016. [arXiv:1609.04036](#).
- [77] R. Bousso and M. Tomašević, *Unitarity From a Smooth Horizon?*, *Phys. Rev. D* **102** (2020), no. 10 106019, [[arXiv:1911.06305](#)].
- [78] P. Chen, Y. C. Ong, D. N. Page, M. Sasaki, and D.-h. Yeom, *Naked Black Hole Firewalls*, *Phys. Rev. Lett.* **116** (2016), no. 16 161304, [[arXiv:1511.05695](#)].
- [79] A. Laddha, S. G. Prabhu, S. Raju, and P. Shrivastava, *The Holographic Nature of Null Infinity*, [arXiv:2002.02448](#).
- [80] S. Raju, *Lessons from the Information Paradox*, [arXiv:2012.05770](#).
- [81] C. Chowdhury, O. Papadoulaki, and S. Raju, *A physical protocol for observers near the boundary to obtain bulk information in quantum gravity*, [arXiv:2008.01740](#).
- [82] G. 't Hooft, *On the quantum structure of a black hole*, *Nuclear Physics B* **256** (1985) 727–745.
- [83] Y. Kiem, H. L. Verlinde, and E. P. Verlinde, *Black hole horizons and complementarity*, *Phys. Rev. D* **52** (1995) 7053–7065, [[hep-th/9502074](#)].
- [84] S. Banerjee, J.-W. Bryan, K. Papadodimas, and S. Raju, *A toy model of black hole complementarity*, *JHEP* **05** (2016) 004, [[arXiv:1603.02812](#)].

- [85] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, *Phys. Rev. Lett.* **23** (Oct, 1969) 880–884.
- [86] J. S. Bell, *On the einstein podolsky rosen paradox*, *Physics Physique Fizika* **1** (Nov, 1964) 195–200.
- [87] B. Toner and F. Verstraete, *Monogamy of Bell correlations and Tsirelson’s bound*, *arXiv e-prints* (Nov., 2006) quant-ph/0611001, [[quant-ph/0611001](#)].
- [88] V. Scarani and N. Gisin, *Quantum communication between N partners and bell’s inequalities*, *Phys. Rev. Lett.* **87** (Aug, 2001) 117901.
- [89] T. Regge and C. Teitelboim, *Role of surface integrals in the hamiltonian formulation of general relativity*, *Annals of Physics* **88** (1974), no. 1 286–318.
- [90] A. Ashtekar, *Asymptotic quantization of the gravitational field*, *Phys. Rev. Lett.* **46** (Mar, 1981) 573–576.
- [91] A. Ashtekar, *Radiative Degrees of Freedom of the Gravitational Field in Exact General Relativity*, *J. Math. Phys.* **22** (1981) 2885–2895.
- [92] A. Ashtekar, *asymptotic quantization: based on 1984 naples lectures*, *Bibliopolis* (1987).
- [93] A. Ashtekar, M. Campiglia, and A. Laddha, *Null infinity, the BMS group and infrared issues*, *Gen. Rel. Grav.* **50** (2018), no. 11 140–163, [[arXiv:1808.07093](#)].
- [94] J. L. F. Barbon and E. Rabinovici, *Long time scales and eternal black holes*, *NATO Sci. Ser. II* **208** (2006) 255–263, [[hep-th/0403268](#)].
- [95] T. Anous, T. Hartman, A. Rovai, and J. Sonner, *Black Hole Collapse in the $1/c$ Expansion*, *JHEP* **07** (2016) 123, [[arXiv:1603.04856](#)].
- [96] A. L. Fitzpatrick and J. Kaplan, *On the Late-Time Behavior of Virasoro Blocks and a Classification of Semiclassical Saddles*, *JHEP* **04** (2017) 072, [[arXiv:1609.07153](#)].
- [97] R. M. Wald, *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.
- [98] K. Papadodimas, *A class of non-equilibrium states and the black hole interior*, [arXiv:1708.06328](#).
- [99] W. Israel, *Thermo field dynamics of black holes*, *Phys. Lett. A* **57** (1976) 107–110.
- [100] Y. Takahashi and H. Umezawa, *Thermo field dynamics*, *Collect. Phenom.* **2** (1975) 55–80.

