Title
Aspects of Generalized Entropy And Quantum Null Energy Condition

Author
Shahbazi-Moghaddam, Arvin

Publication Date
2020

Peer reviewed|Thesis/dissertation

Aspects of Generalized Entropy And Quantum Null Energy Condition

by

Arvin Shahbazi-Moghaddam

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Raphael Bousso, Chair
Professor Petr Hořava
Professor Nicolai Reshetikhin

Summer 2020

Aspects of Generalized Entropy And Quantum Null Energy Condition

Abstract

Aspects of Generalized Entropy And Quantum Null Energy Condition

by

Arvin Shahbazi-Moghaddam

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Raphael Bousso, Chair

There is a deep connection between entanglement and geometry in quantum gravity. One manifestation of this connection is through the generalized entropy: the sum of the area of Cauchy-dividing surfaces in Planck units and the von Neumann entropy of the fields on one side of the surface. The generalized entropy is expected to satisfy fundamental thermodynamical conditions which in certain limits result in quantum field theory (QFT) inequalities relating the energy density and the shape derivatives of the von Neumann entropy, most notably the quantum null energy condition (QNEC). This dissertation is devoted to studying various aspect of the generalized entropy, the QNEC, and other information-theoretic aspects of QFT. First, we will use the QNEC to demonstrate the locality of the logarithm of the vacuum density matrix in certain regions of any holographic conformal field theory (CFT). We will then holographically prove the QNEC in general curved spacetime of less than six spacetime dimensions. Next, we will demonstrate a certain limit of the QNEC where an *equality* between certain components of the stress tensor and the second shape derivative of von Neumann entropy emerges. We will then focus on the study of generalized entropy. Within the AdS/CFT framework, we find a microscopic description for the generalized entropy of black holes using a geometric construction. We then apply this construction to extremal surface in AdS, which constitutes the bulk dual of the Connes cocycle flow on the boundary CFT. Finally, we use propose a novel lower bound on the total energy of spacetime involving the generalized entropy on certain light-sheets.

To my parents,
Afsaneh and Ario,

and

to my sister,
Arezou

Contents

List of Figures

Acknowledgments

First off, I am thankful to my advisor Raphael Bousso whose support, guidance, and encouragement throughout grad school have been essential in my academic pursuit. He went above and beyond what is expected of an advisor to provide me with valuable resources and opportunities. I also appreciate the freedom he gave me to explore my own ideas.

I am extremely indebted to Stefan Leichenauer whose guidance and collaboration helped me find my way in graduate school. Having the privilege to work with him and observe his unique style of doing physics have shaped me professionally.

I would like to thank Tom Faulkner for his collaboration and our many discussions. What I have learned from him has significantly influenced my research interests going forward.

I am grateful to Aron Wall and Tom Hartman who have inspired me with their brilliant minds and taught me a great deal of physics.

Throughout grad school, I have benefitted immensely from continuous discussions with many peers. I would like to especially thank my friends and collaborators Adam Levine, Ven Chandrasekaran, Pratik Rath, Chris Akers, Illan Halpern, Jason Koeller, and Mudassir Moosa. Many ideas in this work were born out of these discussions. I would also like to thank my friend and colleague Hrant Gharibyan with whom I have been discussing physics since the early days of undergrad. Furthermore, I am grateful to Ning Bao and Grant Remmen for essential advice, and collaboration.

I am thankful to my partner Sandra Oseguera whose loving support and friendship made grad school a lot more exciting and enjoyable. I feel so lucky to have such an amazing person in my life.

I would also like to thank Marlou Belyea and Zak Taylor for their warmth and generosity during my years at Berkeley.

Last but not least, I am grateful to Danial Binayi who since middle school has been an incredible friend.

Chapter 1

Introduction

In the past twenty years, quantum gravity research has been dominated by the AdS/CFT correspondence [1, 2]. AdS/CFT states that certain quantum gravity theories in asymptotically $(d+1)$-dimensional Anti-de Sitter (AdS) spacetime are dual to certain conformal field theories (CFT) living on the d-dimensional asymptotic boundary of AdS. More recently, powered by AdS/CFT, there have been exciting developments connecting concepts from quantum information theory like entanglement and quantum error correction to geometry and gravity [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. A central idea behind these new developments is that black holes have entropy proportional to their area [13]:

$$S_{BH} = \frac{\mathcal{A}}{4G\hbar} \tag{1.0.1}$$

where $\mathcal{A}$ is the area of the black hole horizon, G is Newton's constant, and $\hbar$ is Planck's constant.

This idea has been refined by the AdS/CFT correspondence where it has been shown that the von Neumann entropy of certain subregions in the CFT can be computed by the area of the minimal extremal surface in the bulk anchored to the boundary subregion [3]:

$$S_{\text{CFT}} = \frac{\mathcal{A}_{\text{ext}}}{4G\hbar} + \mathcal{O}((G\hbar)^0) \tag{1.0.2}$$

where we work with small $G\hbar$ in the semiclassical bulk expansion. This provides areas with a microscopic information-theoretic interpretation. Furthermore, it is known that the higher order corrections to this formula are given by adding the von Neumann entropy of bulk fields [5, 6]:

$$S_{\text{CFT}} = \frac{\mathcal{A}_{\text{ext}}}{4G\hbar} + S_{\text{bulk}} \tag{1.0.3}$$

The quantity on the RHS above is called the generalized entropy and can be defined for any Cauchy-splitting surface in semiclassical quantum gravity. An interesting feature is that the

divergences in the bulk von-Neumann are absorbed into the renormalization of G, resulting in a finite quantity at the end [14, 11].

It was Bekenstein who realized that the generalized entropy behaves like thermodynamic entropy in quantum gravity [13] and in particular, satisfies the generalized second law (GSL) of thermodynamics. The GSL was proven in [15] where it was shown to follow from the monotonicity of relative entropy in quantum field theory (QFT). This suggests a deep connection between quantum gravity and QFT. The intuition behind this connection is simple: spacetime curvature is related to energy by Einstein field equations, so one can re-interpret statements about the dynamics of generalized entropy as relations between energy and von Neumann entropy in the $G \to 0$ limit.

A recent example of such a connection is given by the quantum focusing conjecture (QFC) [11] which is a statement from which the GSL and other important bounds in quantum gravity follow. The non-gravitational limit of the QFC is the quantum null energy condition (QNEC): a lower bound on certain components of the stress tensor given by the second shape derivative of the von Neumann entropy. This is a novel energy condition predicted by gravity that has since been laboriously proven within QFT [16, 17, 18], bringing together ideas from the conformal bootstrap, quantum information, and algebraic QFT. Therefore, studying spacetime thermodynamics could provide us with shortcuts to learning deep facts in QFT as well as new insights into quantum gravity.

This dissertation is devoted to further studying the generalized entropy, the QNEC, and other information-theoretic aspects of QFT. An important information theoretic quantity that is well-motivated in QFT is the modular Hamiltonian, roughly the logarithm of a given density matrix. The locality of the vacuum modular Hamiltonian is an essential ingredient in connecting horizon physics in quantum gravity to statements about QFT, such as the averaged null energy condition (ANEC) and the monotonicity of relative entropy. In Chapter 2, we show that the vacuum modular Hamiltonian of a region bounded by cuts of Killing horizons is a local integral of the stress tensor if the QNEC is saturated in the vacuum. We then show that for holographic CFTs, the QNEC is indeed saturated in the vacuum, establishing the locality of the vacuum modular Hamiltonian for such CFTs.

In Chapter 3, following [19], we study the consequences of Entanglement Wedge Nesting, an expected condition in general relativity, for holographical CFTs that live on arbitrary curved backgrounds, and derive necessary and sufficient conditions for it to imply the QNEC in $d \leq 5$. We also show that these are the exact same conditions under which the QFT yields the QNEC.

In Chapter 4, we compute the local second variation of the von Neumann entropy of regions in holographic CFTs. For null variations our formula says that the narrow transverse width limit of the QNEC is saturated in every state in $d > 2$, thus providing an equivalence between energy and entropy. In Chapter 5, we demonstrate evidence for the saturation of the QNEC in all QFTs with interacting UV fixed points. We do so by studying the defect operator product expansion (OPE) of displacement operators in free and interacting CFTs using the replica method. We show that as n approaches 1 a contact term can emerge when the OPE contains defect operators of twist $d - 2$. For general states of interacting theories

we give evidence that the only possibility is from the defect operator that approaches the stress tensor in the $n \to 1$ limit. This implies that the QNEC is always saturated for CFTs with a twist gap.

As discussed above, the generalized entropy of extremal surfaces has a well-established microscopic interpretation as the von Neumann entropy of boundary subregions. It is a natural to search for the boundary dual of the generalized entropy of black holes even when there are no extremal surfaces around. In Chapter 6, we find such a microscopic explanation for the generalized entropy of black holes within AdS/CFT. More technically, we propose the existence of semi-classical states where the generalized entropy of a quantum marginally trapped surface could be interpreted in a precise way as a coarse-grained entropy in the fundamental (boundary) theory. We find that the non-gravitational limit of our coarse-graining prescription corresponds to the Connes cocycle flow in algebraic QFT [18].

In chapter 7, we apply a version of the transformation used in Chapter 6 to extremal surfaces in the bulk. The transformation intuitively starts with a Cauchy slice that contains the extremal surface and then adds a "Kink" at the location of the extremal surface. This "Kink' transform' generalizes the notion of a boost in arbitrary spacetimes satisfying gravitational equations of motion. We provide evidence that the boundary dual of this bulk transformation is the Connes cocycle flow [18].

It is natural to ask whether there are connections between total energy and generalized entropy in quantum gravity. In Chapter 8, we propose such a connection in quantum gravity whereby the total energy is lower bounded by the generalized entropy on certain light-sheets. This inequality is the first quantum information bound on the total energy that involves Newton's constant and thus generalizes the classical Penrose inequality [20]. We also find quantum states that violate the Penrose inequality but satisfy this new bound. We discuss and rule out some alternative formulations.

Chapter 2

Local Modular Hamiltonians from the Quantum Null Energy Condition

2.1 Introduction and Summary

The reduced density operator ρ for a region in quantum field theory encodes all of the information about observables localized to that region. Given any ρ, one can define the *modular Hamiltonian* K by

$$\rho = e^{-K}. \tag{2.1.1}$$

Knowledge of this operator is equivalent to knowledge of ρ, but the modular Hamiltonian frequently appears in calculations involving entanglement entropy. In general, i.e. for arbitrary states reduced to arbitrary regions, K is a complicated non-local operator. However, in certain cases it is known to simplify.

The most basic example where K simplifies is the vacuum state of a QFT in Rindler space, i.e. the half-space $t = 0, x \geq 0$. The Bisognano–Wichmann theorem [21] states that in this case the modular Hamiltonian is

$$\Delta K = \frac{2\pi}{\hbar} \int d^{d-2}y \int_0^\infty x\, T_{tt}\, dx \tag{2.1.2}$$

where $\Delta K \equiv K - \langle K \rangle_{\text{vac}}$ defines the vacuum-subtracted modular Hamiltonian, and y are $d-2$ coordinates parametrizing the transverse directions. The vacuum subtraction generally removes regulator-dependent UV-divergences in K. Other cases where the modular Hamiltonian is known to simplify to an integral of local operators are obtained via conformal transformation of Eq. (2.1.2), including spherical regions in CFTs [22], regions in a thermal state of 1+1 CFTs [23], and null slabs [24, 10].

Using conservation of the energy-momentum tensor, one can easily re-express the Rindler modular Hamiltonian in Eq. (2.1.2) as an integral over the future Rindler horizon $u \equiv t - x =$

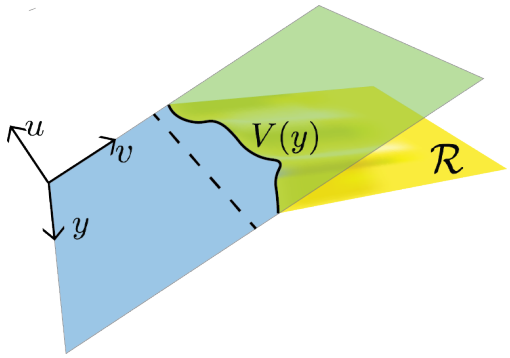

Figure 2.1: This image depicts a section of the plane $u = t - x = 0$. The region $\mathcal{R}$ is defined to be one side of a Cauchy surface split by the codimension-two entangling surface $\partial\mathcal{R} = \{(u = 0, v = V(y), y)\}$. The dashed line corresponds to a flat cut of the null plane.

0 which bounds the future of the Rindler wedge:

$$\Delta K = \frac{2\pi}{\hbar} \int d^{d-2}y \int_0^\infty v\, T_{vv}\, dv, \tag{2.1.3}$$

where $v \equiv t + x$. It is important to note that standard derivations of (2.1.2) or (2.1.3), e.g. [21, 22], do not apply when the entangling surface is defined by a non-constant cut of the Rindler horizon (see Fig. 2.1). One of the primary goals of this paper is to provide such a derivation.

For a large class of quantum field theories satisfying a precise condition specified momentarily, we will show that the vacuum modular Hamiltonian for the region $\mathcal{R}[V(y)]$ above an arbitrary cut $v = V(y)$ of a null plane is given by

$$\Delta K = \frac{2\pi}{\hbar} \int d^{d-2}y \int_{V(y)}^\infty (v - V(y))\, T_{vv}\, dv \tag{2.1.4}$$

This equation has been previously derived by Wall for free field theories [15] building on [25, 26], and to linear order in the deformation away from $V(y) = \text{const}$ in general QFTs by Faulkner et al. [27]. In CFTs, conformal transformations of Eq. (2.1.4) yield versions of the modular Hamiltonian for non-constant cuts of the causal diamond of a sphere.

The condition leading to Eq. (2.1.4) is that the theory should satisfy the *quantum null energy condition* (QNEC) [11, 28, 29, 30] — an inequality between the stress tensor and the von Neumann entropy of a region — and saturate the QNEC in the vacuum for regions defined by cuts of a null plane. We will review the statement of the QNEC in Sec. A.14.

The QNEC has been proven for free and superrenormalizable [28], as well as holographic [29, 30] quantum field theories. We take this as reasonable evidence that the QNEC is a true fact about relativistic quantum field theories in general, and for the purposes of this paper

take it as an assumption. In Sec. A.14 we will show how saturation of the QNEC in a given state leads to an operator equality relating certain derivatives of the modular Hamiltonian of that state to the energy-momentum tensor. Applied to the case outlined above, this operator equality will be integrated to give Eq. (2.1.4).

Given the argument in Sec. A.14, the only remaining question is whether the QNEC is in fact saturated in the vacuum state for entangling surfaces which are cuts of a null plane. This has been shown for free theories in [28]. In Sec. 6.3, we prove that this is the case for holographic theories to all orders in $1/N$. We emphasize that Eq. (2.1.4) holds purely as a consequence of the validity of the QNEC and the saturation in the vacuum for $\mathcal{R}$, two facts which are potentially true in quantum field theories much more generally than free and holographic theories.

Finally, in Sec. 6.6 we will conclude with a discussion of possible extensions to curved backgrounds and more general regions, connections between the relative entropy and the QNEC, and relations to other work.

2.2 Main Argument

Review of QNEC

The von Neumann entropy of a region in quantum field theory can be regarded as a functional of the entangling surface. We will primarily be interested in regions to one side of a cut of a null plane in flat space, for which the entangling surface can be specified by a function $V(y)$ which indicates the v-coordinate of the cut as a function of the transverse coordinates, collectively denoted y. See Fig. 2.1 for the basic setup. Each cut $V(y)$ defines a half-space, namely the region to one side of the cut. We will pick the side towards the future of the null plane. For the purposes of this section we are free to consider the more general situation where the entangling surface is only *locally* given by a cut of a null plane. Thus the von Neumann entropy can be considered as a functional of a profile $V(y)$ which defines the shape of the entangling surface, at least locally.

Suppose we define a one-parameter family of cuts $V(y;\lambda) \equiv V(y;0)+\lambda\dot{V}(y)$, with $\dot{V}(y) > 0$ to ensure that $\mathcal{R}(\lambda_1) \subset \mathcal{R}(\lambda_2)$ if $\lambda_1 > \lambda_2$. If $S(\lambda)$ is the entropy of region $\mathcal{R}(\lambda)$, then the QNEC in integrated form states that

$$\int d^{d-2}y \, \langle T_{vv}(y)\rangle \, \dot{V}(y)^2 \geq \frac{\hbar}{2\pi}\frac{d^2S}{d\lambda^2}. \tag{2.2.1}$$

In general there would be a $\sqrt{h}$ induced metric factor weighting the integral, but here and in the rest of the paper we will assume that the y coordinates have been chosen such that $\sqrt{h} = 1$.

By taking advantage of the arbitrariness of $\dot{V}(y)$ we can derive from this the local form of the QNEC. If we take a limit where $\dot{V}(y')^2 \to \delta(y-y')$, then the l.h.s. reduces to $\langle T_{vv}\rangle$.

We *define* $S''(y)$ as the limit of $d^2S/d\lambda^2$ in the same situation:

$$\frac{d^2S}{d\lambda^2} \to S''(y) \quad \text{when} \quad \dot{V}(y')^2 \to \delta(y-y'). \tag{2.2.2}$$

Taking the limit of the nonlocal QNEC then gives the local one:

$$\langle T_{vv}\rangle \geq \frac{\hbar}{2\pi} S''. \tag{2.2.3}$$

The local QNEC together with strong subadditivity can likewise be used to go backward and derive the nonlocal QNEC [11, 28, 29]. The details of that argument are not important here. In the next section we will discuss the consequences of the saturation of the QNEC, and will have to distinguish whether we mean saturation of the nonlocal inequality Eq. (2.2.1) or the local inequality Eq. (2.2.3), the latter condition being weaker.

The QNEC under state perturbations

In this section we consider how the QNEC behaves under small deformations of the state. We begin with a reference state σ and consider the deformed state $\rho = \sigma + \delta\rho$, with $\delta\rho$ traceless but otherwise arbitrary.

Consider a one-parameter family of regions $\mathcal{R}(\lambda)$ as in the previous section. Define $\overline{\mathcal{R}}(\lambda)$ to be the complement of $\mathcal{R}(\lambda)$ within a Cauchy surface. The reduced density operator for any given region $\mathcal{R}(\lambda)$ given by

$$\rho(\lambda) = \sigma(\lambda) + \delta\rho(\lambda) = \mathrm{Tr}_{\overline{\mathcal{R}}(\lambda)}\sigma + \mathrm{Tr}_{\overline{\mathcal{R}}(\lambda)}\delta\rho. \tag{2.2.4}$$

By the First Law of entanglement entropy, the entropy of $\rho(\lambda)$ is given by

$$S(\rho(\lambda)) = S(\sigma(\lambda)) - \mathrm{Tr}_{\mathcal{R}(\lambda)}\delta\rho(\lambda)\log\sigma(\lambda) + o(\delta\rho^2). \tag{2.2.5}$$

The second term can be written in a more useful way be defining the modular Hamiltonian $K_\sigma(\lambda)$ as

$$K_\sigma(\lambda) \equiv -\mathbb{1}_{\overline{\mathcal{R}}(\lambda)} \otimes \log\sigma(\lambda). \tag{2.2.6}$$

Defining $K_\sigma(\lambda)$ this way makes it a global operator, which makes taking derivatives with respect to λ formally simpler. Using this definition, we can write Eq. (2.2.5) as

$$S(\rho(\lambda)) = S(\sigma(\lambda)) + \mathrm{Tr}\,\delta\rho K_\sigma(\lambda) + o(\delta\rho^2). \tag{2.2.7}$$

Now in the second term the trace is over the global Hilbert space, and the λ-dependence has been isolated to the operator $K_\sigma(\lambda)$. Taking two derivatives, and simplifying the notation slightly, we find

$$\frac{d^2S}{d\lambda^2}(\rho) = \frac{d^2S}{d\lambda^2}(\sigma) + \mathrm{Tr}\,\delta\rho\frac{d^2K_\sigma}{d\lambda^2} + o(\delta\rho^2). \tag{2.2.8}$$

Suppose that the nonlocal QNEC, Eq. (2.2.1), is saturated in the state σ for all profiles $\dot{V}(y)$. Then, using Eq. (2.2.8), the nonlocal QNEC for the state ρ can be written as

$$\int d^{d-2}y \, (\mathrm{Tr}\, \delta\rho \, T_{vv}) \, \dot{V}^2 \geq \frac{\hbar}{2\pi} \mathrm{Tr}\, \delta\rho \frac{d^2 K_\sigma}{d\lambda^2} + o(\delta\rho^2). \tag{2.2.9}$$

The operator $\delta\rho$ was arbitrary, and in particular could be replaced by $-\delta\rho$. Then the only way that Eq. 2.2.9 can hold is if we have the operator equality

$$\frac{d^2 K_\sigma}{d\lambda^2} = C + \frac{2\pi}{\hbar} \int d^{d-2}y \, T_{vv} \dot{V}^2. \tag{2.2.10}$$

Here C is a number that we cannot fix using this method that is present because of the tracelessness of $\delta\rho$.

Eq. (2.2.10) can be integrated to derive the full modular Hamiltonian K_σ if we have appropriate boundary conditions. Up until now we have only made use of local properties of the entangling surface, but in order to provide boundary conditions for the integration of Eq. (2.2.10) we will assume that the entangling surface is globally given by a cut of a null plane, and that $V(y; \lambda = 0) = 0$. We will also make σ the vacuum state. In that situation it is known that the QNEC is saturated for free theories, and in the next section we will show that this is also true for holographic theories at all orders in the large-N expansion.

Our first boundary condition is at $\lambda = \infty$.[1] Since we expect that $K_\sigma(\lambda)$ should have a finite expectation value in any state as $\lambda \to \infty$, it must be that $dK_\sigma/d\lambda \to 0$ as $\lambda \to \infty$. Then integrating Eq. (2.2.10) gives

$$\frac{dK_\sigma}{d\lambda} = -\frac{2\pi}{\hbar} \int d^{d-2}y \int_{V(y;\lambda)}^{\infty} dv \, T_{vv} \dot{V}. \tag{2.2.11}$$

Note that this equation implies that the vacuum expectation value $\langle K_\sigma(\lambda) \rangle_{\mathrm{vac}}$ is actually λ-independent, which makes vacuum subtraction easy.

Our second boundary condition is Eq. (2.1.3), valid at $\lambda = 0$ when $V(y; \lambda) = 0$. Integrating once morenand making use of this boundary condition, we find

$$\Delta K_\sigma(\lambda) = \frac{2\pi}{\hbar} \int d^{d-2}y \int_{V(y;\lambda)}^{\infty} (v - V(y;\lambda)) \, T_{vv} \, dv \tag{2.2.12}$$

which is Eq. (2.1.4). Note that the l.h.s. of this equation is now the vacuum-subtracted modular Hamiltonian.

[1] It is not always possible to consider the $\lambda \to \infty$ limit of a null perturbation to an entangling surface because parts of the entangling surface may become timelike related to each other at some finite value of λ, at which point the surface is no longer the boundary of a region on a Cauchy surface. However, when the entangling surface is globally equal to a cut of a null plane this is not an issue.

Before moving on, we will briefly comment on the situation where the local QNEC, Eq. (2.2.3), is saturated but the nonlocal QNEC, Eq. (2.2.1), is not. Then, analogously to S'' in Eq. (2.2.2), one may define a local second derivative of K_σ:

$$\frac{d^2K_\sigma}{d\lambda^2} \to K''_\sigma(y) \quad \text{when} \quad \dot{V}(y')^2 \to \delta(y-y'). \tag{2.2.13}$$

Very similar manipulations then show that saturation of the local QNEC implies the equality

$$K''_\sigma = \frac{2\pi}{\hbar} T_{vv}. \tag{2.2.14}$$

This equation is weaker than Eq. (2.2.10), which is meant to be true for arbitrary profiles of $\dot{V}(y)$, but it may have a greater regime of validity. We will comment on this further in Sec. 6.6.

2.3 Holographic Calculation

In the previous section we argued that the form of the modular Hamiltonian could be deduced from saturation of the QNEC. In this section we will use the holographic entanglement entropy formula [31, 3, 32, 5] to show that the QNEC is saturated in vacuum for entangling surfaces defined by arbitrary cuts $v = V(y)$ of the null plane $u = 0$ in holographic theories. Our argument applies to any holographic theory defined by a relevant deformation to a holographic CFT, and will be at all orders in the large-N expansion. To reach arbitrary order in $1/N$ we will assume that the all-orders prescription for von Neumann entropy is given by the quantum extremal surface proposal of Engelhardt and Wall [6]. This is the same context in which the holographic proof of the QNEC was extended to all orders in $1/N$ [30].[2]

As before, the entangling surface in the field theory is given by the set of points $\partial\mathcal{R} = \{(u, v, y) : v = V(y), u = 0\}$ with null coordinates $u = t-x$ and $v = t+x$, and the region $\mathcal{R}$ is chosen to lie in the $u < 0$ portion of spacetime. Here y represents $d-2$ transverse coordinates. The bulk quantum extremal surface anchored to this entangling surface is parameterized by the functions $\bar{V}(y, z)$ and $\bar{U}(y, z)$. It was shown in [29, 30] that if we let the profile $V(y)$ depend on a deformation parameter λ, then the second derivative of the entropy is given by

$$\frac{d^2S}{d\lambda^2} = -\frac{d}{4G\hbar}\int d^{d-2}y\, \frac{d\bar{U}_{(d)}}{d\lambda}\ , \tag{2.3.1}$$

to all orders in $1/N$, where $\bar{U}_{(d)}(y)$ is the coefficient of z^d in the small-z expansion of $\bar{U}(z, y)$. We will show that $\bar{U} = 0$ identically for any profile $V(y)$, which then implies that $d^2S/d\lambda^2 = 0$, which is the statement of QNEC saturation in the vacuum.

[2]It is crucial that we demonstrate saturation beyond leading order in large-N. The argument in the previous section used *exact* saturation, and an error that is naïvely subleading when evaluated in certain states may become very large in others.

One way to show that $\bar{U}$ vanishes is to demonstrate that $\bar{U} = 0$ solves the quantum extremal surface equations of motion in the bulk geometry dual to the vacuum state of the boundary theory. The quantum extremal surface is defined by having the sum of the area plus the bulk entropy on one side be stationary with respect to first-order variations of its position. One can show that $\bar{U} = 0$ is a solution to the equations of motion if any only if

$$\frac{\delta S_{\text{bulk}}}{\delta \bar{V}(y,z)} = 0 \tag{2.3.2}$$

in the vacuum everywhere along the extremal surface. This would follow from null quantization if the bulk fields were free [28], but that would only allow us to prove the result at order-one in the $1/N$ expansion.

For an all-orders argument, we opt for a more indirect approach using subregion duality, or entanglement wedge reonstruction [33, 34, 35, 9].[3] A version of this argument first appeared in [30], and we elaborate on it here.

Entanglement wedge reconstruction requires two important consistency conditions in the form of constraints on the bulk geometry which must hold at all orders in $1/N$: The first constraint, *entanglement wedge nesting* (EWN), states that if one boundary region is contained inside the domain of dependence of another, then the quantum extremal surface associated to the first boundary region must be contained within the entanglement wedge of the second boundary region [33, 36]. The second constraint, $\mathcal{C} \subseteq \mathcal{E}$, demands that the causal wedge of a boundary region be contained inside the entanglement wedge of that region [33, 34, 36, 6, 37]. Equivalently, it says that no part of the quantum extremal surface of a given boundary region can be timelike-related to the (boundary) domain of dependence of that boundary region. It was shown in [30] that $\mathcal{C} \subseteq \mathcal{E}$ follows from EWN, and EWN itself is simply the statement that a boundary region should contain all of the information about any of its subregions. We will now explain the consequences of these two constraints for $\bar{U}(y,z)$.

Without loss of generality, suppose the region $\mathcal{R}$ is defined by a coordinate profile which is positive, $V(y) > 0$. Consider a second region $\mathcal{R}_0$ which has an entangling surface at $v = u = 0$ and whose domain of dependence (i.e., Rindler space) contains $\mathcal{R}$. The quantum extremal surface associated to $\mathcal{R}_0$ is given by $\bar{U}_0 = \bar{V}_0 = 0$. This essentially follows from symmetry.[4] The entanglement wedge of $\mathcal{R}_0$ is then a bulk extension of the boundary Rindler space, namely the set of bulk points satisfying $u \leq 0$ and $v \geq 0$. Then EWN implies that $\bar{U} \leq 0$ and $\bar{V} \geq 0$.

The only additional constraint we need from $\mathcal{C} \subseteq \mathcal{E}$ is the requirement that the quantum extremal surface for $\mathcal{R}$ not be in the past of the domain of dependence of $\mathcal{R}$. From the definition of $\mathcal{R}$, it is clear that a bulk point is in the past of the domain of dependence of $\mathcal{R}$ if and only if it is in the past of the region $u < 0$ on the boundary, which is the same as the

[3] The entanglement wedge of a boundary region is the set of bulk points which are spacelike- or null-related to that region's quantum extremal surface on the same side of the quantum extremal surface as the boundary region itself.

[4] One might worry that the quantum extremal surface equations display spontaneous symmetry breaking in the vacuum, but this can be ruled out using $\mathcal{C} \subseteq \mathcal{E}$ with an argument similar to the one we present here.

region $u < 0$ in the bulk. Therefore it must be that $\bar{U} \geq 0$. Combined with the constraint from EWN above, we then conclude that the only possibility is $\bar{U} = 0$. This completes the proof that the QNEC is saturated to all orders in $1/N$.

2.4 Discussion

We conclude by discussing the generality of our analysis, some implications and future directions, and connections with previous work.

Generalizations and Future Directions

General Killing horizons Though we restricted to cuts of Rindler horizons in flat space for simplicity, all of our results continue to hold for cuts of bifurcate Killing horizons for QFTs defined in arbitrary spacetimes, assuming the QNEC is true and saturated in the vacuum in this context. In particular, Eq. (2.1.4) holds with v a coordinate along the horizon. For holographic theories, entanglement wedge nesting (EWN) and the entanglement wedge being outside of the causal wedge ($\mathcal{C} \subseteq \mathcal{E}$) continue to prove saturation of the QNEC. To see this, note that a Killing horizon on the boundary implies a corresponding Killing horizon in the bulk. Now take the reference region $\mathcal{R}_0$ satisfying $V(y) = U(y) = 0$ to be the boundary bifurcation surface. By symmetry, the associated quantum extremal surface lies on the bifurcation surface of the bulk Killing horizon. Then the quantum extremal surface of the region $\mathcal{R}$ defined by $V(y) \geq 0$ must lie in the entanglement wedge of $\mathcal{R}_0$ — inside the bulk horizon — by entanglement wedge nesting, but must also lie on or outside of the bulk horizon by $\mathcal{C} \subseteq \mathcal{E}$. Thus it lies on the bulk horizon, $\bar{U} = 0$, and the QNEC remains saturated by Eq. (2.3.1).

Future work In this work, we have only established the form of $K_{\mathcal{R}}$ for regions $\mathcal{R}$ bounded by arbitrary cuts of a null plane. A natural next direction would be to understand if and how we can extend Eq. (2.2.14) to more general entangling surfaces. As discussed above, the QNEC was shown to hold for locally flat entangling surfaces in holographic, free and super-renormalizable field theories [28, 29, 30]. Thus, if we could prove saturation, i.e. that $S''_{\text{vac}} = 0$ at all orders in $1/N$, then we would establish (2.2.14) for all regions with a locally flat boundary.

One technique to probe this question is to perturb the entangling surface away from a flat cut and compute the contributions to the QNEC order-by-order in a perturbation parameter ϵ. Preliminary calculations [38] have revealed that for holographic theories at leading order in large N, $S''_{\text{vac}} = 0$ at all orders in ϵ.

Another interesting problem is to show that in a *general* QFT vacuum, null derivatives of entanglement entropy across arbitrary cuts of null planes vanish. That, along with a general proof of QNEC will establish (18) as a consequence. We will leave this to future work.

The QNEC as $S(\rho\|\sigma)'' \geq 0$

There is a connection between the QNEC and relative entropy, first pointed out in [30], that we elaborate on here. The relative entropy $S(\rho\|\sigma)$ between two states ρ and σ is defined as

$$S(\rho\|\sigma) = \operatorname{Tr}\rho\log\rho - \operatorname{Tr}\rho\log\sigma \tag{2.4.1}$$

and provides a measure of distinguishability between the two states [39]. Substituting the definition of K, Eq. (2.1.1), into Eq. (2.4.1) provides a useful alternate presentation:

$$S(\rho\|\sigma) = \langle K_\sigma\rangle_\rho - S(\rho). \tag{2.4.2}$$

If Eq. (2.1.4) is valid, then taking two derivatives with respect to a deformation parameter, as in the main text, shows that the nonlocal QNEC, Eq. (2.2.1), is equivalent to

$$\partial_\lambda^2 S(\rho(\lambda)\|\sigma(\lambda)) \geq 0. \tag{2.4.3}$$

For comparison, monotonicity of relative entropy for the types of regions and deformations we have been discussing can be written as

$$\partial_\lambda S(\rho(\lambda)\|\sigma(\lambda)) \leq 0. \tag{2.4.4}$$

Eq. (2.4.3) is a sort of "convexity" of relative entropy.[5] Unlike monotonicity of relative entropy, which says that the first derivative is non-positive, there is no general information-theoretic reason for the *second* derivative to be non-negative. In the event that Eq. (2.2.14) holds but not Eq. (2.1.4), we would still have

$$S(\rho\|\sigma)'' \geq 0. \tag{2.4.5}$$

where the $''$ notation denotes a local deformation as in Sec. A.14.

It would be extremely interesting to characterize what about quantum field theory and null planes makes (2.4.3) true. We can model the null deformation as a non-unitary time evolution in the space of states, with the vacuum state serving as an equilibrium state for this evolution. Then an arbitrary finite-energy state will relax toward the equilibrium state, with the relative entropy $S(\rho\|\sigma)$ characterizing the free energy as a function of time. Monotonicity of relative entropy is then nothing more than the statement that free energy decreases, i.e. the second law of thermodynamics. The second derivative statement gives more information about the approach to equilibrium. If that approach is of the form of exponential decay, then all successive derivatives would alternate in sign. However, for null deformations in quantum field theory we do not expect to have a general bound on the behavior of derivatives of the energy-momentum tensor, meaning that the third derivative of the free energy should not have a definite sign.[6] Perhaps there is some way of characterizing the approach to equilibrium we have here, which is in some sense smoother than the most general possibility but not so constrained as to force exponential behavior.

[5] This is distinct from the well-known convexity of relative entropy, which says that $S(t\rho_1+(1-t)\rho_2\|\sigma) \leq tS(\rho_1\|\sigma) + (1-t)S(\rho_2\|\sigma)$.

[6] We thank Aron Wall for a discussion of this point.

Relation to previous work

Faulkner, Leigh, Parrikar and Wang [27] have discussed results very similar to the ones presented here. They demonstrated that for first-order null deformations $\delta V(y)$ to a flat cut of a null plane, the perturbation to the modular Hamiltonian takes the form

$$\langle K_{\mathcal{R}} \rangle_\psi - \langle K_{\mathcal{R}_0} \rangle_\psi = -\frac{2\pi}{\hbar} \int d^{d-2}y \int_{V(y)} dv\, T_{vv}(y)\, \delta V(y) \tag{2.4.6}$$

This is precisely the form expected from our equation (2.1.4). Faulkner et al. went on to suggest that the natural generalization of the modular Hamiltonian to finite deformations away from a flat cut takes the form of Eq. (2.1.4). In the context of holography they showed that this conclusion applied both on the boundary and in the bulk is consistent with JLMS [7]. In the present paper, we have shown that Eq. (2.1.4) holds for theories which obey the QNEC, and for which the QNEC is saturated in the vacuum. A non-perturbative, field theoretic proof of these assumptions remains a primary goal of future work.

Chapter 3

The Quantum Null Energy Condition, Entanglement Wedge Nesting, and Quantum Focusing

3.1 Introduction and Summary

The Quantum Focusing Conjecture (QFC) is a new principle of semiclassical quantum gravity proposed in [11]. Its formulation is motivated by classical focusing, which states that the expansion θ of a null congruence of geodesics is nonincreasing. Classical focusing is at the heart of several important results of classical gravity [40, 41, 42, 43], and likewise quantum focusing can be used to prove quantum generalizations of many of these results [44, 45, 46, 47].

One of the most important and surprising consequences of the QFC is the Quantum Null Energy Condition (QNEC), which was discovered as a particular nongravitational limit of the QFC [11]. Subsequently the QNEC was proven for free fields [28] and for holographic CFTs on flat backgrounds [29] (and recently extended in [48] in a similar way as we do here). The formulation of the QNEC which naturally comes out of the proofs we provide here is as follows.

Consider a codimension-two Cauchy-splitting surface Σ, which we will refer to as the entangling surface. The Von Neumann entropy $S[\Sigma]$ of the interior (or exterior) or Σ is a functional of Σ, and in particular is a functional of the embedding functions $X^i(y)$ that define Σ. Choose a one-parameter family of deformed surfaces $\Sigma(\lambda)$, with $\Sigma(0) = \Sigma$, such that (i) $\Sigma(\lambda)$ is given by flowing along null geodesics generated by the null vector field k^i normal to Σ for affine time λ , and (ii) $\Sigma(\lambda)$ is either "shrinking" or "growing" as a function of λ, in the sense that the domain of dependence of the interior of Σ is either shrinking or growing. Then for any point on the entangling surface we can define the combination

$$T_{ij}(y)k^i(y)k^j(y) - \frac{1}{2\pi}\frac{d}{d\lambda}\left(\frac{k^i(y)}{\sqrt{h(y)}}\frac{\delta S_{\text{ren}}}{\delta X^i(y)}\right). \tag{3.1.1}$$

Here $\sqrt{h(y)}$ is the induced metric determinant on Σ. Writing this down in a general curved background requires a renormalization scheme both for the energy-momentum tensor T_{ij} and the renormalized entropy S_{ren}. Assuming that this quantity is scheme-independent (and hence well-defined), the QNEC states that it is positive. Our main task is to determine the necessary and sufficient conditions we need to impose on Σ and the background spacetime at the point y in order that the QNEC hold.

In addition to a proof through the QFC, the holographic proof method of [29] is easily adaptable to answering this question in full generality. The backbone of that proof is Entanglement Wedge Nesting (EWN), which is a consequence of subregion duality in AdS/CFT [47]. A given region on the boundary of AdS is associated with a particular region of the bulk, called the entanglement wedge, which is defined as the bulk region spacelike-related to the extremal surface [31, 32, 6, 49] used to compute the CFT entropy on the side toward the boundary region. This bulk region is dual to the given boundary region, in the sense that there is a correspondence between the algebras of operators in the bulk region and the operators in the boundary region which are good semiclassical gravity operators (i.e., they act within the subspace of semiclassical states) [33, 7, 35]. EWN is the statement that nested boundary regions must be dual to nested bulk regions, and clearly follows from the consistency of subregion duality.

While the QNEC can be derived from both the QFC and EWN, there has been no clear connection between these derivations.[1] As it stands, there are apparently two QNECs, the QNEC-from-QFC and the QNEC-from-EWN. We will show in full generality that these two QNECs are in fact the same, at least in $d \leq 5$ dimensions.

Here is a summary of our results:

- The holographic proof of the QNEC from EWN is extended to CFTs on arbitrary curved backgrounds. In $d = 5$ we find necessary that the necessary and sufficient conditions for the ordinary QNEC to hold at a point are that[2]

$$\theta_{(k)} = \sigma^{(k)}_{ab} = D_a\theta_{(k)} = D_a\sigma^{(k)}_{bc} = R_{ka} = 0 \tag{3.1.2}$$

  at that point. For $d < 5$ only a subset of these conditions are necessary. This is the subject of §3.2.

- We also show holographically that under the weaker set of conditions

$$\sigma^{(k)}_{ab} = D_a\theta_{(k)} + R_{ka} = D_a\sigma^{(k)}_{bc} = 0 \tag{3.1.3}$$

  the Conformal QNEC holds. The Conformal QNEC was introduced in [29] as a conformally-transformed version of the QNEC. This is the strongest inequality that we can get out of EWN. This is the subject of §3.2

[1] In [47] it was shown that the QFC in the bulk implies EWN, which in turn implies the QNEC. This is not the same as the connection we are referencing here. The QFC which would imply the boundary QNEC in the sense that we mean is a *boundary* QFC, obtained by coupling the boundary theory to gravity.

[2] Here $\sigma^{(k)}_{ab}$ and $\theta_{(k)}$ are the shear and expansion in the k^i direction, respectively, and D_a is a surface covariant derivative. Our notation is further explained in Appendix A.1.

- By taking the non-gravitational limit of the QFC we are able to derive the QNEC again under the same set of conditions as we did for EWN. This is the subject of §3.3.

- We argue in §3.3 that the statement of the QNEC is scheme-independent whenever the conditions that allow us to prove it hold. This shows that the two proofs of the QNEC are actually proving the same, unambiguous field–theoretic bound.

We conclude in §7.6 with a discussion and suggest future directions. A number of technical Appendices are included as part of our analysis.

Relation to other work While this work was in preparation, [48] appeared which has overlap with our discussion of EWN and the scheme-independence of the QNEC. The results of [48] relied on a number of assumptions about the background: the null curvature condition and a positive energy condition. From this they derive certain sufficient conditions for the QNEC to hold. We do not assume anything about our backgrounds a priori, and include all relevant higher curvature corrections. This gives our results greater generality, as we are able to find both necessary and sufficient conditions for the QNEC to hold.

3.2 Entanglement Wedge Nesting

Subregion Duality

The statement of AdS/CFT includes a correspondence between operators in the semiclassical bulk gravitational theory and CFT operators on the boundary. Moreover, it has been shown [9, 35] that such a correspondence exists between the operator algebras of subregions in the CFT and certain associated subregions in the bulk as follows: Consider a spatial subregion A in the boundary geometry. The extremal surface anchored to ∂A, which is used to compute the entropy of A [31, 32], bounds the so-called entanglement wedge of A, $\mathcal{E}(A)$, in the bulk. More precisely $\mathcal{E}(A)$ is the codimension-zero bulk region spacelike-related to the extremal surface on the same side of the extremal surface as A. Subregion duality is the statement that the operator algebras of $\mathcal{D}(A)$ and $\mathcal{E}(A)$ are dual, where $\mathcal{D}(A)$ denotes the domain of dependence of A.

Entanglement Wedge Nesting The results of this section follow from EWN, which we now describe. Consider two boundary regions A_1 and A_2 such that $\mathcal{D}(A_1) \subseteq \mathcal{D}(A_2)$. Then consistency of subregion duality implies that $\mathcal{E}(A_1) \subseteq \mathcal{E}(A_2)$ as well, and this is the statement of EWN. In particular, EWN implies that the extremal surfaces associated to A_1 and A_2 cannot be timelike-related.

We will mainly be applying EWN to the case of a one-paramter family of boundary regions, $A(\lambda)$, where $\mathcal{D}(A(\lambda_1)) \subseteq \mathcal{D}(A(\lambda_2))$ whenever $\lambda_1 \leq \lambda_2$. Then the union of the one-parameter family of extremal surfaces associated to $A(\lambda)$ forms a codimension-one surface

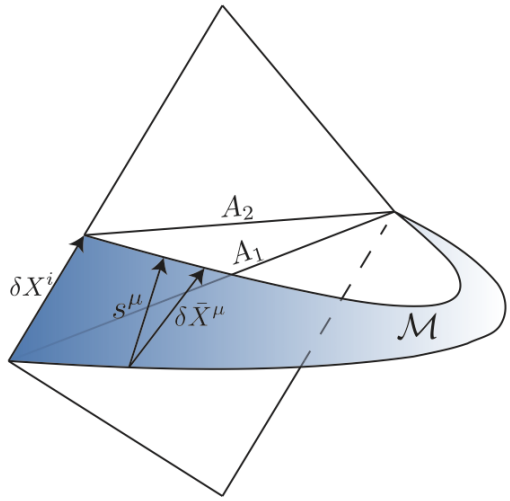

Figure 3.1: Here we show the holographic setup which illustrates Entanglement Wedge Nesting. A spatial region A_1 on the boundary is deformed into the spatial region A_2 by the null vector δX^i. The extremal surfaces of A_1 and A_2 are connected by a codimension-one bulk surface $\mathcal{M}$ (shaded blue) that is nowhere timelike by EWN. Then the vectors $\delta \bar{X}^\mu$ and s^μ, which lie in $\mathcal{M}$, have nonnegative norm.

in the bulk that is nowhere timelike. We denote this codimension-one surface by $\mathcal{M}$. See Fig. 3.1 for a picture of the setup.

Since $\mathcal{M}$ is nowhere timelike, every one of its tangent vectors must have nonnegative norm. In particular, consider the embedding functions $\bar{X}^\mu$ of the extremal surfaces in some coordinate system. Then the vectors $\delta \bar{X}^\mu \equiv \partial_\lambda \bar{X}^\mu$ is tangent to $\mathcal{M}$, and represents a vector that points from one extremal surface to another. Hence we have $(\delta \bar{X})^2 \geq 0$ from EWN, and this is the inequality that we will discuss for most of the remainder of this section.