- [101] J. Hartle and S. Hawking, *Path Integral Derivation of Black Hole Radiance*, *Phys. Rev. D* **13** (1976) 2188–2203.
- [102] J. Hartle and S. Hawking, *Wave Function of the Universe*, *Adv. Ser. Astrophys. Cosmol.* **3** (1987) 174–189.
- [103] S. Lloyd, *Pure state quantum statistical mechanics and black holes*, [arXiv:1307.0378](https://arxiv.org/abs/1307.0378).
- [104] G. Casati, B. V. Chirikov, and I. Guarneri, *Energy-Level Statistics of Integrable Quantum Systems*, *Phys. Rev. Lett.* **54** (Apr, 1985) 1350–1353.
- [105] H. Tasaki, *From quantum dynamics to the canonical distribution: General picture and a rigorous example*, *Phys. Rev. Lett.* **80** (Feb, 1998) 1373–1376.
- [106] D. Stanford and L. Susskind, *Complexity and Shock Wave Geometries*, *Phys. Rev. D* **90** (2014), no. 12 126007, [[arXiv:1406.2678](https://arxiv.org/abs/1406.2678)].
- [107] D. Carmi, S. Chapman, H. Marrochio, R. C. Myers, and S. Sugishita, *On the Time Dependence of Holographic Complexity*, *JHEP* **11** (2017) 188, [[arXiv:1709.10184](https://arxiv.org/abs/1709.10184)].
- [108] L. Susskind, *Three Lectures on Complexity and Black Holes*, 10, 2018. [arXiv:1810.11563](https://arxiv.org/abs/1810.11563).
- [109] J. Chakravarty and S. Raju, *How do black holes manage to look larger from the inside than the outside?*, *Unpublished Essay* (2020).
- [110] L. Motl, *One can't background-independently localize field operators in QG*, <https://motls.blogspot.com/2013/08/one-cant-background-independently.html> (2013).
- [111] A. Giveon and N. Itzhaki, *Stringy Information and Black Holes*, [arXiv:1912.06538](https://arxiv.org/abs/1912.06538).
- [112] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602, [[hep-th/0603001](https://arxiv.org/abs/hep-th/0603001)].
- [113] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045, [[hep-th/0605073](https://arxiv.org/abs/hep-th/0605073)].
- [114] V. E. Hubeny, M. Rangamani, and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062, [[arXiv:0705.0016](https://arxiv.org/abs/0705.0016)].
- [115] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090, [[arXiv:1304.4926](https://arxiv.org/abs/1304.4926)].

- [116] T. Faulkner, A. Lewkowycz, and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074, [[arXiv:1307.2892](#)].
- [117] N. Engelhardt and A. C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073, [[arXiv:1408.3203](#)].
- [118] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, *Local bulk operators in AdS/CFT: A Boundary view of horizons and locality*, *Phys. Rev. D* **73** (2006) 086003, [[hep-th/0506118](#)].
- [119] M. V. Berry and M. Tabor, *Level clustering in the regular spectrum*, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **356** (1977), no. 1686 375–394.
- [120] A. Milekhin, *Quantum error correction and large N* , [arXiv:2008.12869](#).
- [121] O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, and M. Van Raamsdonk, *The Hagedorn - deconfinement phase transition in weakly coupled large N gauge theories*, *Adv. Theor. Math. Phys.* **8** (2004) 603–696, [[hep-th/0310285](#)].
- [122] B. Sundborg, *The Hagedorn transition, deconfinement and $N=4$ SYM theory*, *Nucl. Phys. B* **573** (2000) 349–363, [[hep-th/9908001](#)].
- [123] E. Witten, *Anti-de Sitter space, thermal phase transition, and confinement in gauge theories*, *Adv. Theor. Math. Phys.* **2** (1998) 505–532, [[hep-th/9803131](#)].
- [124] S. Bose, *Quantum communication through an unmodulated spin chain*, *Phys. Rev. Lett.* **91** (Nov, 2003) 207901.
- [125] M. Christandl, N. Datta, T. C. Dorlas, A. Ekert, A. Kay, and A. J. Landahl, *Perfect transfer of arbitrary states in quantum spin networks*, *Phys. Rev. A* **71** (Mar, 2005) 032312.
- [126] A. KAY, *Perfect, efficient, state transfer and its application as a constructive tool*, *International Journal of Quantum Information* **08** (2010), no. 04 641–676, [<https://doi.org/10.1142/S0219749910006514>].
- [127] D. Campo and R. Parentani, *Inflationary spectra and violations of Bell inequalities*, *Phys. Rev. D* **74** (2006) 025001, [[astro-ph/0505376](#)].
- [128] S. Summers and R. Werner, *Bell's inequalities and quantum field theory. I. General setting*, *Journal of Mathematical Physics* **28** (1987) 2440.