Before moving on, we will note that $(\delta \bar{X})^2 \geq 0$ is not necessarily the strongest inequality we get from EWN. At each point on $\mathcal{M}$, the vectors which are tangent to the extremal surface passing through that point are known to be spacelike. Therefore if $\delta \bar{X}^\mu$ contains any components which are tangent to the extremal surface, they will serve to make the inequality $(\delta \bar{X})^2 \geq 0$ weaker. We define the vector s^μ at any point of $\mathcal{M}$ to be the part of $\delta \bar{X}^\mu$ orthogonal to the extremal surface passing through that point. Then $(\delta \bar{X})^2 \geq s^2 \geq 0$. We will discuss the $s^2 \geq 0$ inequality in §3.2 after handling the $(\delta \bar{X})^2 \geq 0$ case.

Near-Boundary EWN

In this section we explain how to calculate the vector $\delta \bar{X}^\mu$ and s^μ near the boundary explicitly in terms of CFT data. Then the EWN inequalities $(\delta \bar{X})^2 > 0$ and $s^2 > 0$ can be given a CFT meaning. The strategy is to use a Fefferman-Graham expansion of both the metric and

extremal surface, leading to equations for $\delta\bar{X}^\mu$ and s^μ as power series in the bulk coordinate z (including possible log terms). In the following sections we will analyze the inequalities that are derived in this section.

Bulk Metric We work with a bulk theory in AdS_{d+1} that consists of Einstein gravity plus curvature-squared corrections. For $d \leq 5$ this is the complete set of higher curvature corrections that have an impact on our analysis. The Lagrangian is[3]

$$\mathcal{L} = \frac{1}{16\pi G_N}\left(\frac{d(d-1)}{\tilde{L}^2} + \mathcal{R} + \ell^2\lambda_1\mathcal{R}^2 + \ell^2\lambda_2\mathcal{R}_{\mu\nu}^2 + \ell^2\lambda_{\rm GB}\mathcal{L}_{\rm GB}\right), \tag{3.2.1}$$

where $\mathcal{L}_{GB} = \mathcal{R}_{\mu\nu\rho\sigma}^2 - 4\mathcal{R}_{\mu\nu}^2 + \mathcal{R}^2$ is the Gauss–Bonnet Lagrangian, ℓ^2 is the cutoff scale, and $\tilde{L}^2$ is the scale of the cosmological constant. The bulk metric has the following near boundary expansion in Fefferman-Graham gauge [50]:

$$ds^2 = \frac{L^2}{z^2}(dz^2 + \bar{g}_{ij}(x,z)dx^i dx^j), \tag{3.2.2}$$

$$\bar{g}_{ij}(x,z) = g_{ij}^{(0)}(x) + z^2 g_{ij}^{(2)}(x) + z^4 g_{ij}^{(4)}(x) + \ldots + z^d \log z\, g_{ij}^{(d,\log)}(x) + z^d g_{ij}^{(d)}(x) + o(z^d). \tag{3.2.3}$$

Note that the length scale L is different from $\tilde{L}$, but the relationship between them will not be important for us. Demanding that the above metric solve bulk gravitational equations of motion gives expressions for all of the $g_{ij}^{(n)}$ for $n < d$, including $g_{ij}^{(d,\log)}(x)$, in terms of $g_{ij}^{(0)}(x)$. This means, in particular, that these terms are all state-independent. One finds that $g_{ij}^{(d,\log)}(x)$ vanishes unless d is even. We provide explicit expressions for some of these terms in Appendix A.3.

The only state-dependent term we have displayed, $g_{ij}^{(d)}(x)$, contains information about the expectation value of the energy-momentum tensor T_{ij} of the field theory. In odd dimensions we have the simple formula [51][4]

$$g_{ij}^{(d=\text{odd})} = \frac{16\pi G_N}{\eta d L^{d-1}}\langle T_{ij}\rangle, \tag{3.2.4}$$

with

$$\eta = 1 - 2\left(d(d+1)\lambda_1 + d\lambda_2 + (d-2)(d-3)\lambda_{\rm GB}\right)\frac{\ell^2}{L^2} \tag{3.2.5}$$

In even dimensions the formula is more complicated. For $d = 4$ we discuss the form of the metric in Appendix A.5

[3]For simplicity we will not include matter fields explicitly in the bulk, but their presence should not alter any of our conclusions.

[4]Even though [51] worked with a flat boundary theory, one can check that this formula remains unchanged when the boundary is curved.

Extremal Surface EWN is a statement about the causal relation between entanglement wedges. To study this, we need to calculate the position of the extremal surface. We parametrize our extremal surface by the coordinate (y^a, z), and the position of the surface is determined by the embedding functions $\bar{X}^\mu(y^a, z)$. The intrinsic metric of the extremal surface is denoted by $\bar{h}_{\alpha\beta}$, where $\alpha = (a, z)$. For convenience we will impose the gauge conditions $\bar{X}^z = z$ and $\bar{h}_{az} = 0$.

The functions $\bar{X}(y^a, z)$ are determined by extremizing the generalized entropy [6, 49] of the entanglement wedge. This generalized entropy consists of geometric terms integrated over the surface as well as bulk entropy terms. We defer a discussion of the bulk entropy terms to §3.4 and write only the geometric terms, which are determined by the bulk action:

$$S_{\text{gen}} = \frac{1}{4G_N}\int \sqrt{\bar{h}}\left[1 + 2\lambda_1\ell^2\mathcal{R} + \lambda_2\ell^2\left(\mathcal{R}_{\mu\nu}\mathcal{N}^{\mu\nu} - \frac{1}{2}\mathcal{K}_\mu\mathcal{K}^\mu\right) + 2\lambda_{\text{GB}}\ell^2\bar{r}\right]. \tag{3.2.6}$$

We discuss this entropy functional in more detail in Appendix A.3. The Euler-Lagrange equations for S_{gen} are the equations of motion for $\bar{X}^\mu$. Like the bulk metric, the extremal surface equations can be solved at small-z with a Fefferman–Graham-like expansion:

$$\bar{X}^i(y, z) = X^i_{(0)}(y) + z^2 X^i_{(2)}(y) + z^4 X^i_{(4)}(y) + \ldots + z^d \log z\, X^i_{(d,\log)}(y) + z^d X^i_{(d)}(y) + o(z^d), \tag{3.2.7}$$

As with the metric, the coefficient functions $X^i_{(n)}$ for $n < d$, including the log term, can be solved for in terms of $X^i_{(0)}$ and $g^{(0)}_{ij}$, and again the log term vanishes unless d is even. The state-dependent term $X^i_{(d)}$ contains information about variations of the CFT entropy, as we explain below.

The z-Expansion of EWN By taking the derivative of (3.2.7) with respect to λ, we find the z-expansion of $\delta\bar{X}^i$. We will discuss how to take those derivatives momentarily. But given the z-expansion of $\delta\bar{X}^i$, we can combine this with the z-expansion of $\bar{g}_{ij}$ in (3.2.3) to get the z-expansion of $(\delta\bar{X})^2$:

$$\frac{z^2}{L^2}(\delta\bar{X})^2 = g^{(0)}_{ij}\delta X^i_{(0)}\delta X^j_{(0)} + z^2\left(2g^{(0)}_{ij}\delta X^i_{(0)}\delta X^j_{(2)} + g^{(2)}_{ij}\delta X^i_{(0)}\delta X^j_{(0)} + X^m_{(2)}\partial_m g^{(0)}_{ij}\delta X^i_{(0)}\delta X^j_{(0)}\right) + \cdots \tag{3.2.8}$$

EWN implies that $(\delta\bar{X})^2 \geq 0$, and we will spend the next few sections examining this inequality using the expansion (3.2.8). From the general arguments given above, we can get a stronger inequality by considering the vector s^μ and its norm rather than $\delta\bar{X}^\mu$. The construction of s^μ is more involved, but we would similarly construct an equation for s^2 at small z. We defer further discussion of s^μ to §3.2.

Now we return to the question of calculating $\delta\bar{X}^i$. Since all of the $X^i_{(n)}$ for $n < d$ are known explicitly from solving the equation of motion, the λ-derivatives of those terms can be taken and the results expressed in terms of the boundary conditions for the extremal surface. The

variation of the state-dependent term, $\delta X^i_{(d)}$, is also determined by the boundary conditions in principle, but in a horribly non-local way. However, we will now show that $X^i_{(d)}$ (and hence $\delta X^i_{(d)}$) can be re-expressed in terms of variations of the CFT entropy.

Variations of the Entropy The CFT entropy $S_{\rm CFT}$ is equal to the generalized entropy $S_{\rm gen}$ of the entanglement wedge in the bulk. To be precise, we need to introduce a cutoff at $z = \epsilon$ and use holographic renormalization to properly define the entropy. Then we can use the calculus of variations to determine variations of the entropy with respect to the boundary conditions at $z = \epsilon$. There will be terms which diverge as $\epsilon \to 0$, as well as a finite term, which is the only one we are interested in at the moment. In odd dimensions, the finite term is given by a simple integral over the entangling surface in the CFT:

$$\delta S_{\rm CFT}|_{\rm finite} = \eta d L^{d-1} \int d^{d-2}y \sqrt{h} g_{ij} X^i_{(d)} \delta X^j. \tag{3.2.9}$$

This finite part of $S_{\rm CFT}$ is the renormalized entropy, $S_{\rm ren}$, in holographic renormalization. Eventually we will want to assure ourselves that our results are scheme-independent. This question was studied in [52], and we will discuss it further in §3.3. For now, the important take-away from (3.2.9) is

$$\frac{1}{\sqrt{h}} \frac{\delta S_{\rm ren}}{\delta X^i(y)} = -\frac{\eta d L^{d-1}}{4G_N} X^i_{(d,\rm odd)}. \tag{3.2.10}$$

The case of even d is more complicated, and we will cover the $d = 4$ case in Appendix A.5.

State-Independent Inequalities

The basic EWN inequality is $(\delta \bar{X})^2 \geq 0$. The challenge is to write this in terms of boundary quantities. In this section we will look at the state-independent terms in the expansion of (3.2.8). The boundary conditions at $z = 0$ are given by the CFT entangling surface and background geometry, which we denote by X^i and g_{ij} without a (0) subscript. The variation vector of the entangling surface is the null vector $k^i = \delta X^i$. We can use the formulas of Appendix A.4 to express the other $X^i_{(n)}$ for $n < d$ in terms of X^i and g_{ij}. This allows us to express the state-independent parts of $(\delta \bar{X})^2 \geq 0$ in terms of CFT data. In this subsection we will look at the leading and subleading state-independent parts. These will be sufficient to fully cover the cases $d \leq 5$.

Leading Inequality From (3.2.8), we see that the first term is actually $k_i k^i = 0$. The next term is the one we call the leading term, which is

$$L^{-2}(\delta \bar{X})^2\big|_{z^0} = 2k_i \delta X^i_{(2)} + g^{(2)}_{ij} k^i k^j + X^m_{(2)} \partial_m g_{ij} k^i k^j. \tag{3.2.11}$$

From (A.3.10), we easily see that this is equivalent to

$$L^{-2}\ (\delta \bar{X}^i)^2\big|_{z^0} = \frac{1}{(d-2)^2} \theta^2_{(k)} + \frac{1}{d-2} \sigma^2_{(k)}, \tag{3.2.12}$$

where $\sigma^{(k)}_{ab}$ and $\theta_{(k)}$ are the shear and expansion of the null congruence generated by k^i, and are given by the trace and trace-free parts of $k_i K^i_{ab}$, with K^i_{ab} the extrinsic curvature of the entangling surface. This leading inequality is always nonnegative, as required by EWN. Since we are in the small-z limit, the subleading inequality is only relevant when this leading inequality is saturated. So in our analysis below we will focus on the $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$ case, which can always be achieved by choosing the entangling surface appropriately. Note that in $d = 3$ this is the only state-independent term in $(\delta\bar{X})^2$, and furthermore we always have $\sigma^{(k)}_{ab} = 0$ in $d = 3$.

Subleading Inequality The subleading term in $(\delta\bar{X})^2$ is order z^2 in $d \geq 5$, and order $z^2 \log z$ in $d = 4$. These two cases are similar, but it will be easiest to focus first on $d \geq 5$ and then explain what changes in $d = 4$. The terms we are looking for are

$$\begin{aligned} L^{-2}(\delta\bar{X})^2\big|_{z^2} = {} & 2k_i\delta X^i_{(4)} + 2g^{(2)}_{ij}k^i\delta X^j_{(2)} + g_{ij}\delta X^i_{(2)}\delta X^j_{(2)} + g^{(4)}_{ij}k^ik^j + X^m_{(4)}\partial_m g_{ij}k^ik^j \\ & + 2X^m_{(2)}\partial_m g_{ij}k^i\delta X^j_{(2)} + X^m_{(2)}\partial_m g^{(2)}_{ij}k^ik^j + \frac{1}{2}X^m_{(2)}X^n_{(2)}\partial_m\partial_n g_{ij}k^ik^j. \end{aligned} \tag{3.2.13}$$

This inequality is significantly more complicated than the previous one. The details of its evaluation are left to Appendix A.4. The result, assuming $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$, is

$$\begin{aligned} L^{-2}(\delta\bar{X})^2\big|_{z^2} = {} & \frac{1}{4(d-2)^2}(D_a\theta_{(k)} + 2R_{ka})^2 \\ & + \frac{1}{(d-2)^2(d-4)}(D_a\theta_{(k)} + R_{ka})^2 + \frac{1}{2(d-2)(d-4)}(D_a\sigma^{(k)}_{bc})^2 \\ & + \frac{\kappa}{d-4}\left(C_{kabc}C_k{}^{abc} - 2C_k{}^c{}_{ca}C_k{}^b{}_b{}^a\right). \end{aligned} \tag{3.2.14}$$

where κ is proportional to $\lambda_{\rm GB}\ell^2/L^2$ and is defined in Appendix A.4. Aside from the Gauss–Bonnet term we have a sum of squares, which is good because EWN requires this to be positive when $\theta_{(k)}$ and $\sigma_{(k)}$ vanish. Since $\kappa \ll 1$, it cannot possibly interfere with positivity unless the other terms were zero. This would require $D_a\theta_{(k)} = D_a\sigma^{(k)}_{bc} = R_{ka} = 0$ in addition to our other conditions. But, following the arguments of [53], this cannot happen unless the components C_{kabc} of the Weyl tensor also vanish at the point in question. Thus EWN is always satisfied. Also noteSecond, the last two terms in middle line of (3.2.14) are each conformally invariant when $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$, which we have assumed. This will become important later.

Finally, though we have assumed $d \geq 5$ to arrive at this result, we can use it to derive the expression for $L^{-2}(\delta\bar{X})^2\big|_{z^2\log z}$ in $d = 4$. The rule, explained in Appendix A.5, is to multiply the RHS by $4 - d$ and then set $d = 4$. This has the effect of killing the conformally non-invariant term, leaving us with

$$L^{-2}(\delta\bar{X})^2\big|_{z^2\log z, d=4} = -\frac{1}{4}(D_a\theta_{(k)} + R_{ka})^2 - \frac{1}{4}(D_a\sigma^{(k)}_{bc})^2. \tag{3.2.15}$$

The Gauss–Bonnet term also disappears because of a special Weyl tensor identity in $d = 4$ [52]. The overall minus sign is required since $\log z < 0$ in the small z limit. In addition, we no longer require that R_{ka} and $D_a\theta_{(k)}$ vanish individually to saturate the inequality: only their sum has to vanish. This still requires that $C_{kabc} = 0$, though.

The Quantum Null Energy Condition

The previous section dealt with the two leading state-independent inequalities that EWN implies. Here we deal with the leading state-*dependent* inequality, which turns out to be the QNEC.

At all orders lower than z^{d-2}, $(\delta\bar{X})^2$ is purely geometric. At order z^{d-2}, however, the CFT energy-momentum tensor enters via the Fefferman–Graham expansion of the metric, and variations of the entropy enter through $X^i_{(d)}$. In odd dimensions the analysis is simple and we will present it here, while in general even dimensions it is quite complicated. Since our state-independent analysis is incomplete for $d > 5$ anyway, we will be content with analyzing only $d = 4$ for the even case. The $d = 4$ calculation is presented in Appendix A.5. Though is it more involved that the odd-dimensional case, the final result is the same.

Consider first the case where d is odd. Then we have

$$L^{-2}(\delta\bar{X})^2\big|_{z^{d-2}} = g^{(d)}_{ij}k^ik^j + 2k_i\delta X^i_{(d)} + X^m_{(d)}\partial_m g_{ij}k^ik^j = g^{(d)}_{ij}k^ik^j + 2\delta\left(k_i\delta X^i_{(d)}\right). \quad (3.2.16)$$

From (3.2.4) and (3.2.10), we find that

$$L^{-2}(\delta\bar{X})^2\big|_{z^{d-2}} = \frac{16\pi G_N}{\eta d L^{d-1}}\left[\langle T_{kk}\rangle - \delta\left(\frac{k^i}{2\pi\sqrt{h}}\frac{\delta S_{\text{ren}}}{\delta X^i}\right)\right]. \quad (3.2.17)$$

The nonnegativity of the term in brackets is equivalent to the QNEC. The case where d is even is more complicated, and we will go over the $d = 4$ case in Appendix A.5.

The Conformal QNEC

As mentioned in §3.2, we can get a stronger inequality from EWN by considering the norm of the vector s^μ, which is the part of $\delta\bar{X}^\mu$ orthogonal to the extremal surface. Our gauge choice $\bar{X}^z = z$ means that $s^\mu \neq \delta\bar{X}^\mu$, and so we get a nontrivial improvement by considering $s^2 \geq 0$ instead of $(\delta\bar{X})^2 \geq 0$.

We can actually use the results already derived above to compute s^2 with the following trick. We would have had $\delta\bar{X}^\mu = s^\mu$ if the surfaces of constant z were already orthogonal to the extremal surfaces. But we can change our definition of the constant-z surfaces with a coordinate transformation in the bulk to make this the case, apply the above results to $(\delta\bar{X})^2$ in the new coordinate system, and then transform back to the original coordinates. The coordinate transformation we are interested in performing is a PBH transformation [54], since it leaves the metric in Fefferman–Graham form, and so induces a Weyl transformation on the boundary.

So from the field theory point of view, we will just be calculating the consequences of EWN in a different conformal frame, which is fine because we are working with a CFT. With that in mind it is easy to guess the outcome: the best conformal frame to pick is one in which all of the non-conformally-invariant parts of the state-independent terms in $(\delta\bar{X})^2$ are set to zero, and when we transform the state-dependent term in the new frame back to the original frame we get the so-called Conformal QNEC first defined in [29]. This is indeed what happens, as we will now see.

Orthogonality Conditions First, we will examine in detail the conditions necessary for $\delta\bar{X}^\mu = s^\mu$, and their consequences on the inequalities derived above. We must check that

$$\bar{g}_{ij}\partial_\alpha \bar{X}^i \delta\bar{X}^j = 0. \tag{3.2.18}$$

for both $\alpha = z$ and $\alpha = a$. As above, we will expand these conditions in z. When $\alpha = z$, at lowest order in z we find the condition

$$0 = k_i X^i_{(2)}, \tag{3.2.19}$$

which is equivalent to $\theta_{(k)} = 0$. When $\alpha = a$, the lowest-order in z inequality is automatically satisfied because k^i is defined to be orthogonal to the entangling surface on the boundary. But at next-to-lowest order we find the condition

$$0 = k_i\partial_a X^i_{(2)} + e_{ai}\delta X^i_{(2)} + g^{(2)}_{ij}e^i_a k^j + X^m_{(2)}\partial_m g_{ij} e^i_a k^j \tag{3.2.20}$$

$$= -\frac{1}{2(d-2)}\left[(D_a - 2w_a)\theta_{(k)} + 2R_{ka}\right]. \tag{3.2.21}$$

Combined with the $\theta_{(k)} = 0$ condition, this tells us that that $D_a\theta_{(k)} = -2R_{ka}$ is required. When these conditions are satisfied, the state-dependent terms of $(\delta\bar{X})^2$ analyzed above become[5]

$$L^{-2}(\delta\bar{X})^2 = \frac{1}{d-2}\sigma^2_{(k)} + \left[\frac{1}{(d-2)^2(d-4)}(R_{ka})^2 + \frac{1}{2(d-2)(d-4)}(D_a\sigma^{(k)}_{bc})^2\right] z^2 + \cdots \tag{3.2.22}$$

Next we will demonstrate that $\theta_{(k)} = 0$ and $D_a\theta_{(k)} = -2R_{ka}$ can be achieved by a Weyl transformation, and then use that fact to write down the $s^2 \geq 0$ inequality that we are after.

Achieving $\delta\bar{X}^\mu = s^\mu$ with a Weyl Transformation Our goal now is to begin with a generic situation in which $\delta\bar{X}^\mu \neq s^\mu$ and use a Weyl transformation to set $\delta\bar{X}^\mu \to s^\mu$. This means finding a new conformal frame with $\hat{g}_{ij} = e^{2\phi(x)}g_{ij}$ such that $\hat{\theta}_{(k)} = 0$ and

[5]We have not included some terms at order z^2 which are proportional to $\sigma^{(k)}_{ab}$ because they never play a role in the EWN inequalities.

$\hat{D}_a\hat{\theta}_{(k)} = -2\hat{R}_{ka}$, which would then imply that $\delta\hat{X}^\mu = s^\mu$ (we omit the bar on $\delta\hat{X}^\mu$ to avoid cluttering the notation, but logically it would be $\delta\hat{\bar{X}}^\mu$).

Computing the transformation properties of the geometric quantities involved is a standard exercise, but there is one extra twist involved here compared to the usual prescription. Ordinarily a vector such as k^i would be invariant under the Weyl transformation. However, for our setup is it is important that k^i generate an affine-parameterized null geodesic. Even though the null geodesic itself is invariant under Weyl transofrmation, k^i will no longer be the correct generator. Instead, we have to use $\hat{k}^i = e^{-2\phi}k^i$. Another way of saying this is that $k_i = \hat{k}_i$ is invariant under the Weyl transformation. With this in mind, we have

$$e^{2\phi}\hat{R}_{ka} = R_{ka} - (d-2)\left[D_a\partial_k\phi - w_a\partial_k\phi - k_j K^j_{ab}\partial^b\phi - \partial_k\phi\partial_a\phi\right], \tag{3.2.23}$$

$$e^{2\phi}\hat{\theta}_{(k)} = \theta_{(k)} + (d-2)\partial_k\phi, \tag{3.2.24}$$

$$e^{2\phi}\hat{D}_a\hat{\theta}_{(k)} = D_a\theta_{(k)} + (d-2)D_a\partial_k\phi - 2\theta_{(k)}\partial_a\phi - 2(d-2)\partial_k\phi\partial_a\phi, \tag{3.2.25}$$

$$\hat{\sigma}^{(k)}_{ab} = \sigma^{(k)}_{ab}, \tag{3.2.26}$$

$$\hat{D}_c\hat{\sigma}^{(k)}_{ab} = D_c\sigma^{(k)}_{ab} - 2\left[\sigma^{(k)}_{c(b}\partial_{a)}\phi + \sigma^{(k)}_{ab}\partial_c\phi - g_{c(a}\sigma^{(k)}_{b)d}\nabla^d\phi\right], \tag{3.2.27}$$

$$\hat{w}_a = w_a - \partial_a\phi. \tag{3.2.28}$$

So we may arrange $\hat{\theta}_{(k)} = 0$ at a given point on the entangling surface by choosing $\partial_k\phi = -\theta_{(k)}/(d-2)$ that that point. Having chosen that, and assuming $\sigma^{(k)}_{ab}$ =0 at the same point, one can check that

$$e^{2\phi}\left(\hat{D}_a\hat{\theta}_{(k)} + 2\hat{R}_{ka}\right) = D_a\theta_{(k)} - 2w_a\theta_{(k)} + 2R_{ka} - (d-2)D_a\partial_k\phi \tag{3.2.29}$$

So we can choose $D_a\partial_k\phi$ to make the combination $\hat{D}_a\hat{\theta}_{(k)} + 2\hat{R}_{ka}$ vanish. Then in the new frame we have $\delta\hat{X}^\mu = s^\mu$.

The $s^2 \geq$ Inequality Based on the discussion above, we were able to find a conformal frame that allows us to compute the s^2. For the state-independent parts we have

$$L^{-2}s^2 = \frac{1}{d-2}\hat{\sigma}^2_{(k)} + \left[\frac{1}{(d-2)^2(d-4)}(\hat{R}_{ka})^2 + \frac{1}{2(d-2)(d-4)}(\hat{D}_a\hat{\sigma}^{(k)}_{bc})^2\right]\hat{z}^2 + \cdots \tag{3.2.30}$$

Here we also have a new bulk coordinate $\hat{z} = ze^\phi$ associated with the bulk PBH transformation. All we have to do now is transform back into the original frame to find s^2. Since $\hat{\theta}_{(k)} = \hat{D}_a\hat{\theta}_{(k)} + 2\hat{R}_{ka} = 0$, we actually have that

$$\hat{R}_{ka} = \hat{D}_a\hat{\theta}_{(k)} - \hat{w}_a\hat{\theta}_{(k)} - \hat{R}_{ka}, \tag{3.2.31}$$

which transforms homogeneously under Weyl transformations when $\sigma^{(k)}_{ab} = 0$. Thus, up to an overall scaling factor, we have

$$L^{-2}s^2 = \frac{1}{d-2}\sigma^2_{(k)} + \left[\frac{1}{(d-2)^2(d-4)}(D_a\theta_{(k)} - w_a\theta_{(k)} - R_{ka})^2 + \frac{1}{2(d-2)(d-4)}(D_a\sigma^{(k)}_{bc})^2\right] z^2 + \cdots, \tag{3.2.32}$$

where we have dropped terms of order z^2 which vanish when $\sigma^{(k)}_{ab} = 0$. As predicted, these terms are the conformally invariant contributions to $(\delta\bar{X})^2$.

In order to access the state-dependent part of s^2 we need the terms in (3.2.32) to vanish. Note that in $d = 3$ this always happens. In that case there is no z^2 term, and $\sigma^{(k)}_{ab} = 0$ always. Though our expression is singular in $d = 4$, comparing to (3.2.22) shows that actually the term in brackets above is essentially the same as the $z^2 \log z$ term in $\delta\bar{X}$. We already noted that this term was conformally invariant, so this is expected. The difference now is that we no longer need $\theta_{(k)} = 0$ in order to get to the QNEC in $d = 4$. In $d = 5$ the geometric conditions for the state-independent parts of s^2 to vanish are identical to those for $d = 4$, whereas in the $(\delta\bar{X})^2$ analysis we found that extra conditions were necessary. These were relics of the choice of conformal frame. Finally, for $d > 5$ there will be additional state-independent terms that we have not analyzed, but the results we have will still hold.

Conformal QNEC Now we analyze the state-dependent part of s^2 at order z^{d-2}. When all of the state-independent parts vanish, the state-dependent part is given by the conformal transformation of the QNEC. This is easily computed as follows:

$$L^{-2}\ s^2\big|_{z^{d-2}} = \frac{16\pi G_N}{\eta d L^{d-1}}\left[2\pi\langle\hat{T}_{ij}\rangle k^i k^j - \delta\left(\frac{k^i}{\sqrt{h}}\frac{\delta\hat{S}_{\rm ren}}{\delta X^i(y)}\right) - \frac{d}{2}\theta_{(k)}\left(\frac{k^i}{\sqrt{h}}\frac{\delta\hat{S}_{\rm ren}}{\delta X^i(y)}\right)\right]. \tag{3.2.33}$$

Of course, one would like to replace $\hat{T}_{ij}$ with T_{ij} and $\hat{S}_{\rm ren}$ with $S_{\rm ren}$. When d is odd this is straightforward, as these quantities are conformally invariant. However, when d is even there are anomalies that will contribute, leading to extra geometric terms in the conformal QNEC [55, 29].

3.3 Connection to Quantum Focusing

The Quantum Foscusing Conjecture

We start by reviewing the statement of the QFC [11, 53] before moving on to its connection to EWN and the QNEC. Consider a codimension-two Cauchy-splitting (i.e. entangling)

surface Σ and a null vector field k^i normal to Σ. Denote by $\mathcal{N}$ the null surface generated by k^i. The generalized entropy, S_{gen}, associated to Σ is given by

$$S_{\text{gen}} = \langle S_{\text{grav}} \rangle + S_{\text{ren}} \tag{3.3.1}$$

where S_{grav} is a state-independent local integral on Σ and S_{ren} is the renormalized von Neumann entropy of the interior (or exterior of Σ. The terms in S_{grav} are determined by the low-energy effective action of the theory in a well-known way [56]. Even though $\langle S_{\text{grav}} \rangle$ and S_{ren} individually depend on the renormalization scheme, that dependence cancels out between them so that S_{gen} is scheme-independent.

The generalized entropy is a functional of the entangling surface Σ, and the QFC is a statement about what happens when we vary the shape of Σ be deforming it within the surface $\mathcal{N}$. Specifically, consider a one-parameter family $\Sigma(\lambda)$ of cuts of $\mathcal{N}$ generated by deforming the original surface using the vector field k^i. Here λ is the affine parameter along the geodesic generated by k^i and $\Sigma(0) \equiv \Sigma$. To be more precise, let y^a denote a set of intrinsic coordinates for Σ, let h_{ab} be the induced metric on Σ, and let $X^i(y, \lambda)$ be the embedding functions for $\Sigma(\lambda)$. With this notation, $k^i = \partial_\lambda X^i$. The change in the generalized entropy is given by

$$\left.\frac{dS_{\text{gen}}}{d\lambda}\right|_{\lambda=0} = \int_\Sigma d^{d-2}y \, \frac{\delta S_{\text{gen}}}{\delta X^i(y)} \partial_\lambda X^i(y) \equiv \frac{1}{4G_N} \int_\Sigma d^{d-2}y \sqrt{h}\, \Theta[\Sigma, y] \tag{3.3.2}$$

This defines the quantum expansion $\Theta[\Sigma, y]$ in terms of the functional derivative of the generalized entropy:

$$\Theta[\Sigma, y] = 4G_N \frac{k^i(y)}{\sqrt{h}} \frac{\delta S_{\text{gen}}}{\delta X^i(y)}. \tag{3.3.3}$$

Note that we have suppressed the dependence of Θ on k^i in the notation, but the dependence is very simple: if $k^i(y) \to f(y)k^i(y)$, then $\Theta[\Sigma, y] \to f(y)\Theta[\Sigma, y]$.

The QFC is simple to state in terms of Θ. It says that Θ is non-increasing along the flow generated by k^i:

$$0 \geq \frac{d\Theta}{d\lambda} = \int_\Sigma d^{d-2}y \, \frac{\delta \Theta[\Sigma, y]}{\delta X^i(y')} k^i(y'). \tag{3.3.4}$$

Before moving on, let us make two remarks about the QFC.

First, the functional derivative $\delta\Theta[\Sigma, y]/\delta X^i(y')$ will contain local terms (i.e. terms proportional to δ-functions or derivatives of δ-functions with support at $y = y'$) as well as non-local terms that have support even when $y \neq y'$. S_{grav}, being a local integral, will only contribute to the local terms of $\delta\Theta[\Sigma, y]/\delta X^i(y')$. The renormalized entropy S_{ren} will contribute both local and non-local terms. The non-local terms can be shown to be nonpositive using strong subadditivity of the entropy [11], while the local terms coming from S_{ren} are in general extremely difficult to compute.

Second, and more importantly for us here, the QFC as written in (3.3.4) does not quite make sense. We have to remember that S_{grav} is really an operator, and its expectation value

$\langle S_{\text{grav}} \rangle$ is really the thing that contributes to Θ. In order to be well-defined in the low-energy effective theory of gravity, this expectation value must be smeared over a scale large compared to the cutoff scale of the theory. Thus when we write an inequality like (3.3.4), we are implicitly smearing in y against some profile. The profile we use is arbitrary as long as it is slowly-varying on the cutoff scale. This extra smearing step is necessary to avoid certain violations of (3.3.4), as we will see below [53].

QNEC from QFC

In this section we will explicitly evaluate the QFC inequality, (3.3.4), and derive the QNEC in curved space from it as a nongravitational limit. We consider theories with a gravitational action of the form

$$I_{\text{grav}} = \frac{1}{16\pi G_N} \int \sqrt{g} \left(R + \ell^2 \lambda_1 R^2 + \ell^2 \lambda_2 R_{ij} R^{ij} + \ell^2 \lambda_{\text{GB}} \mathcal{L}_{\text{GB}} \right) \tag{3.3.5}$$

where $\mathcal{L}_{GB} = R^2_{ijmn} - 4R^2_{ij} + R^2$ is the Gauss-Bonnet Lagrangian. Here ℓ is the cutoff length scale of the effective field theory, and the dimensionless couplings λ_1, λ_2, and λ_{GB} are assumed to be renormalized.

The generalized entropy functional for these theories can be computed using standard replica methods [56] and takes the form

$$S_{\text{gen}} = \frac{A[\Sigma]}{4G_N} + \frac{\ell^2}{4G_N} \int_\Sigma \sqrt{h} \left[2\lambda_1 R + \lambda_2 \left(R_{ij} N^{ij} - \frac{1}{2} K_i K^i \right) + 2\lambda_{\text{GB}} r \right] + S_{\text{ren}}. \tag{3.3.6}$$

Here $A[\Sigma]$ is the area of the entangling surface, N^{ij} is the projector onto the normal space of Σ, K^i is the trace of the extrinsic curvature of Σ, and r is the intrinsic Ricci scalar of Σ.

We can easily compute Θ by taking a functional derivative of (3.3.6), taking care to integrate by parts so that the result is proportional to $k^i(y)$ and not derivatives of $k^i(y)$. One finds

$$\begin{aligned} \Theta = \theta_{(k)} + \ell^2 \Bigg[& 2\lambda_1 (\theta_{(k)} R + \nabla_k R) + \lambda_2 \left((D_a - w_a)^2 \theta_{(k)} + K_i K^{iab} K^k_{ab} \right. \\ & + \theta_{(k)} R_{klkl} + \nabla_k R - 2\nabla_l R_{kk} + \theta_{(k)} R_{kl} - \theta_{(l)} R_{kk} + 2K^{kab} R_{ab} \Big) \\ & \left. - 4\lambda_{GB} \left(r^{ab} K^k_{ab} - \frac{1}{2} r \theta_{(k)} \right) \right] + 4G_N \frac{k^i}{\sqrt{h}} \frac{\delta S_{\text{ren}}}{\delta X^i} \end{aligned} \tag{3.3.7, 3.3.8}$$

Now we must compute the λ-derivative of Θ. When we do this, the leading term comes from the derivative of $\theta_{(k)}$, which by Raychaudhuri's equation contains the terms $\theta^2_{(k)}$ and $\sigma^2_{(k)}$. Since we are ultimately interested in deriving the QNEC as the non-gravitational limit of the QFC, we need to set $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$ so that the nongravitational limit is not dominated

by those terms. So for the rest of this section we will set $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$ at the point of evaluation (but not globally!). Then we find

$$\begin{aligned}\frac{d\Theta}{d\lambda} &= -R_{kk} + 2\lambda_1\ell^2\left(\nabla_k^2 R - RR_{kk}\right)\\ &+ \lambda_2\ell^2\Big[2D_a(w^aR_{kk}) + \nabla_k^2R - D_aD^aR_{kk} - \frac{d}{d-2}(D_a\theta_{(k)})^2 - 2R_{kb}D^b\theta_{(k)} - 2(D_a\sigma_{bc})^2\\ &- 2\nabla_k\nabla_l R_{kk} - 2R_{kakb}R^{ab} - \theta_{(l)}\nabla_k R_{kk}\Big] - 2\lambda_{\rm GB}\ell^2\left[\frac{d(d-3)(d-4)}{(d-1)(d-2)^2}RR_{kk}\right.\\ &\left.- 4\frac{(d-4)(d-3)}{(d-2)^2}R_{kk}R_{kl} - \frac{2(d-4)}{d-2}C_{klkl}R_{kk} - \frac{2(d-4)}{d-2}R^{ab}C_{akbk} + 4C^{kalb}C_{kakb}\right]\\ &+ 4G_N\frac{d}{d\lambda}\left(\frac{k^i}{\sqrt{h}}\frac{\delta S_{\rm ren}}{\delta X^i}\right)\end{aligned} \tag{3.3.9}$$

This expression is quite complicated, but it simplifies dramatically if we make use of the equation of motion coming from (3.3.5) plus the action of the matter sector. Then we have $R_{kk} = 8\pi GT_{kk} - H_{kk}$ where [57]

$$\begin{aligned}H_{kk} &= 2\lambda_1\left(RR_{kk} - \nabla_k^2R\right) + \lambda_2\Big(2R_{kikj}R^{ij} - \nabla_k^2R + 2\nabla_k\nabla_lR_{kk} - 2R_{klki}R^i_k\\ &+D_cD^cR_{kk} - 2D_c(w^cR_{kk}) - 2(D_b\theta_{(k)} + R_{bmkj}P^{mj})R^b_k + \theta_{(l)}\nabla_kR_{kk}\Big)\\ &+ 2\lambda_{\rm GB}\left(\frac{d(d-3)(d-4)}{(d-1)(d-2)^2}RR_{kk} - 4\frac{(d-4)(d-3)}{(d-2)^2}R_{kk}R_{kl} - 2\frac{d-4}{d-2}R^{ij}C_{kikj} + C_{kijm}C_k{}^{ijm}\right)\end{aligned} \tag{3.3.10}$$

For the Gauss-Bonnet term we have used the standard decomposition of the Riemann tensor in terms of the Weyl and Ricci tensors. Using similar methods to those in Appendix A.4, we have also exchanged $k^ik^j\Box R_{ij}$ in the R^2_{ij} equation of motion for surface quantities and ambient curvatures.

After using the equation of motion we have the relatively simple formula

$$\begin{aligned}\frac{d\Theta}{d\lambda} &= -\lambda_2\ell^2\left(\frac{d}{d-2}(D_a\theta_{(k)})^2 + 4R^b_kD_b\theta_{(k)} + 2R_{bk}R^b_k + 2(D_a\sigma^{(k)}_{bc})^2\right)\\ &+ 2\lambda_{\rm GB}\ell^2\left(C_{kabc}C_k{}^{abc} - 2C_{kba}{}^bC_{kc}{}^{ac}\right) + 4G_N\frac{d}{d\lambda}\left(\frac{k^i}{\sqrt{h}}\frac{\delta S_{\rm ren}}{\delta X^i}\right) - 8\pi G_N\left\langle T_{kk}\right\rangle\end{aligned} \tag{3.3.11}$$

The Gauss-Bonnet term agrees with the expression derived in [52]. However unlike [52] we have not made any perturbative assumptions about the background curvature.

At first glance it seems like (3.3.11) does not have definite sign, even in the non-gravitational limit, due to the geometric terms proportional to λ_2 and $\lambda_{\rm GB}$. The difficulty posed by the Gauss-Bonnet term, in particular, was first pointed out in [48]. However, this is where we have to remember the smearing prescription mentioned in §3.3. We must integrate (3.3.11)

over a region of size larger than ℓ before testing its nonpositivity. The crucial point, used in [53], is that we must also remember to integrate the terms $\theta_{(k)}^2$ and $\sigma_{(k)}^2$ that we dropped earlier over the same region. When we integrate $\theta_{(k)}^2$ over a region of size ℓ centered at a point where $\theta_{(k)} = 0$, the result is $\xi\ell^2(D_a\theta_{(k)})^2 + o(\ell^2)$, where $\xi \gtrsim 10$ is a parameter associated with the smearing profile. A similar result holds for $\sigma_{ab}^{(k)}$. Thus we arrive at

$$\begin{aligned}\frac{d\Theta}{d\lambda} &= -\frac{\xi}{d-2}\ell^2(D_a\theta_{(k)})^2 - \xi\ell^2(D_a\sigma_{bc}^{(k)})^2 \\ &\quad - \lambda_2\ell^2\left(\frac{d}{d-2}(D_a\theta_{(k)})^2 + 4R_k^bD_b\theta_{(k)} + 2R_{bk}R_k^b + 2(D_a\sigma_{bc}^{(k)})^2\right) \\ &\quad + 2\lambda_{\rm GB}\ell^2\left(C_{kabc}C_k{}^{abc} - 2C_{kba}{}^bC_{kc}{}^{ac}\right) \\ &\quad + 4G_N\frac{d}{d\lambda}\left(\frac{k^i}{\sqrt{h}}\frac{\delta S_{\rm ren}}{\delta X^i}\right) - 8\pi G_N\langle T_{kk}\rangle + o(\ell^2) \end{aligned} \tag{3.3.12}$$

Since the size of ξ is determined by the validity of the effective field theory, by construction the terms proportional to ξ in (3.3.12) dominate over the others. Thus in order to take the non-gravitational limit, we must eliminate these smeared terms.

Clearly we need to be able to choose a surface such that $D_a\theta_{(k)} = D_a\sigma_{bc}^{(k)} = 0$. Then smearing $\theta_{(k)}^2$ and $\sigma_{(k)}^2$ would only produce terms of order ℓ^4 (terms of that order would also show up from smearing the operators proportional to λ_2 and $\lambda_{\rm GB}$). As explained in [53], this is only possible given certain conditions on the background spacetime at the point of evaluation. We must have

$$C_{kabc} = \frac{1}{d-2}h_{ab}R_{kc} - \frac{1}{d-2}h_{ac}R_{kb}. \tag{3.3.13}$$

This can be seen by using the Codazzi equation for Σ. Imposing this condition, which allows us to set $D_a\theta_{(k)} = D_a\sigma_{bc}^{(k)} = 0$, we then have.

$$\begin{aligned}\frac{d\Theta}{d\lambda} &= -2\ell^2\left(\lambda_2 + 2\frac{(d-3)(d-4)}{(d-2)^2}\lambda_{\rm GB}\right)R_{bk}R_k^b \\ &\quad + 4G_N\frac{d}{d\lambda}\left(\frac{k^i}{\sqrt{h}}\frac{\delta S_{\rm ren}}{\delta X^i}\right) - 8\pi G_N\langle T_{kk}\rangle + o(\ell^3). \end{aligned} \tag{3.3.14}$$

This is the quantity which must be negative according to the QFC. In deriving it, we had to assume that $\theta_{(k)} = \sigma^{(k)} = D_a\theta_{(k)} = D_a\sigma_{bc}^{(k)} = 0$.

We make two observations about (3.3.14). First, if we assume that $R_{ka} = 0$ as an additional assumption and take $\ell \to 0$, then we arrive at the QNEC as long as $G_N > o(\ell^3)$. This is the case when ℓ scales with the Planck length and $d \leq 5$. These conditions are similar to the ones we found previously from EWN, and below in §3.3 we will discuss that in more detail.

The second observation has to do with the lingering possibility of a violation of the QFC due to the terms involving the couplings. In order to have a violation, one would need the

linear combination

$$\lambda_2 + 2\frac{(d-3)(d-4)}{(d-2)^2}\lambda_{\rm GB} \tag{3.3.15}$$

to be negative. Then if one could find a situation where the first line of (3.3.14) dominated over the second, there would be a violation. It would be interesting to interpret this as a bound on the above linear combination of couplings coming from the QFC, but it is difficult to find a situation where the first line of (3.3.14) dominates. The only way for R_{ka} to be large compared to the cutoff scale is if T_{ka} is nonzero, in which case we would have $R_{ka} \sim G_N T_{ka}$. Then in order for the first line of (3.3.14) to dominate we would need

$$G_N \ell^2 T_{ka} T^a_k \gg T_{kk}. \tag{3.3.16}$$

As an example, for a scalar field Φ this condition would say

$$G_N \ell^2 (\partial_a \Phi)^2 \gg 1. \tag{3.3.17}$$

This is not achievable within effective field theory, as it would require the field to have super-Planckian gradients. We leave a detailed and complete discussion of this issue to future work.

Scheme-Independence of the QNEC

We take a brief interlude to discuss the issue of the scheme-dependence of the QNEC, which will be important in the following section. It was shown in [52], under some slightly stronger assumptions than the ones we have been using, that the QNEC is scheme-independent under the same conditions where we expect it to hold true. Here we will present our own proof of this fact, which actually follows from the manipulations we performed above involving the QFC.

In this section we will take the point of view of field theory on curved spacetime without dynamical gravity. Then each of the terms in $I_{\rm grav}$, defined above in (3.3.5), are completely arbitrary, non-dynamical terms we can add to the Lagrangian at will.[6] Dialing the values of those various couplings corresponds to a choice of *scheme*, as even though those couplings are non-dynamical they will still contribute to the definitions of quantities like the renormalized energy-momentum tensor and the renormalized entropy (as defined through the replica trick). The QNEC is scheme-independent if it is insensitive to the values of these couplings.

To show the scheme-independence of the QNEC, we will begin with the statement that $S_{\rm gen}$ is scheme-independent. We remarked on this above, when our context was a theory with dynamical gravity. But the scheme-independence of $S_{\rm gen}$ does not require use of the equations of motion, so it is valid even in a non-gravitational theory on a fixed background.

[6] We should really be working at the level of the quantum effective action, or generating functional, for correlation functions of T_{ij} [48]. The geometrical part has the same form as the classical action $I_{\rm grav}$ and so does not alter this discussion.

In fact, only once in the above discussion did we make use of the gravitational equations of motion, and that was in deriving (3.3.11). Following the same steps up to that point, but without imposing the gravitational equations of motion, we find instead

$$\frac{d\Theta}{d\lambda} = -\lambda_2 \ell^2 \left(\frac{d}{d-2} (D_a \theta_{(k)})^2 + 4 R^b_k D_b \theta_{(k)} + 2 R_{bk} R^b_k + 2 (D_a \sigma_{bc})^2 \right) + 2\lambda_{\rm GB} \ell^2 \left(C_{kabc} C_k{}^{abc} - 2 C_{kba}{}^b C_{kc}{}^{ac} \right) + 4 G_N \frac{d}{d\lambda} \left(\frac{k^i}{\sqrt{h}} \frac{\delta S_{\rm ren}}{\delta X^i} \right) - k_i k_j \frac{16\pi G_N}{\sqrt{g}} \frac{\delta I_{\rm grav}}{\delta g_{ij}}. \tag{3.3.18}$$

Since the theory is not gravitational, we would not claim that this quantity has a sign. However, it is still scheme-independent.

To proceed, we will impose all of the additional conditions that are necessary to prove the QNEC. That is, we impose $D_b \theta_{(k)} = R^b_k = D_a \sigma_{bc} = 0$, as well as $\theta_{(k)} = \sigma^{(k)}_{ab} = 0$, which in turn requires $C_{kabc} = 0$. Under these conditions, we learn that the combination

$$\frac{d}{d\lambda} \left(\frac{k^i}{\sqrt{h}} \frac{\delta S_{\rm ren}}{\delta X^i} \right) - k_i k_j \frac{4\pi}{\sqrt{g}} \frac{\delta I_{\rm grav}}{\delta g_{ij}} \tag{3.3.19}$$

is scheme-independent. The second term here is one of the contributions to the renormalized $2\pi \langle T_{kk} \rangle$ in the non-gravitational setup, the other contribution being $k_i k_j \frac{4\pi}{\sqrt{g}} \frac{\delta I_{\rm matter}}{\delta g_{ij}}$. But $I_{\rm matter}$ is already scheme-independent in the sense we are discussing, in that it is independent of the parameters appearing in $I_{\rm grav}$. So adding that to the terms we have above, we learn that

$$\frac{d}{d\lambda} \left(\frac{k^i}{\sqrt{h}} \frac{\delta S_{\rm ren}}{\delta X^i} \right) - 2\pi \langle T_{kk} \rangle \tag{3.3.20}$$

is scheme-independent. This is what we wanted to show.

QFC vs EWN

As we have discussed above, by taking the non-gravitational limit of (3.3.14) under the assumptions $D_b \theta_{(k)} = R^b_k = D_a \sigma_{bc} = \theta_{(k)} = \sigma^{(k)}_{ab} = 0$ we find the QNEC as a consequence of the QFC (at least for $d \leq 5$). And under the same set of geometric assumptions, we found the QNEC as a consequence of EWN in (3.2.17). The discussion of the previous section demonstrates that these assumptions also guarantee that the QNEC is scheme-independent. So even though these two QNEC inequalities were derived in different ways, we know that at the end of the day they are the same QNEC. It is natural to ask if there is a further relationship between EWN and the QFC, beyond the fact that they give the same QNEC. We will begin to investigate that question in this section.

The natural thing to ask about is the state-independent terms in the QFC and in $(\delta\bar{X})^2$. We begin by writing down all of the terms of $(\delta\bar{X})^2$ in odd dimensions that we have computed:

$$\begin{aligned}(d-2)L^{-2}(\delta\bar{X}^i)^2 = &\frac{1}{(d-2)}\theta_{(k)}^2 + \sigma_{(k)}^2 \\ &+ z^2\frac{1}{4(d-2)}(D_a\theta_{(k)} + 2R_{ka})^2 \\ &+ z^2\frac{1}{(d-2)(d-4)}(D_a\theta_{(k)} + R_{ka})^2 + z^2\frac{1}{2(d-4)}(D_a\sigma_{bc}^{(k)})^2 \\ &+ z^2\frac{\kappa}{d-4}\left(C_{kabc}C_k^{\ abc} - 2C_k^{\ c}{}_{ca}C_k^{\ b}{}_b^{\ a}\right) \\ &+ \cdots + z^{d-2}\frac{16\pi(d-2)G_N}{\eta d L^{d-1}}\left[\langle T_{kk}\rangle - \delta\left(\frac{k^i}{2\pi\sqrt{h}}\frac{\delta S_{\text{ren}}}{\delta X^i}\right)\right]. \end{aligned} \tag{3.3.21}$$

The first line looks like $-\dot{\theta}$, which would be the leading term in $d\Theta/d\lambda$, except it is missing an R_{kk}. Of course, we eventually got rid of the R_{kk} in the QFC by using the equations of motion. Suppose we set $\theta_{(k)} = 0$ and $\sigma_{ab}^{(k)} = 0$ to eliminate those terms, as we did with the QFC. Then we can write $(\delta\bar{X})^2$ suggestively as

$$\begin{aligned}(d-2)L^{-2}(\delta\bar{X}^i)^2 = &z^2\tilde{\lambda}_2\Big(\frac{d}{(d-2)}(D_a\theta_k)^2 + 4R_k^a D_a\theta + \frac{4(d-3)}{(d-2)}R_{ka}R_k^a + 2(D_a\sigma_{bc}^{(k)})^2\Big) \\ &- 2z^2\tilde{\lambda}_{\text{GB}}\left(C_{kabc}C_k^{\ abc} - 2C_k^{\ c}{}_{ca}C_k^{\ b}{}_b^{\ a}\right) \\ &+ \cdots + 8\pi\tilde{G}_N\langle T_{kk}\rangle - 4\tilde{G}_N\delta\left(\frac{k^i}{\sqrt{h}}\frac{\delta S_{\text{ren}}}{\delta X^i}\right). \end{aligned} \tag{3.3.22}$$

where

$$\tilde{G}_N = G_N\frac{2(d-2)z^{d-2}}{\eta d L^{d-1}}, \tag{3.3.23}$$

$$\tilde{\lambda}_2 = \frac{1}{4(d-4)}, \tag{3.3.24}$$

$$\tilde{\lambda}_{\text{GB}} = -\frac{\kappa}{2(d-4)}. \tag{3.3.25}$$

Written this way, it almost seems like $(d-2)L^{-2}(\delta\bar{X}^i)^2 \sim -d\Theta/d\lambda$ in some kind of model gravitational theory. One discrepancy is in the coefficient of the $R_{ka}R^{ka}$ term, unless $d = 4$. It is also intriguing that the effective coefficients $\tilde{G}_N$, $\tilde{\lambda}_2$, and $\tilde{\lambda}_{\text{GB}}$ are close to, but not exactly the same as, the effective braneworld induced gravity coefficients found in [58]. This is clearly something that deserves further study.

3.4 Discussion

We have displayed a strong similarity between the state-independent inequalities in the QFC and the state-independent inequalities from EWN. We now discuss several possible future

directions and open questions that follow naturally from these results.

Bulk Entropy Contributions

We ignored the bulk entropy S_{bulk} in this work, but we know that it produces a contribution to CFT entropy [5] and plays a role in the position of the extremal surface [6, 49]. The bulk entropy contributions to the entropy are subleading in N^2 and do not interfere with the gravitational terms in the entropy. We could include the bulk entropy as a source term in the equations determining $\bar{X}$, which could lead to extra contributions to the $X_{(n)}$ coefficients. However, it does not seem possible for the bulk entropy to have an effect on the state-independent parts of the extremal surface, namely on $X_{(n)}$ for $n < d$, which means the bulk entropy would not affect the conditions we derived for when the QNEC should hold.

Another logical possibility is that the bulk entropy term could affect the statement of the QNEC itself, meaning that the schematic form $T_{kk} - S''$ would be altered. This would be problematic, especially given that the QFC always produces a QNEC of that same form. It was argued in [47] that this does not happen, and that argument holds here as well.

Smearing of EWN

We were careful to include a smearing prescription for defining the QFC, and it was an important ingredient in the analysis of §3.3. But what about smearing of EWN? Of course, the answer is that we *should* smear EWN appropriately, but as we will see now it would not make a difference to our analysis,

The issue is that the bulk theory is a low-energy effective theory of gravity with a cutoff scale ℓ, and the quantities that we use to probe EWN, like $(\delta\bar{X})^2$, are operators in that theory. As such, these operators need to be smeared over a region of proper size ℓ on the extremal surface. Of course, due to the warp factor, such a region has coordinate size $z\ell/L$. We can ask what effect such a smearing would have on the inequality $(\delta\bar{X})^2$.

When we performed our QNEC derivation, we assumed that $\theta_{(k)} = 0$ at the point of evaluation, so that the $\theta_{(k)}^2$ term in $(\delta\bar{X})^2\big|_{z^0}$ would not contribute. However, after smearing this term would contribute a term of the form $\ell^2(D_a\theta_{(k)})^2/L^2$ to $(\delta\bar{X})^2\big|_{z^2}$. But we already had such a term at this order, so all this does is shift the coefficient. Furthermore, the coefficient is shifted only by an amount of order ℓ^2/L^2. If the cutoff ℓ is of order the Planck scale, then this is suppressed in powers of N^2. In other words, this effect is negligible for the analysis. A similar statement applies for $\sigma_{ab}^{(k)}$. So in summary, EWN should be smeared, but the analysis we performed was insensitive to it.

Future Work

There are a number of topics that merit investigation in future work. We will touch on a few of them to finish our discussion.

Relevant Deformations Perhaps the first natural extension of our work is to include relevant deformations in the EWN calculation. There are a few reasons why this is interesting. First, one would like to test the continued correspondence between the QFC and EWN when it comes to the QNEC. The QFC arguments do not care whether relevant deformations are turned on, so one would expect that the same is true in EWN. This is indeed the case when the boundary theory is formulated on flat space [29], and one would expect similar results to hold when the boundary is curved.

Another reason to add in relevant deformations is to test the status of the Conformal QNEC when the theory is not a CFT. To be more precise, the $(\delta\bar{X})^2$ and s^2 calculations we performed differed by a Weyl transformation on the boundary, and since our boundary theory was a CFT this was a natural thing to do. When the boundary theory is not a CFT, what is the relationship between $(\delta\bar{X})^2$ and s^2? One possibility, perhaps the most likely one, is that they simply reduce to the same inequality, and the Conformal QNEC no longer holds. It would be good to know the answer.

Finally, and more speculatively, having a relevant deformation turned on when the background is curved allows for interesting state-independent inequalities from EWN. We saw that for a CFT the state-independent terms in both $(\delta\bar{X})^2$ and s^2 were trivially positive. Perhaps when a relevant deformation is turned on then more nontrivial things might happen, such as the possibility of a c-theorem hiding inside of EWN. We are encouraged by the similarity of inequalities used in recent proofs of the c-theorems to inequalities obtained from EWN [59].

Higher Dimensions Another pressing issue is extending our results to $d = 6$ and beyond. This is an algebraically daunting task using the methods we have used for $d \leq 5$. Considering the ultimate simplicity of our final expressions, especially compared to the intermediate steps in the calculations, it is likely that there are better ways of formulating and performing the analyses we performed here. It is hard to imagine performing the full $d = 6$ analysis without such a simplification.

Further Connections Between EWN and QFC Despite the issues outlined in §3.3, we are still intrigued by the similarities between EWN and the QFC. It is extremely natural to couple the boundary theory in AdS/CFT to gravity using a braneworld setup [60, 61, 62, 58]. Upon doing this, one can formulate the QFC on the braneworld. However, at the same time near-boundary EWN becomes lost, or at least changes form: extremal surfaces anchored to a brane will in general not be orthogonal to the brane, and in that case a null deformation on the brane will induce a timelike deformation of the extremal surface in the vicinity of the brane. Of course, one has to be careful to take into account the uncertainty in the position of the brane, which complicates things. We hope that such an analysis could serve to unify the QFC with EWN, or at least illustrate their relationship with each other.

Conformal QNEC from QFC While we emphasized the apparent similarity between the EWN-derived inequality $(\delta\bar{X})^2 \geq 0$ and the QFC, the stronger EWN inequality $s^2 \geq 0$ is nowhere to be found in the QFC discussion. It would be inteesting to see if there was some direct QFC-like way to derive the Conformal QNEC (rather than first deriving the ordinary QNEC and then performing a Weyl transformation). In particular, the Conformal QNEC applies even in cases where $\theta_{(k)}$ is nonzero, while in those cases the QFC is dominated by classical effects. Perhaps there is a useful change of variables that one can do in the semiclassical gravity when the matter sector is a CFT which makes the Conformal QNEC manifest from the QFC point of view. This is worth exploring.

Chapter 4

Energy is Entanglement

4.1 Introduction

The connection between quantum information and energy has been an emerging theme of recent progress in quantum field theory. Causality combined with universal inequalities like positivity and monotonicity of relative entropy can be used to derive many interesting energy-entropy bounds. Examples include the Bekenstein bound [63], the quantum Bousso bound [24, 64], the Averaged Null Energy Condition (ANEC) [27, 65], and the Quantum Null Energy Condition (QNEC) [16, 29, 17, 66]. Here we strengthen the energy-entropy connection, moving from bounds to equalities.

The key insight of the QNEC, which we will exploit, is that one should look at variations of the entropy S of a region as the region is deformed. Consider the entropy as a functional of the entangling surface embedding functions X^μ. Then one can compute the functional derivative $\delta^2 S/\delta X^\mu(y)\delta X^\nu(y')$ which encodes how the entropy depends on the shape of the region. In general, this second variation will contain contact, or "diagonal," terms, proportional to δ-functions and derivatives of δ-functions, as well as "off-diagonal" terms. Our interest here is in the δ-function contact term, and we introduce $S''_{\mu\nu}$ as the coefficient of the δ-function:

$$\frac{\delta^2 S}{\delta X^\mu(y)\delta X^\nu(y')} = S''_{\mu\nu}(y)\delta^{(d-2)}(y-y') + \cdots \tag{4.1.1}$$

Null Variations First consider the null-null component of the second variation, $S''_{vv}(y)$, where v is a null coordinate in a direction orthogonal to the entangling surface at the point y.[1] Suppose the entangling surface is locally restricted to lie in the null plane orthogonal to v near the point y. With this setup we can apply the QNEC, which says $S''_{vv} \geq 2\pi\langle T_{vv}\rangle$. Our main conjecture is that this inequality is always saturated:[2]

$$S''_{vv} = 2\pi\langle T_{vv}\rangle. \tag{4.1.2}$$

[1] We are restricting attention to field theories in Minkowski space throughout the main text.

[2] In [67] the issue of QNEC saturation was also investigated, but this is a different notion of saturation. Their analysis did not isolate the δ-function component, and instead considered the total variation in the

We believe this holds for all relativistic quantum field theories with an interacting UV fixed point in $d > 2$ dimensions. For the special case of an interacting CFT this fully specifies the stress tensor in terms of entropy variations: by considering (4.1.2) for all entangling surfaces passing through a point, $\langle T_{\mu\nu}\rangle$ is completely determined up to a trace term. In a CFT the trace of the stress tensor vanishes, and so the entropy variations determine the full stress tensor in that case. This is the sense in which energy comes from entanglement.

Our primary evidence for (4.1.2) is holographic, as explained below. But if we restrict attention to quantities that can be built out of local expectation values of operators and the local surface geometry there is no other possibility for S''_{vv}. A significant constraint comes from considering the vacuum modular Hamiltonian, K, which is defined by

$$S(\sigma + \delta\sigma) - S(\sigma) = \mathrm{Tr}\,(K\delta\sigma) + O\big(\delta\sigma^2\big)\,, \tag{4.1.3}$$

where σ is the vacuum state reduced to the region under consideration and $\delta\sigma$ is an arbitrary perturbation of the state. If we had a general formula for S in terms of expectation values of operators, we would be able to read off the modular Hamiltonian from the terms in that formula linear in expectation values.[3] For a region bounded by an entangling surface restricted to a null plane the modular Hamiltonian has a known formula in terms of the stress tensor [68], and in particular we have

$$K''_{vv} = 2\pi\langle T_{vv}\rangle. \tag{4.1.4}$$

That is why $\langle T_{vv}\rangle$ is the only possible linear term we could have had in (4.1.2).

A nonlinear contribution to S''_{vv}, such as a product of expectation values, is restricted by dimensional analysis and unitarity bounds: the only possibility is if the theory contains a free field. Then we can take the classical expression for T_{vv}, which is quadratic in the field, and replace each of those fields with expectation values to get an expression quadratic in expectation values with the right dimensionality to contribute to S''_{vv}. For interacting fields, nonzero anomalous dimensions prevent this from working. We will say more about free theories in Appendix A.7, where we will see that this possibility is realized by a term $\sim \langle\partial_v\phi\rangle^2$ for a free scalar field, which is why we limit ourselves to interacting theories in the main text. The substance of (4.1.2), then, is the statement that there are no non-local contributions to S''_{vv}.

Relative Entropy There is a natural interpretation of (4.1.2) in terms of relative entropy. The relative entropy of a state ρ and a reference state σ—for us, the vacuum—is a measure of the distinguishability of the two states. We will denote the relative entropy of ρ and the vacuum by $S_{\mathrm{rel}}(\rho)$. By definition, the relative entropy is

$$S_{\mathrm{rel}}(\rho) = \Delta\langle K\rangle - \Delta S, \tag{4.1.5}$$

entropy including the contribution of off-diagonal terms. So the examples in [67] where the QNEC is not "saturated" are not in contradiction with our results.

[3]For simplicity of the discussion we set all vacuum expectation values to zero.

where $\Delta\langle K\rangle$ and ΔS denote the vacuum-subtracted modular energy and vacuum-subtracted entropy, respectively. A consequence of (4.1.2) is that $\Delta S''_{vv} = \Delta\langle K''_{vv}\rangle$, so we can say that

$$S''_{\text{rel},vv} = 0. \tag{4.1.6}$$

This equation is implied by (4.1.2) but is weaker, since it does not require us to know what the modular Hamiltonian actually is. The extra information of (4.1.2) is the expression (4.1.4) for the second variation of the modular Hamiltonian. It can be useful to formulate our results in terms of relative entropy instead of entropy itself because relative entropy is generally free from UV divergences, at least for nice states.[4]

Non-Null Deformations Now let us move beyond the null case. Our goal in doing this is to understand the simplest setup where non-null deformations can be analyzed, and so we will make several additional restrictions that we do not make in the null case. As explained in [69, 52] and below in Section 4.2, (4.1.2) for the null case is a well-defined, finite equation in field theory. Local stationarity conditions on the entangling surface are enough to eliminate state-independent geometric divergences in the entropy, and the remaining state-dependent divergences cancel between the entropy and stress tensor. In the non-null case, eliminating divergences is more difficult. State-independent divergences can be dealt with by considering the vacuum-subtracted entropy ΔS rather than just S. State-dependent divergences associated with low-lying operators in the theory are more problematic. To eliminate these divergences, it is enough to restrict our attention theories where all relevant couplings have mass dimension greater than $d/2$, and to states where operators of dimension $\Delta \leq d/2$ have vanishing expectation values near the entangling surface. The idea of these restrictions is to make sure there are no parameters with scaling dimension small enough to contribute to divergences. We will make the further restriction in the non-null case to planar entangling surfaces, and this last restriction is made purely to simplify the analysis and presentation. With these assumptions in place we find

$$\Delta S''_{\mu\nu} = 2\pi \left(n^\rho_\mu n^\sigma_\nu \langle T_{\rho\sigma}\rangle + \frac{d^2-3d-2}{2(d+1)(d-2)} n_{\mu\nu} h^{ab} \langle T_{ab}\rangle \right), \tag{4.1.7}$$

where $n_{\mu\nu}$ is the normal projector to the entangling surface and h_{ab} is the intrinsic metric on the entangling surface.[5] Note that (6.3.8) implies that $S''_{\text{rel},\mu\nu} = 0$.

[4] It is possible for relative entropy to be infinite, for instance if we take our region to be the whole space and consider two orthogonal pure states. This is an expected and understood type of infinity, and not dependent on a choice of UV regulator.

[5] In [66], a quantum version of the dominant energy condition which involved spacelike deformations of entropy was proposed for $d = 2$ dimensions. In that inequality, timelike components of the stress tensor were bounded by spacelike components of the entropy variation, whereas in (6.3.8) timelike components of the stress tensor are related to timelike components of $\Delta S''_{\mu\nu}$ (ignoring the second term of (6.3.8), which is absent in two dimensions). Our techniques are not directly applicable to two dimensions, and a naïve extrapolation of (6.3.8) is probably incorrect, but it would interesting to investigate this issue further in the future.

Consequences for Field Theory and Gravity We view (4.1.2) and (6.3.8) as deep truths about interacting quantum field theories, worthy of further study. At present, our evidence for these conjectures comes from holography. We will calculate $S''_{\mu\nu}$ directly and prove that (4.1.2) and (6.3.8) hold precisely at leading order in large-N for all bulk states. We will also argue that subleading corrections in $1/N$ do not alter these conclusions. While this does not amount to a full proof, it is enough evidence for us to posit that (4.1.2) is true universally, and that (6.3.8) holds with relatively few additional assumptions.

An immediate application, which we discuss in Section 4.6, is to gravity. If we couple our field theory to gravity, then we can effectively isolate the δ-function part of the null second variation by deforming the entangling surface over a Planck-sized, or slightly larger, domain. According to the Raychaudhuri equation, if the surface is locally stationary then the leading change in its area due to this deformation is determined by R_{vv}, the null-null component of the Ricci tensor. Using (4.1.2) together with Einstein's equations, $R_{vv} = 8\pi G_N T_{vv}$, we learn that this change in area is precisely canceled by $4G_N S''_{vv}$. This means that the leading-order change in generalized entropy—area in Planck units plus entropy—is actually zero under such a deformation. In Section 4.6 we will show how this argument can also be reversed, demonstrating that this leading-order cancellation in the variation of the generalized entropy can be taken as a fundamental principle and used to derive Einstein's equations. This is essentially an update of the thermodynamic derivation of Einstein's equations by Jacobson [70].

Outline In Section 4.2 we review some of the basic concepts of entropy, relative entropy, and the holographic setup that will be relevant for our calculation. In Section 4.3 we prove (4.1.2) for situations where it is sufficient to consider linear perturbations of the bulk geometry. This includes any state where gravitational backreaction in the bulk is small. In Section 4.4 we extend this proof to any bulk state. The idea is that S''_{vv} is related to near-boundary physics in the bulk, and for any state the near-boundary geometry is approximately vacuum. So the proof reduces to the linear case. In Section 4.5 we move away from null deformations to prove (6.3.8) using the same techniques. We conclude in Section 4.6 with a discussion of extensions and implications of our work. Several appendices are included discussing closely related topics.

4.2 Setup and Conventions

In this section we will make some general remarks about the known relations between entropy and energy, and the implications of our conjecture.

The Field Theory Setup

Let $u = (t - x)/\sqrt{2}$ and $v = (t + x)/\sqrt{2}$ be null coordinates, and let y denote the other $d - 2$ spatial coordinates. For now, and for most of the rest of the paper, we will take

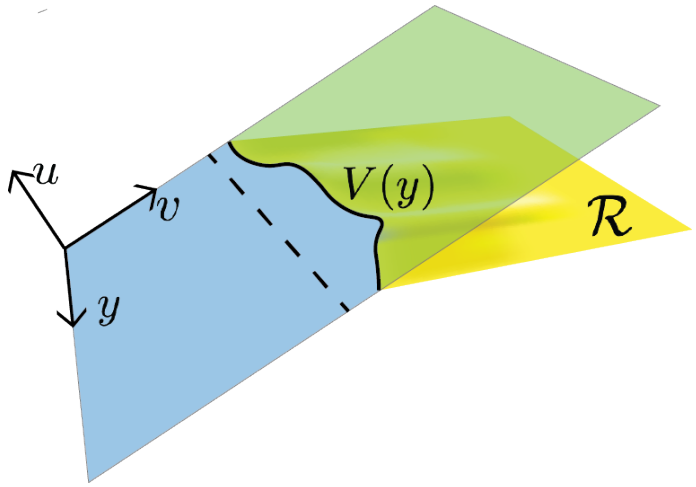

Figure 4.1: Most of our work concerns the variations of entanglement entropy for the yellow region $\mathcal{R}$ whose boundary $\partial\mathcal{R}$ lies on the null plane $u = 0$. The entangling surface is specified by the function $V(y)$.

the boundary of our region $\partial\mathcal{R}$ to be a section of the null plane $u = 0$. This boundary is specified by the equation $v = V(y)$. We take the region $\mathcal{R}$ to be a surface lying witin the "right quadrant," having $u < 0$ and $v > V(y)$ (marked in yellow in Fig 4.1). A one-parameter family of functions $V_\lambda(y)$ specifies a one-parameter family of regions $\mathcal{R}(\lambda)$. We always take the one-parameter family to be of the form $V_\lambda(y) = V_0(y) + \lambda\dot{V}(y)$ with $\dot{V} \geq 0$, so that λ plays the roll of an affine parameter along a future-directed null geodesic located at position y.

Given any global state of the theory, we can compute the von Neumann entropy S of the region $\mathcal{R}$. Keeping the state fixed, the entropy becomes a functional of the boundary of the region, $S = S[V(y)]$. When we have a one-parameter family of regions, then we can write $S(\lambda) = S[V_\lambda(y)]$. Throughout the rest of this work we will be interested in the derivatives of S with respect to λ, as well as the functional derivatives of S with respect to $V(y)$. These are related by the chain rule:

$$\frac{dS}{d\lambda} = \int d^{d-2}y \, \frac{\delta S}{\delta V(y)}\dot{V}(y), \tag{4.2.1}$$

$$\frac{d^2S}{d\lambda^2} = \int d^{d-2}y d^{d-2}y' \, \frac{\delta^2 S}{\delta V(y)\delta V(y')}\dot{V}(y)\dot{V}(y'). \tag{4.2.2}$$

We can parametrize the second functional derivative as follows:

$$\frac{\delta^2 S}{\delta V(y)\delta V(y')} = S''_{vv}(y)\delta^{(d-2)}(y - y') + \left(\frac{\delta^2 S}{\delta V(y)\delta V(y')}\right)_{\text{od}}. \tag{4.2.3}$$

We have extracted the δ-function term explicitly, which we sometimes refer to as the "diagonal" part, and the remainder carries the label "od" for "off-diagonal." Note that the off-diagonal part of the variation does not have to vanish at $y = y'$. The quantity S''_{vv} is the same as S'' in [16, 19, 11].

In addition to the entropy of the region $\mathcal{R}$, we can define the vacuum-subtracted modular energy, $\Delta\langle K\rangle$, and relative entropy with respect to the vacuum, $S_{\rm rel}$, associated to the region $\mathcal{R}$. The modular energy is given by the boost energy along each generator of the null plane [68]:

$$\Delta\langle K\rangle = 2\pi \int d^{d-2}y \int_{V(y)}^{\infty} dv\ (v - V(y))\langle T_{vv}\rangle. \tag{4.2.4}$$

The relative entropy is defined as the difference between the vacuum-subtracted modular energy and the vacuum-subtracted entropy:

$$S_{\rm rel} = \Delta\langle K\rangle - \Delta S. \tag{4.2.5}$$

For the regions we are talking about, the entropy of the vacuum is stationary and so drops out when we take derivatives of $S_{\rm rel}$. Then for a one-parameter family of regions we have the relations

$$\frac{dS_{\rm rel}}{d\lambda} = -\int d^{d-2}y \left[\frac{\delta S}{\delta V(y)} + 2\pi \int_{V(y)}^{\infty} dv\ \langle T_{vv}\rangle\right] \dot{V}(y), \tag{4.2.6}$$

$$\frac{d^2S_{\rm rel}}{d\lambda^2} = \int d^{d-2}y\ \left(2\pi\langle T_{vv}\rangle - S''_{vv}\right)\dot{V}(y)^2 - \int d^{d-2}y d^{d-2}y'\ \left(\frac{\delta^2 S}{\delta V(y)\delta V(y')}\right)_{\rm od} \dot{V}(y)\dot{V}(y'). \tag{4.2.7}$$

Note here that our conjectured equation (4.1.2) can be restated as saying that the diagonal second variation of the relative entropy is zero. These equations will be mirrored holographically in Section 4.3 below.

The Bulk Setup

While we have a few remarks on the free-field and weakly-interacting cases in Appendix A.7, most of our nontrivial evidence for (4.1.2) and (6.3.8) comes from holography. In this section we will describe the holographic setup for the calculations outlined above. We are actually able to do without much of this machinery in Section 4.3, though it will become important afterward.

The boundary theory is a quantum field theory in d-dimensional Minkowski space obtained by deforming a CFT with relevant couplings. We take the bulk metric to be in Fefferman-Graham gauge (at least near the boundary) and choose to set the AdS length to one:

$$ds^2_{d+1} = \frac{1}{z^2}\left(dz^2 - 2dudv + d\vec{y}^2_{d-2} + \gamma_{\mu\nu}dx^\mu dx^\nu\right). \tag{4.2.8}$$

Here x^μ stands for u, v, or y. In the small-z expansion, the metric $\gamma_{\mu\nu}$ is given by [71][6]

$$\gamma_{\mu\nu} = \sum_\alpha \gamma_{\mu\nu}^{(\alpha)} z^\alpha \tag{4.2.9}$$

In a fully-quantum treatment, $\gamma_{\mu\nu}$ is an operator in the bulk theory and we would need to take the expectation value of any geometric expression to extract a numerical result. Then there would be a difference between, say, $\langle\gamma_{\mu\nu}\rangle^2$ and $\langle\gamma_{\mu\nu}^2\rangle$ that we would have to resolve in order to move beyond leading order in a semiclassical expansion. A consequence of our analysis below is that only expressions which are linear $\gamma_{\mu\nu}$ end up being important for proving (4.1.2) and (6.3.8), and thus this potential difficulty is avoided. With that in mind, we will treat the bulk geometry as classical for ease of presentation.

The term at order z^d in (4.2.9), $\gamma_{\mu\nu}^{(d)}$, contains information about $\langle T_{\mu\nu}\rangle$ [72]. We will review the dictionary below. The terms at lower orders than z^d are associated with low-dimension operators in the theory [71]. If $\mathcal{O}$ is a relevant operator of dimension Δ and coupling g, then possible such terms that we need to be aware of include

$$\langle\mathcal{O}^m\rangle\eta_{\mu\nu}z^{m\Delta}, \qquad g^m\eta_{\mu\nu}z^{m(d-\Delta)}, \qquad g\langle\mathcal{O}\rangle\eta_{\mu\nu}z^d, \tag{4.2.10}$$

with $m \geq 2$. The coupling g, when present, is a constant. With only a single operator, terms involving derivatives of $\mathcal{O}$ will always be of higher order than z^d as long as the unitarity bound $\Delta > (d-2)/2$ is obeyed. When there is more than one low-dimension operator then we can also have terms with different combinatorial mixes of couplings and expectation values [73]. In this case, there could also be terms of the form

$$g_1^l\langle\mathcal{O}_2\rangle\eta_{\mu\nu}z^{l(d-\Delta_1)+\Delta_2}, \qquad g_1^l\partial_\mu\partial_\nu\langle\mathcal{O}_2\rangle z^{l(d-\Delta_1)+\Delta_2+2} \tag{4.2.11}$$

where $\mathcal{O}_1$ and $\mathcal{O}_2$ are two operators and g_1 is a relevant coupling associated to $\mathcal{O}_1$. There are other possibilities as well, but we will not need to enumerate them. In order to demonstrate the cancellation of divergences explicitly in (4.1.2), we would need to make use of certain relationships among the various parts of the small-z expansion of the metric. Since there are general arguments for the finiteness of (4.1.2), we will be content to show that the leading state-dependent divergences cancel.[7] To that end, we will need the following fact. Suppose that in the sum (4.2.9) there is a term of the form $\gamma_{\mu\nu}^{(\alpha)} = \gamma^{(\alpha)}\eta_{\mu\nu}$. Then, assuming that α cannot be written as $\alpha_1+\alpha_2$ for some other α_1, α_2 occuring in the sum, there will be another term $\gamma_{\mu\nu}^{(\alpha+2)}$ with a null-null component given by

$$\gamma_{vv}^{(\alpha+2)} = \frac{d-2}{(\alpha+2)(d-2-\alpha)}\partial_v^2\gamma^{(\alpha)}. \tag{4.2.12}$$

[6]For the purposes of this discussion, we will assume all operators have generic scaling dimensions. In the generic case on a flat background a $\log z$ term in the metric expansion is unnecessary.

[7]In other words, we will only explicitly demonstrate the finiteness of (4.1.2) given some conditions on the operator dimensions which make the terms we display the only ones that are around.

This equation is obtained by solving Einstein's equations at small-z [72, 71]. Four-derivative terms are also possible, at order $\alpha + 4$, but if $d \leq 6$ then the unitarity bound ensures that $\alpha + 4 > d$. For simplicity we will ignore those terms in this section, but with a little more effort they can also be accounted for.

Holographic Entropy and its Variations Our tool for computing the entropy is the Ryu-Takayanagi holographic entropy formula [31, 32] including the first quantum corrections [5],[8]

$$S = \frac{A_{\text{ext}}}{4G_N} + S_{\text{bulk}}. \tag{4.2.13}$$

A_{ext} refers to the area of the extremal area surface anchored to $\partial\mathcal{R}$ at $z = 0$. The dictionary for computing variations in the entropy as a function of $V(y)$ was laid out in [19] as follows. Let the bulk location of the extremal surface be given by

$$x^\mu = \bar{X}^\mu(y, z) = X^\mu(y) + z^2 X^\mu_{(2)}(y) + \cdots + z^d \log z X^\mu_{\log} + z^d X^\mu_{(d)} + \cdots , \tag{4.2.14}$$

where the log term is important for even dimensions and the in the case of relevant deformations with particular operator dimensions. $X^\mu(y)$ are the embedding functions of $\partial\mathcal{R}$ and $\bar{X}^\mu(y, z)$ satisfies the extremal surface equation,

$$\frac{1}{\sqrt{H}}\partial_\alpha \left(\sqrt{H} H^{\alpha\beta}\partial_\beta \bar{X}^\mu\right) + \Gamma^\mu_{\rho\sigma} H^{\alpha\beta}\partial_\alpha \bar{X}^\rho \partial_\beta \bar{X}^\sigma = 0, \tag{4.2.15}$$

where H is the induced metric on the extremal surface and Γ are bulk Christoffel symbols. Note that we have introduced the notation $\bar{X}^\mu$ for the bulk extremal surface coordinates which approach X^μ on the boundary. We will be interested in computing $\delta A_{\text{ext}}/\delta X^\mu(y)$, which by extremality is a pure boundary term evaluated at a $z = \epsilon$ cutoff surface:

$$\delta A_{\text{ext}} = \delta \int d^{d-2}y dz\ \sqrt{H} = -\int_{z=\epsilon} d^{d-2}y\ \sqrt{H} H^{zz} g_{\mu\nu}\partial_z \bar{X}^\mu \delta \bar{X}^\nu. \tag{4.2.16}$$

All of the factors appearing in the integrand need to be expanded in ϵ. The result will be a power series in ϵ containing divergent terms as well as finite terms:

$$\frac{\delta A_{\text{ext}}}{\delta X^\mu} = -\frac{K_\mu}{(d-2)\epsilon^{d-2}} + (\text{lower-order divergences in } \epsilon) - (d\, X^{(d)}_\mu + X^{(\log)}_\mu) + O(\epsilon). \tag{4.2.17}$$

Here K_μ is the extrinsic curvature of the entangling surface. We need to ensure that all divergences cancel or otherwise vanish in (4.1.2) and (6.3.8) in order that these be well-defined statements. So here we will explain the structure of the divergences in the entropy variations, as well as how to extract the finite part.

[8] In this section and in our main analysis we are only working to next-to-leading order so that the prescriptions of [5] and [74, 49] agree. If we wanted to work to higher orders in $1/N$, we would need to use the quantum extremal surface prescription instead [74, 49]. We discuss this further in Section 4.6.

Null Variations First, we will consider the special case $X^\mu(y) = V(y)$, which is the relevant case for (4.1.2). If there are no terms of the form (4.2.11) in the metric, then the situation reduces to that of [19], in which it was shown that the divergent terms in (4.2.17) are absent as long as the entangling surface $\partial\mathcal{R}$ is locally constrained to lie in a null plane. If there are state-dependent terms of the form (4.2.11) in the metric, then there will be non-vanishing divergent contributions to $\delta A_{\text{ext}}/\delta V(y)$ proportional to, e.g., $g_1\partial_v\langle\mathcal{O}_2\rangle$. In general, an extra term at order z^α in the metric leads to a contribution at order $\alpha+2$ in $\bar{X}^\mu$ that we can obtain by solving (4.2.15) at small z. We only need to concern ourselves with terms that have $\alpha+2<d$, as those are the ones which lead to divergences. As mentioned above, for $d\leq 6$ the only terms in the metric at order α such that $\alpha+2<d$ are those of the form $\gamma^{(\alpha)}_{\mu\nu}=\gamma^{(\alpha)}\eta_{\mu\nu}$. After solving the extremal surface equation in the presence of such a term we find

$$(\alpha+2)(\alpha+2-d)X^\mu_{(\alpha+2)}=\frac{2(d-2)-\alpha d}{2(d-2)}K^\mu\gamma^{(\alpha)}+\frac{d-2}{2}\partial^\mu\gamma^{(\alpha)}. \tag{4.2.18}$$

Plugging this in to (4.2.16) leads to

$$\frac{\delta A_{\text{ext}}}{\delta V(y)}=\frac{d-2}{2(d-2-\alpha)\epsilon^{d-2-\alpha}}\partial_v\gamma^{(\alpha)}(y)+dU_{(d)}(y)+\frac{\delta S_{\text{bulk}}}{\delta V(y)}, \tag{4.2.19}$$

where we have eliminated a potential log term by restricting ourselves to the case of generic operator dimensions. The non-generic case can be recovered later as a limit. Using this, we can find the leading-order contribution to the second variation of the entropy:

$$\frac{\delta^2 S}{\delta V(y)\delta V(y')}=\frac{d-2}{8G_N(d-2-\alpha)\epsilon^{d-2-\alpha}}\partial_v^2\gamma^{(\alpha)}(y)\delta^{(d-2)}(y-y')+\frac{d}{4G_N}\frac{\delta U_{(d)}(y)}{\delta V(y')}+\frac{\delta^2 S_{\text{bulk}}}{\delta V(y)\delta V(y')}. \tag{4.2.20}$$

Even though this is a very complicated expression in general, we will be able to extract the δ-function contribution and see that it is given by $\langle T_{vv}\rangle$ as in (4.1.2).

Non-Null Variations For a general non-null variation we lose some of the simplifications present in the non-null case. One additional assumption we will make in Section 4.5 is to consider entangling surfaces which are planar prior to being deformed, which simplifies some of the geometric expressions. More importantly, however, notice that (6.3.8) only makes reference to the vacuum-subtracted entropy variation, $\Delta S''_{\mu\nu}$, and not $S''_{\mu\nu}$ itself. So any state-independent terms in (4.2.17) can be ignored. Furthermore, for the discussion of the non-null variations we are only going to consider theories where relevant couplings (if present) have mass dimension greater than $d/2$, and states where operators of dimension $\Delta\leq d/2$ have vanishing expectation values in the vicinity of the entangling surface. The result of these restrictions is that terms like (4.2.11) will not be present in the metric up to order z^d, and so there will be no state-dependent entropy divergences. Thus for our analysis of non-null deformations, it follows from (4.2.17) that

$$\frac{\delta^2\Delta S}{\delta X^\mu(y)\delta X^\nu(y')}=-\frac{d}{4G_N}\Delta\left(\frac{\delta X^{(d)}_\mu(y)}{\delta X^\nu(y')}\right)+\frac{\delta^2\Delta S_{\text{bulk}}}{\delta X^\mu(y)\delta X^\nu(y')}. \tag{4.2.21}$$

In Section 4.5 we will also not deal explicitly with the bulk entropy term, but we expect its contributions to be qualitatively similar to the bulk entropy term in the null case.

Identification of the Stress Tensor We will also need a holographic formula for the stress tensor, $\langle T_{\mu\nu}\rangle$. Normally a renormalization procedure is required to define a finite stress tensor. Since our conjectures (4.1.2) and (6.3.8) are meant to be finite equations, it will be enough to regulate the stress tensor with a cutoff as we did with the entropy above.[9]

By definition, the (regulated) stress tensor is computed as the derivative of the regulated action:

$$\langle T_{\mu\nu}\rangle = \frac{2}{\sqrt{g}}\frac{\delta I_{\text{reg}}}{\delta g^{\mu\nu}} - (\text{vacuum energy})\ . \tag{4.2.22}$$

In holography, the regulated action is defined as the action of the bulk spacetime within the $z = \epsilon$ cutoff surface, plus additional boundary terms (like the Gibbons-Hawking-York term) which are necessary to make the variational principle well-defined. [72, 75]. For Einstein gravity in the bulk with minimally-coupled matter fields, the regulated stress tensor is then given by the Brown-York stress tensor evaluated on the $z = \epsilon$ cutoff surface [76]:[10]

$$\begin{aligned}\frac{2}{\sqrt{g}}\frac{\delta I_{\text{reg}}}{\delta g^{\mu\nu}} &= \frac{-1}{8\pi G_N \epsilon^{d-2}}\left(K_{\mu\nu} - \frac{1}{2}K g_{\mu\nu}(x,\epsilon)\right)\\ &= \frac{-1}{8\pi G_N \epsilon^{d-2}}\left(-\frac{1}{2\epsilon}\partial_\epsilon \gamma_{\mu\nu}(x,\epsilon) + \frac{1}{2\epsilon}\eta_{\mu\nu}\eta^{\rho\sigma}\partial_\epsilon\gamma_{\rho\sigma}(x,\epsilon) + \frac{1-d}{\epsilon^2}\eta_{\mu\nu}\right)\end{aligned} \tag{4.2.23}$$

Any state-dependent terms in the metric that occur at order z^α with $\alpha < d$ will contribute to divergences in the stress tensor. In particular, when we discuss null variations we will find contributions from terms of the form (4.2.12). In total we find

$$\begin{aligned}\langle T_{vv}\rangle &= \frac{\alpha+2}{16\pi G_N \epsilon^{d-2-\alpha}}\gamma_{vv}^{(\alpha+2)} + \frac{d}{16\pi G_N}\gamma_{vv}^{(d)}\\ &= \frac{d-2}{16\pi G_N (d-2-\alpha)\epsilon^{d-2-\alpha}}\partial_v^2\gamma^{(\alpha)} + \frac{d}{16\pi G_N}\gamma_{vv}^{(d)}.\end{aligned} \tag{4.2.24}$$

In the second line we used (4.2.12). Comparing this to (4.2.20), we see that the divergences indeed cancel out in (4.1.2).

For the non-null case we have additional difficulties. One can easily see that, in general, there are state-dependent divergences in $\langle T_{\mu\nu}\rangle$ that do not appear in $S''_{\mu\nu}$. For example, if there are operators of dimension $\Delta < d/2$ in the theory then there will be a term in $\gamma_{\mu\nu}$ at

[9] We still want to define the stress tensor so that $\langle T_{\mu\nu}\rangle = 0$ in vacuum, so the constant vacuum energy term will be subtracted.

[10] Care must be taken to impose the correct boundary conditions at $z = \epsilon$. Since we are interested in a flat-space result, we must place a flat metric boundary condition at $z = \epsilon$ before taking $\epsilon \to 0$. This is the only way to get the divergences to cancel out properly between the entropy and the energy in (4.1.2), and this treatment of the boundary condition is especially important if one wants to extend the analysis to curved space [69].

order $z^{2\Delta}$ proportional to $\langle\mathcal{O}^2\rangle\eta_{\mu\nu}$. By the unitary bound, $2\Delta > d-2$, such a term will not contribute divergences to $S''_{\mu\nu}$, but it will contribute divergences to the stress-tensor of the form

$$\langle T_{\mu\nu}\rangle\,|_{\epsilon^{2\Delta-d}} \propto \epsilon^{2\Delta-d}\langle\mathcal{O}^2\rangle\eta_{\mu\nu}. \tag{4.2.25}$$

Thus, when we derive the relationship (6.3.8) in Section 4.5, we will put sufficient restrictions on the theory and the states in consideration so that both sides of the equality are finite and well-defined. As in the case of the entropy variation, all divergences in $\langle T_{\mu\nu}\rangle$ can be eliminated by restricting the theory so that any nonzero relevant couplings have mass dimension greater than $d/2$, and by restricting the state so that operators of dimension $\Delta \leq d/2$ have vanishing expectation values (at least locally near the entangling surface). When this is true, the metric perturbation $\gamma_{\mu\nu}$ starts at order z^d, and so $\langle T_{\mu\nu}\rangle$ will be finite. Furthermore, we can treat the stress tensor as being effectively traceless even though we are not in a CFT. That is because in general the trace is proportional to products of couplings and scalar expectation values, $g\langle\mathcal{O}\rangle$, but with our restrictions on the theory and state there is no pair of nonzero coupling and operator expectation value with total dimension adding up to d. The end result is the standard formula for the stress tensor familiar from holographic renormalization [72]:

$$\langle T_{\mu\nu}\rangle = \frac{d}{16\pi G_N}\gamma^{(d)}_{\mu\nu}. \tag{4.2.26}$$

We will make use of this formula in Section 4.5.