- [129] K. Banaszek and K. Wódkiewicz, *Nonlocality of the einstein-podolsky-rosen state in the wigner representation*, *Phys. Rev. A* **58** (Dec, 1998) 4345–4347.
- [130] H. Bondi, *Gravitational Waves in General Relativity*, *Nature* **186** (1960), no. 4724 535–535.
- [131] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, *Gravitational waves in general relativity. 7. Waves from axisymmetric isolated systems*, *Proc. Roy. Soc. Lond. A* **269** (1962) 21–52.
- [132] R. Sachs, *Asymptotic symmetries in gravitational theory*, *Phys. Rev.* **128** (1962) 2851–2864.
- [133] R. K. Sachs, *Gravitational waves in general relativity. 8. Waves in asymptotically flat space-times*, *Proc. Roy. Soc. Lond. A* **270** (1962) 103–126.
- [134] A. Strominger, *Lectures on the Infrared Structure of Gravity and Gauge Theory*, [arXiv:1703.05448](https://arxiv.org/abs/1703.05448).
- [135] G. Compère and A. Fiorucci, *Advanced Lectures on General Relativity*, [arXiv:1801.07064](https://arxiv.org/abs/1801.07064).
- [136] R. F. Streater and A. S. Wightman, *PCT, spin and statistics, and all that*. 1989.
- [137] R. Haag, *Local quantum physics: Fields, particles, algebras*. 1992.
- [138] J. C. Baez, I. E. Segal, and Z. Zhou, *Introduction to Algebraic and Constructive Quantum Field Theory*. Princeton University Press, 2014.
- [139] A. JAFFE, *CONSTRUCTIVE QUANTUM FIELD THEORY*, pp. 111–127. 2000. https://www.worldscientific.com/doi/pdf/10.1142/9781848160224_0007.
- [140] A. Ashtekar and A. Magnon-Ashtekar, *Energy-momentum in general relativity*, *Phys. Rev. Lett.* **43** (Jul, 1979) 181–184.
- [141] R. L. Arnowitt, S. Deser, and C. W. Misner, *The Dynamics of general relativity*, *Gen. Rel. Grav.* **40** (2008) 1997–2027, [[gr-qc/0405109](https://arxiv.org/abs/gr-qc/0405109)].
- [142] H. Reeh and S. Schlieder, *Bemerkungen zur unitäräquivalenz von lorentzinvarianten feldern*, *Nuovo Cim.* **22** (1961), no. 5 1051–1068.
- [143] E. Witten, *APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory*, *Rev. Mod. Phys.* **90** (2018), no. 4 045003, [[arXiv:1803.04993](https://arxiv.org/abs/1803.04993)].

- [144] D. Kapec, V. Lysov, S. Pasterski, and A. Strominger, *Higher-dimensional supertranslations and Weinberg's soft graviton theorem*, *Ann. Math. Sci. Appl.* **02** (2017) 69–94, [[arXiv:1502.07644](#)].
- [145] S. Hollands, A. Ishibashi, and R. M. Wald, *BMS Supertranslations and Memory in Four and Higher Dimensions*, *Class. Quant. Grav.* **34** (2017), no. 15 155005, [[arXiv:1612.03290](#)].
- [146] A. Aggarwal, *Supertranslations in Higher Dimensions Revisited*, *Phys. Rev. D* **99** (2019), no. 2 026015, [[arXiv:1811.00093](#)].
- [147] M. Campiglia and L. Coito, *Asymptotic charges from soft scalars in even dimensions*, *Phys. Rev. D* **97** (2018), no. 6 066009, [[arXiv:1711.05773](#)].
- [148] T. He and P. Mitra, *Asymptotic symmetries in $(d + 2)$ -dimensional gauge theories*, *JHEP* **10** (2019) 277, [[arXiv:1903.03607](#)].
- [149] M. Campiglia and A. Laddha, *Sub-subleading soft gravitons: New symmetries of quantum gravity?*, *Phys. Lett. B* **764** (2017) 218–221, [[arXiv:1605.09094](#)].
- [150] B. Sahoo and A. Sen, *Classical and Quantum Results on Logarithmic Terms in the Soft Theorem in Four Dimensions*, *JHEP* **02** (2019) 086, [[arXiv:1808.03288](#)].
- [151] S. Ghosh and S. Raju, *Loss of locality in gravitational correlators with a large number of insertions*, *Phys. Rev. D* **96** (2017), no. 6 066033, [[arXiv:1706.07424](#)].
- [152] G. Akemann, J. Baik, and P. D. Francesco, *The Oxford Handbook of Random Matrix Theory*. Oxford University Press, New York, 2011.
- [153] D. Gross and E. Witten, *Possible Third Order Phase Transition in the Large N Lattice Gauge Theory*, *Phys. Rev. D* **21** (1980) 446–453.
- [154] S. R. Wadia, *$N = \infty$ phase transition in a class of exactly soluble model lattice gauge theories*, *Physics Letters B* **93** (1980), no. 4 403 – 410.
- [155] S. R. Wadia, *A study of $u(n)$ lattice gauge theory in 2-dimensions*, [arXiv:1212.2906](#).
- [156] T. Suzuki, S. Tameike, and E. Yamada, *Some Remarks on Null-Plane Quantization*, *Progress of Theoretical Physics* **55** (03, 1976) 922–928.
- [157] L. S. Brown, *Quantum Field Theory*. Cambridge University Press, 1992.