4.3 Null Deformations and Perturbative Geometry

In this section we will prove the relation $S''_{vv} = 2\pi\langle T_{vv}\rangle$ for states with geometries corresponding to perturbations of vacuum AdS where it suffices to work to linear order in the metric perturbation. This includes classical as well as quantum states. Below in Section 4.4 we will extend our results to non-perturbative geometries.

The arguments presented here can be repeated for linearized perturbations to a non-AdS vacuum, i.e., the vacuum of a non-CFT. We restrict ourselves to the AdS case because explicit solutions to the equations are available, and the AdS case also suffices for nearly all applications in the following sections. We will see in Section 4.4 that in certain situations an appeal to the non-AdS vacuum case is necessary, but because of general arguments (like the known form of the modular Hamiltonian as discussed in the Introduction) we know that the non-AdS case should not behave differently than the AdS case.

Bulk and Boundary Relative Entropies

In [7] it was argued that bulk and boundary relative entropies are identical:

$$S_{\text{rel}} = S_{\text{rel,bulk}}, \tag{4.3.1}$$

where $S_{\text{rel,bulk}}$ is calculated using the bulk quantum state restricted to the entanglement wedge of the boundary region $\mathcal{R}$ — the region of the bulk bounded by the extremal surface and $\mathcal{R}$.[11]

We already discussed in Section 4.2 the form of S_{rel} for the regions we are considering, but to leading order in bulk perturbation theory there is an analogous simple formula for $S_{\text{rel,bulk}}$. We only need to know two simple facts. First, if $\partial\mathcal{R}$ is restricted to lie in the $u = 0$ plane on the boundary then, to leading order, the extremal surface in the bulk also lies in the $u = 0$ plane. Second, to leading order the bulk modular energy corresponding to such a region is given by the AdS analogue of (4.2.4):

$$\Delta K_{\text{bulk}} = 2\pi \int \frac{dz d^{d-2}y}{z^{d-1}} \int_{\bar{V}(y)}^{\infty} dv\ (v - \bar{V}(y,z)) \langle T^{\text{bulk}}_{vv} \rangle. \tag{4.3.2}$$

In keeping with our earlier notation, $\bar{V}(y,z)$ gives the location of the bulk extremal surface with $\bar{V}(y, z = 0) = V(y)$. Now we simply solve (4.3.1) for the vacuum-subtracted boundary entropy ΔS,

$$\Delta S = \Delta\langle K \rangle - \Delta\langle K_{\text{bulk}} \rangle + \Delta S_{\text{bulk}}, \tag{4.3.3}$$

and take two derivatives with respect to a deformation parameter λ to find

$$\frac{d^2 S}{d\lambda^2} = 2\pi \int d^{d-2}y\ \langle T_{vv} \rangle \dot{V}^2 - 2\pi \int \frac{dz d^{d-2}y}{z^{d-1}}\ \langle T^{\text{bulk}}_{vv} \rangle \dot{\bar{V}}^2 + \frac{d^2 S_{\text{bulk}}}{d\lambda^2}. \tag{4.3.4}$$

The first term represents a contribution of $2\pi\langle T_{vv}\rangle$ to S''_{vv}. So (4.1.2), $S''_{vv} = 2\pi\langle T_{vv}\rangle$, amounts to showing that the remaining two terms do not contribute to S''_{vv}. We examine them both in the next section.

Proof of the Conjecture

From the discussion around (4.3.4), the conjecture $S''_{vv} = 2\pi\langle T_{vv}\rangle$ amounts to the statement that the terms

$$-2\pi \int \frac{dz d^{d-2}y}{z^{d-1}}\ \langle T^{\text{bulk}}_{vv} \rangle \dot{\bar{V}}^2 + \frac{d^2 S_{\text{bulk}}}{d\lambda^2}. \tag{4.3.5}$$

do not contribute a δ-function to the second variation of S. Together these terms comprise the second derivative of the bulk relative entropy. We treat the two terms individually.

Bulk Modular Energy The modular energy term is simple to evaluate. Note that (4.3.2) depends on the entangling surface $V(y)$ through the extremal surface $\bar{V}(y,z)$. So functional derivatives of that expression with respect to $V(y)$ involves factors of $\delta\bar{V}(y,z)/\delta V(y')$. This

[11] At higher orders in $1/N$ this equation is corrected [5, 56, 6]. We will not go into these corrections in detail, but will make a few comments below in Section 4.6.

is the boundary-to-bulk propagator of the extremal surface equation in pure AdS. The result, which can be extracted from our discussion in later sections, is [77]

$$\frac{\delta \bar{V}(y,z)}{\delta V(y)} = \frac{2^{d-2}\Gamma(\frac{d-1}{2})}{\pi^{\frac{d-1}{2}}} \frac{z^d}{(z^2+(y-y')^2)^{d-1}}. \tag{4.3.6}$$

Then we have

$$\frac{\delta^2 K_{\text{bulk}}}{\delta V(y_1)\delta V(y_2)} = 2\pi \left(\frac{2^{d-2}\Gamma(\frac{d-1}{2})}{\pi^{\frac{d-1}{2}}}\right)^2 \int \frac{dz d^{d-2}y}{z^{d-1}} \langle T^{\text{bulk}}_{vv}\rangle \frac{z^{2d}}{(z^2+(y-y_1)^2)^{d-1}(z^2+(y-y_2)^2)^{d-1}} \tag{4.3.7}$$

We can diagnose the presence of a δ-function by integrating with respect to y_1 over a small neighborhood of y_2. If the result remains finite as the size of the neighborhood goes to zero, then we have a δ-function. Whether or not this happens depends on the falloff conditions on $\langle T^{\text{bulk}}_{vv}\rangle$ near $z=0$, which in turn depends on the matter content of the bulk theory. If we suppose $\langle T^{\text{bulk}}_{vv}\rangle \sim z^\beta$ as $z \to 0$, then it is easy to see that there is no δ-function so long as

$$\beta > d-2. \tag{4.3.8}$$

For scalar fields in the bulk, $T^{\text{bulk}}_{vv} \sim (\partial_v \phi)^2 \sim z^{2\Delta}$ where Δ is the dimension of the dual operator. This is even true when the non-normalizable mode $\phi \sim g z^{d-\Delta}$ is turned on, as long as the coupling g is constant. For bulk Dirac fields, $T^{\text{bulk}}_{vv} \sim \bar{\psi}\Gamma_v \nabla_v \psi \sim z^{2\Delta-1}$. In either case, equation (4.3.8) reduces to the unitarity bound on the dual operator dimension, $\Delta > (d-2)/2 + s$, where $s = 0, 1/2$ is the spin. In the limiting case where the unitarity bound is saturated and the dual operator is a free scalar or free fermion, one may find a δ-function in (4.3.7). Indeed, in Appendix A.7 we find extra contributions to S''_{vv} besides $2\pi\langle T_{vv}\rangle$ for a free scalar field, so the appearance of an additional δ-function in this case is an expected feature. The case of a free fermion has not yet been worked out in the field theory, but methods similar to those in Appendix A.7 should be applicable. For operators which do not saturate the unitarity bound, we have shown that ΔK_{bulk} does not contribute to S''_{vv}.

Bulk Entropy It is much more difficult to make statements about $d^2 S_{\text{bulk}}/d\lambda^2$. In a coherent bulk state we know that $d^2 S_{\text{bulk}}/d\lambda^2 = 0$, so for that class of states we are done.[12] More generally, we can write

$$\frac{\delta^2 S_{\text{bulk}}}{\delta V(y_1)\delta V(y_2)} = \left(\frac{2^{d-2}\Gamma(\frac{d-1}{2})}{\pi^{\frac{d-1}{2}}}\right)^2 \int d^{d-2}y dz d^{d-2}y' dz' \frac{\delta^2 S_{\text{bulk}}}{\delta \bar{V}(y,z)\delta \bar{V}(y',z')} \frac{(zz')^d}{(z^2+(y-y_1)^2)^{d-1}(z'^2+(y'-y_2)^2)^{d-1}} \tag{4.3.9}$$

[12] In this section we treat the bulk matter fields as free. If we turn on weak interactions, then the comments of Appendix A.7 apply. Qualitatively nothing changes.

and ask what sort of behavior would be required of $\delta^2 S_{\text{bulk}}/\delta\bar{V}(y,z)\bar{V}(y',z')$ in order to lead to a δ-function in $y_1 - y_2$.

As a toy model, we can imagine a collection of particles on the $u = 0$ surface which are entangled in a way that depends on their distance from each other. This is a fairly general ansatz for the state of a free theory in the formalism of null quantization [15]. At small z (which is the dominant part for our calculation) this would correspond to a second variation of the form

$$\frac{\delta^2 S_{\text{bulk}}}{\delta\bar{V}(y,z)\delta\bar{V}(y',z')} \sim \frac{(zz')^{\Delta}}{(zz')^{d-1}} F\left(\frac{zz'}{(z-z')^2+(y-y')^2}\right). \tag{4.3.10}$$

The factor $(zz')^{\Delta}/(zz')^{d-1}$ reflects that entropy variations should be proportional to the amount of matter present at locations z and z'. The numerator encodes the falloff conditions on the density of particles in a way that is consistent with the falloff conditions for a bosonic matter field, and the denominator is a measure factor that converts coordinate areas to physical areas. The function F is arbitrary.

With the assumption of (4.3.10), a constant rescaling of all coordinates by α leads to an overall factor of $\alpha^{4-2d+2\Delta}$ in (4.3.9). A δ-function in $y_1 - y_2$ would scale like α^{2-d}, and anything that scales with a power of α less than $2-d$ would correspond to a more-divergent distribution, like the derivative of a δ-function. As long as $\Delta > (d-2)/2$ this is avoided, and a δ-function is only present when the unitarity bound $\Delta = (d-2)/2$ is saturated. This is consistent with what we found previously for the modular energy, and with our general expectations for free theories.

4.4 Non-Perturbative Bulk Geometry

Now we turn to a proof that applies for a general bulk geometry, still restricting the deformations to be null on the boundary. We will use the techniques outlined in Section 4.2, which relate the entropy variations to changes in the bulk extremal surface location. At first we will stick to boundary regions where $\partial\mathcal{R}$ is restricted to a null plane, leaving a generalization to regions where $\partial\mathcal{R}$ only satisfies certain local conditions for Section 4.6.

Extremal Surface Equations

Small z, Large k The extremal surface equation (4.2.15) for $\bar{U}$ and $\bar{V}$ is a very complicated equation. If we perturb the boundary conditions by taking $V \to V + \delta V$, then the responses $\delta\bar{U}$ and $\delta\bar{V}$ will satisfy the linearized extremal surface equation, which is a bit simpler. It may be that the coordinates we have chosen are not well-suited to describing the surface perturbations deep into the bulk. That problem is solved by only aiming to analyze the equations in the range $z < z_*$ for some small but finite z_*. In fact, by choosing z_* small enough we can say that the spacetime is perturbatively close to vacuum AdS, with the perturbation given by the Fefferman-Graham expansion (4.2.9). Since the corrections to the vacuum geometry are small when z_* is small, the extremal surface equation reduces to the

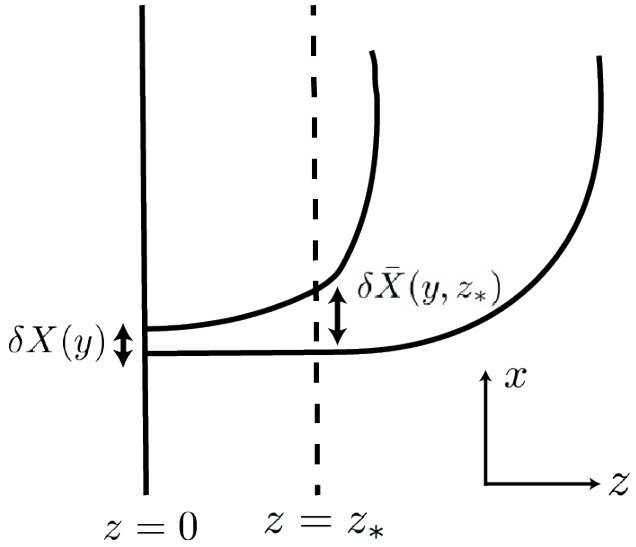

Figure 4.2: By restricting attention to $z < z_*$ the geometry is close to pure AdS, and we can solve for $\delta\bar{X}$ perturbatively. All of the $z < z_*$ data imprints itself as boundary conditions at $z = z_*$. We show that these boundary conditions are unimportant for our analysis, which means that a perturbative calculation is enough.

vacuum extremal surface equation plus perturbative corrections. All of the deep-in-the-bulk physics is encoded in boundary conditions at $z = z_*$. The situation is illustrated in Fig. 4.2

The boundary conditions at $z = z_*$ are essentially impossible to find in the general case, so the restriction to $z < z_*$ does not make the problem of finding the extremal surface any easier. However, according to (4.2.20) all we are interested in is the δ-function part of $\delta U_{(d)}$. It will turn out that this quantity is actually independent of those boundary conditions.

The idea is very simple. In Fourier space a δ-function has constant magnitude. That means it does not go to zero at large values of k, unlike the Fourier transform of a smooth function. So the strategy will be to analyze the extremal surface equation in Fourier space at large k. We will see that the large-k response of $\bar{U}$ (and hence $U_{(d)}$) is completely determined by near-boundary physics, and in particular will match the results we found in previous sections. This will establish that $S''_{vv} = 2\pi\langle T_{vv}\rangle$ for very general bulk states.

Integral Equation for $\bar{U}$ We will begin by finding an integral equation for $\bar{U}$ in the range $z < z_*$. Since $\bar{U}$ vanishes at $z = 0$ it must remain small throughout $z < z_*$, as long as z_* is small enough, and so we can use perturbation theory to find $\bar{U}$ in that range. Then we will compute the response of $\bar{U}$ to variations of the boundary conditions V at $z = 0$. Expanding (4.2.15) in small z, we can write the equation for $\bar{U}$ as

$$\partial_a^2\bar{U} + \partial_z^2\bar{U} + \frac{1-d}{z}\partial_z\bar{U} = J[\gamma_{\mu\nu}, \bar{V}, \bar{U}], \tag{4.4.1}$$

where $\gamma_{\mu\nu}/z^2$ is the deviation of the metric from vacuum AdS, as in (4.2.9). To solve this equation perturbatively we require a Green's function $G(z, y|z', y')$ of the linearized extremal surface equation that vanishes when $z = 0$ or $z = z_*$. Then the solution to (4.4.1) can be written as

$$\bar{U}(y,z) = \int \frac{d^{d-2}y'}{z_*^{d-1}}\, \partial_{z'} G(y,z|y',z_*)\bar{U}(y',z_*) + \int_{z<z_*} \frac{d^{d-2}y'dz'}{z'^{d-1}} G(y,z|y',z')J(y',z') \quad (4.4.2)$$

It is important to remember that $J(y, z)$ is itself a functional of $\bar{U}$, and the usual methods of perturbation theory would involve solving for $\bar{U}$ iteratively. It will be more useful for us to look at the Fourier transform of this equation:

$$\bar{U}(k,z) = z_*^{1-d}\partial_{z'} G_k(z|z_*)\bar{U}(k,z_*) + \int_0^{z_*} \frac{dz'}{z'^{d-1}} G_k(z|z')J(k,z'). \quad (4.4.3)$$

The Green's function with the correct boundary conditions is easily obtained from the standard Green's function G^{AdS} by adding a particular solution of the vacuum extremal surface equation. In Fourier space, the answer is

$$G_k(z|z') = G_k^{\text{AdS}}(z|z') + (zz')^{d/2} I_{d/2}(kz) I_{d/2}(kz') \frac{K_{d/2}(kz_*)}{I_{d/2}(kz_*)} \quad (4.4.4)$$

where

$$G_k^{\text{AdS}}(z|z') = -\begin{cases} (zz')^{d/2} I_{d/2}(kz) K_{d/2}(kz'), & z < z', \\ (zz')^{d/2} I_{d/2}(kz') K_{d/2}(kz), & z > z'. \end{cases} \quad (4.4.5)$$

In the limit of large k, the first term of (4.4.3) becomes exponentially suppressed. So we see that the boundary conditions at $z = z_*$ do not matter. Furthermore, the integration range $z' \gtrsim 1/k$ in the second term also becomes exponentially suppressed. So only the small-z part of the source J contributes at leading order in the large-k limit.

Terms in the Source

Let us consider the form of the source in position space in more detail. We know that $J = J[\bar{U}, \bar{V}, \gamma]$ is a functional of the extremal surface coordinates and the metric perturbation. We can treat J as a double power series in γ and $\bar{U}$ since we are doing perturbation theory in those two parameters. We will repeatedly take advantage of the "boost" symmetry of the equation: under the coordinate transformation $u \to \alpha u$, $v \to \alpha^{-1} v$, the source must transform as $J \to \alpha J$ in order for the whole equation to be covariant. Since every occurrence of $\bar{V}$ must be accompanied by either a γ or $\bar{U}$ to preserve the boost symmetry, $J[\bar{U}, \bar{V}, \gamma]$ is actually a triple power series in all three of its parameters. Another important fact is dimensional analysis, which comes from scaling all coordinates together: J has length dimension -1, while $\bar{U}$ and $\bar{V}$ have dimension 1 and γ has dimension zero. This will also be used to restrict the types of terms we can find.

The variation $\delta\bar{U}$ satisfies an integral equation similar to that of $\bar{U}$ except with the source, J, replaced by the variation of the source, δJ. Like J, we can treat δJ as a power series. Each term in the δJ power series contains a single $\delta\bar{U}$, $\delta\gamma$, or $\delta\bar{V}$, multiplied by some number of $\bar{U}$, $\bar{V}$, and γ factors (and their derivatives). It is important to note that these unvaried $\bar{U}$, $\bar{V}$, and γ factors are smooth, and therefore their Fourier transforms decay at large k. So the Fourier transform of a term in δJ looks schematically like

$$\delta J(k) \sim \int_{k'<<k} dk' \; h(k')\delta\Psi(k-k'), \tag{4.4.6}$$

where Ψ is either γ, $\bar{V}$, $\bar{U}$, or their derivatives and h is the Fourier transform of a smooth function. The k-dependence at large k of a given term in δJ is completely determined by the factor $\delta\Psi$ being varied. The case where $\Psi = \gamma$ can be reduced immediately to the other two, because $\delta\gamma = \delta\bar{V}\partial_v\gamma + \delta\bar{U}\partial_u\gamma$.

In Fourier space, we can write $\delta J(k,z)$ as a sum of terms of the form $\delta J_{mn} z^m k^n$ at small z and large k.[13] Since the effect of z_* is exponentially suppressed at large k, we can drop the first term in (4.4.3) and push the limit in the second term off to infinity. Additionally, the difference between $G_k(z|z')$ and $G_k^{\rm AdS}(z|z')$ is exponentially suppressed. Thus for our purposes we have

$$\begin{aligned}\delta\bar{U}(k,z) &= \sum_{m,n}\int_0^\infty G_k^{\rm AdS}(z|z')\delta J_{mn} z^m k^n + O(e^{-kz_*}) \qquad (4.4.7)\\ &= \sum_{m,n}\delta J_{mn}\left(\frac{k^n z^{2+m}(d-2(m+2))}{d(m+2)(d-m-2)} - z^d 2^{m-d}k^{n-m-2+d}\frac{\Gamma\left(1+\frac{m}{2}\right)\Gamma\left(\frac{m-d+2}{2}\right)}{\Gamma(1+d/2)}\right) + \mathcal{O}(z^{d+1})\end{aligned}$$

If $m < d-2$ then the first term in (4.4.7) represents a contribution to $\bar{U}$ that could have been obtained by doing the small-z expansion of the extremal surface equation. In a CFT these would consist only of geometric terms that depend on extrinsic curvatures of the entangling surface, but our boundary condition $U = 0$ guarantees that those vanish. Still, when a relevant deformation is turned on there may be terms proportional to $g_1^l\partial_v\langle\mathcal{O}_2\rangle$ which enter $\bar{U}$ at low orders in z. An important fact, enforced by the unitarity bound, is that these low-order terms are all linear in expectation values. When $m = d-2$ each of the terms in (4.4.7) becomes singular, but actually the combination above remains finite and generates at $z^d \log z$ term. Since (4.4.7) is well-behaved in this limit, we can treat the non-generic case $m = d-2$ as a limiting case of generic m. Thus throughout our discussion below m is assumed to be generic. Finally, for $d > 6$ another term proportional to z^{4+m} (and z^{6+m} in $d > 8$, etc.) should be included, but for simplicity we have not written it down. Qualitatively it has the same properties as the z^{2+m} term.

[13] There may also be terms in the source of the form $z^m \log(z)$. Qualitatively these terms behave similarly to the z^m terms as far as the δ-function part of the entropy variation is concerned, so we will not explicitly keep track of them.

Our focus is on the z^d term, as this is where the finite contributions to the entropy variation come from, as in (4.2.20). From (4.4.7), we see that the δ-function is determined by source terms with $n-m = 2-d$, which corresponds to k^0 behavior at large k. So our task is simply to enumerate the possible terms in δJ which have this behavior. We will see that such terms are completely accounted for by the linearized analysis of the previous section,[14] which completes the proof.

Ingredients Before diving into the terms of the source, we will collect all of the facts we need about the function $\bar{U}$, $\bar{V}$, γ, and their variations. In particular, we will need to know what powers of k and z we can expect them to contribute to the source.

We begin with $\bar{V}$. Unlike $\bar{U}$, $\bar{V}$ does not have any particular boundary condition at $z = 0$. Thus the Fefferman-Graham expansion for $\bar{V}$ contains low powers of z that depend on geometric data of the entangling surface. In particular, the boundary condition itself enters $\bar{V}$ at order z^0, which is neutral in terms of the $n - m$ counting. That same behavior extends to the variation $\delta\bar{V}$: in Fourier space, the state-independent parts of $\delta\bar{V}$ are functions of the combination kz. In other words, we find schematically

$$\delta\bar{V} \sim (1 + k^2z^2 + k^4z^4 + \cdots)\delta V. \tag{4.4.8}$$

The boundary condition δV itself is taken to go like k^0 at large k (i.e., a δ-function variation). So in terms of our power counting, which only depends on $n - m$, these terms are all completely neutral. So a factor of $\delta\bar{V}$ in the source is "free" as far as the power counting is concerned. There will be other terms in $\delta\bar{V}$, even at low powers of z, but the terms in (4.4.8) are the ones which dominate the $n - m$ counting.

$\bar{U}$ is also an extremal surface coordinate, but it has the restricted boundary condition $U = 0$. That means it does not possess terms like those in (4.4.8). The lowest-order-in-z terms that can be present are of the form $g_1^l\partial_v\langle\mathcal{O}_2\rangle z^{2+l(d-\Delta_1)+\Delta_2}$. It is only terms like this which contain a single factor of $\mathcal{O}$ that can show up at lower orders than z^d, because of the unitarity bound $\Delta > (d-2)/2$. Taking a variation, we find a term in $\delta\bar{U}$ of the form

$$\delta\bar{U} \sim g_1^l\partial_v^2\langle\mathcal{O}_2\rangle\delta V z^{2+l(d-\Delta_1))+\Delta_2}, \tag{4.4.9}$$

which has $n - m = -(2 + l(d - \Delta_1) + \Delta_2)$.

The final ingredient is the metric perturbation γ. We don't have to consider variations of γ directly, since they can be re-expressed in term s of variations of $\bar{U}$ and $\bar{V}$. γ itself has a Fefferman-Graham expansion which in includes information about the stress tensor at order z^d, but can have lower-order terms as well that depend on couplings and expectation values of operators. We will see that the important terms in the source that affect the δ-function response are those which are linear in γ.

[14] As mentioned in the previous section, for simplicity of presentation we are performing our perturbation theory around empty AdS, whereas in complete generality one would want to perform the analysis based around the vacuum of the theory in question. The difference is that some terms which are linear in expectation values $\langle\mathcal{O}\rangle$ might appear at higher orders in perturbation theory around empty AdS even though they are fully accounted for in the linearized analysis about the correct vacuum.

Terms with $\delta\bar{U}$ Now we will analyze the possible terms in the source which can be obtained by piecing together the above ingredients. We begin with terms proportional to $\delta\bar{U}$. As stated above, there are dominant contributions to $\bar{U}$ in terms of the $n-m$ counting which are proportional to derivatives of expectation values of operators.

But $\bar{U}$ does not occur alone in the source J: since all terms with $\bar{U}$ alone in the equation of motion are part of the linearized equation of motion on the left-hand-side of (4.4.1). An additional factor of $\bar{V}$ does not affect the dominant $n-m$ value of the term, but the combination $\bar{U}\bar{V}$ is also prevented from appearing in J by boost symmetry. We need to have at least another factor of $\bar{U}$, or else a factor of γ. The dominant possibility without using γ is something of the form $\partial\bar{U}\partial\bar{V}\partial^2\delta\bar{U}$, where derivatives have been inserted to enforce the correct total dimensionality. Taking into account the derivatives, a term like this can have at most $n-m = 3-2(2+l(d-\Delta_1)+\Delta_2) < 1-d-2l(d-\Delta_1) < 2-d$, using the unitarity bound. So this sort of term will not matter for the δ-function response.

Making use of γ allows for more possibilities. Terms of the schematic form $\gamma\delta\bar{U}$ in the source can have $n-m > 2-d$, and if we allow fine-tuning of operator dimensions we can even reach $n-m = 2-d$. These sources are obtained by taking a state-independent term in γ which is proportional to some power of g_1 and a term in $\delta\bar{U}$ which is proportional to $\partial_v^2\langle\mathcal{O}_2\rangle$. We can even multiply by more factors of γ, giving $\gamma^l\delta\bar{U}$ schematically, as well as factors of $\bar{V}$, as long as we don't involve more factors of $\bar{U}$. A second factor of $\bar{U}$ brings with it a large z-scaling, so we run into the same problem we had above in the $\bar{U}\bar{V}\delta\bar{U}$ case. The end result is that all of the potentially-important terms in this analysis are linear in the expectation value $\langle\mathcal{O}\rangle$. That means they are subject to restrictions on the modular Hamiltonian as mentioned in the Introduction, which means that they will actually not show up in (4.1.2) despite being allowed by dimensional analysis.

Terms with $\delta\bar{V}$ Now we consider terms in δJ that are proportional to a variation $\delta\bar{V}$. As discussed above, $\delta\bar{V}$ has several state-independent terms which are neutral in the $n-m$ counting. Due to the boost symmetry, $\delta\bar{V}$ cannot occur alone in δJ. It must be accompanied by at least two factors of $\bar{U}$ or one factor of γ. We have already discussed how two factors of $\bar{U}$ have a large-enough z-scaling to make the term uninteresting, so it remains to consider factors of γ.

Terms in the source proportional to $\delta\bar{V}$ with only a single factor of γ are those present in the theory of linearized gravity about vacuum AdS. Furthermore, since we argued that boundary conditions at $z = z_*$ do not affect the answer, the Green's function we use to compute the effects of the source is *also* the same as we would use in linearized gravity about vacuum AdS. We already considered the linearized gravity setup in Section 4.3, even though we didn't solve it using the methods of this section. In Section 4.3 we saw that $S''_{vv} = 2\pi\langle T_{vv}\rangle$, and so it is enough for us now to prove that the general computation of the δ-function terms reduces to the linearized gravity case. There is only one more loose end to consider: terms in δJ proportional to $\delta\bar{V}$ that have more than one factor of γ.

With more than a single factor of γ, it is clear that the only contributions that could

possibly be important at large k are those coming from the powers of z less than z^d in (4.2.9). These terms are made up of couplings g, operator expectation values $\langle\mathcal{O}\rangle$, and their derivatives. In order to have the correct boost scaling, we need to include v-derivatives acting on operator expectation values. As we have discussed many times, the unitarity bound prevents any term with more than one factor of $\langle\mathcal{O}\rangle$ from being important. So just as with the $\delta\bar{U}$ terms discussed previously, all of these terms are subject to constraints from the modular Hamiltonian and hence do not appear in (4.1.2)

Our analysis so far has been very simple , but we have reached an important conclusion that bears repeating: the source terms which give the k^0 behavior for $\delta U_{(d)}$ were already present in the linearized gravity calculation of the previous section, and we are allowed to use the ordinary Green's function $G^{\rm AdS}$ to compute their effects. In other words, for the purpose of calculating the δ-function response we have reduced the problem to linearized gravity. We have shown previously that the linearized gravity setup leads to $S''_{vv} = 2\pi\langle T_{vv}\rangle$, and so our proof is complete.

4.5 Non-Null Deformations

Having established $S''_{vv} = 2\pi\langle T_{vv}\rangle$ for deformations of entangling surfaces restricted to lie in the plane $u = 0$, we will now analyze arbitrary deformations of the entangling surface to prove (6.3.8). The technique is very similar to that of the previous section. As discussed in Sec 4.2, there are additional assumptions and restrictions we make in this case to help us deal with divergences and to simplify the analysis. First, we restrict attention to theories where all relevant couplings, if present, have mass dimension greater than $d/2$. Second, we restrict the state so that operators with scaling dimension $\Delta \leq d/2$ have vanishing expectation value near the entangling surface. Finally, we restrict the entangling surface itself to be planar prior to taking any variations.

New Boundary Conditions

Above we analyzed deformations within the null plane $u = 0$ at small z and large k. These limits allowed us to show that the perturbation theory for $\delta U_{(d)}$ reduced to linearized gravity, which we had already studied in Section 4.3. There strategy here is the same, except we want to be able to perform perturbation theory on both $\bar{U}$ and $\bar{V}$ in order to get more than just the null-null variations. The simplest case, which is all that we will analyze in this work, is to start with the boundary condition $V = 0$ at $z = 0$ in addition to $U = 0$. In other words, we take our undeformed entangling surface to be the $v = u = 0$ plane. That is a severe restriction on the type of surface we are considering, but we gain the flexibility of being able to do perturbation theory in both $\bar{U}$ and $\bar{V}$. From (4.2.21),

$$\frac{\delta^2\Delta S}{\delta X^\mu(y)\delta X^\nu(y')} = -\frac{d}{4G_N}\Delta\left(\frac{\delta X^{(d)}_\mu(y)}{\delta X^\nu(y')}\right) + \frac{\delta^2\Delta S_{\rm bulk}}{\delta X^\mu(y)\delta X^\nu(y')}, \tag{4.5.1}$$

where ΔS refers to the vacuum-subtracted entropy. Vacuum subtraction removes all state-independent terms from the entropy, including divergences. blueFor the remainder of the section, we will drop the bulk entropy contribution.

With the $U = V = 0$ boundary conditions, we can again write down our perturbative extremal surface equation for the $z < z_*$ part of the bulk. Since the null direction is no longer preferred, we will use a covariant form of the linearized equation:

$$\partial_a^2 \bar{X}^\mu + \partial_z^2 \bar{X}^\mu + \frac{1-d}{z}\partial_z \bar{X}^\mu = J^\mu[\gamma, \bar{X}] \tag{4.5.2}$$

Following the same steps as in the previous section, we can use Green's functions to solve this equation in Fourier space. There is one new ingredient that we did not have before. When we computed the variation of $U_{(d)}$ with respect to V, we were changing the boundary conditions of $\bar{V}$ and computing the response in $\bar{U}$. In particular, the boundary condition of $\bar{U}$ itself remained zero. In the more general setup of this section, we need to compute the response of a particular component of $\bar{X}^\mu$ when its own boundary conditions at $z = 0$ are varied.

Since we only care about the δ-function contribution to the entropy variation, we will immediately use $\delta X^\mu(k) = e^{iky_0}\xi^\mu$ as the boundary condition for $\delta\bar{X}^\mu$. Here ξ^μ is just a constant vector which tells us the direction of the perturbation. The presence of this boundary condition at $z = 0$ is simple to account for with one additional term in the integral equation for $\bar{X}^\mu$ compared to (4.4.3) in the previous section. In total, we now have

$$\begin{aligned}\delta X^\mu(k,z) = {} & z^{d/2}K_{d/2}(kz)\frac{dk^{d/2}}{2^{d/2}\Gamma(1+d/2)}\xi^\mu e^{iky_0} \\ & + z_*^{1-d}\partial_{z'}G(z|z_*)\delta\bar{X}^\mu(k,z_*) + \int_0^{z_*}\frac{dz'}{z'^{d-1}}G_k(z|z')\delta J^\mu(k,z')\end{aligned} \tag{4.5.3}$$

As above, in the large-k limit the term coming from boundary conditions at $z = z_*$ (the first term in the second line of (4.5.3)) will drop out and so can be ignored completely. The term from boundary conditions at $z = 0$ (the first line of (4.5.3)) will not drop out automatically, and so will contribute to the second entropy variation. This contribution to the entropy variation is known as the entanglement density in the literature and was previously computed in [78, 79]. From (4.5.3) it is clear that the entanglement density is completely determined by the AdS Green's function and is therefore state-independent. By restricting attention to the vacuum-subtracted entropy the entanglement density will drop out, and in any case is not proportional to a δ-function.

Terms in the Source

As in the null deformation discussion of Section 4.4, we need to compute the effects of the source δJ^μ. As we did there, we will accomplish this by cataloging the various terms which can appear in the power series expansion of J^μ as a function of $\bar{X}$ and γ. Again, terms

which scale like $k^n z^m$ ultimately lead to $k^{n-m+d-2}$ dependence at large k for $\delta X^\mu_{(d)}$. Any term in δJ^μ will look like $\delta \bar{X}^\nu$ multiplied by some function of γ and $\bar{X}$. For the purposes of computing δJ^μ only the state-independent parts of $\delta \bar{X}^\nu$, represented by the first line of (4.5.3), will matter. That is because these terms are a function of the combination kz, which means they have $n - m = 0$. Now we just have to consider all of the possible combinations of γ and $\bar{X}$ which multiply $\delta \bar{X}$.

There cannot be any terms in δJ^μ that are schematically of the form $\bar{X}\delta\bar{X}$ with some derivatives but no factors of γ. Such a term would have to come from nonlinearities in the vacuum AdS extremal surface equation. That equation is invariant under $\bar{X} \to -\bar{X}$, so all terms have to have odd parity like the linear terms. Anything of the form $\bar{X}\bar{X}\delta\bar{X}$, or higher powers of $\bar{X}$, will not contribute at large k because of power counting: The vanishing boundary condition means that $\bar{X}$ starts at order z^d, which means that the most favorable possible term of this type, $(\partial_z \bar{X})^2 \partial_z^2 \delta\bar{X}$, still only amounts to a contribution to the entropy variation which scales like k^{2-d}.

Now we consider terms which have at least one factor of γ. Because we have assumed that all couplings have dimension greater than $d/2$ and that expectation values of operators with dimension $\Delta \leq d/2$ vanish, the leading order piece of γ scales like z^d. Thus we can get contributions to $\delta X_{(d)}$ which go like k^0 from source terms which are schematically of the form $\gamma\partial^2\delta\bar{X}$, as well as other combinations. Given their importance, we will analyze terms of the form $\gamma\delta\bar{X}$ below in more detail.

Terms with additional factors of $\bar{X}$ or γ beyond the first power of γ will not lead to non-decaying behavior at large k because of power counting. So we see that only the linear gravitational backreaction is necessary to completely characterize $\Delta S''_{\mu\nu}$. We will now calculate those terms explicitly.

Linearized Geometry

We have reduced our task to computing J^μ to linear order in γ and $\bar{X}^\mu$ (the latter condition comes from our choice of a planar undeformed entangling surface). This is a simple exercise in expanding (4.2.15). The result in position space is

$$\begin{aligned} J^\mu = &-\frac{1}{2}\partial_z\gamma_{cc}\partial_z\bar{X}^\mu + \partial_a(\gamma_{ab}\partial_b\bar{X}^\mu) - \eta^{\mu\nu}\partial_z\gamma_{\nu\rho}\partial_z\bar{X}^\rho \\ &- \eta^{\mu\nu}(\partial_a\gamma_{\nu\rho} + \partial_\rho\gamma_{\nu a} - \partial_\nu\gamma_{a\rho})\partial_a\bar{X}^\rho - \frac{1}{2}\eta^{\mu\nu}(2\partial_a\gamma_{\nu a} - \partial_\nu\gamma_{aa}) - \frac{1}{2}\partial_a\gamma_{cc}\partial_a\bar{X}^\mu. \end{aligned} \tag{4.5.4}$$

a, b, c indices represent the y-directions and repeated indices are summed over. Taking the variation and evaluating at $\bar{X}^\mu = 0$ gives

$$\begin{aligned} \delta J^\mu = &-\frac{1}{2}\partial_z\gamma_{cc}\partial_z\delta\bar{X}^\mu + \partial_a(\gamma_{ab}\partial_b\delta\bar{X}^\mu) - \eta^{\mu\nu}\partial_z\gamma_{\nu\rho}\partial_z\delta\bar{X}^\rho \\ &- \eta^{\mu\nu}(\partial_a\gamma_{\nu\rho} + \partial_\rho\gamma_{\nu a} - \partial_\nu\gamma_{a\rho})\partial_a\delta\bar{X}^\rho \\ &- \frac{1}{2}\eta^{\mu\nu}(2\partial_\rho\partial_a\gamma_{\nu a} - \partial_\rho\partial_\nu\gamma_{aa})\delta\bar{X}^\rho - \frac{1}{2}\partial_a\gamma_{cc}\partial_a\delta\bar{X}^\mu. \end{aligned} \tag{4.5.5}$$

The only terms in (4.5.5) that will contribute at k^0 are those with two y derivatives acting on $\delta\bar{X}^\mu$ or with z derivatives, i.e., the first line of (4.5.5). Then the result for $\delta X^\mu_{(d)}$ at large k is obtained from (4.5.3) as

$$\begin{aligned}\delta X^\mu_{(d)}(k) &= \frac{-1}{2^{d-2}\Gamma(d/2)^2}\left[\left(\langle\gamma^{(d)\mu}_\nu\rangle + \frac{1}{2}h^{ab}\langle\gamma^{(d)}_{ab}\rangle\eta^\mu_\nu\right)\left(\lim_{z\to 0}\frac{1}{2}z^d K_{d/2}(z)^2\right)\right.\\ &\qquad\left. - \left(\eta^\mu_\nu\frac{k^a k^b}{k^2}\langle\gamma^{(d)}_{ab}\rangle\right)\left(\int_0^\infty dz z^{d+1}K_{d/2}(z)^2\right)\right]e^{iky_0}\xi^\nu\\ &= -\frac{8\pi G_N}{d}\left[\langle T^\mu_\nu\rangle + \frac{1}{2}h^{ab}\langle T_{ab}\rangle\eta^\mu_\nu - \frac{d}{d+1}\eta^\mu_\nu\frac{k^a k^b}{k^2}\langle T_{ab}\rangle\right]e^{iky_0}\xi^\nu\end{aligned} \tag{4.5.6}$$

Here we have explicitly included factors of the entangling surface metric h^{ab} (which is equal to δ^{ab}) rather than using repeated a, b indices for added clarity. In the last line, we have used the dictionary (4.2.26) to replace $\gamma^{(d)}_{\mu\nu}$ with $\langle T_{\mu\nu}\rangle$.

The first two terms of (4.5.6) correspond to δ-functions in position space. The final term clearly contains a δ-function piece which will end up being proportional to the trace of $\langle T_{ab}\rangle$, but it also contains off-diagonal contributions. We can use the identity

$$\int d^{d-2}k\ \frac{k^a k^b}{k^2}e^{ik(y-y_0)} \propto \partial_a\partial_b\frac{1}{|y-y_0|^{d-4}} \propto \frac{\delta_{ab} - (d-2)(y-y_0)^a(y-y_0)^b/(y-y_0)^2}{|y-y_0|^{d-2}}. \tag{4.5.7}$$

to see the full effect in position space. However, for our purposes we are only interested in the δ-function contribution. Isolating this part and combining it with the first two terms of (4.5.6), we ultimately find

$$\Delta S''_{\mu\nu} = 2\pi\left(n^\rho_\mu n^\sigma_\nu\langle T_{\rho\sigma}\rangle + \frac{d^2-3d-2}{2(d+1)(d-2)}n_{\mu\nu}h^{ab}\langle T_{ab}\rangle\right) \tag{4.5.8}$$

where $n_{\mu\nu}$ is the normal projector of the entangling surface. This completes our derivation of (6.3.8).

4.6 Discussion

We have found formulas for the δ-function piece of the second variation of entanglement entropy in terms of the expectation values of the stress tensor. In this section we conclude by discussing a number of possible extensions and future applications of this result.

Higher Orders in $1/N$

Since we believe (4.1.2) and (6.3.8) to be valid at finite N, it must be that our calculations are not affected by higher-order corrections within holography.

One potential source of higher-order corrections comes from incorporating quantum fluctuations in the geometry, rather than treating the geometry as a classical background. We have already addressed this issue in Section 4.2, but we will repeat it here. The problem of a fluctuating geometry arises because the metric fluctuation $\gamma_{\mu\nu}$ is actually a quantum operator, and as such a classical expression which is nonlinear in $\gamma_{\mu\nu}$ has an ambiguous quantum interpretation because, in general, $\langle\gamma_{\mu\nu}^2\rangle \neq \langle\gamma_{\mu\nu}\rangle^2$. However, our analysis has shown that the δ-function part of the second entropy variation is determined entirely by terms which are linear in $\gamma_{\mu\nu}$, and so this problem is avoided.

There are two other classes of higher-order corrections we can consider: those coming form higher-curvature corrections to the bulk gravity, and those coming from the bulk entropy. These corrections can be encapsulated in the all-orders formula [6, 49]

$$S = S_{\text{gen}}[e(\mathcal{R})] = S_{\text{Dong}}[e(\mathcal{R})] + S_{\text{bulk}}[e(R)]. \tag{4.6.1}$$

The first term here is the Dong entropy functional [56], which is an integral of geometric data over the surface $e(\mathcal{R})$,[15] and the second term is the bulk entropy lying within the region bounded by $e(\mathcal{R})$. Finally, the surface $e(\mathcal{R})$ is the one that extremizes the S_{gen} functional.

If we ignore the S_{bulk} term for a moment, then S_{Dong} behaves qualitatively the same way as the area in the Ryu-Takayanagi formula. The coordinates $\bar{X}^\mu$ of $e(\mathcal{R})$ obey a certain differential equation, and the variations in the entropy are still related to $\delta X^\mu_{(d)}$ as before. One change is that the overall coefficient of $\delta X^\mu_{(d)}$ relative to the entropy will change in a way that depends on the bulk higher curvature couplings. However, the dictionary relating $\gamma_{\mu\nu}$ to $T_{\mu\nu}$ also changes in a way that precisely preserves (4.1.2) and (6.3.8) [69].

Incorporating the S_{bulk} term is simple in principle but difficult in practice to deal with. Since it is S_{gen} that must be extremized, we have to include an extra term in the extremal surface equation of motion proportional to $\delta S_{\text{bulk}}/\delta\bar{X}^\mu(y)$. That means the bulk entropy itself plays a role in determining the position of the surface. It was argued in [47] (assuming some mild falloff conditions on variations of the bulk entropy) that the presence of this source could be incorporated to all orders simply by removing the explicit bulk entropy term from (4.2.20). In other words, calculating $\delta X^\mu_{(d)}$ using the correct quantum extremal surface equation is enough to properly account for all bulk entropy contributions to the total entropy variation. At order-one in the large-N expansion this prescription agrees with our analysis above, as it must. Beyond this, the most we can say about the contributions of the entropy are arguments of the type given above in Section 4.3. While this is a potential loophole in our arguments, we still believe that our evidence suggests that new contributions to (4.1.2) and (6.3.8) do not appear.

Local Conditions On $\partial\mathcal{R}$ Are Enough

We now briefly discuss why we expect that we can relax the stationarity conditions on the entangling surface to hold just in the vicinity of the deformation point. We will focus on the

[15] Really S_{Dong} is the expectation value of geometric data, but we have already argued that it is enough to treat the geometry classically for our purposes.

null-null case, but a similar result should hold in the non-null case (where it should also be true that our restriction on expectation values for operators with $\Delta < d/2$ is allowed to be local).

We can analyze the source (4.4.6) in a little more detail in the case where we only impose local stationarity near $y = y_0$. Even though in position space $\bar{U}(y_0, z)$ does not contain any state-independent terms at low orders in the z-expansion near, the inherent non-locality of the Fourier transform $\bar{U}(k, z)$ will contain those terms. There are two ways this could affect (4.4.6): through $\delta\Psi = \delta\bar{U}$ or through the h-factor. In either case, the large k limit reduces to the problem back to the globally-stationary setup.

For example, by setting $\delta V(k) = e^{iky_0}$ we can isolate the part of $\delta U_{(d)}$ that gives a δ-function localized at $y = y_0$. Then the important part of $\delta\bar{V}$ (i.e., the state-independent part) is

$$\delta\bar{V}(k, z) = e^{iky_0} 2^{\frac{d-2}{2}} \Gamma(d/2)(kz)^{d/2} K_{d/2}(kz). \tag{4.6.2}$$

Then we can organize (4.4.6) as a derivative expansion of h, with the leading term given by

$$\delta J(k, z) \sim e^{iky_0} h(z, y_0)(kz)^{d/2} K_{d/2}(kz), \tag{4.6.3}$$

and the remaining terms suppressed by powers of k. In other words, the integral over k' in (4.4.6) combined with the $(k - k')$-dependence of δV essentually returns h to position space localized near $y = y_0$. Only the first d derivatives of h at $y = y_0$ will be relevant at large k, so only the first d derivatives of U need to be set equal to zero at $y = y_0$ in order for the large-k behavior to match the case where U vanishes identically. Thus it is enough to have entangling surfaces which are in the $u = 0$ plane up to order d in $y - y_0$.

Note, this crude analysis does not strictly apply if the entangling surface cannot be globally written in terms of functions $U(y), V(y)$. For example, an entangling surface which is topologically a sphere does not fall within the regime of our arguments. We leave an analysis of those types of regions for future work.

Curved Backgrounds

It is interesting to ask what happens to this proof when the boundary spacetime is curved. Our arguments make it clear that $S''_{\mu\nu}$ is completely determined by local properties of the state in the bulk and on the boundary. So naturally one would expect that there is a curved-space analogue of the same formula. In [69, 48], several local conditions on the entangling surface and spacetime curvature were found such that the QNEC would hold in curved space and be manifestly scheme-independent. We would expect that under those same conditions one could show that $S''_{vv} = 2\pi\langle T_{vv}\rangle$. Non-null variations in a curved background have yet to be explored, and it would be interesting to investigate aspects of the curved background setup in more detail.

Connections to the QFC and Gravity

An interesting application of our result is to the interpretation of Einstein's equations. Combining (6.3.8) with Einstein's equations leads to an explicit formula relating geometry to entropy. This result is the latest in a growing trend of connections between geometry and entanglement [80, 81, 82, 83, 84, 85, 86].

We can make a direct connection with the deep result by Jacobson of the Einstein equation of state [70]. There it was argued that Einstein's equations were equivalent to a statement of thermal equilibrium across an arbitrary local Rindler horizon, namely the equation $\delta Q = T\delta S$, together with an assumption that S is proportional to area. This argument used a thermodynamic definition of the entropy without mentioning quantum entanglement. We can give this result a modern interpretation with the equation $S''_{vv} = 2\pi\langle T_{vv}\rangle$.

The connection to our result is most easily phrased in terms of the generalized entropy for a field theory coupled to gravity, which is defined as

$$S_{\text{gen}} = S_{\text{Dong}} + S_{\text{ren}}. \tag{4.6.4}$$

Here G_N is the renormalized Newton's constant, and S_{ren} is the renormalized entropy of the field theory system restricted to a region, and S_{Dong} is the same geometric functional of the boundary of the region introduced in Section 4.6, and which at leading order is Area$/4G_N$. Variations of this quantity were considered in [11], where the conjecture $S''_{\text{gen},vv} \leq 0$ was dubbed the Quantum Focusing Conjecture (QFC).

Inspired by the arguments of [70], we will consider evaluating $S''_{\text{gen},vv}$ on a surface passing through a given point in an arbitrary spacetimem where v now denotes a null direction of our choosing. We will want to make sure that the surface is as close to stationary as possible in the v direction. It is always possible to make the expansion and shear of our surface vanish at the chosen point, but generically these quantities will have nonzero derivatives along the surface. In order to keep our calculations well-defined, and avoid potential violations of the QFC [52], we should consider deformations which are integrated over at least a Planck-sized region of the surface [87]. While not strictly a δ-function, if the mass scales governing the matter sector are must less than the Planck scale then for all practical purposes this is the same as a δ-function deformation from the point of view of the matter entropy. The result of doing this type of deformation is [88]

$$4G_N S''_{\text{gen},vv} = -R_{vv} + 4G_N S_{\text{ren},vv} + O(\ell^2/L^4), \tag{4.6.5}$$

where L is the characteristic scale of the background geometry and ℓ is the Planck scale (or whatever other cutoff scale is appropriate for the effective gravitational theory). The corrections at order ℓ^2/L^4 come both from higher curvature corrections present in S_{Dong} beyond the Area$/4G_N$ term, as well as from the generic non-zero derivatives of the expansion and shear at the central point of the deformation.

Now suppose we imposed the principle that $4G_N S''_{\text{gen},vv}$ is always of order ℓ^2/L^4, which is much smaller than the size $1/L^2$ of the first term $-R_{vv}$. Then it must be that this large

contribution is canceled by $4G_N S_{\text{ren},vv}$, which by our result above (or, more precisely, by the appropriate curved-space generalization) is equal to $8\pi G_N \langle T_{vv} \rangle$. In other words, we would be imposing

$$R_{vv} = 8\pi G_N \langle T_{vv} \rangle + O(\ell^2/L^4). \tag{4.6.6}$$

This is the leading-order part of the full gravitational equations of motion, up to an unknown cosmological constant term coming from our restriction to null variations. The argument can also be run the other way, so that Einstein's equations, interpreted as the leading order part of the gravitational equations of motion, become equivalent to the statement

$$4G_N S''_{\text{gen},vv} = O(\ell^2/L^4). \tag{4.6.7}$$

We have essentially retraced the steps of [70], replacing the Jacobson's original assumption of $\delta Q = T\delta S$ with the this statement about the generalized entropy, together with (4.1.2).

Proof for General CFTs

We view our results as sufficient motivation to look for a proof of (6.3.8) and (4.1.2) in general field theories. In conformal field theories, entanglement entropy can be calculated using the replica trick. A replicated CFT is equivalent to a CFT with a twist defect. Within the technology of defect CFTs, shape deformations of entropy is generated by displacement operators (see [17] for a review of these concepts). The variation $\delta^2 S/\delta V(y)\delta V(y')$ then is related to the OPE structure of displacement operators in this setup. Since the coefficient of the delta function piece in (4.1.1) is fixed to have dimension d and spin 2, one might be able to see that only the stress tensor could appear as a local operator in S''_{vv}. It further needs to be shown that no other non-linear (in the state) contributions could appear in S''_{vv}. Results in that direction will be reported in future work [89].

Chapter 5

Entropy Variations and Light Ray Operators from Replica Defects

5.1 Introduction

Despite much progress in understanding entanglement entropy using bulk geometric methods in holographic field theories [31, 3, 32], significantly less progress has been made on the more difficult problem of computing entanglement entropy directly in field theory. Part of what makes entanglement entropy such a difficult object to study in field theory is its inherently non-local and state-dependent nature.

One way to access the structure of entanglement in field theories is to study its dependence on the shape of the entangling surface. Such considerations have led to important results on the nature of entanglement in quantum field theories [27, 28, 29, 69, 16, 17, 90, 78, 91]. To study the shape dependence of entanglement entropy for QFTs in $d > 2$ dimensions, consider a Cauchy slice Σ containing a subregion $\mathcal{R}$ with entangling surface $\partial\mathcal{R}$ in a general conformal field theory. By unitary equivalence of Cauchy slices which intersect the same surface $\partial\mathcal{R}$, the entanglement entropy for some fixed global state can be viewed as a functional of the entangling surface embedding coordinates $X^\mu(y^i)$ where the y^i with $i = 1, ..., d-2$ are internal coordinates on $\partial\mathcal{R}$. We write:

$$S_\mathcal{R} = S[X(y)]. \tag{5.1.1}$$

The shape dependence of the entanglement entropy can then be accessed by taking functional derivatives. In particular, we can expand the entanglement entropy about some background entangling surface $X(y) = X_0(y) + \delta X(y)$ as

$$\begin{aligned} S[X] = S[X_0] &+ \int d^{d-2}y \left.\frac{\delta S_\mathcal{R}}{\delta X^\mu(y)}\right|_{X_0} \delta X^\mu(y) \\ &+ \int d^{d-2}y d^{d-2}y' \left.\frac{\delta^2 S_\mathcal{R}}{\delta X^\mu(y)\delta X^\nu(y')}\right|_{X_0} \delta X^\mu(y)\delta X^\nu(y') + ... \,. \end{aligned} \tag{5.1.2}$$

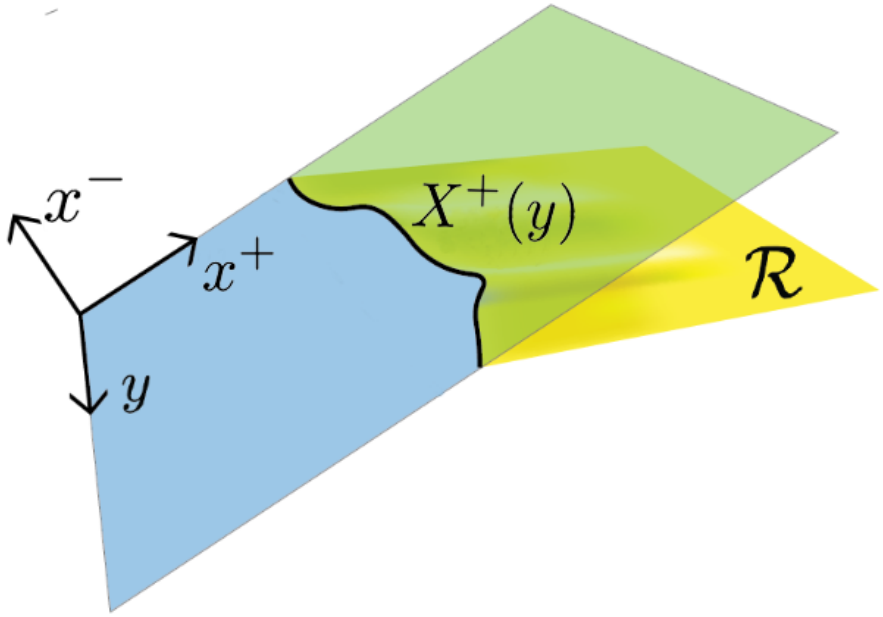

Figure 5.1: We consider the entanglement entropy associated to a spatial subregion $\mathcal{R}$. The entangling surface lies along $x^- = 0$ and $x^+ = X^+(y)$. In this work, we study the dependence of the entanglement entropy on the profile $X^+(y)$.

This second variation has received a lot of attention in part because it is an essential ingredient in defining the *quantum null energy condition* (QNEC) [11, 16]. The QNEC bounds the null-null component of the stress tensor at a point by a specific contribution from the second shape variation of the entanglement entropy. More specifically, this second variation can be naturally split into two pieces - the *diagonal* term which is proportional to a delta function in the internal coordinates y^i and the *off-diagonal* terms[1]

$$\frac{\delta^2 S_{\mathcal{R}}}{\delta X^+(y)\delta X^+(y')} = S''(y)\delta^{(d-2)}(y-y') + (\text{off-diagonal}). \tag{5.1.3}$$

where (X^+, X^-) are the null directions orthogonal to the defect. The QNEC states that the null energy flowing past a point must be lower bounded by the diagonal second variation

$$\langle T_{++}(y)\rangle \geq \frac{\hbar}{2\pi} S''(y), \tag{5.1.4}$$

[1]Note that the entanglement entropy, being UV divergent, will typically have divergent contributions that are local to the entangling surface. These will show up as a limited set of diagonal/contact terms in (5.1.3). For deformations about a sufficiently flat entangling surface these terms do not contribute to the contact term that is the subject of the QNEC. The divergent terms will not be the subject of investigation here.

where we are taking $\mathcal{R}$ to be a Rindler wedge. This inequality was first proposed as the $G_N \to 0$ limit of the quantum focussing conjecture [11], and was first proven in free and super-renormalizable field theories in [28]. The proof for general QFTs with an interacting UV fixed point was given in [17]. More recently, yet another proof was given using techniques from algebraic quantum field theory [18].

The method of proof in the free case involved explicitly computing S''_{++} where it was found that

$$S'' = \frac{2\pi}{\hbar} \langle T_{++} \rangle - Q \tag{5.1.5}$$

where for general states $Q \geq 0$. In contrast, the proof in general QFTs relied on relating the inequality (5.1.4) to the causality of a certain correlation function involving modular flow. This left open the question of whether S'' could be explicility computed in more general field theories.

In [92] the diagonal term S'' was computed in large N QFTs in states with a geometric dual. Remarkably, the result was

$$S''(y) = 2\pi \langle T_{++}(y) \rangle \tag{5.1.6}$$

where we have now set $\hbar = 1$. In other words, $Q = 0$ for such theories. In that work, it was argued that neither finite coupling nor finite N corrections should affect this formula. This led the authors of [92] to conjecture (5.1.6) for all interacting CFTs. The main goal of this paper is to provide evidence for (5.1.6) in general CFTs with a twist gap.

The method of argument will follow from the replica trick for computing entanglement entropy. The replica trick uses the formula

$$S[\mathcal{R}] = \lim_{n \to 1} (1 - n\partial_n) \log \mathrm{Tr}[\rho_{\mathcal{R}}^n] \tag{5.1.7}$$

to relate the entanglement entropy to the partition function of the CFT on a replicated manifold [93, 94] (see also [95, 96, 97, 4])

$$\mathrm{Tr}[\rho_{\mathcal{R}}^n] = Z_n / (Z_1)^n. \tag{5.1.8}$$

At integer n, Z_n can be computed via a path integral on a branched manifold with n-sheets. Alternatively, one can compute this as a path integral on an unbranched manifold but in the presence of a twist defect operator Σ_n of co-dimension 2 that lives at the entangling surface [98]. Doing so allows us to employ techniques from defect CFTs. See [99, 100, 101, 102] for a general introduction to these tools.

In particular, shape deformations of the defect are controlled by a defect operator, namely the displacement operator, with components $\hat{D}_+, \hat{D}_-$. This operator is universal to defect CFTs. Its importance in entanglement entropy computations was elucidated in [99, 17, 98]. Consequently, the second variation of the entanglement entropy is related to the two-point function of displacement operators

$$\frac{\delta^2 S}{\delta X^+(y) \delta X^+(y')} = \lim_{n \to 1} \frac{-2\pi}{n-1} \langle \Sigma_n^\psi \hat{D}_+(y) \hat{D}_+(y') \rangle \,, \tag{5.1.9}$$

where the notation Σ_n^ψ will be explained in the next section.

Since we are interested in the delta function contribution to this second variation, we can take the limit where the two displacement operators approach each other, $y \to y'$. This suggests that we should study the OPE of two displacement operators and look for terms which produce a delta function, at least as $n \to 1$.

It might seem strange to look for a delta function in an OPE since the latter, without further input, results in an expansion in powers of $|y - y'|$. We will find a delta function can emerge from a delicate interplay between the OPE and the replica limit $n \to 1$.

An obvious check of our understanding of (5.1.6) is to explain how this formula can be true for interacting theories while there exist states for which $Q > 0$ in free theories. This is a particularly pertinent concern in, for example, $\mathcal{N} = 4$ super-Yang Mills where one can tune the coupling to zero while remaining at a CFT fixed point. We will find that in the free limit certain terms in the off-diagonal contributions of (5.1.3) become more singular and "condense" into a delta function in the zero coupling limit. In a weakly interacting theory it becomes a question of resolution as to whether one considers Q to be zero or not.

In fact this phenomenon is not unprecedented. The authors of [103] studied energy correlation functions in a so called conformal collider setup. The statistical properties of the angular distribution of energy in excited states collected at long distances is very different for free and interacting CFTs. We conjecture that these situations are controlled by the same physics. Explicitly, in certain special "near vacuum" states, there is a contribution to the second variation of entanglement that can be written in terms of these energy correlation functions.

Schematically, we will find

$$\frac{\delta^2 S}{\delta X^+(y)\delta X^+(y')} - \frac{2\pi}{\hbar}\langle T_{++}\rangle\, \delta^{(d-2)}(y-y') \sim \int ds e^s \langle \mathcal{O}\hat{\mathcal{E}}_+(y)\hat{\mathcal{E}}_+(y')e^{iKs}\mathcal{O}\rangle \tag{5.1.10}$$

where

$$\hat{\mathcal{E}}_+(y) = \int_{-\infty}^{\infty} d\lambda\, \langle T_{++}(x^+ = \lambda, x^- = 0, y)\rangle \tag{5.1.11}$$

is the averaged null energy operator discussed in [103] and the $\mathcal{O}$'s should be thought of as state-creation operators. The operator K is the boost generator about the undeformed entangling surface.

The singularities in $|y - y'|$ of the correlator in (5.1.10) are then understood by taking the OPE of two averaged null energy operators. This OPE was first discussed in [103] where a new non-local "light ray" operator of spin 3 was found to control the small $y - y'$ limit.

In the free limit, we will show that this non-local operator has the correct scaling dimension to give rise to a new delta function term in (5.1.10). In the interacting case this operator picks up an anomalous dimension and thus lifts the delta function.

In other words, the presence of an extra delta function in the second variation of the entanglement entropy in free theories can be viewed as a manifestation of the singular behavior

of the conformal collider energy correlation functions in free theories. This is just another manifestation of the important relationship between entanglement and energy density in QFT.

The presence of this spin-3 light ray operator in the shape variation of entanglement in specific states however points to an issue with our defect OPE argument. In particular one can show that this contribution cannot come directly from one of the local defect operators that we enumerated in order to argue for saturation. Thus one might worry that there are other additional non-trivial contributions to the OPE that we miss by simply analyzing this local defect spectrum. The main issue seems to be that the $n \to 1$ limit does not commute with the OPE limit. Thus in order to take the limit in the proper order we should first re-sum a subset of the defect operators in the OPE before taking the limit $n \to 1$. For specific states we can effectively achieve this resummation (by giving a general expression valid for finite $|y - y'|$) however for general states we have not managed to do this. Thus, we are not sure how this spin-3 light ray operator will show up for more general states beyond those covered by (5.1.10). Nevertheless we will refer to these non-standard contributions as arising from "nonlocal defect operators."

The basic reason it is hard to make a general statement is that entanglement can be thought of as a state dependent observable. This state dependence shows up in the replica trick as a non-trivial n dependence in the limit $n \to 1$ so the order of limits issue discussed above is linked to this state dependence. We are thus left to compute the OPE of two displacement operators for some specific states and configurations. This allows us to check the power laws that appear in the $|y_1 - y_2|$ expansion for possible saturation violations. Given this we present two main pieces of evidence that the nonlocal defect operators do not lead to violations of QNEC saturation. The first is the aforementioned near vacuum state calculation. The second is a new calculation of the fourth shape variation of *vacuum* entanglement entropy which is also sensitive to the displacement operator defect OPE. In both cases we find that the only new operator that shows up is the spin-3 light ray operator. The outline of the paper is as follows.

- In Section 5.2, we begin by reviewing the basics of the replica trick and the relevant ideas from defect conformal field theory. We review the spectrum of local operators that are induced on the defect, including the infinite family of so-called higher spin displacement operators. We show that, in an interacting theory, these higher spin operators by themselves cannot contribute to the diagonal QNEC. We also present a present a certain conjecture about the nonlocal defect operators.

- In Section 5.3, we discuss how a delta function appears in the OPE of two displacement operators. We focus on a specific defect operator that limits to T_{++} as $n \to 1$. For this defect operator we derive a prediction for the ratio of the D_+D_+ OPE coefficient and its anomalous defect dimension. In Section 5.4, we check this prediction by making use of a modified Ward identity for the defect theory. In Appendix A.10-A.11 we also explicitly compute the anomalous dimension and the OPE coefficient to confirm this prediction.

- In Section 5.5, we take up the concern that there could be other operators which lead to delta functions even for interacting CFTs. To do this, we compute the defect four point function $\mathcal{F}_n := \langle \Sigma_n^0 \hat{D}_+(y_1)\hat{D}_+(y_2)\hat{D}_-(y_3)\hat{D}_-(y_4)\rangle$ in the limit $n \to 1$. From this we can read off the spectrum by analyzing the powers of $|y_1 - y_2|$ that appear as $y_1 \to y_2$. We will find that these powers arise from the light-ray OPE of two averaged null energy operators.

- Finally, in Section 5.6, we check our results by explicitly computing the entanglement entropy second variation in near-vacuum states. By using null quantiation for free theories, we show that our results agree with that of [16].

- In Section 5.7, we end with a discussion of our results.

5.2 Replica Trick and the Displacement Operator

In this section, we will review the replica trick and discuss the connection between entanglement entropy and defect operators. This naturally leads to the displacement operator, which will be the key tool for studying (5.1.6).

As outlined in the introduction, the replica trick instructs us to compute the partition function $Z_n/(Z_1)^n = \mathrm{Tr}[\rho_{\mathcal{R}}^n]$, which can be understood as a path integral on a branched manifold $\mathcal{M}_n(\mathcal{R})$, where taking the product of density matrices acts to glue each consecutive sheet together. Using the state operator correspondence, a general state can be represented by the insertion of of a scalar operator in the Euclidean section, so that

$$Z_n = \langle \psi^{\dagger \otimes n} \psi^{\otimes n} \rangle_{\mathcal{M}_n(\mathcal{R})} \tag{5.2.1}$$

where each ψ is inserted on cyclicly consecutive sheets. Alternatively, we can view this $2n$-point correlation function as being computed not on an n-sheeted manifold but on a manifold with trivial topology in the presence of a codimension 2 twist defect operator

$$Z_n = \langle \Sigma_n^0 \psi^{\dagger \otimes n} \psi^{\otimes n} \rangle_{\mathrm{CFT}^{\otimes n}/\mathbb{Z}_n} \equiv \langle \Sigma_n^\psi \rangle \tag{5.2.2}$$

where we have used a compact notation for the twist operator that includes the state operator insertions: $\Sigma_n^\psi \equiv \Sigma_n^0 \psi^{\dagger \otimes n} \psi^{\otimes n}$. It is convenient (and possible) to orbifold the $\mathrm{CFT}^{\otimes n}$ which projects onto states in the singlet of $\mathbb{Z}_n$. This allows us to work with a CFT that for example has only one conserved stress tensor.

We take the defect Σ_n^0 to be associated to a flat cut of a null plane in Minkowski space. We take the metric to be

$$ds^2 = dz d\bar{z} + d\vec{y}^2 \tag{5.2.3}$$

where z and $\bar{z}$ are complexified lightcone coordinates. That is, on the Lorentzian section we have $z = -x^- = x + i\tau$ and $\bar{z} = x^+ = x - i\tau$. Thus, we take the defect to lie at $x^- = X^-(y) = 0$ and $x^+ = X^+(y) = 0$.

For the case of a flat defect, the operator Σ_n^0 breaks the conformal symmetry group down to $SO(2) \times SO(d-1,1)$, with the $SO(2)$ corresponding to the rotations of the plane orthogonal to the defect. This symmetry group suggests that a bulk dimension-d CFT descends to a dimension $d-2$ defect CFT, which describes the excitations of the defect. We can thus use the language of boundary CFTs to analyze this problem. We will only give a cursory overview of this rich subject. For a more thorough review of the topic see [17, 98, 99], and for additional background see [104, 105, 106, 107]. The important aspect for us will be the spectrum of operators that live on the defect.

The spectrum of operators associated to the twist defect was studied in [17]. In that work, techniques were laid out to understand how bulk primary operators induce operators on the defect. This can be quantitatively understood by examining the two-point function of bulk scalar operators in the limit that they both approach the defect. We imagine that as a bulk operator approaches the defect, we can expand in the transverse distance $|z|$ in a bulk to defect OPE so that

$$\lim_{|z|\to 0} \sum_{k=0}^{n-1} \mathcal{O}^{(k)}(z,\bar{z},y)\Sigma_n^0 = z^{-(\Delta_\mathcal{O}+\ell_\mathcal{O})}\bar{z}^{-(\Delta_\mathcal{O}-\ell_\mathcal{O})} \sum_j C_\mathcal{O}^j z^{(\hat{\Delta}_j+\ell_j)/2}\bar{z}^{(\hat{\Delta}_j-\ell_j)/2}\hat{\mathcal{O}}_j(y)\Sigma_n^0 \tag{5.2.4}$$

where $\Delta_\mathcal{O}$ is the dimension of the bulk operator, while $\hat{\Delta}_j$ is the dimension of the jth defect operator $\hat{\mathcal{O}}_j$. Every operator is also now labeled by its spin, ℓ, under the $SO(2)$ rotations $z \to ze^{-i\phi}$. From the defect CFT point of view, the $SO(2)$ spin is an internal symmetry and the ℓ_j's are the defect operators' associated quantum numbers. Notice that the $\mathbb{Z}_n$ symmetry has the effect of projecting out operators of non-integer spin. This is another reason for why the $\mathbb{Z}_n$ orbifolding is needed for treating the theory on the defect as a normal Euclidean CFT.

Equation (5.2.4) suggests an easy way to obtain defect operators in terms of the bulk operators. Consider the lowest dimension defect operator $\hat{\Delta}_\ell$ of a fixed spin ℓ. Then we can extract the defect operator via a residue projection,

$$\hat{\mathcal{O}}_\ell(0)\Sigma_n^0 = \lim_{|z|\to 0} \frac{|z|^{-\hat{\tau}_\ell+\tau_\alpha}}{2\pi i} \oint \frac{dz}{z} z^{-\ell+\ell_\alpha} \sum_{k=0}^{n-1} \mathcal{O}_\alpha^{(k)}(z,|z|^2/z,0)\Sigma_n^0 \tag{5.2.5}$$

where $\hat{\tau}_\ell$ and τ_α are the twists of the defect and bulk operators respectively. Note that these leading twist operators are necessarily defect primaries.

Note that in general, due to the breaking of full conformal symmetry, $\hat{\Delta}_\ell$ will contain an anomalous dimension $\gamma_\ell(n)$. In this paper we will mainly be interested in the defect spectrum near $n=1$ so after analytically continuing in n we can expand $\gamma_\ell(n)$ around $n=1$ as $\gamma(n) = \gamma^{(0)} + \gamma^{(1)}(n-1) + \mathcal{O}((n-1)^2)$. We now give a brief review of the various defect operators discovered in [17].[2]

[2]See [108] for a complementary method for computing the defect spectrum from the bootstrap and an appropriate Lorentzian inversion formula. It would be interesting to derive some of the results presented here in that language.

Operators induced by bulk scalars or spin one primaries

Associated to each bulk scalar ϕ, or spin-one primary V_μ, of dimension Δ_ϕ, Δ_V, the authors of [17] found a family of defect operators of dimension $\hat{\Delta}^\ell_{\phi,V} = \Delta_{\phi,V} - J_{\phi,V} + \ell + \gamma^{(1)}_{\phi,V}(n-1) + \mathcal{O}((n-1)^2)$ with $SO(2)$ spin ℓ along with their defect descendants. Here $J_{\phi,V} = 0, 1$ for ϕ and V respectively and importantly $\ell \geq J$. The anomalous dimensions for the operators induced by bulk scalars, γ_ϕ, are given in formula (3.25) of [17]. We will not be concerned with these two families in this paper.

Operators induced by bulk primaries of spin $J \geq 2$

For primary operators of spin $J \geq 2$, the authors of [17] again found a similar family of operators with dimensions $\hat{\Delta}^\ell_J = \Delta_J - J + \ell + \gamma^{(1)}_{J,\ell}(n-1) + \mathcal{O}((n-1)^2)$ where $\ell \geq J$.

For a primary of spin $J \geq 2$, there are also $J-1$ "new" operators with $SO(2)$ charge $J-1 \geq \ell \geq 1$. These "displacement operators" can be written at integer n as

$$\hat{D}^J_\ell = i \oint d\bar{z} \frac{\bar{z}^{J-\ell-1}}{|z|^{\gamma_{J,\ell}(n)}} \sum_{k=0}^{n-1} \mathcal{J}^{(k)}_{+\ldots+}(|z|^2/\bar{z}, \bar{z}) \tag{5.2.6}$$

where J is the spin of the bulk primary $\mathcal{J}_{+\ldots+}$ and $1 \leq \ell \leq J-1$ is the $SO(2)$ spin of the defect operator. The power of $|z|^\gamma$ accounts for the dependence of the defect operator dimension on n.

We will primarily be interested in the spectrum of T_{++} on the defect for which there is only one displacement operator, $\hat{D}_+$. The displacement operator can also be equivalently defined in terms of the diffeomorphism Ward identity in the presence of the defect [99]

$$\nabla^\mu \langle \Sigma^\psi_n T_{\mu\nu} \rangle = \delta(z, \bar{z}) \langle \Sigma^\psi_n \hat{D}_\nu \rangle. \tag{5.2.7}$$

This implies that $\hat{D}_+$ corresponds to a null deformation of the orbifold partition function with respect to the entangling surface. In particular, entropy variations are given by $\hat{D}_+$ insertions in the limit $n \to 1$:

$$\langle \Sigma^\psi_n \hat{D}_+(y) \rangle = (n-1) \langle \Sigma^\psi_n \rangle \frac{\delta S_\psi}{\delta x^+(y)} + \mathcal{O}((n-1)^2) \tag{5.2.8}$$

The generalization to two derivatives is then just

$$\langle \Sigma^\psi_n \hat{D}_+(y) \hat{D}_+(y') \rangle = (n-1) \langle \Sigma^\psi_n \rangle \frac{\delta^2 S_\psi}{\delta X^+(y) X^+(y')} + \mathcal{O}((n-1)^2). \tag{5.2.9}$$

We see importantly that statements about entropy variations can be related directly to displacement operator correlation functions.

5.3 Towards saturation of the QNEC

With the displacement operator in hand, we can now describe an argument for QNEC saturation. As just described, second derivatives of the entanglement entropy can be computed via two point functions of the defect CFT displacement operator. Thus, we are interested in proving the following identity:

$$\lim_{n\to 1}\frac{1}{n-1}\langle\Sigma_n^{\psi}\hat{D}_+(y)\hat{D}_+(y')\rangle = 2\pi\,\langle\hat{T}_{++}(y)\rangle_{\psi}\,\delta^{d-2}(y-y') + (\text{less divergent in } |y-y'|) \tag{5.3.1}$$

where $|\psi\rangle$ is any well-defined state in the CFT.

Since we are only interested in the short distance behavior of this equality - namely the delta function piece - we can examine the OPE of the displacement operators

$$\frac{1}{n-1}\hat{D}_+(y)\hat{D}_+(y') = \frac{1}{n-1}\sum_{\alpha}\frac{c_\alpha(n)\hat{\mathcal{O}}^{\alpha}_{++}(y)}{|y-y'|^{2(d-1)-\Delta_\alpha+\gamma_\alpha(n)}} + \text{descendants} \tag{5.3.2}$$

where Δ_α is the dimension of the defect primary $\hat{\mathcal{O}}_\alpha$ at $n=1$ and $\gamma_\alpha(n)$ gives the n dependence of the dimension away from $n=1$. We will refer to $\gamma_\alpha(n)$ as an anomalous dimension. Note that this is an OPE defined purely in the defect CFT. The $++$ labels denote the $SO(2)$ spin of the defect operator, which must match on both sides of the equation. The dimension of the displacement operators themselves are independent of n and fixed by a Ward identity to be $d-1$.

At first glance, this equation would suggest that there are no delta functions in the OPE, only power law divergences. In computing the entanglement entropy, however, we are interested in the limit as $n\to 1$. In this limit, it is possible for a power law to turn into a delta function as follows:

$$\lim_{n\to 1}\frac{n-1}{|y-y'|^{d-2-\gamma^{(1)}(n-1)}} = \frac{S_{d-3}}{\gamma^{(1)}}\delta^{(d-2)}(y-y') \tag{5.3.3}$$

where $\gamma=\gamma^{(1)}(n-1)+\mathcal{O}((n-1)^2)$ and S_{d-3} in the area of the $d-3$ sphere. Comparison of equations (5.3.3) and (5.3.2) shows that a delta function can "condense" in the $\hat{D}_+\times\hat{D}_+$ OPE only if the OPE coefficient and anomalous dimension obey

$$c_\alpha(n)/\gamma_\alpha(n) \sim (n-1)+\mathcal{O}((n-1)^2) \tag{5.3.4}$$

as n approaches 1.

This is, however, not sufficient for a delta function to appear in (5.3.2) as $n\to 1$. We also need to have

$$\Delta_\alpha = d \tag{5.3.5}$$

at $n = 1$. In other words, the defect operators we are looking for must limit to an operator of $SO(2)$ spin two and dimension d as the defect disappears. Clearly, the $\ell = 2$ operator induced by the bulk stress tensor, $\hat{T}_{++}$, satisfies these conditions. Indeed, the first law of entanglement necessitates the appearance of $\hat{T}_{++}$ in the $\hat{D}_+ \times \hat{D}_+$ OPE with a delta function (see Section 5.4 below).

Our main claim, (5.3.1), is the statement that no other operator can show up in (5.3.2) whose contribution becomes a delta function in the $n \to 1$ limit. In the rest of this section, we enumerate all the possible operators that could appear in the $\hat{D}_+ \times \hat{D}_+$ OPE (5.3.2).

Defect operators induced by low-dimension scalars

If there exists a scalar operator of dimension $\Delta = d - 2$, then the associated defect operator with $SO(2)$ spin $\ell = 2$ will have dimension $\Delta = d$ at leading order in $n - 1$. This possibility was discussed in [92]. The contribution of such an operator was found to drop out of the final quantity $\langle T_{++} \rangle - \frac{1}{2\pi} S''_{++}$ for holographic CFTs. We expect the same thing to happen in general CFTs in the presence of such an operator, so we ignore this possibility.

$\ell = 2$ operators induced by spin one primaries

As discussed earlier, these defect operators have dimension $\hat{\Delta} = \Delta_V + 1 + \mathcal{O}(n-1)$. We see that for spin one primaries not saturating the unitarity bound, i.e. $\Delta_V > d-1$, these cannot contribute delta functions. Actually, since these operators exist in the CFT at $n = 1$, we will argue in the next section that the first law of entanglement forces their OPE coefficients to be of order $(n-1)^2$.

For spin-one primaries saturating the unitarity bound, V_μ is then the current associated to some internal symmetry. The entropy is uncharged under all symmetries, so such operators cannot contribute to $\hat{D}_+ \times \hat{D}_+$.

$\ell = 2$ higher spin displacement operators

The most natural candidate for contributions to the $\hat{D}_+ \times \hat{D}_+$ OPE are the $\ell = 2$ higher spin displacement operators discussed in the previous section. These operators are given by equation (5.2.6).

To show that such operators do not contribute delta functions to $\hat{D}_+ \times \hat{D}_+$, we need to argue that their dimensions $\Delta_n(\ell = 2, J)$ do not limit to d as $n \to 1$. As discussed in the previous section, the dimensions of the higher spin displacement operators are given by

$$\Delta_n(\ell, J) = \Delta_J - J + \ell + \mathcal{O}(n-1). \tag{5.3.6}$$

The anomalous dimensions have not yet been computed but we expect them to be of order $n - 1$, although we will not need this calculation here. The important point for us will be that in a CFT with a twist gap, the leading order dimension of these operators is

$$\Delta_n(2, J) = \tau_J + 2 + \mathcal{O}(n-1) > d \tag{5.3.7}$$

assuming the twist of the bulk primaries satisfies $\tau_J > d-2$. Here we are using a result on the convexity of twist on the leading Regge trajectory for all J proven in [109]. We see that the bulk higher spin operators would need to saturate the unitarity bound to contribute a delta function. Furthermore, there could be defect descendants of the form $(\partial_y^i \partial_y^i)^k \hat{D}^J_{++}(y)$. But such operators will necessarily contribute to the OPE with larger, positive powers of $|y-y'|$, hence they cannot produce delta functions.

Nonlocal defect operators

So far we have focused on the individual contribution of local defect operators and by power counting we see that these operators cannot appear in the diagonal QNEC. At fixed n, it is reasonable to conjecture that this list we just provided is complete. However we have not fully concluded that something more exotic does not appear in the OPE. As discussed in the introduction this possibility arises because the $n \to 1$ limit may not commute with the OPE.

Indeed, we will find evidence that something non-standard does appear in the displacement OPE. In Section 5.5 and Section 5.6 we will present some computations of correlation functions of the displacement operator for particular states and entangling surfaces. In these specific cases we will be able to make the analytic continuation to $n \to 1$ before taking the OPE. In both cases, we find that the power laws as $y_1 \to y_2$ are controlled by the dimensions associated to non-local spin-3 light ray operators [110]. In the discussion section we will come back to the possibility that these contributions come from an infinite tower of the local defect operators that we have thus far enumerated. We conjecture that when this tower is appropriately re-summed, we will find these non-standard contributions to the entanglement entropy.

We will refer to these operators as *nonlocal defect operators*, and we further conjecture that a complete list of such operators and dimensions is determined by the nonlocal $J=3$ lightray operators that appear in the lightray OPE of two averaged null energy operators as studied in [103, 111] for the CFT *without* a defect. In order to give further evidence for this conjecture, in Section 5.5 we will compute the analytic continuation of the spectrum of operators appearing around $n=1$ in the $\hat{D}_+ \times \hat{D}_+$ OPE by computing a fourth order shape variation of vacuum entanglement. Our answer is consistent with the above conjecture. While this relies on a specific continuation in n (a specific choice of "state dependence") we think this is strong evidence that we have not missed anything.

Before studying this nonlocal contribution further, we return to the local defect contribution where we would like to check that the ratio of $c(n)/\gamma(n)$ for $\hat{T}_{++}$ obeys (5.3.4).

5.4 Contribution of $\hat{T}_{++}$

In this section, we will review the first law argument which fixes the coefficient of the stress tensor defect operator to leading order in $n-1$. We will then use defect methods to demon-

strate that the stress tensor does contribute with the correct ratio of $c(n)$ and $\gamma(n)$ to produce a delta function with the right coefficient demanded by the first law. To do this, we will make use of a slightly modified form of the usual diffeomorphism Ward identity in the presence of a twist defect that will compute $c(n)/\gamma(n)$. In Appendices A.10 and A.11, we also explicitly calculate $c(n)$ and $\gamma(n)$ separately for the stress tensor and show that they agree with the result of this sub-section.

The First Law

A powerful guiding principle for constraining which defect operators can appear in the OPE (5.3.2) is the first law of entanglement entropy. The entanglement entropy $S(\rho) = -\operatorname{Tr}[\rho \log \rho]$, when viewed as the expectation value of the operator $-\log \rho$, is manifestly non-linear in the state. The first law of entanglement says that if one linearizes the von Neumann entropy about a reference density matrix - σ - then the change in the entropy is just equal to the change in the expectation value of the vacuum modular Hamiltonian. Specifically it says that

$$\delta \operatorname{Tr}[\rho \log \rho] = \operatorname{Tr}[\delta\rho \log \sigma] \tag{5.4.1}$$

where $\rho = \sigma + \delta\rho$.

The case we will be interested in here is when σ is taken to be the vacuum density matrix for the Rindler wedge. The first law then tells us that the *only* contributions to $\langle \Sigma_n^{\psi} \hat{D}_+(y)\hat{D}_+(y')\rangle$ that are linear in the state as $n \to 1$ must come from the shape variations of the vacuum modular Hamiltonian.

The second shape derivative of the Rindler wedge modular Hamiltonian is easy to compute from the form of the vacuum modular Hamiltonian associated to generalized Rindler regions [15, 27, 112, 68]. Defining $\Delta \langle H^{\sigma}_{\mathcal{R}} \rangle_{\psi} = -\operatorname{Tr}[\rho_{\mathcal{R}} \log \sigma_{\mathcal{R}}] + \operatorname{Tr}[\sigma_{\mathcal{R}} \log \sigma_{\mathcal{R}}]$ to be the vacuum subtracted modular Hamiltonian for a general region $\mathcal{R}$ bounded by a cut of the $x^- = 0$ null plane, then we have the simple universal formula

$$\frac{\delta^2 \Delta \langle H^{\sigma}_{\mathcal{R}} \rangle_{\psi}}{\delta X^+(y) \delta X^+(y')} = \frac{2\pi}{\hbar} \langle T_{++} \rangle_{\psi} \delta^{(d-2)}(y - y'). \tag{5.4.2}$$

This is a simple but powerful constraint on the displacement operator OPE; it tells us that the only operator on the defect which is manifestly linear in the state as $n \to 1$ and appears in $\hat{D}_+ \times \hat{D}_+$ at $n = 1$ is the stress tensor defect operator

$$\hat{T}_{++} = \oint \frac{d\bar{z}}{\bar{z}|z|^{\gamma_n}} \sum_{j=0}^{n-1} T^{(j)}_{++}(|z|^2/\bar{z}, \bar{z}). \tag{5.4.3}$$

Thus, any other operator which appears in the OPE around $n = 1$ must contribute in a manifestly non-linear fashion. Examining the list of local defect operators discussed in Section 5.3 the only operators that are allowed by the above argument, aside from $\hat{T}_{++}$, are

the higher spin displacement operators. As shown in [17] the limit $n \to 1$ of the expectation value of these operators give a contribution that is non-linear in the state.

We will return to these state dependent operators in later sections. Now we check that indeed the stress tensor contributes with the correct coefficient.

Using the modified Ward identity

In Appendix A.8, we prove the following intuitive identity:

$$\int d^{d-2}y' \langle \Sigma_n^0 \hat{D}_+(y') \hat{D}_+(y) T_{--}(w, \bar{w}, 0) \rangle = -\partial_{\bar{w}} \langle \Sigma_n^0 \hat{D}_+(y) T_{--}(w, \bar{w}, 0) \rangle. \tag{5.4.4}$$

We now show that the identity (5.4.4) allows us to compute the stress tensor contribution to the $\hat{D}_+ \times \hat{D}_+$ OPE, which can be written as:

$$\hat{D}_+(y) \hat{D}_+(y') \supset \frac{c(n)}{|y - y'|^{d-2-\gamma(n)}} \hat{T}_{++}(y) + \dots \tag{5.4.5}$$

where we have focused on the $\hat{T}_{++}$ contribution and the ellipsis stand for the defect descendants of $\hat{T}_{++}$. We are free to ignore other defect primaries since they get projected out by the $T_{--}(w, \bar{w}, 0)$ insertion in (5.4.4). Of course, since (5.4.4) involves a y integral, one might worry that we are using the OPE outside its radius of convergence. For now, we will follow through with this heuristic computation using the OPE. At the end of this subsection, we will say a few words about why this is justified.

Inserting (5.4.5) into (5.4.4) and ignoring the descendants, we find

$$\int d^{d-2}y' \frac{c(n)}{|y - y'|^{d-2-\gamma(n)}} \langle \Sigma_n^0 \hat{T}_{++}(y) T_{--}(w, \bar{w}, 0) \rangle = \frac{c(n)}{\gamma(n)} S_{d-3} \langle \Sigma_n^0 \hat{T}_{++}(y) T_{--}(w, \bar{w}, 0) \rangle \tag{5.4.6}$$

where S_n is the area of the unit n-sphere. We can write $\hat{T}_{++}(y)$ in terms of T_{++} integrated around the defect:

$$\hat{T}_{++}(y) = -\frac{1}{2\pi i} \sum_{k=0}^{n-1} \oint \frac{d\bar{z}}{\bar{z} |z|^{\gamma(n)}} T_{++}^{(k)}(|z|^2/\bar{z}, \bar{z}, y) \tag{5.4.7}$$

We now take the $n \to 1$ limit of equation (5.4.4). Since the right hand side starts at order $(n-1)$, we see that $c(n)$ must begin at one higher order in $n-1$ than $\gamma(n)$. Generically we expect $\gamma(n)$ to begin at order $n-1$ and in Appendix A.11 we will see that it does. We thus get the relation

$$\frac{c^{(2)}}{\gamma^{(1)}} \langle \Sigma_1^0 \hat{T}_{++}(y) T_{--}(w, \bar{w}, 0) \rangle = -\partial_n \big|_{n=1} \partial_{\bar{w}} \langle \Sigma_n^0 \hat{D}_+(y) T_{--}(w, \bar{w}, 0) \rangle \tag{5.4.8}$$

where $c(n) = c^{(1)}(n-1) + c^{(2)}(n-1)^2 + ...$ and $\gamma(n) = \gamma^{(1)}(n-1) + ...$.

At $n = 1$, $\langle \Sigma_1^0 \hat{T}_{++}(y) T_{--}(w, \bar{w}, 0)\rangle$ is just the usual stress tensor 2-point function. Moreover, we can evaluate the right hand side of (5.4.4) at order $(n-1)$ by following the steps leading up to eq. (3.31) of [17]. This leads to

$$\begin{aligned}\partial_{\bar{w}} \langle \hat{D}_+(y) T_{--}(w, \bar{w}, 0)\rangle\Big|_{|w|\to 0} &= i(n-1) \oint d\bar{z}\ \partial_{\bar{w}} \left(\int_0^{-\infty} \frac{d\lambda\ \lambda^2}{(\lambda-1)^2} \frac{c_T y^4}{4(w\bar{w} - w\bar{z}\lambda + y^2)^{d+2}} \right)\Big|_{|w|,|z|\to 0} \\ &= -2\pi(n-1)\frac{c_T}{4} y^{-2d} \end{aligned} \tag{5.4.9}$$

We are then left with the following expressions for c_1 and c_2:

$$c^{(2)} = \frac{2\pi\gamma^{(1)}}{S_{d-3}},\ c^{(1)} = 0 \tag{5.4.10}$$

This is exactly what is needed in order to write (5.4.5) near $y = y'$ as $\hat{D}_+(y)\hat{D}_+(y') \supset \delta^{(d-2)}(y-y')\hat{T}_{++}(y)$.

We now comment on the justification for using the $\hat{D}_+ \times \hat{D}_+$ OPE. Since the left hand side of (5.4.4) involves a y integral over the whole defect, one might worry that the we have to integrate outside the radius of convergence for the $\hat{D}_+ \times \hat{D}_+$ OPE. We see, however, that the y integral produces an enhancement in $(n-1)$ only for the T_{++} primary. In particular, this enhancement does not happen for the descendants of T_{++}. This suggests that if we were to plug in the explicit form of the defect-defect-bulk 3 point function into equation (5.4.4) we would have seen that the $(n-1)$ enhancement comes from a region of the y integral where $\hat{D}_+$ and $\hat{D}_+$ approach each other. We could then effectively cap the integral over y so that it only runs over regions where the OPE is convergent and still land on the same answer. As a check of our reasoning, in Appendices A.10 and A.11, we also compute the $c(n)$ and $\gamma(n)$ coefficients separately and check that they have the correct ratio.

5.5 Higher order variations of vacuum entanglement

In this section, we return to the possibility mentioned in Section 5.3 that something non-standard might appear in the displacement operator OPE. The authors of [17] argued that they had found a complete list of all local defect operators. This leaves open the possibility that the $n \to 1$ limit behaves in such a way that forces us to re-sum an infinite number of defect operators. In this Section and the next, we will find evidence that indeed this does occur. We will also give evidence that we have found a complete list of such nonlocal operators important for the $\hat{D}_+ \times \hat{D}_+$ OPE. In interacting theories with a twist gap this list does not include an operator with the correct dimension and spin that would contribute a delta function and violate saturation.

To get a better handle on what such a re-summed operator might be, we turn to explicitly computing the spectrum of operators in the $\hat{D} \times \hat{D}$ OPE. To do this, we consider the defect

four point function

$$\mathcal{F}_n(y_1,y_2,y_3,y_4) = \langle \Sigma_n^0 \hat{D}_+(y_1)\hat{D}_+(y_2)\hat{D}_-(y_3)\hat{D}_-(y_4)\rangle \,. \tag{5.5.1}$$

We will consider configurations where $|y_1 - y_2| = |y_3 - y_4|$ are small but $|y_1 - y_4|$ is large. With these kinematics, we can use the $\hat{D}\times\hat{D}$ OPE twice and re-write the four point function as a sum over defect two point functions

$$\mathcal{F}_n = \sum_{\mathcal{O},\mathcal{O}'} \frac{c^{\mathcal{O}}_{++}(n)c^{\mathcal{O}'}_{--}(n)\,\langle \Sigma_n^0 \hat{\mathcal{O}}_{++}(y_2)\hat{\mathcal{O}}'_{--}(y_4)\rangle}{|y_1-y_2|^{2(d-1)+\hat{\Delta}^{\mathcal{O}}_n}|y_3-y_4|^{2(d-1)+\hat{\Delta}^{\mathcal{O}'}_n}} \tag{5.5.2}$$

where $\mathcal{O},\mathcal{O}'$ denote the local defect primaries and their descendants appearing in $\hat{D}\times\hat{D}$. We immediately see that by examining the powers of $|y_1 - y_2|$ appearing in $\mathcal{F}_n$, we can read off the spectrum of operators we are after. That is, at least before taking the limit $n \to 1$. We have not attempted to compute the OPE coefficients explicitly for all the local defect operators. This is left as an important open problem that would greatly clarify some of our discussion, but this is beyond the scope of this paper.

If we assume that the $n \to 1$ limit commutes with the OPE limit $y_1 \to y_2$ we can now find a contradiction. To see this contradiction, we can compute $\lim_{n\to 1}\mathcal{F}_n$ in an alternate manner holding y_1, y_2 fixed and compare to (5.5.2). The main result we will find is that the divergences in $|y_1 - y_2|$ appear to arise from defect operators of dimension $\Delta_{J_*} - J_* + 2$ where $J_* = 3$ and Δ_{J_*} is defined by analytically continuing the dimensions in (5.3.6) to odd J (recall that (5.3.6) was only considered for even spins previously.) Generically we do not expect these particular dimensions to appear in the list of operator dimensions of the local defect operators that we enumerated. However we conjecture that by including such operator dimensions we complete the list of possible powers that can appear in the displacement OPE at $n = 1$.

This discussion further suggests that the final non-local defect operator that makes the leading contribution beside T_{++} should be an analytic continuation in spin of the local higher spin displacement operators. We will come back to this possibility in the discussion.

We now turn to computing $\mathcal{F}_n$ without using the defect OPE. In Appendix A.12, we explicitly do the analytic continuation of $\mathcal{F}_n$, but here we simply state the answer. We find that $\mathcal{F}_n$ takes the form

$$\begin{aligned}\mathcal{F}_n \sim (n-1)\int ds e^{-s}\Big\langle T_{--}(x^+=0,x^-=-1,y_3)\hat{\mathcal{E}}_+(y_1)\hat{\mathcal{E}}_+(y_2)T_{--}(x^+=0,x^-=-e^{-s},y_4)\Big\rangle \\ + \mathcal{O}\left((n-1)^2\right),\end{aligned} \tag{5.5.3}$$

which can also be written as:

$$\mathcal{F}_n \sim (n-1)\frac{\Big\langle \mathcal{E}_-(y_3)\hat{\mathcal{E}}_+(y_1)\hat{\mathcal{E}}_+(y_2)\mathcal{E}_-(y_4)\Big\rangle}{\operatorname{vol} SO(1,1)}. \tag{5.5.4}$$

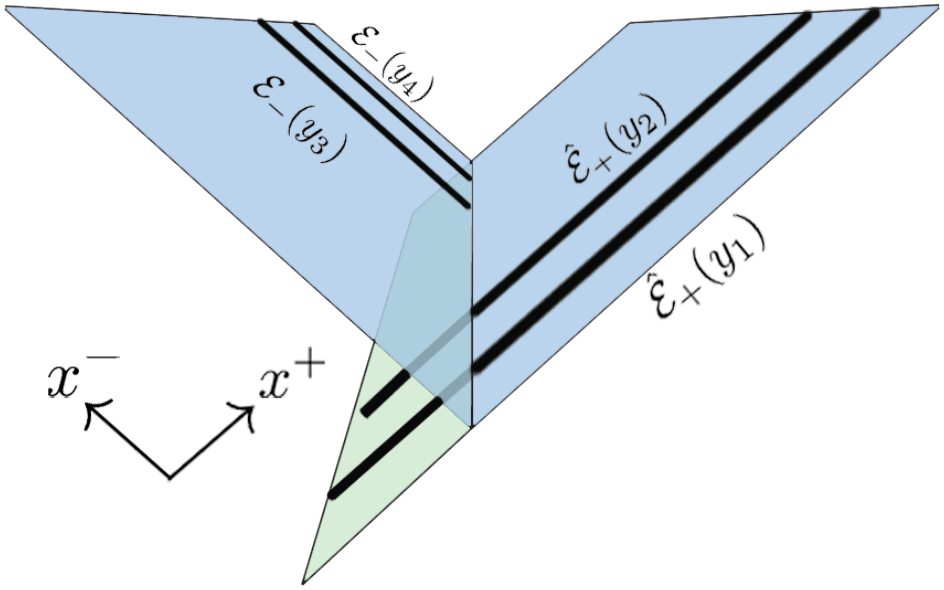

Figure 5.2: The answer for the defect four point function $\mathcal{F}_n$ upon analytic continuation to $n = 1$. We find that there are two insertions of half-averaged null energy operators, $\mathcal{E}_-$, as well as two insertions of $\hat{\mathcal{E}}_+$. Note that strictly speaking, in (5.5.3), the half-averaged null energy operators are inserted in the right Rindler wedge, but by CRT invariance of the vacuum, we can take the half-averaged null energy operators to lie in the left Rindler wedge instead, as in the figure.

The later division by the infinite volume of the 1 dimensional group of boosts is necessary to remove an infinity arising from an overall boost invariance of the four light-ray integrals. See for example [113]. The un-hatted $\mathcal{E}_-$ operators represent half averaged null energy operators, integrated from the entangling surface to infinity. Similar modifications to light-ray operators were used in [111] in order to define their correlation functions and it is necessary here since otherwise the full light-ray operator would annihilate the vacuum.

We see that the effect of two $\hat{D}_+$ insertions was to create two $\hat{\mathcal{E}}_+$ insertions in the limit $n \to 1$. Thus considering the OPE of two displacement operators leads us to the OPE of two null energy operators. This object was studied in [103] and more recently [111]. These authors found that the two averaged null energy insertions can be effectively replaced by a sum over spin 3 "light-ray" operators, one for each Regge trajectory. In other words,

$$\hat{\mathcal{E}}_+(y_1)\hat{\mathcal{E}}_+(y_2) \sim \sum_i \frac{c_i \hat{\mathbb{O}}_i(y_2)}{|y_1 - y_2|^{2(d-2)-\tau^i_{\text{even},J=3}}} \tag{5.5.5}$$

where $\tau^i_{\text{even},J=3}$ is the twist of the even J primaries on the ith Regge trajectory analytically continued down to $J = 3$. A delta function can appear in this expression if $\tau^i_{\text{even},J=3} = d-2$, i.e. if the dimensions saturate the unitarity bound.

Using the recent results in [109] again, we know that the twists on the leading Regge trajectory obey $\frac{d\tau(J)}{dJ} \geq 0$ and $\frac{d^2\tau(J)}{dJ^2} \leq 0$. Since the stress tensor saturates the unitarity bound, for a theory with a twist gap we know that $\tau^i_{\text{even},J=3} > d-2$, therefore there cannot be a delta function in $y_1 - y_2$. By the previous discussion then, formula (5.5.3) suggests that there are no extra operators besides the stress tensor that produce a delta function. To give further evidence for this we next explicitly work out another case where we can compute the $n \to 1$ limit before we do the OPE and we find the same spectrum of operators.

5.6 Near Vacuum States

We have just seen that the OPE of two displacement operators appears to be controlled by defect operators of dimension $\Delta_{J=3} - 1$. As a check of this result, we will now independently compute the second variation of the entanglement entropy for a special class of states. In these states, we will again see the appearance of the OPE of two null energy operators $\hat{\mathcal{E}}_+(y)\hat{\mathcal{E}}_+(y')$. This again implies a lack of a delta function for theories with a twist gap.

This computation is particularly illuminating in the case of free field theory where we can use the techniques of null quantization (see Appendix A.13 for a brief review). Null quantization allows us to reduce a computation in a general state of a free theory to a near-vacuum computation. In this way we will also reproduce the computations in [16] using a different method.

The state we will consider is a near vacuum state reduced to a right half-space

$$\rho(\lambda) = \sigma + \lambda\delta\rho + \mathcal{O}(\lambda^2) \tag{5.6.1}$$

where σ is the vacuum reduced to the right Rindler wedge. We can imagine $\rho(\lambda)$ as coming from the following pure state reduced to the right wedge

$$|\psi(\lambda)\rangle = \left(1 + i\lambda \int dr d\theta d^{d-2}y g(r,\theta,y)\mathcal{O}(r,\theta,y)\right)|\Omega\rangle + \mathcal{O}(\lambda^2) \tag{5.6.2}$$

where (r,θ,y) are euclidean coordinates centered around the entangling surface and

$$\mathcal{O}(r,\theta,y) = \exp\left(iH^\sigma_R\theta\right)\mathcal{O}(r,0,y)\exp\left(-iH^\sigma_R\theta\right) \tag{5.6.3}$$

where H^σ_R is the Rindler Hamiltonian for the right wedge.

From this expression for $|\Psi(\lambda)\rangle$, we have the formula

$$\delta\rho = \sigma \int dr d\theta d^{d-2}y f(r,\theta,y)\mathcal{O}(r,\theta,y) \tag{5.6.4}$$

where

$$f(r,\theta,y) = i\left(g(r,\theta,y) - g(r,2\pi-\theta,y)^*\right). \tag{5.6.5}$$

Note that f obeys the reality condition $f(r,\theta,y) = f(r,2\pi-\theta,y)^*$.

We are interested in calculating the shape variations of the von-Neumann entropy. To this aim, since the vacuum has trivial shape variations we can compute the vacuum-subtracted entropy ΔS instead. We start by using the following identity

$$\Delta S = \mathrm{Tr}\left((\rho(\lambda)-\sigma)\,H^\sigma\right) - S_{\mathrm{rel}}(\rho(\lambda)|\sigma). \tag{5.6.6}$$

We can now obtain ΔS to second order in λ. The vacuum modular Hamiltonian of the Rindler wedge is just the boost energy

$$\mathrm{Tr}\left[(\rho(\lambda)-\sigma)H^\sigma\right] = \int d^{d-2}y \int dv v\,\mathrm{Tr}\left[\rho(\lambda)T_{++}(u=0,v,y)\right] \tag{5.6.7}$$

where the computation of $S_{\mathrm{rel}}(\rho(\lambda)|\sigma)$ was done in Appendix B of [114]. There it was demonstrated that

$$S_{\mathrm{rel}}(\rho(\lambda)|\sigma) = -\frac{\lambda^2}{2}\int \frac{ds}{4\sinh^2(\frac{s+i\epsilon}{2})}\,\mathrm{Tr}\left[\sigma^{-1}\delta\rho\sigma^{\frac{is}{2\pi}}\delta\rho\sigma^{\frac{-is}{2\pi}}\right] + \mathcal{O}(\lambda^3) \tag{5.6.8}$$

For a pure state like (5.6.2), we can instead write the above expression as a correlation function

$$S_{\mathrm{rel}}(\rho|\sigma) = -\frac{\lambda^2}{2}\int d\mu \int \frac{ds}{4\sinh^2(\frac{s+i\epsilon}{2})}\langle \mathcal{O}(r_1,\theta_1,y_1)e^{is\hat{K}}\mathcal{O}(r_2,\theta_2,y_2)\rangle \tag{5.6.9}$$

where we have used the shorthand

$$\int d\mu = \int dr_{1,2}d\theta_{1,2}d^{d-2}y_{1,2}f(r_1,\theta_1,y_1)f(r_2,\theta_2,y_2) \tag{5.6.10}$$

and $\hat{K} = H^\sigma_R - H^\sigma_L$ is the full modular Hamiltonian associated to Rindler space. This formula (5.6.9) and generalizations has been applied and tested in various contexts [115, 116, 86, 117]. Most of these papers worked with perturbations about a state and a cut with associated to a modular Hamiltonian with a local flow such as the Rindler case. However it turns out that this formula can be applied more widely where $\hat{K}$ need not be local.[3]

We can thus safely replace the Rindler Hamiltonian in (5.6.9) with the Hamiltonian associated to an arbitrary cut of the null plane. This allows us to take shape deformations directly from (5.6.9); by using the algebraic relation for arbitrary-cut modular Hamiltonians [68]

$$e^{-i\hat{K}(X^+)s}e^{i\hat{K}(0)s} = e^{i(e^s-1)\int dy\int dx^+ X^+(y)T_{++}(x^+)} \tag{5.6.11}$$

[3]The only real subtlety is the angular ordering of the insertion of $\mathcal{O}$ in Euclidean. This can be dealt with via an appropriate insertion of the modular conjugation operator - a detail that does not affect the final result. We plan to work out these details in future work.

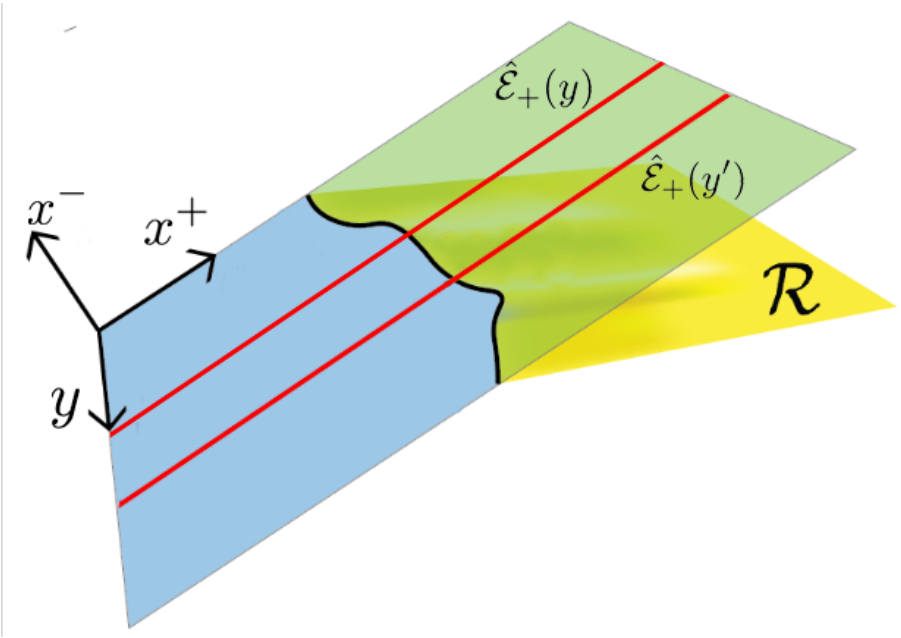

Figure 5.3: For near vacuum states, the insertions of displacement operators limit to two insertions of the averaged null energy operators $\hat{\mathcal{E}}_+$.

we have

$$\frac{\delta^2 S_{\text{rel}}(\rho|\sigma)}{\delta X^+(y)\delta X^+(y')} = \frac{\lambda^2}{2}\int d\mu \int ds e^s \langle \mathcal{O}(r_1,\theta_1,y_1)\mathcal{E}_+(y)\mathcal{E}_+(y')e^{is\hat{K}(X^+)}\mathcal{O}(r_2,\theta_2,y_2)\rangle \quad (5.6.12)$$

where the states ρ, σ depend implicitly on $X^+(y)$.[4] Notice that upon taking the variations the double poles in the $1/\sinh^2(s/2)$ kernel of (5.6.8) were precisely canceled by the factors of $e^s - 1$ in the exponent of equation (5.6.11).

This equation is the main result of this section. We see that taking shape derivatives of the entropy can for this class of states be accomplished by insertions of averaged null energy operators. This helps to explain the appearance and disappearance of extra delta functions as we change the coupling in a CFT continuously connected to a free theory. For example, in a free scalar theory, one can show that the OPE contains a delta function,

$$\hat{\mathcal{E}}_+(y)\hat{\mathcal{E}}_+(y') \supset \delta^{d-2}(y-y'). \quad (5.6.13)$$

This is consistent with the findings of [28] where this extra delta function contribution to the QNEC was computed explicitly. To this aim, in Appendix A.13, we explicitly reproduce the answer in [28] using the above techniques.

[4]Note the similarity between (5.6.12) and (A.12.6). This is because one can view the defect four point function in (5.5.3) as going to second order in a state-deformation created by stress tensors with a particular smearing profile.

5.7 Discussion

In this discussion, we briefly elaborate on the possible origin of the non-local operators whose dimensions we found in the displacement operator OPE considered in Sections 5.5 and 5.6. As mentioned in the main text, the appearance of new operators is a bit puzzling since the authors in [17] found a complete set of defect operators as $n \to 1$. In other words, at fixed $n > 1$, it should in principle be possible to expand these new operators as a (perhaps infinite) sum of $\ell = 2$ defect operators.

In particular, we expect them to be representable as an infinite sum over the higher spin displacement operators. We believe that it is necessary to do such an infinite sum before taking the $n \to 1$ limit, which entails that the OPE and replica limits do not commute. This is why [17] did not find such operators. It also seems, given the non-trivial re-derivation of the results in [17] using algebraic tecniques in [18], that these new non-local defect operators are not necessary for the limit $n \to 1$ limit of the bulk to defect OPE used in [17] to compute modular flow correlation functions.

We give the following speculative picture for how the nonlocal defect operators might arise:

$$\hat{D}_+(y_1)\hat{D}_+(y_2) = \frac{c_{J=2}(n)\hat{T}_{++}}{|y_1 - y_2|^{2(d-1)-\Delta_n^{J=2}}} + \sum_{J=3}^{\infty} \frac{c_J(n)\hat{D}^{(J)}_{++}}{|y_1 - y_2|^{2(d-1)-\Delta_n^J}} \tag{5.7.1}$$

where we have suppressed the contribution of defect descendants. The latter sum in (5.7.1) comes from the spin 2 displacement operators that come from the spin J CFT operator. This is a natural infinite class of operators that one could try to re-sum should that prove necessary.

In our calculations, we did not see any powers in $|y_1 - y_2|$ that could be associated to any individual higher spin displacement operator (as in the second term in (5.7.1)). Instead, in Section 5.5 and Section 5.6 after taking the $n \to 1$ limit we observed dimensions that did not belong to any of the known local defect operators. One possibility is that the higher spin operators in (5.7.1) re-sum into a new term that has a non-trivial interplay with the $n \to 1$ limit. One way this might happen is if the OPE coefficients of the higher spin displacement operators take the form

$$c_{J=2k}(n) \sim \frac{1}{(J-3)(n-1)^{J-3}} \tag{5.7.2}$$

so that they diverge as n approaches 1. Such a divergent expansion is highly reminiscent of the Regge limit for four point functions where instead the divergence appears from the choice of kinematics. This pattern of divergence where the degree increases linearly with spin can be handled using the Sommerfeld-Watson trick for re-summing the series. The basic idea is to re-write the sum as a contour integral in the complex J-plane. One then unwraps the contour and picks up various other features depending on the correlator.

Our conjecture in (5.7.2) is that the other features which one encounters upon unwrapping the J contour is quite simple: there is just one pole at $J = 3$. Upon unwrapping the

contour in the J-plane, we pick up the pole at $J = 3$, which suggests that indeed these new divergences in $|y_1 - y_2|$ are associated to operators which are analytic continuations in spin of the higher spin displacement operators. In this way we would reproduce the correct power law in $|y_1 - y_2|$ as predicted for near vacuum states.

Note that this needs to be true for *any* CFT - not just at large N or large coupling. The universality of this presumably comes from the universality of three point functions. Indeed, one can try to compute these OPE coefficients. We should consider the following three point function:

$$\langle \Sigma_n^0 \hat{D}_+(y_1)\hat{D}_+(y_2)\hat{D}^{(J)}_{--}(y_3)\rangle \sim \frac{c_J(n)\,\langle \Sigma_n^0 \hat{D}^{(J)}_{++}(y_2)\hat{D}^{(J)}_{--}(y_3)\rangle}{|y_1-y_2|^{2(d-1)-\hat{\Delta}_n(J)}} \tag{5.7.3}$$

Via calculations based on the results in Appendix A.9, we find the three point function above in the the replica limit is:

$$\sim (n-1)\oint dw w^{J-3}\,\langle \mathcal{J}_{-\ldots-}(w,\bar{w}=0,y_3)\hat{\mathcal{E}}_+(y_1)\mathcal{E}_+(y_2)\rangle + \mathcal{O}((n-1)^2). \tag{5.7.4}$$

Naively, the full null energy operator $\hat{\mathcal{E}}_+(y_1)$ commutes with the half null energy operator $\mathcal{E}_+(y_2)$ and one can use the fact that $\hat{\mathcal{E}}_+(y_1)\,|\Omega\rangle = 0$ to conclude that $c_J(n=1)$ vanishes. This seems to be incorrect however due to a divergence that arrises in the null energy integrals. Rather we claim that this coefficient diverges. The way to see this is to write

$$\begin{aligned}&\langle \mathcal{J}_{-\ldots-}(w,\bar{w}=0,y_3)\hat{\mathcal{E}}_+(y_1)\mathcal{E}_+(y_2)\rangle = \\ &\int_{-\infty}^{\infty} dx_1^+ \int_0^{\infty} dx_2^+\,\langle \mathcal{J}_{-\ldots-}(w,\bar{w}=0,y_3)T_{++}(0,x_1^+,y_1)T_{++}(0,x_2^+,y_2)\rangle\,.\end{aligned} \tag{5.7.5}$$

We can now attempt to apply the bulk OPE between the two T_{++}'s which in these kinematics must become[5]

$$T_{++}(x^-=0,x_1^+,y_1)T_{++}(x^-=0,x_2^+,y_2) = \sum_{J=2}^{\infty}\frac{(x_{12}^+)^{J-4}\mathcal{J}^J_{+\ldots+}(x_2^+,y_2)}{|y_1-y_2|^{2(d-1)-\hat{\Delta}_1(J)}} + (\text{descendants}). \tag{5.7.6}$$

where $\hat{\Delta}_1(J) = \Delta_J - J + 2$. Plugging (5.7.6) into (5.7.5) and re-labeling $x_1 \to \lambda_1 x_2$, we see that for even $J \geq 3$, the λ_1 integral has an IR divergence

[5]To get the exact answer, one needs to account for all of the $SO(2)$ descendants in this OPE as well since they contribute equally to the higher spin displacement operator. We expect all of these descendants to have the same scaling behavior with $n-1$ and $J-3$.

One can cut-off the integral over λ_1 at some cutoff Λ. The answer will then diverge like

$$\frac{\left(\int_{-\Lambda}^{\Lambda} d\lambda_1\, \lambda_1^{J-4}\right)}{|y_1-y_2|^{2(d-1)-\hat{\Delta}_1(J)}} \times \int_0^\infty dx_2 x_2^{J-3}\, \langle \mathcal{J}_{-\ldots-}(w,\bar{w}=0,y_3)\mathcal{J}_{+\ldots+}(z=0,\bar{z}=x_2^+,y_2)\rangle$$
$$\sim \frac{\Lambda^{J-3}}{J-3}\int_0^\infty dx_2\, x_2^{J-3}\, \langle \mathcal{J}_{-\ldots-}(w,\bar{w}=0,y_3)\mathcal{J}_{+\ldots+}(z=0,\bar{z}=x_2^+,y_2)\rangle \times \frac{1}{|y_1-y_2|^{2(d-1)-\hat{\Delta}_1(J)}}. \tag{5.7.7}$$

The $\mathcal{J}-\mathcal{J}$ correlator on the right is precisely the order $n-1$ piece in $\langle \Sigma_n^0 \hat{D}^J_{++}\hat{D}^{(J)}_{--}\rangle$ so we find that the OPE coefficient scales like $c(n=1)\sim \frac{\Lambda^{J-3}}{J-3}$.

Since Λ is some auxiliary parameter, it is tempting to assign $\Lambda\sim 1/(n-1)$; we then find the conjectured behavior in (5.7.2). This is ad hoc and we do not have an argument for this assignmennt, except to say that the divergence is likely naturally regulated by working at fixed n close to 1. This is technically difficult so we leave this calculation to future work.

Chapter 6

Ignorance is Cheap: From Black Hole Entropy To Energy-Minimizing States In QFT

6.1 Introduction and Summary

There is a remarkable interplay between testable low-energy properties of quantum field theory (QFT), and certain conjectures about quantum gravity, in which the area of surfaces is associated to an entropy. For example, the classical focussing theorem in General Relativity relies on the Null Energy Condition and so can fail in the presence of quantum matter. A Quantum Focussing Conjecture (QFC) was proposed to hold in the semiclassical regime; it implements a quantum correction to the classical statement by replacing the area with the area plus exterior entropy, i.e., the "generalized entropy." This was a guess about quantum gravity, but it led to a new result in QFT. Namely, the Quantum Null Energy Condition (QNEC) was discovered as the QFT limit of the QFC [11].

The QNEC has since been laboriously proven within relativistic quantum field theory [16, 17, 18]. The fact that the QNEC arises more directly and simply from a hypothesis about quantum gravity is striking. Experimental tests of the QNEC may be viable and should be regarded as test of this hypothesis.

Here we will discover a related but distinct connection of this type. We begin again with a classical gravity construction, though one motivated by quantum gravity. The notion that black holes carry Bekenstein-Hawking entropy (proportional to their area) has been fruitful and widely explored, but we stress here that it is a hypothesis that has not been experimentally tested. This hypothesis leads to a puzzle: if the black hole was formed from a pure state, then the entropy should vanish. Thus the Bekenstein-Hawking entropy must be the von Neumann entropy of another quantum state, presumably one that is obtained by an appropriate coarse-graining of the original state. What characterizes this coarse-grained state?

This question was the subject of a recent conjecture by Engelhardt and Wall (EW) [118]. The EW conjecture applies to a class of surfaces that may lie on or inside the event horizon. The Bekenstein-Hawking entropy associated with a "minimar" surface σ is the area of the extremal (Ryu-Takayanagi [31] or HRT [32]) surface, maximized over all spacetimes that agree with the given solution outside of σ. (The input spacetime may have no such surface and thus no entropy.) Engelhardt and Wall showed that the coarse-grained entropy so defined does indeed agree with the area of σ. The interpretation of extremal surface area as an entropy in the quantum gravity theory is well-motivated by the success of the RT proposal in asymptotically Anti-de Sitter spacetimes. We review the EW coarse-graining procedure in Sec. 6.2.

However, the EW construction and proof are purely classical. In particular, the construction fails when quantum matter is included, because it relies on the Null Energy Condition. Moreover, there is considerable evidence that in semi-classical gravity, it is the generalized entropy [119] (and not the area) that is naturally associated with thermal states of the underlying quantum gravity theory [5, 74].

Here, we will formulate a semi-classical extension of the EW coarse-graining proposal for black hole states; that is, we include effects that are suppressed by one power of $G\hbar$ compared to the classical construction. In Sec. 6.3, we consider a suitably defined quantum version of a "minimar" surface. At this order, we must hold fixed not only its exterior geometry but also the exterior state of the quantum fields. We conjecture a construction that explains the generalized entropy of the quantum minimar surface σ in terms of a suitably coarse-grained state: one can find an interior completion of the geometry and quantum state that contains a quantum stationary surface [5, 74, 49] with equal generalized entropy, but none with larger generalized entropy. Moreover, we propose that saturation is obtained by extending σ along a stationary null hypersurface whose classical and quantum expansions both vanish.

Unlike the classical EW construction, we cannot prove our conjecture. But in Sec. 6.4, following the example of the QFC $\rightarrow$ QNEC derivation, we are able to extract a pure quantum field theory limit. We apply our construction to states on a fixed background black hole spacetime with a complete Killing horizon. In this limit, coarse-graining requires the existence of QFT states with specific and somewhat surprising properties, which we list. The most striking property of the coarse-grained state is that the energy flux across the horizon has delta-function support on σ; and that it vanishes at all earlier times on the horizon. (At later times the state agrees with the input state by construction.) The strength of the delta function is set by the derivative of the von Neumann entropy along the horizon in the input state, $\hbar S'/2\pi$.

In particular, the existence of a quantum state with these properties would imply a new result in QFT, Wall's "ant conjecture" [66] concerning the minimum energy of global completions of a half-space quantum state. (We review the ant conjecture in Appendix A.14. The QNEC follows from this conjecture, but it has also been directly proven.) Our proposal thus implies that a state that maximizes the generalized entropy minimizes the nongravitational energy inside of a cut of a Killing horizon, subject to holding fixed the state on the outside. Roughly speaking, ignorance saves energy.

In fact, Wall's ant conjecture was recently proven by Ceyhan and Faulkner (CF) [18]. The CF construction takes as input a state on a Killing horizon and a cut at some surface σ on the horizon. Connes cocycle flow then generates a family of states that differ only to the past of the cut. In the limit of infinite flow, a state is approached whose properties prove the ant conjecture.

In greater than 1+1 dimensions, the requirements we derive appear to be stronger than those demanded by the ant conjecture; see Appendix A.14. Thus it is not immediately obvious that the quantum states required for our coarse-graining proposal exist. However, in Sec. 6.5 we show that the CF family of states attains all of the properties required by our conjecture. In particular, a delta function shock appears at the cut, with precisely the predicted strength. It is interesting that this feature arises in an algebraic construction whereas in the black hole setting, it arose geometrically from requiring a source for a discontinuity in the metric derivative. Thus, the CF construction proves the QFT limit of our conjecture, even though it was originally designed to prove the ant conjecture.

We briefly discuss some future directions in Sec. 6.6.

6.2 Classical coarse-graining of black hole states

In this section we review a classical geometric construction by Engelhardt and Wall (EW) [118, 120]. In Sec. 6.2, we provide definitions of (classically) marginally trapped, "minimar", stationary, and HRT surfaces.

In Sec. 6.2, we summarize the EW proposal for the outer entropy of a "minimar" surface, a marginally trapped surface σ that satisfies certain addition conditions. EW define this entropy in terms of geometries that agree with in the exterior of σ but differ in the interior. For any such auxiliary geometry, inspired by the Ryu-Takayanagi proposal, the von Neumann entropy is assumed to be given by the area of a stationary surface. Maximizing this area over all possible auxiliary geometries, EW show that it agrees with the area of σ, which thus represents a coarse-grained entropy in agreement with the Bekenstein-Hawking formula.

Classical marginal, minimar, and stationary surfaces

We begin by fixing some notations and conventions; see Sec. 2 of [120] for details. Let σ be a *Cauchy splitting surface*, that is, σ is an achronal codimension two compact surface that divides a Cauchy surface Σ into two sides, Σ_{in} and Σ_{out}.

Let k^a, l^a be the two future-directed null vector fields orthogonal to σ, normalized so that $k_a l^a = -1$; and let θ_k, θ_l be their expansions.

If exactly one null expansion vanishes, we shall take this to be the k-expansion. Then σ is called *marginally outer trapped*, with k defining the "outside." If $\theta_l < 0$ everywhere on a marginally outer trapped σ, we call σ *marginally trapped*.

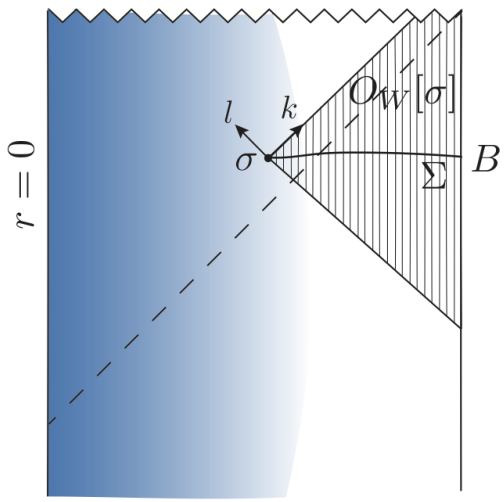

Figure 6.1: Penrose diagram of a black hole formed from collapse in Anti-de Sitter space, showing a minimar surface σ and its outer wedge $\mathcal{O}_W[\sigma]$ with Cauchy surface Σ.

The *outer wedge* $O_W[\sigma]$ of a marginally trapped surface σ is the set of spacelike separated events on the outside of σ (the side that k points towards, see above): $O_W[\sigma] \equiv D[\Sigma_{\text{out}}]$, where D denotes the domain of dependence. See Fig. 6.1.

A *minimar surface* is a marginally trapped surface σ that satisfies two additional restrictions:

- $O_W[\sigma]$ contains a connected component B of an asymptotic conformal boundary (as would be the case, for example, if σ lies in a single black hole formed from collapse in asymptotically anti-de Sitter or flat spacetime). Moreover, $O_W[\sigma]$ admits a Cauchy surface on which σ is the surface homologous to B that minimizes the area; see Fig. 6.1.
- $k^a \nabla_a \theta_{(l)} < 0$

A *stationary surface* X is a surface whose expansion vanishes in both null directions, k and l:

$$\theta_k = \theta_l = 0 \text{ everywhere on } X \ . \tag{6.2.1}$$

A *Hubeny-Rangamani-Takayanagi (HRT) surface* X is a stationary surface that satisfies additional requirements: it is the stationary surface with the smallest area, subject to a homology condition [32, 31]. Here, we will require that X be homologous to a minimar surface σ, and hence to a connected component B of a conformal boundary.

Bekenstein-Hawking entropy from coarse-graining behind minimar surfaces

Engelhardt and Wall [120] argued that the area of a minimar surface σ can be understood as a coarse-grained entropy. For geometries with a CFT dual, an explicit prescription for this coarse-graining can be formulated in the CFT. Here, we will be interested in the bulk definition of this coarse-graining, which can be discussed in more general geometries.

In the bulk, the coarse-graining consists of holding fixed the outer wedge of σ, $O_W[\sigma]$, while erasing the spatial interior of σ and replacing it with an auxiliary geometry. One seeks the auxiliary geometry with the largest possible HRT surface X behind σ. The coarse-grained entropy of σ is defined as $A[X]/4G\hbar$.

So far, we have reviewed the definition of the outer entropy. The EW proposal is the conjecture that

- $S_{\text{outer}} \equiv A[X]/4G\hbar$ represents the von Neumann entropy of a well-defined state in a quantum gravity theory; and
- $A[X] = A[\sigma]$.

EW proved the first part of the conjecture for the special case where B lies on the conformal boundary of an asymptotically AdS spacetime, and σ lies on a perturbed Killing horizon; moreover the proof assumes the Ryu-Takayanagi [31] and HRT [32] proposals for the von Neumann entropy of the boundary CFT. In this case, it is possible to construct the dual CFT state explicitly, and to show that its entropy agrees with S_{outer}.

The second part of the conjecture was proven more generally [120]. Using the maximin definition of the HRT surface [36], it can be shown that

$$A[X] \leq A[\sigma] \ . \tag{6.2.2}$$

This argument assumes the Null Energy Condition (NEC), that the stress tensor satisfies

$$T_{ab}k^a k^b \geq 0 \tag{6.2.3}$$

for any null vector k^a.

EW explicitly construct an interior geometry that saturates the inequality (6.2.2). This implies

$$S_{\text{outer}}[\sigma] \equiv \frac{A[X]}{4G\hbar} = \frac{A[\sigma]}{4G\hbar} \ . \tag{6.2.4}$$

The interior geometry with $A[X] = A[\sigma]$ is constructed by specifying initial conditions on the null hypersurface N_k^- orthogonal to σ towards the interior and past. Appropriate initial data is generated by null-translating the intrinsic geometry of σ, thus generating a stationary null hypersurface:

$$\theta_k = 0 \ \text{ on } N_k^- \ . \tag{6.2.5}$$

This ensures that all cross sections of N_k^-—in particular, X—have the same intrinsic metric and area as σ. This construction is consistent with the relevant constraint, the Raychaudhuri equation,

$$k^a \nabla_a \theta_k = -\frac{1}{2}\theta_k^2 - \varsigma^2 - 8\pi G\, T_{kk} \ , \tag{6.2.6}$$

if one sets

$$\varsigma = 0 \text{ and } T_{kk} = 0 \ \text{ on } N_k^- \ . \tag{6.2.7}$$

on N_k^-. EW [120] show that this choice is always possible. Since θ_k vanishes on σ, Eqs. (A.16.13) and (6.2.7) ensure that the entire extrinsic curvature tensor in the k-direction vanishes everywhere on N_k^-, achieving the desired stationarity of N_k^-.

Moreover, it is important to show that there exists a stationary (HRT) surface X on N_k^-. The outgoing expansion θ_k vanishes on any cut of N_k^-, by the above construction. The question is whether there exists a cut X on which the ingong expansion θ_l vanishes as well. This is accomplished in the following sequence of steps.

The minimar assumption dictates that on σ, $\theta_l < 0$ and $k^a \nabla_a \theta_l < 0$. One can choose initial conditions on N_k^- such that along every null generator of N_k^-, $k^a \nabla_a \theta_l$ is constant and equal to its value on σ: by the cross-focussing equation,

$$k^a \nabla_a \theta_l = -\frac{1}{2}\mathcal{R} - \theta_k \theta_l + \chi^2 + \nabla \cdot \chi + 8\pi G\, T_{kl} \ , \tag{6.2.8}$$

this can be accomplished by choosing all terms on the right hand side to be constant on N_k^-. This is already ensured for the intrinsic curvature scalar $\mathcal{R}$ and for the (vanishing) $\theta_k \theta_l$ term, by stationarity of N_k^-. The twist, or normal 1-form, is defined by

$$\chi_a = h^c{}_a l^d \nabla_c k_d \ , \tag{6.2.9}$$

where $h_{ab} = g_{ab} + 2 l_{(a} k_{b)}$ is the induced metric on a cut. The twist evolves according to

$$k^a \nabla_a \chi_i = 8\pi T_{ik} (+ \text{ terms that vanish when } \theta_k = \varsigma = 0) \ . \tag{6.2.10}$$

To summarize, one can accomplish $k^a \nabla_a \theta_l = k^a \nabla_a \theta_l|_\sigma$ on N_k^- by choosing Eqs. (6.2.5) and (6.2.7) and in addition, along each null generator of N_k^-,

$$T_{kl} = T_{kl}|_\sigma \ \text{ and } T_{ik} = 0 \ \text{ on } N_k^- \ . \tag{6.2.11}$$

Again, EW argue that these choices are always possible.

Let v be the affine parameter associated to k^a, and let y be the transverse coordinates (angular coordinates) on σ. The location of a stationary surface X, $v = f(y)$, is determined by the differential equation

$$L^a[f] = -\theta_l|_\sigma \ , \tag{6.2.12}$$

where L^a is the stability operator (see Ref. [120] for details). This can be shown to have a solution with $-\infty < f < 0$, so the HRT surface exists and lies on N_k^-.

EW then glue the geometry exterior to X (that is, N_k^- and the outer wedge) to its CPT image across X. This constructs a "two-sided" geometry in which X functions as a kind of bifurcation surface of a two-sided black hole/white hole pair. (However, the stationary auxiliary portion N_k^- does not in general correspond to the horizon of a Kerr-Newman black hole, as its intrinsic metric can differ.)

In a final step, EW show that X is not just stationary but is an HRT surface, i.e., that X is the smallest-area stationary surface homologous to σ. This step uses the NEC as well as the second part of the minimar property of σ.

This concludes our summary of the EW coarse-graining prescription. Again, we refer the interested reader to Ref. [120] for more detailed definitions and arguments.

6.3 Semiclassical coarse-graining of black hole states

In this section, we formulate a semiclassical extension of the Engelhardt-Wall construction, starting from a quantum marginally trapped surface σ. We conjecture that the semiclassical state invoked in our construction exists in the full quantum gravity theory; and that in this theory this state has a von Neumann entropy given by the generalized entropy of σ.

In Sec. 6.3, we introduce relevant concepts such as generalized entropy, quantum expansion, quantum marginally trapped surfaces, and quantum HRT surfaces.

In Sec. 6.3, we state our quantum extension of the EW coarse-graining proposal.

In Sec. 6.3, we refine our conjecture by describing key properties that the coarse-grained state is expected to satisfy at the level of semiclassical gravity. (These properties will be shown to have an interesting nongravitational limit in Sec. 6.4. In Sec. 6.5 we will show that a recent construction by Ceyhan and Faulkner [18] generates quantum field theory states which achieve these properties in a certain limit.)

Quantum marginal, minimar, and stationary surfaces

Before we turn to the question of why and how the EW construction should be extended to the semiclassical regime, we introduce here the relevant concepts: generalized entropy, quantum expansion, quantum (marginally) trapped surfaces, and quantum extremal surfaces. More details can be found, e.g., in Refs. [45, 6, 16, 15].

The notion of *generalized entropy* was originally introduced by Bekenstein [119] as an extension of ordinary entropy that includes the contribution from black holes, $S_{\text{out}} \to S_{\text{out}} + \frac{A}{4G\hbar}$. But in an expansion in $G\hbar$, it is the exterior entropy that should be regarded as a quantum correction:

$$S_{\text{gen}} = \frac{A}{4G\hbar} + S_{\text{out}} + \dots \ , \tag{6.3.1}$$

Equivalently, $4G\hbar S_{\text{gen}}$ represents a quantum-corrected area.

In Bekenstein's original proposal, A represented the area of a cut of a black hole event horizon; and S_{out} represented the entropy in the black hole's exterior. However, the general-

ized entropy can be defined for any Cauchy-splitting surface σ, with $S_{\rm out}$ the von Neumann entropy of the quantum fields restricted to one side of σ. $A/4G\hbar$ should be regarded as the leading counterterm that cancels divergences in the entropy; we suppress subleading terms here. Given its wide applicability, the notion of generalized entropy can be used to define quantum-corrected notions of trapped, stationary, etc., as follows.

Recall that the classical expansion of a surface $\hat{\sigma}$ at a point $y \in \hat{\sigma}$ is the trace of the null extrinsic curvature at y. It can also be defined as a functional derivative,

$$\theta[\hat{\sigma}; y] = h(y)^{-1/2} \frac{\delta A[V]}{\delta V(y)} \ , \tag{6.3.2}$$

where h is the area element on $\hat{\sigma}$. Here $V(y)$ defines a surface that lies an affine parameter distance V from $\hat{\sigma}$ along the null geodesic emanating from $\hat{\sigma}$ at y.

The above definition is overkill, as the classical expansion depends only on the local geometry near y. But it generalizes directly to the *quantum expansion*, Θ, which depends on $\hat{\sigma}$ nonlocally:

$$\Theta[\hat{\sigma}; y] = \frac{4G\hbar}{\sqrt{h(y)}} \frac{\delta S_{\rm gen}[V]}{\delta V(y)} \ . \tag{6.3.3}$$

A *quantum marginally outer trapped surface* is a surface whose quantum expansion in one of the two null directions (say, k) vanishes at every point. Let σ be such a surface:

$$\Theta_k[\sigma; y] \equiv 0 \ . \tag{6.3.4}$$

It follows that

$$\theta_k(y) = -\frac{4G\hbar}{\sqrt{h(y)}} \frac{\delta S_{\rm out}}{\delta V(y)} \tag{6.3.5}$$

at every point on σ.

A *quantum marginally trapped surface* is a quantum marginally outer trapped surface for which in addition

$$\Theta_l[\sigma; y] < 0 \ . \tag{6.3.6}$$

(As usual, *anti-trapped* corresponds to the opposite inequality on the l-expansion.)

The *outer wedge* $O_W[\sigma]$ of a quantum marginal surface σ is the set of spacelike separated events on the "marginal" side of σ, i.e., the side that k points towards: $O_W[\sigma] = D[\Sigma_{\rm out}]$; see Fig. 6.1.

A *quantum minimar surface*, is a quantum marginally trapped surface σ that satisfies two additional restrictions:

- $O_W[\sigma]$ contains a connected component of an asymptotic conformal boundary (as would be the case, for example, if σ lies in a single black hole formed from collapse in asymptotically anti-de Sitter or flat spacetime). Moreover, $O_W[\sigma]$ admits a Cauchy surface on which σ is the surface homologous to B that minimizes the generalized entropy; see Fig. 6.1.

- $k^a \nabla_a \theta_l < 0$.

Note that we impose the second condition on the classical expansion, not the quantum expansion. Since the inequality is strict, the classical expansion θ_l will dominate in the semiclassical expansion in $G\hbar$.

A *quantum stationary surface*[1] is a surface whose quantum expansions vanish in both null directions, k and l. We will demand that X be such a surface:

$$\Theta_k[X; y] \equiv 0 \ , \quad \Theta_l[X; y] \equiv 0 \ . \tag{6.3.7}$$

A *quantum HRT surface* satisfies additional requirements: it is the quantum stationary surface with the smallest generalized entropy; and it must obey a homology condition. Here, we will require that it be homologous to a quantum minimar surface σ, and and hence to a connected component B of a conformal boundary.

Generalized entropy from coarse-graining behind quantum marginally trapped surfaces

We will now motivate and formulate a quantum extension of the EW proposal. To see that such an extension is needed, note that the classical EW construction relies on the Null Energy Condition, Eq. (8.1.2). The NEC guarantees that no HRT surface with area greater than that of the marginally trapped surface can be constructed. It also guarantees that the stationary surface with equal area is an HRT surface. But the NEC is known to fail in any relativistic quantum field theory, so none of these conclusions survive at the semiclassical level.

Indeed, one does not expect any quantum state of the full quantum gravity theory to correspond to just the area of a surface (as is implicit in the classical EW construction). Rather, one expects its von Neumann entropy to match the generalized entropy. That is, to the extent that a quantum state corresponds to a surface, one expects it to also describe the surface's exterior.

There is significant evidence supporting this expectation from the AdS/CFT correspondence [1]. Consider the quantum state ρ_B on a region B, where B can be all or part of the boundary. This state is expected [5] to describe the entire *entanglement wedge* of B, i.e., the spacetime region enclosed by B and the HRT surface $X[B]$. The $1/N$ expansion on the boundary (with N the rank of the CFT's gauge group) corresponds to the $G\hbar$ expansion in the bulk. In particular, the von Neumann entropy $S(\rho_B)$ can be expanded in this way, with the leading $O(N^2)$ piece corresponding to the area of $X[B]$, and the subleading $O(1)$ piece corresponding to the exterior bulk entropy S_{out}. When expanding to higher orders, X_B should be taken to be the quantum HRT surface of B [6].

We thus seek a proposal in which the generalized entropy of a surface σ is explained as a coarse-grained entropy. The coarse-graining should correspond to maximizing the generalized

[1] In an abuse of language, this is sometimes referred to as extremal rather than stationary.

entropy of a quantum HRT surface X, subject to holding fixed the outer wedge $O_W[\sigma]$ (now including the quantum state of bulk fields in $O_W[\sigma]$). The coarse-graining prescription will be successful if $S_{\text{gen}}[X] = S_{\text{gen}}[\sigma]$.

The remaining question is what characterizes a surface σ that we may consider for coarse-graining. In the classical case, the appropriate criterion was that σ be minimar. In the quantum case, the natural candidates are minimar surfaces or quantum minimar surfaces. In the EW construction of the maximally coarse-grained state, the HRT surface X of the coarse-grained state lies on a stationary null surface N_k^- extended to the past and inwards from σ. Our construction will share this feature. This excludes (classically) minimar as the relevant criterion for σ. The variation of S_{out} does not have definite sign on such surfaces, and so their quantum expansion would not have a definite sign. However, if $\Theta[\sigma] > 0$ then by the quantum focussing conjecture, it would be impossible to find an X with $\Theta[X] = 0$ on $N_k^-[\sigma]$. Therefore, we will require that σ be quantum minimar; in particular, $\Theta[\sigma] = 0$.

We now state our **proposal**. Let σ be a quantum minimar surface homologous to a boundary region B, with generalized entropy $S_{\text{gen}}[\sigma]$ and outer wedge $O_W[\sigma]$. Let $\bar{X}$ be a quantum HRT surface in any geometry such that:

- $O_W[\bar{X}] \supset O_W[\sigma]$.
- $\bar{X}$ is homologous to σ.
- Both the geometry and the quantum state of $O_W[\bar{X}]$ agree with that of $O_W[\sigma]$ upon restriction of $O_W[\bar{X}]$ to $O_W[\sigma]$. (To be precise, let $\Sigma_{\text{out}}[\bar{X}]$ be a Cauchy surface of $O_W[\bar{X}]$ such that $\Sigma_{\text{out}}[\bar{X}] \cap O_W[\sigma]$ is a Cauchy surface of $O_W[\sigma]$, $\Sigma_{\text{out}}[\sigma]$, and let $\rho_{\bar{X}}$ and ρ_σ be the state of the quantum fields on $\Sigma_{\text{out}}[\bar{X}]$ and $\Sigma_{\text{out}}[\sigma]$, respectively. We require that $\text{Tr}_{\Sigma_{\text{out}}[\bar{X}]-\Sigma_{\text{out}}[\sigma]}\, \rho_{\bar{X}} = \rho_\sigma$.)

We claim that

$$\sup_{\bar{X}} S_{\text{gen}}[\bar{X}] = S_{\text{gen}}[\sigma] \ . \tag{6.3.8}$$

Moreover, let X be a surface $\bar{X}$ that achieves the supremum. (This should be taken as a limiting statement if no such X exists.) Then $O_W[X]$ represents a coarse-graining of the original geometry, with respect to the quantum minimar surface σ. In particular, in AdS/CFT the quantum state on B dual to the entanglement wedge $O_W[X]$ has von Neumann entropy $S_{\text{gen}}[\sigma]$.

Unlike the classical case, we will not prove this conjecture, but we will provide some evidence supporting its plausibility. We proceed in two steps as in the classical case: first, we will argue that

$$S_{\text{gen}}[\bar{X}] \leq S_{\text{gen}}[\sigma] \tag{6.3.9}$$

for any $\bar{X}$ satisfying the conditions in our proposal. We then refine our conjecture by detailing the properties of a semiclassical geometry and quantum state that would achieve equality.

In order to show Eq. (6.3.9), we generalize the result in [120] to the quantum case. This involves two main assumptions. The first assumption is the quantum focusing conjecture [11]

which asserts that in the semi-classical limit the derivative of the quantum expansion of codimension 2 surfaces under any null deformation is non-negative:

$$\frac{\delta\Theta_k[X;y]}{\delta V(y)} \leq 0 \ . \tag{6.3.10}$$

The second assumption is a slightly weaker quantum generalization of the classical maximin construction [36]. More precisely, we assume that the quantum extremal surface $\bar{X}$ is also the surface of minimal generalized entropy on some Cauchy slice Σ.

By global hyperbolicity, the congruence of null geodesics orthogonal to σ in the $\pm k$ directions intersect Σ at some Cauchy splitting surface $\bar{\sigma}$. (The congruence should be terminated at conjugate points or self-intersections [121, 122]. Since σ is a quantum marginally trapped surface, quantum focusing ensures that

$$S_{\rm gen}[\bar{\sigma}] \leq S_{\rm gen}[\sigma] \ . \tag{6.3.11}$$

The quantum maximin assumption further implies

$$S_{\rm gen}[\bar{X}] \leq S_{\rm gen}[\bar{\sigma}] \ , \tag{6.3.12}$$

which establishes Eq. (6.3.9).

Properties of a Generalized Entropy Maximizing Bulk State

We will now describe a geometry and quantum state with a quantum extremal surface X whose generalized entropy saturates the inequality (6.3.9). The existence of a state with the properties we describe would imply our conjecture, Eq. (6.3.8).

By asserting the existence of this semiclassical state, we are refining our conjecture. In Sec. 6.4, we will explore the implications of this refinement in a pure field theory limit. In Sec. 6.5, we will show that these implications are realized in a recent construction by Ceyhan and Faulkner [18].

Our construction will be analogous to the classical one, in that we will approach X along the null hypersurface $N_k^-[\sigma]$. Since we require $\Theta_k[\sigma] = \Theta_k[X] = 0$, the quantum focussing conjecture ($\Theta_k' \leq 0$) requires that $\Theta_k = 0$ everywhere on N_k^-. That is, $S_{\rm gen}$ must be constant along N_k^-. (This is analogous to classical focussing and the null energy condition requiring that N_k^- have constant area in the classical case.)

In the classical case, all relevant quantities could be chosen to be constant on N_k^-. In other words, the surface N_k^- is truly stationary. This would not be the case if θ and the derivative of the entropy varied along N_k^-, with only their sum Θ_k vanishing. Motivated by this observation, we conjecture that a state can be found such that the two terms in Θ_k vanish *separately* on N_k^-:

$$\theta_k = 0 \quad \text{and} \quad \frac{\delta S_{\rm out}}{\delta V(y)} = 0 \ . \tag{6.3.13}$$

In analogy with the classical construction we also take the shear tensor to vanish at all orders in $\hbar$ along N_k^-:

$$\varsigma = 0 \ . \tag{6.3.14}$$

These considerations place nontrivial constraints on the limit state we seek. For θ_k and ς to vanish everywhere on N_k^-, the stress tensor component T_{kk} must vanish on N_k^-. Moreover, note that θ_k need not vanish on σ, where only $\Theta_k = 0$ is required. It follows that generically, θ_k must jump discontinuously, by an amount

$$\Delta\theta_k|_\sigma = -\frac{4G\hbar}{\sqrt{h(y)}} \left.\frac{\delta S_{\rm out}}{\delta V(y)}\right|_\sigma \ . \tag{6.3.15}$$

By Raychaudhuri's equation, this implies the presence of a delta function term in the stress tensor, at σ. Combining these results, we conclude that

$$T_{vv} = \frac{\hbar}{2\pi} \left.\frac{\delta S_{\rm out}}{\delta V(y)}\right|_\sigma \delta(v) \ \ , v \leq 0 \ , \tag{6.3.16}$$

i.e., in the region $N_k^- \cup \sigma$.

To summarize, we conjecture the existence of a state with

$$T_{vv} = \frac{\hbar}{2\pi} \left.\frac{\delta S_{\rm out}}{\delta V(y)}\right|_\sigma \delta(v) \ \ , \ v \leq 0 \ , \tag{6.3.17}$$

$$\varsigma = 0 \ \ , \ v < 0 \ , \tag{6.3.18}$$

$$\frac{\delta S_{\rm out}}{\delta V(y)} = 0 \ \ , \ v < 0. \tag{6.3.19}$$

Eq. (6.3.17) trivially implies that

$$\int_{-\infty}^{v} dv\, T_{vv} = 0 \ , \tag{6.3.20}$$

and we will use this property in Sec. 6.4.[2] In addition, we assume that the remaining EW conditions listed in Eq. (6.2.11) can be met at the classical level.

With these assumptions, the existence of a classical HRT surface on N_k^- is guaranteed by the argument summarized around Eq. (6.2.12). This surface satisfies $\theta_l = 0$. A quantum stationary surface X can be found nearby (in the $G\hbar \to 0$ limit), by solving iteratively for $\theta_l = -\frac{4G\hbar}{\sqrt{h(y)}} \frac{\delta S}{\delta U(y)}$, where the functional derivative refers to the shape deformation along the l-congruence.

Finally, we need to show that X is quantum HRT, i.e., that it is the quantum stationary surface homologous to σ with smallest generalized entropy. This proceeds in exact analogy with the classical argument [120], with the QFC replacing the NEC, so we will not spell out the argument here. See [123] for details.

[2] Strictly, we must allow for the possibility that a state with the properties we conjecture does not itself exist. It suffices that the properties we require can be arbitrarily well approximated by some family or sequence of states (as in the example of Sec. 6.5). In this case, Eq. (6.3.17) need not imply Eq. (6.3.20), so the latter property should be considered explicitly as part of our refined conjecture.

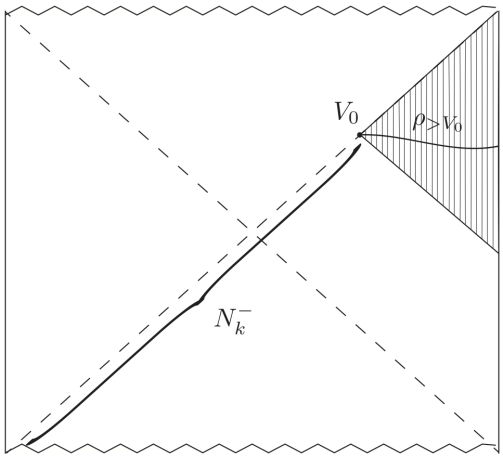

Figure 6.2: Coarse-graining behind a Killing horizon. Any cut V_0 can be viewed as a quantum marginally trapped surface in the limit as $G \to 0$. The state $\rho_{>V_0}$ on the Cauchy surface Σ of the outer wedge is held fixed. The coarse-grained geometry is the original geometry. The stationary null surface N_k^- is the past of V_0 on the Killing horizon. The coarse-grained quantum state demanded by our proposal lives on $N_k^- \cup \sigma \cup \Sigma$. We identify the properties the state must have, and we show that the Ceyhan and Faulkner "ant states" satisfy these.

6.4 Quantum field theory limit of coarse-grained quantum gravity states

In this section, we study the implications of our conjecture for quantum field theory decoupled from gravity. We will apply our proposal to input states that are small perturbations of the Killing horizon of a maximally extended vacuum solution such as Kruskal; see Fig. 6.2.

In the perturbative setting, any quantum marginally trapped surface σ will be at a distance of order G from the Killing horizon, and so will lie on the horizon as $G \to 0$. We can think of the area and null expansion of σ as fields defined on the unperturbed Killing horizon whose changes are sourced by the state of the matter fields on the horizon. Thus, every cut of the Killing horizon can be viewed as quantum marginally trapped, and our conjecture can be applied.

We will first establish notation and review some standard results in Sec. 6.4. In Sec. 6.4, we will derive some interesting additional properties of the coarse-graining states that must hold in the perturbative setting. In the limit as $G \to 0$, our conjecture thus implies the existence of states with both the properties established in the previous section, and the additional properties derived here, in quantum field theory on a fixed background. This is an in-principle testable conjecture about quantum field theory.

Notation, definitions, and standard results

Consider a quantum field theory on a background with a Killing horizon and an arbitrary global state ρ defined on the horizon. Let v be the affine parameter on the Killing horizon, u the affine parameter that moves off of the Killing horizon (associated with null vectors k and l respectively), and take y to be the transverse coordinates on a cut $V(y)$ of the horizon. The cut defines a surface σ, which we assume to be Cauchy-splitting as usual.

Let the *right half-space state* $\rho_{>V_0}$ be the restriction of ρ to the half-space $v > V_0(y)$ as in Fig. 6.2:

$$\rho_{>V_0} \equiv \mathrm{Tr}_{\leq V_0}\rho \ . \tag{6.4.1}$$

where the trace is over the algebra associated with the complement region. Let us denote the *von Neumann entropy* of $\rho_{>V_0}$ by

$$S(V_0) = -\mathrm{Tr}\,\rho_{>V_0}\log\rho_{>V_0} \ . \tag{6.4.2}$$

Let $\sigma \equiv |\Omega\rangle\langle\Omega|$ be the global vacuum, which can be reduced to the *right vacuum* $\sigma_{>V_0} = \mathrm{Tr}_{\leq V_0}\sigma$. The *vacuum-subtracted von Neumann entropy* of $\rho_{>V_0}$ is

$$\Delta S(V_0) = S(V_0) + \mathrm{Tr}\,\sigma_{>V_0}\log\sigma_{>V_0} \ . \tag{6.4.3}$$

The *right (half-)modular Hamiltonian* K is defined by the relation

$$\sigma_{>V_0} = \frac{e^{-K(V_0)}}{\mathrm{Tr}\,e^{-K(V_0)}} \ . \tag{6.4.4}$$

The *right modular energy* in a global state ρ is $\langle K(V_0)\rangle \equiv \mathrm{Tr}\,[K(V_0)\rho_{>V_0}]$, and the *vacuum-subtracted right modular energy* is

$$\begin{aligned} \Delta K(V_0) &\equiv \langle K(V_0)\rangle - \mathrm{Tr}\,[\sigma_{>V_0}K(V_0)] &(6.4.5)\\ &= \frac{2\pi}{\hbar}\int dy\int_{V_0(y)}^{\infty} dv\,[v-V_0(y)]T_{vv} \ , &(6.4.6)\end{aligned}$$

where the explicit expression is due to Bisognano and Wichmann [124] and its generalization to arbitrary cuts of Killing horizons [27, 125]. The *relative entropy* of $\rho_{>v_0}$ with respect to the reduced global vacuum, $\sigma_{>v_0}$, is defined as

$$\begin{aligned} S_{\mathrm{rel}}(V_0) &\equiv S(\rho_{>V_0}|\sigma_{>V_0}) &(6.4.7)\\ &\equiv \mathrm{Tr}\,\rho_{>V_0}\log\rho_{>V_0} - \mathrm{Tr}\,\rho_{>V_0}\log\sigma_{>V_0} \ . &(6.4.8)\end{aligned}$$

It follows from this definition that

$$S_{\mathrm{rel}}(V) = \Delta K(V) - \Delta S(V) \ . \tag{6.4.9}$$

We will often be interested in derivatives, where the vacuum-subtraction drops out. For example,

$$\frac{\delta K}{\delta V(y)} = \frac{\delta\Delta K}{\delta V(y)} = -\frac{2\pi}{\hbar}\int_v^{\infty} d\tilde{v}\ T_{vv}(y) \ . \tag{6.4.10}$$

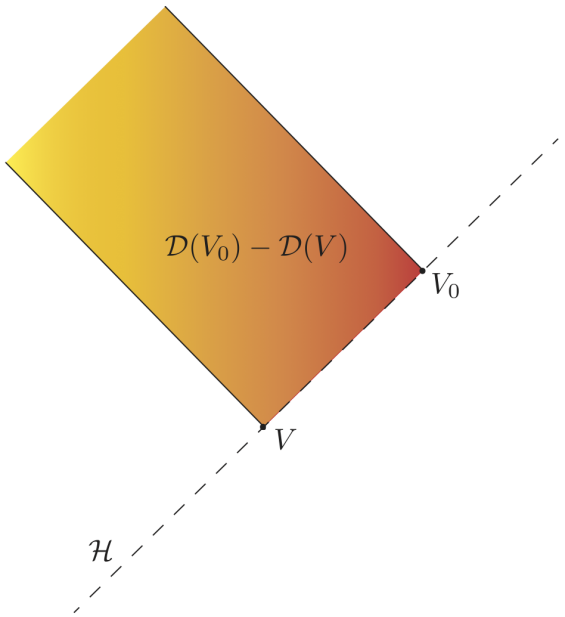

Figure 6.3: The spacetime region associated to the interval $V < v < V_0$ on the null surface for which all observables in the algebra should register vacuum values in the coarse-graining state.

Similar definitions apply to the region $v < V_0$; we denote the associated "left" quantities with an overbar. Strictly, we define the left and right quantities in terms of the limit as $\epsilon \to 0$ of the open intervals $(-\infty, V_0(y) + \epsilon)$ and $(V_0(y) + \epsilon, \infty)$, respectively. The small shift ensures that any distributional sources at $V_0(y)$ contribute asymmetrically to the left but not to the right quantities. (We will see that in the minimum energy states of interest in this paper, the stress tensor generically has a delta function at $V_0(y)$. Our choice resolves an associated ambiguity, attributing this energy entirely to the left.)

The relative entropy satisfies positivity and monotonicity:

$$S_{\rm rel} \geq 0 \ , \quad \frac{\delta S_{\rm rel}}{\delta V} \leq 0 \ . \tag{6.4.11}$$

Via Eq. (6.4.9), monotonicity implies

$$\frac{\delta \bar{K}}{\delta V} \geq \frac{\delta \bar{S}}{\delta V} \geq \frac{\delta S}{\delta V} \ . \tag{6.4.12}$$

The second inequality follows from the strong subadditivity of the von Neumann entropy,

$$S_{BC} + S_{CD} \geq S_B + S_D \ , \tag{6.4.13}$$

applied to the intervals $B = (-\infty, v_0), C = [v_0, v_0 + \delta], D = (v_0 + \delta, \infty)$ in the limit as $\delta \to 0$ [66].

Additional properties of the coarse-graining states

Our conjecture says that the coarse-grained state will have vanishing T_{vv} and constant right entropy in the left region:

$$\langle T_{vv}\rangle = \frac{\hbar}{2\pi}\frac{\delta S}{\delta V(y)}\bigg|_\sigma \delta(v - V_0(y)) \ , \ v \le V_0 \ , \tag{6.4.14}$$

$$\frac{\delta S}{\delta V(y)} = 0 \ , \ v < 0 \ . \tag{6.4.15}$$

In particular, in the strong form of Eq. (6.3.20), these properties imply the ant conjecture (see Appendix A.14).

But additionally, on the Killing horizon, the nested inequalities (6.4.12) hold. Combined with the above equations, this implies that the left von Neumann entropy is also constant:

$$0 = \int_{-\infty}^{V(y)} \langle T_{vv}\rangle \ge \frac{\delta \bar{S}}{\delta V(y)} \ge \frac{\delta S}{\delta V(y)} = 0 \tag{6.4.16}$$

$$\implies \frac{\delta \bar{S}}{\delta V(y)} = 0, \ v < V_0(y) \ . \tag{6.4.17}$$

By Eqs. (6.3.20), (6.4.9) and (6.4.10), it follows that the left relative entropy is constant:

$$\frac{\delta \bar{S}_{\rm rel}(\rho_{<V}|\sigma_{<V})}{\delta V(y)} = 0, \ v < V_0(y) \ . \tag{6.4.18}$$

But the relative entropy is a measure of the distinguishability of the state $\rho_{<V}$ from the vacuum $\sigma_{<V}$. Suppose that by moving up the cut V, i.e., by gaining access to a larger region, one could perform some measurement that would better distinguish $\rho_{<V}$ from the vacuum. Then the relative entropy of the larger region would have to be greater. Thus, Eq. (6.4.18) implies that all observables restricted to the difference between the left domains of dependence associated to cuts $V_0(y)$ and $V(y)$ (as in Fig. 6.3) need to register vacuum values. In particular, the stress tensor one-point function must vanish:

$$\langle T_{\mu\nu}(x)\rangle = 0, \quad x \in \mathcal{D}(V_0) - \mathcal{D}(V) \ . \tag{6.4.19}$$

It is more subtle to draw conclusions about $\langle T_{\mu\nu}(x)\rangle$ when x is on the boundary of the region (marked by red in Fig. 6.3), $u = 0, v < V_0$. Because $T_{\mu\nu}$ does not exist as an operator unless it is smeared to both sides of this boundary, it will not be in the left operator algebra, and it cannot be used to distinguish $\rho_{<V}$ from the vacuum $\sigma_{<V}$.

We will now give a rough physical argument that certain components of $\langle T_{\mu\nu}(x)\rangle$ must vanish also on the Killing horizon below the cut, $u = 0, v < V_0$. We emphasize that this argument is not rigorous, as it borrows from classical intuition. (In forthcoming work we will explore a more detailed coarse-graining proposal involving a family of states; in that setting a rigorous argument can be given.)

Physically, $\langle T_{vv}\rangle$ can be thought of as the momentum orthogonal to an observer's world-line in the (u, v) plane, in the limit as the observer moves at the speed of light in the v-direction. Similarly, T_{iv} is the transverse momentum seen by such an observer. Since all observables in the algebra associated to $\mathcal{D}(V_0) - \mathcal{D}(V)$ have to register vacuum values, no excitations can enter this region. By causality, therefore, the state on the null surface $u = 0, v < V_0$ can only differ from the vacuum by matter moving *along* it, i.e., purely in the v-direction. This implies $\langle T_{vv}\rangle = 0$, consistent with Eq. (6.4.15) above. It also implies the new result

$$\langle T_{iv}\rangle = 0 \ , \quad v < V_0 \ . \tag{6.4.20}$$

Conservation of the stress tensor,

$$-\partial_v \langle T_{uv}\rangle - \partial_u \langle T_{vv}\rangle + \partial_i \langle T_{iv}\rangle = 0 \ , \tag{6.4.21}$$

combined with (6.4.14) then yields

$$\langle T_{uv}\rangle = \text{const} \ . \tag{6.4.22}$$

We conclude that coarse-grained states on Killing horizons must satisfy not only Eqs. (6.4.14) and (6.4.15) but also Eqs. (6.4.17), (6.4.18), (6.4.20), and (6.4.22).

Crucially, these results pertain to quantum field theory on a fixed background, so they can be checked in a rigorous setting. In the next section we will see that all of the above properties are indeed satisfied by the "ant states" constructed by Ceyhan and Faulkner [18]. This proves our conjecture in the Killing horizon limit.

6.5 Existence of coarse-graining states in QFT limit

In this section we show that the "predictions" of the previous section have already been confirmed. We consider a recent explicit construction of states in QFT by Ceyhan and Faulkner (CF) [18]. CF constructed these states in order to prove a conjecture by Wall [126] that we will discuss in detail in Appendix A.14 below. For now, we merely verify that they satisfy the properties we found for the coarse-graining state on Killing horizons in the non-gravitational limit: Eqs. (6.4.14), (6.4.15), (6.4.17), (6.4.18), (6.4.20), and (6.4.22).

Consider a cut $V_0(y)$ of the Rindler horizon $u = 0$ and let $\mathcal{A}_{V_0}, \mathcal{A}'_{V_0}$ be the algebra of operators associated to the region $\{u = 0, v > V_0(y)\}$ and its complement respectively. Given a global state $|\psi\rangle$ we can consider its restriction to $\mathcal{A}_{V_0}$. One can then purify this restriction in different ways, including the trivial purification. We will be interested in the purification introduced in [18], which is based on modular flow.

For the global vacuum $|\Omega\rangle$ recall that the full modular Hamiltonian associated to the cut V_0 defines a modular operator via $K_{V_0} = -\log \Delta_{\Omega;\mathcal{A}_{V_0}}$ and that $\Delta^{is}_{\Omega;\mathcal{A}_{V_0}}$ simply acts as the boost that fixes V_0. We note that $\Delta_{\Omega;\mathcal{A}_{V_0}}$ is related to the reduced density matrix in Eq. (6.4.4) by $\Delta_{\Omega;\mathcal{A}_{V_0}} = \log \sigma_{>V_0} \otimes \mathbb{1}_{<V_0} - \mathbb{1}_{>V_0} \otimes \log \sigma_{<V_0}$.

For a general state $|\psi\rangle$ that is cyclic and separating, one can define the relative modular operator as [127, 124, 128]

$$\Delta_{\psi|\Omega;\mathcal{A}_{V_0}} = S^\dagger_{\psi|\Omega;\mathcal{A}_{V_0}} S_{\psi|\Omega;\mathcal{A}_{V_0}} \ , \tag{6.5.1}$$

where

$$S_{\psi|\Omega;\mathcal{A}_{V_0}} \alpha|\psi\rangle = \alpha^\dagger|\Omega\rangle, \ \forall \alpha \in \mathcal{A}_{V_0} \tag{6.5.2}$$

defines the Tomita operator.

We then purify $|\psi\rangle$ restricted to $\mathcal{A}_{V_0}$ using the Connes cocycle

$$|\psi_s\rangle = u'_s|\psi\rangle, \ u'_s = (\Delta'_\Omega)^{is}(\Delta'_{\Omega|\psi})^{-is} \in \mathcal{A}'_{V_0} \ . \tag{6.5.3}$$

The Connes cocycle can roughly be thought of as a half-sided boost that fixes the state restricted to $\mathcal{A}_{V_0}$ but stretches all of the excited modes in the complement region. Specifically, expectation values of operators in $\mathcal{A}_{V_0}$ are left invariant whereas expectation values of operators in $\mathcal{A}'_{V_0}$ are equivalent to those evaluated in the state $\Delta_\Omega^{-is}|\psi\rangle$. This follows (restricting to cyclic and separating states for simplicity) from the relation $(\Delta')^{is}_{\psi|\Omega}\Delta^{-is}_{\Omega|\psi} = 1$, which implies

$$|\psi_s\rangle = \Delta_\Omega^{-is} u_s|\psi\rangle \ . \tag{6.5.4}$$

If we consider an operator $\mathcal{O}' \in \mathcal{A}'_{V_0}$ then $[u_s, \mathcal{O}'] = 0$ so

$$\langle\psi_s|\mathcal{O}'|\psi_s\rangle = \langle\psi|\Delta_\Omega^{is}\mathcal{O}'\Delta_\Omega^{-is}|\psi\rangle \ . \tag{6.5.5}$$

Note that $v = V_0(y)$ is a fixed point of the boost.

In the limit $s \to \infty$ all of these excitations become soft. More specifically,

$$\begin{aligned}\langle T_{vv}\rangle_s|_{v<V_0(y)} &\equiv \langle\psi_s|T_{vv}(v)|\psi_s\rangle|_{v<V_0(y)} \\ &= e^{-4\pi s}\langle\psi|T_{vv}(V_0 + e^{-2\pi s}(v - V_0))|\psi\rangle|_{v<V_0(y)}\end{aligned} \tag{6.5.6}$$

which just follows from the usual algebra of half-sided modular inclusions. Hence $\langle T_{vv}\rangle_s \to 0$ as $s \to \infty$ for $v < V_0(y)$.

Not only that but also

$$\lim_{s\to\infty}\int_{-\infty}^{v} dv \ \langle T_{vv}\rangle_s \to 0, \ v < V_0(y) \ . \tag{6.5.7}$$

To see what this implies about the energy of the boosted side, we make use of the sum rule derived in [18] for null derivatives of the relative entropy:

$$2\pi\left(P_s - e^{-2\pi s}P\right) = \left(e^{-2\pi s} - 1\right)\frac{\delta S_{\rm rel}(\psi|\Omega;\mathcal{A}_V)}{\delta V}\Big|_{V_0} \ , \tag{6.5.8}$$

where

$$P = \int_{-\infty}^{\infty} dv \ \langle T_{vv} \rangle_\psi \tag{6.5.9}$$

is the average null energy of the original state, P_s is the average null energy of $|\psi_s\rangle$, and

$$S_{\text{rel}}(\psi|\Omega; \mathcal{A}_V) = -\langle \psi| \log \Delta_{\psi|\Omega;\mathcal{A}_V} |\psi\rangle \tag{6.5.10}$$

is the relative entropy of the original state for some general cut $V(y)$.

The relative entropy can also be written as

$$S_{\text{rel}}(\psi|\Omega; \mathcal{A}_V) = \langle K_V \rangle_\psi - S(V) \tag{6.5.11}$$

and moreover [18]

$$\frac{\delta \langle K_V \rangle_\psi}{\delta V}\Big|_{V_0} = -2\pi \int_{V_0(y)}^{\infty} dv \ \langle T_{vv} \rangle_\psi \ . \tag{6.5.12}$$

Thus in the limit $s \to \infty$ we find, using Eq. (6.5.7),

$$\langle T_{vv} \rangle|_{v \leq V_0(y)} = \frac{1}{2\pi} \frac{\delta S}{\delta V}\Big|_{V_0} \ \delta(v - V_0(y)) \tag{6.5.13}$$

as desired. This reproduces both Eq. (6.4.14) and Eq. (6.4.15).

As a final point, note that under the Connes cocycle we also have the following properties:

$$\langle T_{uv} \rangle_{s\to\infty} = \langle T_{uv}(V_0) \rangle_\psi \ , \tag{6.5.14}$$

$$\langle T_{iv} \rangle_{s\to\infty} \to 0 \ . \tag{6.5.15}$$

This very easily reproduces the properties Eq. (6.4.20) and Eq. (6.4.22).

6.6 Discussion

We end by discussing the boundary interpretation of the generalized entropy of a QMT surface. We will also briefly describe future work on a systematic algorithm for constructing the states we conjectured in Sec. 6.3.

Boundary dual

Within AdS/CFT, it is natural to ask whether the coarse-graining prescription for S_{outer} in Sec. 6.3 has a boundary dual. In other words, there must exist a boundary state dual to the bulk coarse-grained semiclassical state of Sec. 6.3. Based on Eq. (6.3.12), we know that the boundary dual to this state is a mixed state that maximizes the boundary von

Neumann entropy subject to fixing the semiclassical state in $O_W[\mu]$. Since in Sec. 6.3 we only considered a case where we have reflecting boundary conditions at infinity, fixing $O_W[\mu]$ amounts to fixing the past boundary of $O_W[\mu]$, labelled $N_{-l}(t_i)$ in Fig. 6.4.

Therefore, the question of whether there is a natural boundary dual to our bulk coarse-graining prescription reduces to that of whether fixing the semiclassical state on $N_{-l}(t_i)$ has a natural interpretation in the boundary. Our answer to this question is very similar to the simple entropy S_{simple} prescription of [118, 120].

Since we would like to refer to the bulk as little as possible, we define the QMT surface μ associated to a time slice t_i of the boundary by constructing an ingoing null surface from t_i and marking the first QMT surface on it. In general, this surface could reach caustics before reaching μ; Ref. [120] deals with this technicality. Here we ignore this issue by restricting to special classes of states (e.g. perturbations to Killing horizons).

Let $\rho(t_i)$ be the original boundary state at time t_i. We would like to construct a boundary state with maximum von Neumann entropy, which agrees with the semiclassical bulk state on $N_{-l}(t_i)$. In order to accomplish this, we must find a boundary definition of $\mathcal{F}$, the set of density matrices dual to the semiclassical state on $N_{-l}(t_i)$.

Let us first consider $\mathcal{F}$ to be the states that agree with $\rho(t_i)$ on simple boundary observables $\mathcal{A}$ on $t > t_i$. Simple observables are defined to be boundary operators whose associated excitations propagate causally in the bulk [118, 120], so this data fixes the bulk causal wedge of $t > t_i$ ($C[t_i]$ in Fig. 6.4). However, $C[t_i] \subseteq O_W[\sigma_{t_i}]$, so in general this set $\mathcal{F}$ would not be constrained enough to fix all of the data on $N_{-l}(t_i)$.

The discrepancy between $C[t_i]$ and $O_W[\sigma_{t_i}]$ arises from matter that enters the black hole to the future of σ_{t_i}. This causes the event horizon to grow and lie properly inside of the outer wedge. To fix all of $O_W[\sigma_{t_i}]$ given $\rho(t_i)$, one must turn on boundary sources that will absorb the future infalling excitations and achieve $C[t_i] = O_W[\sigma_{t_i}]$. This may seem acausal, but so is the definition of simple operator as an operator that can be represented by local boundary operators smeared over space *and time*.

Therefore, the coarse-graining set $\mathcal{F}$ should consist of the states such that the simple boundary observables $\mathcal{A}$ agree with those of $\rho(t_i)$ even after both states have been subject to turning on various simple sources on the boundary:

$$S_{\text{simple}}(t_i) = \max_{\tilde{\rho} \in \mathcal{F}} \; S(\tilde{\rho}) \tag{6.6.1}$$

with

$$\mathcal{F} = \{\rho : \langle E\mathcal{A}E^\dagger \rangle_{\tilde{\rho}(t)} = \langle E\mathcal{A}E^\dagger \rangle_\rho, \; t \geq t_i; \; \forall E\} \tag{6.6.2}$$

where $\mathcal{A}$ is the set of simple observables and E denotes unitaries associated with turning on various simple boundary sources.

Note that $C[t_i] \subseteq O_W[\sigma_{t_i}]$ in all semiclassical states [45]. Therefore, subjecting the states to various simple sources is never going to make a slice larger than $N_{-l}(t_i)$ causally accessible

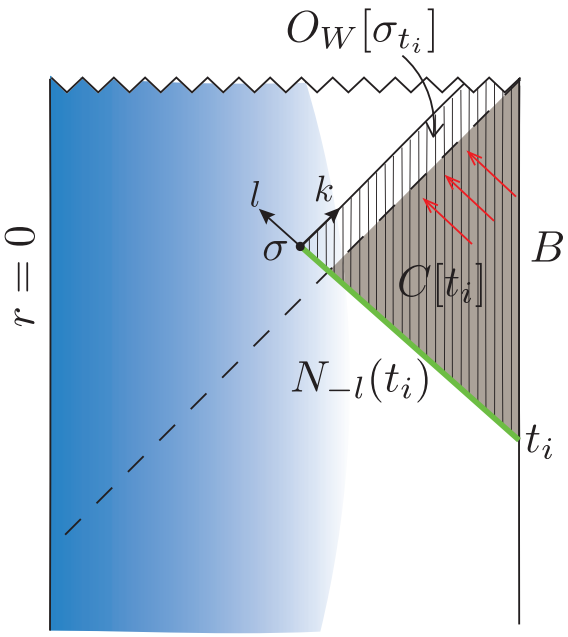

Figure 6.4: We would like to fix the data on $N_{-l}(t_i)$ (green thick line), while coarse-graining in the interior of the QMT surface. Simple data in the boundary region $t > t_i$ fixes the causal wedge $C[t_i]$ and thus fixes only a portion of $N_{-l}(t_i)$. In order to fix all of $N_{-l}(t_i)$ one must allow for sources that remove the excitations (red arrows) that enter the black hole after σ; this can cause the causal wedge to grow to include N_{-l}. In the coarse-graining set $\mathcal{F}$, the simple data must agree for all allowed sources.

from the boundary. Given the state $\rho(t_i)$, there exists a fine-tuned choice of sources that will make $C[t_i] = O_W[\sigma_{t_i}]$. But since this choice is state-dependent and difficult to specify from a pure boundary perspective, we choose the boundary coarse-graining family $\mathcal{F}$ to agree with $\rho(t_i)$ on simple data subject to *all* simple sources turned on.

So far we have defined $\mathcal{A}$ as the set of boundary observers that correspond to bulk excitations that propagate causally. The classical analysis of Refs. [118, 120] further specified $\mathcal{A}$ to consist only of one point functions of all local operators on the boundary. This will fix the states of the classical fields in the bulk that are causally determined by the boundary region $t \geq t_i$. Since here we are interested in fixing the quantum state of the bulk fields on $N_{-l}(t_i)$, our set $\mathcal{A}$ needs to include higher point function of local bulk operators.

However, we are still interested in maintaining locality in the bulk and therefore want to disallow a large density of local probes in any bulk region. This is following the expectation that such excitations would cause large backreaction and therefore a breakdown of locality [8]. From a boundary perspective, a local bulk operator in the causal wedge is dual to a smeared boundary operator [129]. Therefore, our set $\mathcal{A}$ needs to include all products of smeared boundary operator as long as there is not an $\mathcal{O}(N)$ number of overlap in the support of

the smeared operators. This choice of $\mathcal{A}$ in Eq. (6.6.2) is a natural candidate for fixing the quantum state on $N_{-l}(t_i)$; we leave a thorough investigation of this issue to future work.

We refer the reader to [120] for a careful demonstration of $S_{\text{simple}} = S_{\text{outer}}$ in the bulk classical limit.

Semiclassical Stretched States

In this paper, we started from a classical construction in general relativity, whose quantum interpretation is the coarse-graining of a quantum state so that its entropy matches the area of a marginally trapped surface. We elevated this to a semi-classical conjecture that we interpret as a coarse-graining that will match the generalized entropy of a quantum marginally trapped surface, while holding fixed the exterior quantum state. In the QFT limit, our conjecture is confirmed by the limit of the CF sequence of states [18].

Thus, we were able to derive a nontrivial, testable property of QFT from a hypothetical assumption about quantum gravity. This is similar to how the QNEC was derived from the QFC, a hypothetical extension of the classical focussing property of general relativity. This is a satisfying connection. QFT has not been directly probed in this limit, and a direct verification of the CF limit or of the QNEC would constitute a test of our ideas about quantum gravity.

Interestingly, there appears to be a larger set of relations of the type we explored here. Our starting point, the EW construction, is essentially unique. However, the CF construction produces a one-parameter family of states, given an input state and a cut on a Killing horizon. Here we only made use of the limit approached by these states as the flow parameter diverges. But we expect that there exists a classical construction (which may limit to the EW construction) that matches the entire one-parameter CF family.

In the special case where the cut is a bifurcation surface of the Killing horizon, the CF construction admits an interesting intuitive interpretation: all correlators of operators restricted to the left (or to the right) behave as if we had boosted the state on the left side of the cut (but not on the right). In QFT, a one-sided boost would result in a divergent-energy shock at the cut, because it would destroy the vacuum. But the CF flow is more subtle; in a sense it boosts only the "excited part" of the state on N_k^-, while leaving the vacuum entanglement across the cut intact.

This suggests a simple classical analogue of CF flow. At the classical level, a half-sided boost is innocuous. It can be applied to initial data on the null surface N_k^- with no ill consequences at the cut. However, a generic cut of a Killing horizon is not a bifurcation surface and hence is not a fixed point of the Killing flow.

Nonetheless, one can construct a sequence of geometries by a construction we will call the *left stretch*. Given the state and affine parameter v on the entire Killing horizon N_k, rescale $V \to V' = e^s V$ on the left side N_k^-, and do nothing on the right: $V \to V'$ on N_k^+. This will rescale all v-derivatives of classical fields by e^{-s}. To preserve the inner product $k^a l_a \equiv g_{ab}(\partial_v)^a(\partial_u)^b = -1$, rescale the u-derivatives at constant (v, y) by e^s. Then glue the two halves back together, treating V' as a true affine parameter.

For the full initial data on N, we need to know not only the intrinsic geometry but also θ_l, the expansion in the null direction off of N_k. This is obtained by holding θ_l fixed on N_k^+ and integrating the cross-focussing equation,

$$k^a \nabla_a \theta_l = -\frac{1}{2}\mathcal{R} - \theta_k \theta_l + \chi^2 + \nabla \cdot \chi + 8\pi G\, T_{kl} \ , \tag{6.6.3}$$

to obtain θ_l on N_k^-. Since all terms on the right hand side scale trivially, this rescales the difference $\theta_l - \theta_l|_{V_0}$ by e^s.

Because θ_l is not given by a simple rescaling unless $\theta_l|_{V_0} = 0$, the left stretch results in physically inequivalent initial data *even in the left exterior of σ alone*. The intrinsic data on N_k^- are stretched, as measured by a ruler defined by the evolution of the extrinsic curvature θ_l.

Interestingly, the left stretch is physically sensible if and only if the cut is a trapped surface. This is because the expansion θ_k along N_k is determined not only by the left stretch itself, but also by the Raychaudhuri equation, and the two methods must agree. Let the inaffinity κ be defined by $k^b \nabla_b k^a = \kappa k^a$. Affine parametrization corresponds to $\kappa = 0$ everywhere. The left stretch implements

$$V(y) \to e^{sH[-V(y)+V_0(y)]} V(y) \ , \tag{6.6.4}$$

where $H(v)$ is the Heaviside step function and $v = V_0(y)$ is the marginally trapped surface σ. This generates a non-zero inaffinity

$$\kappa = (1 - e^{-s})\delta[V(y)] \ . \tag{6.6.5}$$

The Raychaudhuri equation for non-affine parametrization reads

$$k^a \nabla_a \theta_k = -\frac{1}{2}\theta_k^2 - \varsigma_k^2 - \kappa\theta_k - 8\pi G\, T_{kk} \ . \tag{6.6.6}$$

We insist that the new parameter V' be treated as affine, which means we are demanding that the inaffinity term $\kappa\theta$ vanishes even after the left stretch. By Eq. (6.6.5), this will be the case if and only if $\theta_k = 0$ at the cut.

Importantly, Eqs. (6.2.5), (6.2.7) and (6.2.11) become satisfied in the limit as $s \to \infty$. These are precisely the conditions imposed by EW for the classical coarse-graining construction. In this sense the left stretch can be viewed as generating a one-parameter interpolation from the original initial data to the coarse-grained data.[3]

We close with two brief remarks. At the level of semiclassical gravity, the left stretch should naturally combine with the CF construction, so that not only the geometric and classical data, but also the quantum initial data are stretched. Moreover, we expect that the left stretch (applied classically to the RT or semiclassically to the quantum RT surface) is the gravity dual of the CF flow applied to the boundary of Anti-de Sitter space.

[3]However, there are interesting differences to the EW analysis. For example, the left stretch yields divergent T_{uu}, as does the CF limit. Yet, EW argue that this can be avoided. There may be a larger family of relevant states.

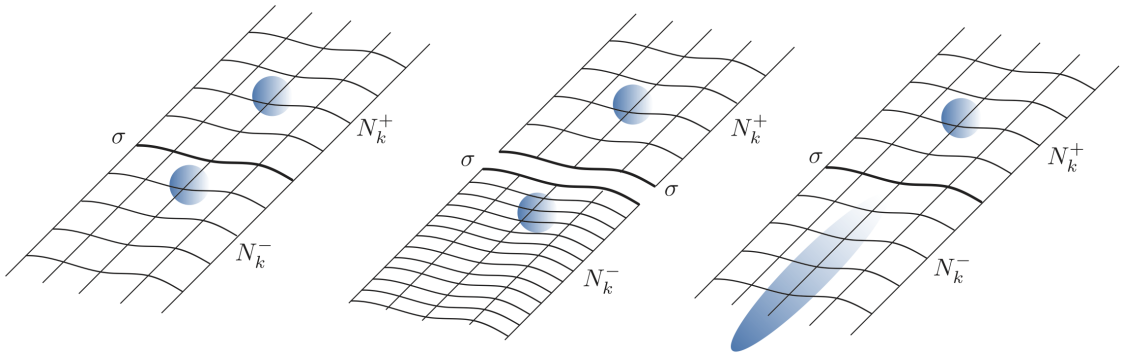

Figure 6.5: The left stretch is a classical analogue of the CF flow that generalizes it to nontrivial geometries. Left: The null surface N_k split by the marginally trapped surface σ. Middle: The affine parameter is rescaled on N_k^- but held fixed on N_k^+. This is the same initial data in nonaffine parametrization. Right: The two pieces are glued back together, treating the new parameter as affine. This yields inequivalent initial data.

Chapter 7

Gravity Dual of Connes Cocycle Flow

7.1 Introduction

The AdS/CFT duality [1, 2, 130] has led to tremendous progress in the study of quantum gravity. However, our understanding of the holographic dictionary remains limited. In recent years, quantum error correction was found to play an important role in the emergence of a gravitating ("bulk") spacetime from the boundary theory [8, 9, 131]. The study of modular operators led to the result that the boundary relative entropy in a region A equals the bulk relative entropy in its entanglement wedge $\mathrm{EW}(A)$ [7]. The combination of these insights was used to derive subregion duality: bulk operators in $\mathrm{EW}(A)$ can in principle be reconstructed from operators in the subregion A [35].

The relation between bulk modular flow in $\mathrm{EW}(A)$ and boundary modular flow in A has been used to explicitly reconstruct bulk operators both directly [132, 133], and indirectly via the Petz recovery map and its variants [134, 135, 136]. Thus, modular flow has shed light on the emergence of the bulk spacetime from entanglement properties of the boundary theory.

Modular flow has also played an important role in proving various properties of quantum field theory (QFT), such as the averaged null energy condition (ANEC) and quantum null energy condition (QNEC) [27, 137]. Tomita-Takesaki theory, the study of modular flow in algebraic QFT, puts constraints on correlation functions that are otherwise hard to prove directly [138].

Recently, an alternate proof of the QNEC was found using techniques from Tomita-Takesaki theory [18]. The key ingredient was Connes cocyle (CC) flow. Given a subregion A and global pure state ψ, Connes cocycle flow acts with a certain combination of modular operators to generate a sequence of well-defined global states ψ_s. In the limit $s \to \infty$, these states saturate Wall's "ant conjecture" [126] on the minimum amount of energy in the complementary region A'. This proves the ant conjecture, which, in turn, implies the QNEC.

CC flow also arises from a fascinating interplay between quantum gravity, quantum information, and QFT. Recently, the classical black hole coarse-graining construction of Engelhardt and Wall [118] was conjecturally extended to the semiclassical level [139]. In the

non-gravitational limit, this conjecture requires the existence of flat space QFT states with properties identical to the $s \to \infty$ limit of CC flowed states. This is somewhat reminiscent of how the QNEC was first discovered as the nongravitational limit of the quantum focusing conjecture [28]. Clearly, CC flow plays an important role in the connection between QFT and gravity. Our goal in this paper is to investigate this connection at a deeper level within the setting of AdS/CFT.

In Sec. 7.2, we define CC flow and discuss some of its properties. If ∂A lies on a null plane in Minkowski space, operator expectation values and subregion entropies within the region A remain the same, whereas those in A' transform analogously to a boost [18]. Further, CC flowed states ψ_s exhibit a characteristic stress tensor shock at the cut ∂A, controlled by the derivative of the von Neumann entropy of the region A in the state ψ under shape deformations of ∂A [139].

As is familiar from other examples in holography, bulk duals of complicated boundary objects are often much simpler [31, 29]. Motivated by the known properties of CC flow, we define a bulk construction in Sec. 7.3, which we call the "kink transform." This is a one-parameter transformation of the initial data of the bulk spacetime dual to the original boundary state ψ. We consider a Cauchy surface Σ that contains the Ryu-Takayanagi surface $\mathcal{R}$ of the subregion A. The kink transform acts as the identity except at $\mathcal{R}$, where an s-dependent shock is added to the extrinsic curvature of Σ. We show that this is equivalent to a one-sided boost of Σ in the normal bundle to $\mathcal{R}$. We prove that the new initial data satisfies the gravitational constraint equations, thus demonstrating that the kink transform defines a valid bulk spacetime $\mathcal{M}_s$. We show that $\mathcal{M}_s$ is independent of the choice of Σ.

We propose that $\mathcal{M}_s$ is the holographic dual to the CC-flowed state ψ_s, if the boundary cut ∂A is (conformally) a flat plane in Minkowski space.

In Sec. 7.4, we provide evidence for this proposal. The kink transform separately preserves the entanglement wedges of A and A', but it glues them together with a relative boost by rapidity $2\pi s$. This implies the one-sided expectation values and subregion entropies of the CC flowed state ψ_s are correctly reproduced when they are computed holographically in the bulk spacetime $\mathcal{M}_s$. We then perform a more nontrivial check of this proposal. By computing the boundary stress tensor holographically in $\mathcal{M}_s$, we reproduce the stress tensor shock at ∂A in the CC-flowed state ψ_s.

Having provided evidence for kink transform/CC flow duality, we use the duality to make a novel prediction for CC flow in Sec. 7.5. The kink transform fully determines all independent components of the shock at ∂A in terms of shape derivatives of the entanglement entropy. Strictly, our results only apply only to the CC flow of a holographic CFT across a planar cut. However, their universal form suggests that they will hold for general QFTs under CC flow. Moreover, the shocks we find agree with properties required to exist in quantum states under the coarse-graining proposal of Ref. [140]. Thus, our new results may also hold for CC flow across general cuts of a null plane.

In Sec. 7.6, we discuss the relation of our construction to earlier work on the role of modular flow in AdS/CFT [141, 7, 142]. The result of Jafferis *et al.* (JLMS) [7] has conventionally been understood as a relation that holds for a small code subspace of bulk states on

a fixed background spacetime. However, results from quantum error correction suggest that this code subspace could be made much larger to include different background geometries [9, 143, 144, 145]. Our proposal then follows from such an extended version of the JLMS result which includes non-perturbatively different background geometries. Equipped with this understanding, we can distinguish our proposal from the closely related bulk duals of one-sided modular flowed states [141, 142]. We provide additional evidence for our proposal based on two sided correlation functions of heavy operators, and we discuss generalizations and applications of the proposed kink transform/CC flow duality.

In Appendix A.15 we derive the null limit of the kink transform, and show that it generates a Weyl shock, which provides intuition for how the kink transform modifies gravitational observables.

7.2 Connes Cocycle Flow

In this section, we review Connes cocycle flow and its salient properties; for more details see [18, 139]. We then reformulate Connes cocycle flow in as a simpler map to a state defined on a "precursor" slice. This will prove useful in later sections.

Definition and General Properties

Consider a quantum field theory on Minkowski space $\mathbb{R}^{d-1,1}$ in standard Cartesian coordinates $(t, x, y_1, \ldots, y_{d-2})$. Consider a Cauchy surface $\mathcal{C}$ that is the disjoint union of the open regions A_0, A_0' and their shared boundary ∂A_0. Let $\mathcal{A}_0, \mathcal{A}_0'$ denote the associated algebras of operators. Let $|\psi\rangle$ be a cyclic and separating state on $\mathcal{C}$, and denote by $|\Omega\rangle$ the global vacuum (the assumption of cyclic and separating could be relaxed for $|\psi\rangle$, at the cost of complicating the discussion below). The Tomita operator is defined by

$$S_{\psi|\Omega;\mathcal{A}_0}\alpha|\psi\rangle = \alpha^\dagger|\Omega\rangle, \forall\alpha \in \mathcal{A}_0 \ . \tag{7.2.1}$$

The relative modular operator is defined as

$$\Delta_{\psi|\Omega} \equiv S^\dagger_{\psi|\Omega;\mathcal{A}_0} S_{\psi|\Omega;\mathcal{A}_0} \ , \tag{7.2.2}$$

and the vacuum modular operator is

$$\Delta_\Omega \equiv \Delta_{\Omega|\Omega} \ . \tag{7.2.3}$$

Note that we do not include the subscript $\mathcal{A}_0$ on Δ; instead, for modular operators, we indicate whether they were constructed from $\mathcal{A}_0$ or $\mathcal{A}_0'$ by writing Δ or Δ'.

Connes cocycle (CC) flow of $|\psi\rangle$ generates a one parameter family of states $|\psi_s\rangle$, $s \in \mathbb{R}$, defined by

$$|\psi_s\rangle = (\Delta'_\Omega)^{is}(\Delta'_{\Omega|\psi})^{-is}|\psi\rangle \ . \tag{7.2.4}$$

Thus far the definitions have been purely algebraic. In order to elucidate the intuition behind CC flow, let us write out the modular operators in terms of the left and right density operators, $\rho^{\psi}_{A_0} = \mathrm{Tr}_{A_0'}|\psi\rangle\langle\psi|$ and $\rho^{\psi}_{A_0'} = \mathrm{Tr}_{A_0}|\psi\rangle\langle\psi|$:[1]

$$\Delta_{\psi|\Omega} = \rho^{\Omega}_{A_0} \otimes (\rho^{\psi}_{A_0'})^{-1} \ . \tag{7.2.5}$$

One finds that the CC operator acts only in $\mathcal{A}_0'$:

$$(\Delta'_{\Omega})^{is}(\Delta'_{\Omega|\psi})^{-is} = (\rho^{\Omega}_{A_0'})^{is}(\rho^{\psi}_{A_0'})^{-is} \in \mathcal{A}_0' \ . \tag{7.2.6}$$

It follows that the reduced state on the right algebra satisfies

$$\rho^{\psi_s}_{A_0} = \rho^{\psi}_{A_0} \ . \tag{7.2.7}$$

Therefore, expectation values of observables $\mathcal{O} \in \mathcal{A}_0$ remain invariant under CC flow. These heuristic arguments would be valid only for finite-dimensional Hilbert spaces [146]; but Eq. (7.2.6) can be derived rigorously [18].

It can also be shown that $(\Delta'_{\psi|\Omega})^{is}\Delta^{is}_{\Omega|\psi} = 1$. Hence for operators $\mathcal{O}' \in \mathcal{A}_0'$, one finds that CC flow acts as Δ^{is}_{Ω} inside of expectation values:

$$\begin{aligned}
\langle\psi_s|\mathcal{O}'|\psi_s\rangle &= \mathrm{Tr}_{\mathcal{A}_0'}\left[\rho^{\psi}_{A_0'}(\Delta^{-is}_{\psi|\Omega}\Delta^{is}_{\Omega})\mathcal{O}'(\Delta^{-is}_{\Omega}\Delta^{is}_{\psi|\Omega})\right] \ , \\
&= \mathrm{Tr}_{\mathcal{A}_0'}\left(\rho^{\psi}_{A_0'}(\rho^{\Omega}_{A_0'})^{-is}\mathcal{O}'(\rho^{\Omega}_{A_0'})^{is}\right) \qquad (7.2.8) \\
&= \mathrm{Tr}\left[|\psi\rangle\langle\psi|\Delta^{is}_{\Omega}(\mathbf{1}\otimes\mathcal{O}')\Delta^{-is}_{\Omega}\right] \ , \qquad (7.2.9)
\end{aligned}$$

where we have used the cyclicity of the trace.

To summarize, expectation values of one-sided operators transform as follows:

$$\langle\psi_s|\mathcal{O}|\psi_s\rangle = \langle\psi|\mathcal{O}|\psi\rangle \ , \tag{7.2.10}$$

$$\langle\psi_s|\mathcal{O}'|\psi_s\rangle = \langle\psi|\Delta^{is}_{\Omega}\mathcal{O}'\Delta^{-is}_{\Omega}|\psi\rangle \ . \tag{7.2.11}$$

There is no simple description of two-sided correlators in $|\psi_s\rangle$ such as $\langle\psi_s|\mathcal{O}\mathcal{O}'|\psi_s\rangle$; we discuss such objects in Sec. 7.6.

CC Flow from Cuts on a Null Plane

Let us now specialize to the case where ∂A_0 corresponds to a cut $v = V_0(y)$ of the Rindler horizon $u = 0$. We have introduced null coordinates $u = t - x$ and $v = t + x$ and denoted the transverse coordinates collectively by y. It can be shown that the modular operator Δ^{is}_{Ω}

[1] We follow the conventions in [146] where complement operators are written to the right of the tensor product.

acts locally on each null generator y of $u = 0$ as a boost about the cut $V_0(y)$ [125]. More explicitly, one can define the *full* vacuum modular Hamiltonian $\widehat{K}_{V_0}$ by

$$\widehat{K}_{V_0} = -\log \Delta_{\Omega;\mathcal{A}_{V_0}} \ . \tag{7.2.12}$$

We can write the full modular Hamiltonian as

$$\widehat{K}_{V_0} = K_{V_0} \otimes \mathbf{1}' - \mathbf{1} \otimes K'_{V_0} \ . \tag{7.2.13}$$

Let Δ denote vacuum subtraction, $\Delta\langle\mathcal{O}\rangle = \langle\mathcal{O}\rangle_\psi - \langle\mathcal{O}\rangle_\Omega$. Then, for arbitrary cuts of the Rindler horizon, we have [125]

$$\Delta\langle K'_{V_0}\rangle = -2\pi \int dy \int_{-\infty}^{V_0} dv[v - V_0(y)]\langle T_{vv}\rangle_\psi \ , \tag{7.2.14}$$

and similarly for K_{V_0}. Thus K'_{V_0} is simply the boost generator about the cut $V_0(y)$ in the left Rindler wedge. That is, it generates a y-dependent dilation,

$$v \to V_0(y) + [v - V_0(y)]e^{2\pi s} \ . \tag{7.2.15}$$

This allows us to evaluate Eq. (7.2.11) explicitly for local operators at $u = 0$. For example, the CC flow of the stress tensor is

$$\langle\psi_s|T_{vv}|\psi_s\rangle|_{v<V_0} = e^{-4\pi s}\langle\psi|T_{vv}\left(V_0 + e^{-2\pi s}(v - V_0)\right)|\psi\rangle|_{v<V_0} \ , \tag{7.2.16}$$

and similarly for the other components of $T_{\mu\nu}$. There is a slight caveat here since Δ_Ω^{is} only acts as a boost strictly at $u = 0$. This would be sufficient for free theories, where T_{vv} can be defined through null quantization on the Rindler horizon with a smearing that only needs support on $u = 0$ [15]. More generally, $T_{\mu\nu}$ must be smeared in an open neighborhood of $u = 0$. However, if $V_0(y)$ is a perturbation of a flat cut then one can show that inside correlation functions Δ_Ω^{is} approximately acts as a boost with subleading errors that vanish as $u \to 0$, to all orders in the perturbation [137, 147]. In the non-perturbative case, evidence comes from the fact that classically the vector field on the Rindler horizon which generates boosts about $V_0(y)$ can be extended to an approximate Killing vector field in a neighborhood of the horizon [148, 149]. Therefore we expect Eq. (7.2.16) to hold on the null surface even after smearing.

Now consider a second cut $V(y)$ of the Rindler horizon which lies entirely below $V_0(y)$, so $V < V_0$ for all y. The cut defines a surface ∂A_V that splits a Cauchy surface $\mathcal{C}_V = A'_V \cup \partial A_V \cup A_V$; we take A'_V to be the "left" side ($v < V$), with operator algebra $\mathcal{A}'_V$. The Araki definition of relative entropy is [146]

$$S'_{\rm rel}(\psi|\Omega; V) = -\langle\psi|\log \Delta_{\psi|\Omega;\mathcal{A}'_V}|\psi\rangle \ . \tag{7.2.17}$$

It has the following transformation properties [18]:

$$S_{\rm rel}(\psi_s|\Omega;V) = S_{\rm rel}(\psi|\Omega;V_0+e^{-2\pi s}(V-V_0)) \ , \tag{7.2.18}$$

$$\frac{\delta S_{\rm rel}(\psi_s|\Omega;V)}{\delta V} = e^{-2\pi s}\frac{\delta S_{\rm rel}(\psi|\Omega;V_0+e^{-2\pi s}(V-V_0))}{\delta V} \ . \tag{7.2.19}$$

Moreover, the "left" von Neumann entropy is defined as

$$S'(\psi,V) = -{\rm tr}_{A'_V}\rho^{\psi}_{A'_V}\log\rho^{\psi}_{A'_V} \ . \tag{7.2.20}$$

With these definitions in hand, one can decompose the relative entropy as

$$S'_{\rm rel}(\psi|\Omega;V) = \Delta\langle K'_V\rangle - \Delta S'(V) \ . \tag{7.2.21}$$

At this point we drop the explicit vacuum subtractions, as we will only be interested in shape derivatives of the vacuum subtracted quantities, which automatically annihilate the vacuum expectation values. In particular, one can directly compute shape derivatives of K'_V:

$$\frac{\delta\langle K'_V\rangle_\psi}{\delta V}\Big|_{V_0} = 2\pi\int_{-\infty}^{V_0} dv\ \langle T_{vv}\rangle_\psi \ . \tag{7.2.22}$$

Hence the transformations of both K'_V and its derivative simply follow from Eq. (7.2.16).

Combining Eq. (7.2.18) and Eq. (7.2.14), as well as Eq. (7.2.19) and Eq. (7.2.22), we see that $S'(\psi,V)$ and its derivative transform as

$$S'(\psi_s,V) = S'(\psi,V_0+e^{-2\pi s}(V-V_0)) \ , \tag{7.2.23}$$

$$\frac{\delta S'}{\delta V}\Big|_{\psi_s,V} = e^{-2\pi s}\frac{\delta S'}{\delta V}\Big|_{\psi,V_0+e^{-2\pi s}(V-V_0)} \ . \tag{7.2.24}$$

The respective properties of the complement entropy follow from purity.

Stress Tensor Shock at the Cut

CC flow generates a stress tensor shock at the cut V_0, proportional to the jump in the variation of the one-sided von Neumann entropy under deformations, at the cut [139]. To see this, let us start with the sum rule derived in [18] for null variations of relative entropy:[2]

$$2\pi(P_s - e^{-2\pi s}P_0) = (e^{-2\pi s}-1)\frac{\delta S'_{\rm rel}(\psi|\Omega;V)}{\delta V}\Big|_{V_0} \ , \tag{7.2.25}$$

where

$$P \equiv \int_{-\infty}^{\infty} dv\ T_{vv} \tag{7.2.26}$$

[2]For type I algebras, one can derive the analogous sum rule from simpler arguments [107].

is the averaged null energy operator at $u = 0$, and $P_s \equiv \langle\psi_s|P|\psi_s\rangle$, so in particular $P_0 \equiv \langle\psi|P|\psi\rangle$. (There is one such operator for every generator, *i.e.*, for every y.)

Inserting Eq. (7.2.21) and Eq. (7.2.22) into Eq. (7.2.25), and making use of Eq. (7.2.16), we see that there must exist a shock at $v = V_0(y)$:

$$\langle\psi_s|T_{vv}|\psi_s\rangle = (1 - e^{-2\pi s})\frac{1}{2\pi}\frac{\delta S'}{\delta V}\Big|_{V_0}\delta(v - V_0) + o(\delta) \ . \tag{7.2.27}$$

Here $o(\delta)$ designates the finite (non-distributional) terms. These are determined by Eq. (7.2.16), and by its trivial counterpart in the $v > V_0$ region.

This s-dependent shock is a detailed characteristic of the CC flowed state. As such, reproducing it through the holographic dictionary will be the key test of our proposal of the bulk dual of CC flow (see Sec. 7.4).

Flat Cuts and the Precursor Slice

For the remainder of the paper we further specialize to flat cuts of the Rindler horizon, so that ∂A_0 corresponds to $u = v = 0$. We therefore set $V_0 = 0$ in what follows. We take $\mathcal{C}$ to be the Cauchy surface $t = 0$, so that A_0 ($t = 0$, $x > 0$) and A_0' ($t = 0$, $x < 0$) are partial Cauchy surfaces for the right and left Rindler wedges.

In this case Δ_Ω^{is} is a global boost by rapidity s about ∂A_0 [21]. Thus, it has a simple geometric action not only on the null plane $u = 0$, but everywhere. CC flow transforms observables in $\mathcal{A}_0'$ by Δ_Ω^{is} and leaves invariant those in $\mathcal{A}_0$. For a flat cut, this action can be represented as a geometric boost in the entire left Rindler wedge. This allows us to characterize the CC flowed state $|\psi(s)\rangle$ on $\mathcal{C}$ very simply in terms of a different state on a different Cauchy surface which we call the "precursor slice". This description will motivate the formulation of our bulk construction in Sec. 7.3.

By Eq. (7.2.11), the CC flowed state on the slice $\mathcal{C}$,

$$|\psi_s(\mathcal{C})\rangle = (\Delta_\Omega')^{is}(\Delta'_{\Omega|\psi})^{-is}\,|\psi(\mathcal{C})\rangle \ , \tag{7.2.28}$$

satisfies

$$\langle\psi_s(\mathcal{C})|\,\mathcal{O}_A\,|\psi_s(\mathcal{C})\rangle = \langle\psi(\mathcal{C})|\,\mathcal{O}_A\,|\psi(\mathcal{C})\rangle \ , \tag{7.2.29}$$

$$\langle\psi_s(\mathcal{C})|\,\Delta_\Omega^{-is}\mathcal{O}_{A'}\Delta_\Omega^{is}\,|\psi_s(\mathcal{C})\rangle = \langle\psi(\mathcal{C})|\,\mathcal{O}_{A'}\,|\psi(\mathcal{C})\rangle \ , \tag{7.2.30}$$

where $\mathcal{O}_A$ and $\mathcal{O}_{A'}$ denote an arbitrary collection of local operators that act on spacelike half-slices A and A' of $\mathcal{C}$ respectively.[3] In the second equality above, we used the fact that Δ_Ω^{is} acts as a global boost to move it to the other side of the equality, compared to Eq. (7.2.11).

[3]More precisely, one would have to smear the operator in a codimension 0 neighborhood of points on the slices.

We work in the Schrödinger picture where the argument $\mathcal{C}$ should be interpreted as the time variable. The fact that Δ_Ω^{is} acts as a boost around ∂A_0 motivates us to consider the time slice

$$\mathcal{C}_s = A'_s \cup \partial A_0 \cup A_0 \ , \tag{7.2.31}$$

where

$$A'_s = \{t = (\tanh 2\pi s)x, \ x < 0\} \ . \tag{7.2.32}$$

By Eqs. (7.2.29) and (7.2.30), each side of the CC-flowed state $|\psi_s(\mathcal{C}_s)\rangle$ is simply related to the left and right restrictions of the original state on the original slice:

$$\langle\psi_s(\mathcal{C}_s)|\, \mathcal{O}_A\, |\psi_s(\mathcal{C}_s)\rangle = \langle\psi(\mathcal{C})|\, \mathcal{O}_A\, |\psi(\mathcal{C})\rangle \ , \tag{7.2.33}$$

$$\langle\psi_s(\mathcal{C}_s)|\, \mathcal{O}_{A'_s}\, |\psi_s(\mathcal{C}_s)\rangle = \langle\psi(\mathcal{C})|\, \mathcal{O}_{A'}\, |\psi(\mathcal{C})\rangle \ . \tag{7.2.34}$$

In the second equation, $\mathcal{O}_{A'_s}$ denotes local operators on A'_s which are analogous to $\mathcal{O}_{A'}$ on A'. More precisely, because the intrinsic metric of A' and A'_s are the same, there exists a natural map between local operators on A' and A'_s.

In words, Eqs. (7.2.33) and (7.2.34) say that correlation functions in each half of $\mathcal{C}$ in the state $|\psi(\mathcal{C})\rangle$ are equal to the analogous correlation functions on each half of $\mathcal{C}_s$ in the state $|\psi_s(\mathcal{C}_s)\rangle$. This justifies calling $\mathcal{C}_s$ the precursor slice since the CC flowed state on $\mathcal{C}$ arises from it by time evolution.

We find it instructive to repeat this point in the less rigorous language of density operators. In the density operator form of CC flow,

$$|\psi_s(\mathcal{C})\rangle = (\rho^\Omega_{A'_0})^{is}(\rho^\psi_{A'_0})^{-is}\, |\psi(\mathcal{C})\rangle \ , \tag{7.2.35}$$

it is evident that the action of $(\rho^\Omega_{A'_0})^{is}$ can be absorbed into a change of time slice $\mathcal{C} \to \mathcal{C}_s$:

$$|\psi_s(\mathcal{C}_s)\rangle = (\rho^\psi_{A'_0})^{-is}\, |\psi(\mathcal{C})\rangle \ . \tag{7.2.36}$$

Tracing out each side of ∂A_0 implies

$$\rho^{\psi_s}_{A_0} = \rho^\psi_{A_0} \ , \tag{7.2.37}$$

$$\rho^{\psi_s}_{A'_s} = \rho^\psi_{A'_0} \ . \tag{7.2.38}$$

The first equality is trivial and was already discussed in Eq. (7.2.7). The second equality follows because $(\rho^\psi_{A'_0})^{is}$ commutes with $(\rho^\psi_{A'_0})$. This is the density operator version of Eqs. (7.2.33) and (7.2.34).

Eq. (7.2.36) should be contrasted with the one-sided modular-flowed state $|\phi(\mathcal{C})\rangle = (\rho^\psi_{A'_0})^{-is}\, |\psi(\mathcal{C})\rangle$. The latter state would live on the original slice $\mathcal{C}$, but it is not well-defined since it would have infinite energy at the entangling surface.

It will be useful to define new coordinates adapted to the precursor slice $\mathcal{C}_s$. Let

$$\tilde{v} = v\, \Theta(v) + e^{-2\pi s}\, v\, (1 - \Theta(v)) \ , \tag{7.2.39}$$

$$\tilde{u} = e^{2\pi s} u\, \Theta(u) + \ u\, (1 - \Theta(u)) \ , \tag{7.2.40}$$

where $\Theta(.)$ is the Heaviside step function. Let $\tilde{t} = \frac{1}{2}(\tilde{v} + \tilde{u})$ and $\tilde{x} = \frac{1}{2}(\tilde{v} - \tilde{u})$. In these coordinates, the Minkowski metric takes the form

$$ds^2 = \left[\Theta(\tilde{t}+\tilde{x}) + e^{2\pi s}(1-\Theta(\tilde{t}+\tilde{x}))\right]\left[e^{-2\pi s}\Theta(\tilde{t}-\tilde{x}) + (1-\Theta(\tilde{t}-\tilde{x}))\right](-d\tilde{t}^2 + d\tilde{x}^2) + d^{d-2}y \ , \tag{7.2.41}$$

and the precursor slice corresponds to $\tilde{t} = 0$.

In these "tilde" coordinates, the stress tensor shock of Eq. (7.2.27) takes the form[4]

$$\langle\psi_s|T_{\tilde{v}\tilde{v}}|\psi_s\rangle = \frac{1}{2\pi}\left(\frac{\partial v}{\partial\tilde{v}}\right)^2(1-e^{-2\pi s})\frac{\delta S}{\delta V}\Big|_{V=0}\delta(v) + o(\delta) \ . \tag{7.2.42}$$

Recall that the entropy variation is evaluated in the state $|\psi\rangle$. By Eq. (7.2.24),

$$\frac{\delta S}{\delta V}\Big|_{\psi} = \frac{\delta S}{\delta\widetilde{V}}\Big|_{\psi_s} \ , \tag{7.2.43}$$

where $\widetilde{V}(y)$ is a cut of the Rindler horizon in the $\tilde{v}$ coordinates. Thus we may instead evaluate the entropy variation in the state $|\psi_s\rangle$ on the precursor slice. This will be convenient when matching the bulk and boundary.

The Jacobian in Eq. (7.2.42) has a step function in it, as will the Jacobian coming from $\delta(v)$. A step function multiplying a delta function is well-defined if one averages the left and right derivatives:

$$\left(\frac{\partial v}{\partial\tilde{v}}\right)^2\delta(v) = \frac{1}{2}\left(\frac{\partial v}{\partial\tilde{v}}\Big|_{0^-} + \frac{\partial v}{\partial\tilde{v}}\Big|_{0^+}\right)\delta(\tilde{v}) \ . \tag{7.2.44}$$

Thus Eq. (7.2.42) becomes

$$\langle\psi_s|T_{\tilde{v}\tilde{v}}(\tilde{v})|\psi_s\rangle = \frac{1}{2\pi}\sinh(2\pi s)\frac{\delta S}{\delta\widetilde{V}}\Big|_{\psi_s,\widetilde{V}=0}\delta(\tilde{v}) + o(\delta). \tag{7.2.45}$$

Since we are dealing with a flat cut, the symmetry $s \leftrightarrow -s, v \leftrightarrow u$ implies that CC flow also generates a T_{uu} shock in the state $|\psi_s\rangle$ at $u = v = 0$:

$$\langle\psi_s|T_{\tilde{u}\tilde{u}}|\psi_s\rangle = \frac{1}{2\pi}\sinh(2\pi s)\frac{\delta S}{\delta\widetilde{U}}\Big|_{\psi_s,\widetilde{V}=0}\delta(\tilde{u}) + o(\delta) \ . \tag{7.2.46}$$

(Note that $\delta/\delta V$ goes to $-\delta/\delta U$.) The linear combination

$$\langle\psi_s|T_{\tilde{t}\tilde{x}}(\tilde{t},\tilde{x})|\psi_s\rangle = \frac{1}{2\pi}\sinh(2\pi s)\frac{\delta S}{\delta\widetilde{X}}\Big|_{\psi_s,\widetilde{X}=0}\delta(\tilde{x}) + \langle\psi|T_{tx}(t=\tilde{t},x=\tilde{x})|\psi\rangle \ . \tag{7.2.47}$$

will be useful in Sec. 7.4. The last term was obtained from Eqs. (7.2.37) and (7.2.38); it makes the finite piece explicit. Note that these equations are valid in the entire left and right wedges, not just on $\mathcal{C}_s$.

[4]We remind the reader that $o(\delta)$ refers to any finite (non-distributional) terms.

7.3 Kink Transform

In this section, we introduce a novel geometric transformation called the kink transform. The construction is motivated by thinking about what the bulk dual of the boundary CC flow would be in the context of AdS/CFT. As we discussed in Sec. 7.2, CC flow boosts observables in $D(A')$ and leaves observables in $D(A)$ unchanged. Subregion duality in AdS/CFT then implies that the bulk dual of the state $|\psi_s\rangle$ has to have the property that the entanglement wedges of $D(A)$ and $D(A')$ will be diffeomorphic to those of the state $|\psi\rangle$, but are glued together with a "one-sided boost" at the HRT surface. In a general geometry, a boost Killing symmetry need not exist. The kink transform appropriately generalizes the notion of a one-sided boost to any extremal surface.

In Sec. 7.3, we formulate the kink transform. In Sec. 7.3, we describe a different but equivalent formulation of the kink transform and show that the kink transform results in the same new spacetime, regardless of which Cauchy surface containing the extremal surface is used for the construction. In Sec. 7.4, we will describe the duality between the bulk kink transform and the boundary CC flow in AdS/CFT and provide evidence for it.

Formulation

Consider a $d+1$ dimensional spacetime $\mathcal{M}$ with metric $g_{\mu\nu}$ satisfying the Einstein field equations. (We will discuss higher curvature gravity in Sec. 7.6.) Let Σ be a Cauchy surface of $\mathcal{M}$ that contains an extremal surface $\mathcal{R}$ of codimension 1 in Σ. (That is, the expansion of both sets of null geodesics orthogonal to $\mathcal{R}$ vanishes.)

Initial data on Σ consist of [150] the intrinsic metric $(h_\Sigma)_{ab}$ and the extrinsic curvature,

$$(K_\Sigma)_{ab} = P^\mu_a P^\nu_b \nabla_{(\mu} t_{\nu)} \ . \tag{7.3.1}$$

Here P^μ_a is the projector from $\mathcal{M}$ onto Σ, and t^μ is the unit norm timelike vector field orthogonal to Σ. Indices $a, b, \ldots$ are reserved for directions tangent to Σ. For matter fields, initial data consist of the fields and normal derivatives, for example $\phi(w^a)$ and $[t^\mu \nabla_\mu \phi](w^a)$, where ϕ is a scalar field and w^a are coordinates on Σ.

By the Einstein equations, the initial data on Σ must satisfy the following constraints:

$$r_\Sigma + K_\Sigma^2 - (K_\Sigma)_{ab}(K_\Sigma)^{ab} = 16\pi G\, T_{\mu\nu} t^\mu t^\nu \ , \tag{7.3.2}$$

$$D^a (K_\Sigma)_{ab} - D_b K_\Sigma = 8\pi G\, T_{b\nu} t^\nu \ , \tag{7.3.3}$$

where $D_a = P^\mu_a \nabla_\mu$ is the covariant derivative that Σ inherits from $(\mathcal{M}, g_{\mu\nu})$; r_Σ is the Ricci scalar intrinsic to Σ; and K_Σ is the trace of the extrinsic curvature: $K_\Sigma = (h_\Sigma)^{ab}(K_\Sigma)_{ab}$.

Let Σ be a Cauchy slice of $\mathcal{M}$ containing $\mathcal{R}$ and smooth in a neighborhood of $\mathcal{R}$. The kink transform is then a map of the initial data on Σ to a new initial data set, parametrized by a real number s analogous to boost rapidity. The transform acts as the identity on all data except for the extrinsic curvature, which is modified only at the location of the extremal

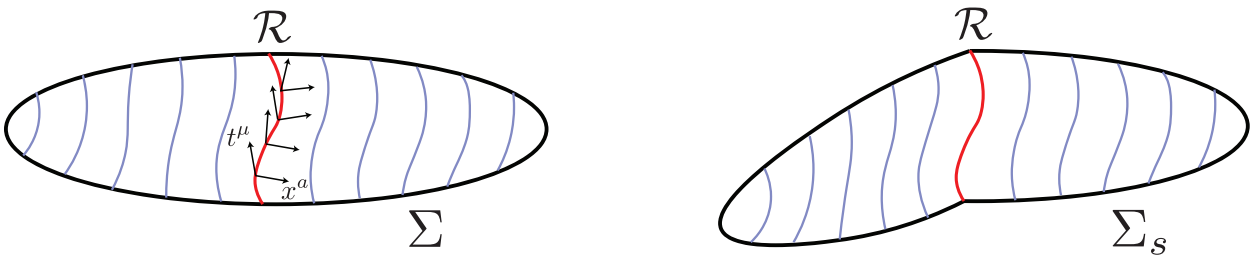

Figure 7.1: Kink transform. Left: a Cauchy surface Σ of the original bulk $\mathcal{M}$. An extremal surface $\mathcal{R}$ is shown in red. The orthonormal vector fields t^a and x^a span the normal bundle to $\mathcal{R}$; x^a is tangent to Σ. Right: The kink transformed Cauchy surface Σ_s. As an initial data set, Σ_s differs from Σ only in the extrinsic curvature at $\mathcal{R}$ through Eq. (7.3.4). Equivalently, the kink transform is a relative boost in the normal bundle to $\mathcal{R}$, Eq. (7.3.21).

surface $\mathcal{R}$, as follows:

$$(K_\Sigma)_{ab} \to (K_{\Sigma_s})_{ab} = (K_\Sigma)_{ab} - \sinh(2\pi s)\; x_a x_b\; \delta(\mathcal{R})\ . \tag{7.3.4}$$

Here x^a is a unit norm vector field orthogonal to $\mathcal{R}$ and tangent to Σ, and we define

$$\delta(\mathcal{R}) \equiv \delta(x)\ , \tag{7.3.5}$$

where x is the Gaussian normal coordinate to $\mathcal{R}$ in Σ ($\partial_x = x^a$). Thus, the only change in the initial data is in the component of the extrinsic curvature normal to $\mathcal{R}$. An equivalent transformation exists for initial choices of Σ that are not smooth around $\mathcal{R}$ though the transformation rule will be more complicated than Eq. (7.3.4). We will discuss this later in the section.

Let Σ_s be a time slice with this new initial data, as in Fig. 7.1, and let $\mathcal{M}_s$ be the Cauchy development of Σ_s. That is, $\mathcal{M}_s$ is the new spacetime resulting from the evolution of the kink-transformed initial data. Since the intrinsic metric of Σ_s and Σ are the same, they can be identified as d-manifolds with metric; the subscript s merely reminds us of the different extrinsic data they carry. In particular the surface $\mathcal{R}$ can be so identified; thus $\mathcal{R}_s$ has the same intrinsic metric as $\mathcal{R}$. It also trivially has identical extrinsic data with respect to Σ_s. In fact, we will find below that like $\mathcal{R}$ in $\mathcal{M}$, $\mathcal{R}_s$ is an extremal surface in $\mathcal{M}_s$. However, the trace-free part of the extrinsic curvature of $\mathcal{R}_s$ in $\mathcal{M}_s$ may have discontinuities.

We will now show that the constraint equations hold on Σ_s; that is, the kink transform generates valid initial data. This need only be verified at $\mathcal{R}$ since the transform acts as the identity elsewhere. Here we will make essential use of the extremality of $\mathcal{R}$ in $\mathcal{M}$, which we express as follows.

The extrinsic curvature of $\mathcal{R}$ with respect to $\mathcal{M}$ has two independent components. Often these are chosen to be the two orthogonal null directions, but we find it useful to consider

$$(B^{(t)}_{\mathcal{R}})_{ij} = P^\mu_i P^\nu_j \nabla_{(\mu} t_{\nu)} \ , \tag{7.3.6}$$

$$(B^{(x)}_{\mathcal{R}})_{ij} = P^\mu_i P^\nu_j \nabla_{(\mu} x_{\nu)} \ . \tag{7.3.7}$$

Here i, j represent directions tangent to $\mathcal{R}$, and P^μ_i is the projector from $\mathcal{M}$ to $\mathcal{R}$. Extremality of $\mathcal{R}$ in $\mathcal{M}$ is the statement that the trace of each extrinsic curvature component vanishes:

$$B^{(t)}_{\mathcal{R}} = (\gamma_{\mathcal{R}})^{ij}(B^{(t)}_{\mathcal{R}})_{ij} = 0 \ , \tag{7.3.8}$$

$$B^{(x)}_{\mathcal{R}} = (\gamma_{\mathcal{R}})^{ij}(B^{(x)}_{\mathcal{R}})_{ij} = 0 \ , \tag{7.3.9}$$

where $(\gamma_{\mathcal{R}})_{ij} = P^a_i P^b_j (h_\Sigma)_{ab}$ is the intrinsic metric on $\mathcal{R}$.

Orthogonality of t^μ and x^μ implies that $P^\mu_i = P^a_i P^\mu_a$, and hence

$$(B^{(t)}_{\mathcal{R}})_{ij} = P^a_i P^b_j (K_\Sigma)_{ab} \ . \tag{7.3.10}$$

Since x^a is the unit norm orthogonal vector field at $\mathcal{R}$, the trace of $(K_\Sigma)_{ab}$ at $\mathcal{R}$ can be written as:

$$K_\Sigma|_{\mathcal{R}} = x^a x^b (K_\Sigma)_{ab} + (\gamma_{\mathcal{R}})^{ij}(B^{(t)}_{\mathcal{R}})_{ij} = x^a x^b (K_\Sigma)_{ab} \ . \tag{7.3.11}$$

A little algebra then implies

$$(K_{\Sigma_s})^2 - (K_{\Sigma_s})_{ab}(K_{\Sigma_s})^{ab} = (K_\Sigma)^2 - (K_\Sigma)_{ab}(K_\Sigma)^{ab} \ . \tag{7.3.12}$$

Moreover, we have $r_\Sigma = r_{\Sigma_s}$ since the two initial data slices have the same intrinsic metric. Thus Eq. (7.3.2) implies that the kink-transformed slice satisfies the scalar constraint equation:

$$r_{\Sigma_s} + (K_{\Sigma_s})^2 - (K_{\Sigma_s})_{ab}(K_{\Sigma_s})^{ab} = 16\pi G T_{\mu\nu} t^\mu t^\nu \ . \tag{7.3.13}$$

To check the vector constraint Eq. (7.3.3), we separately consider the two cases of $b = x$ and $b = i$ where i, j represent directions tangent to $\mathcal{R}$:

$$\begin{aligned} D_a(K_{\Sigma_s})^a_x - D_x K_{\Sigma_s} &= D_a(K_\Sigma)^a_x - D_x K_\Sigma + B^{(x)}_{\mathcal{R}} \sinh(2\pi s)\delta(x) \\ &= D_a(K_\Sigma)^a_x - D_x K_\Sigma = 8\pi G\, T_{x\nu} t^\nu \ , \end{aligned} \tag{7.3.14}$$

$$D_a(K_{\Sigma_s})^a_i - D_i K_{\Sigma_s} = D_a(K_\Sigma)^a_i - D_i K_\Sigma = 8\pi G\, T_{i\nu} t^\nu \ , \tag{7.3.15}$$

where the second line of the first equation follows from the extremality of $\mathcal{R}$.

We conclude that the kink transform is a valid modification to the initial data. For both constraints to be satisfied after the kink, it was essential that $\mathcal{R}$ is an extremal surface. Thus

the kink transform is only well-defined across an extremal surface. Note also that $\mathcal{R}_s \subset \Sigma_s$ is an extremal surface in $\mathcal{M}_s$. By Eq. (7.3.10),

$$(B^{(t)}_{\mathcal{R}_s})_{ij} = P^a_i P^b_j \, (K_{\Sigma_s})_{ab}|_{\mathcal{R}_s} = (B^{(t)}_{\mathcal{R}})_{ij} \implies B^{(t)}_{\mathcal{R}_s} = 0 \ . \tag{7.3.16}$$

In the second equality we used Eq. (7.3.4) as well as the fact that all relevant quantities are intrinsic to Σ_s, so $\mathcal{R}_s$ can be identified with $\mathcal{R}$. Moreover,

$$(B^{(x)}_{\mathcal{R}_s})_{ij} = (B^{(x)}_{\mathcal{R}})_{ij} \implies B^{(x)}_{\mathcal{R}_s} = 0 \ , \tag{7.3.17}$$

since this quantity depends only on the intrinsic metrics of Σ and Σ_s, which are identical.

Properties

We will now establish important properties and an equivalent formulation of the kink transform.

Let us write Σ as the disjoint union

$$\Sigma = a' \cup \mathcal{R} \cup a \ . \tag{7.3.18}$$

The spacetime $\mathcal{M}$ contains $D(a)$ and $D(a')$ where $D(.)$ denotes the domain of dependence. The kink transformed slice Σ_s contains regions a and a' with identical initial data, so $\mathcal{M}_s$ also contains $D(a)$ and $D(a')$. Because Σ_s has different extrinsic curvature at $\mathcal{R}$, the two domains of dependence will be glued to each other differently in $\mathcal{M}_s$, so the full spacetime will differ from $\mathcal{M}$ in the future and past of $\mathcal{R}$. This is depicted in Fig. 7.2.

We will now derive an alternative formulation of the kink transform as a one-sided local Lorentz boost at $\mathcal{R}$. The unit vector field $t^\mu_{\Sigma_s}$ normal to Σ_s is discontinuous at $\mathcal{R}$ due to the kink. Let

$$(t^\mu_{\Sigma_s})_R = \lim_{x \to 0^+} t^\mu_{\Sigma_s} \ , \tag{7.3.19}$$

$$(t^\mu_{\Sigma_s})_L = \lim_{x \to 0^-} t^\mu_{\Sigma_s} \tag{7.3.20}$$

be the left and right limits to $\mathcal{R}$. The metric of $\mathcal{M}_s$ is continuous since it arises from valid initial data on Σ_s. Therefore, the normal bundle of 1+1 dimensional normal spacetimes to points in $\mathcal{R}$ is well-defined. The above vector fields $(t^\mu_{\Sigma_s})_R$ and $(t^\mu_{\Sigma_s})_L$ belong to this normal bundle. Therefore at each point on $\mathcal{R}$, the two vectors can only differ by a Lorentz boost acting in 1+1 dimensional Minkowski space. The kink transform, Eq. (7.3.4), implies:

$$(t^\mu_{\Sigma_s})_R = (\Lambda_{2\pi s})^\mu_\nu (t^\nu_{\Sigma_s})_L \ , \tag{7.3.21}$$

where $(\Lambda_{2\pi s})^\mu_\nu$ is a Lorentz boost of rapidity $2\pi s$. In this sense, the kink transform resembles a local boost around $\mathcal{R}$. Alternatively, we can view Eq. (7.3.21) as the definition of the kink transform. This definition can be applied to Cauchy slices that are not smooth around $\mathcal{R}$, but it reduces to Eq. (7.3.4) in the smooth case.

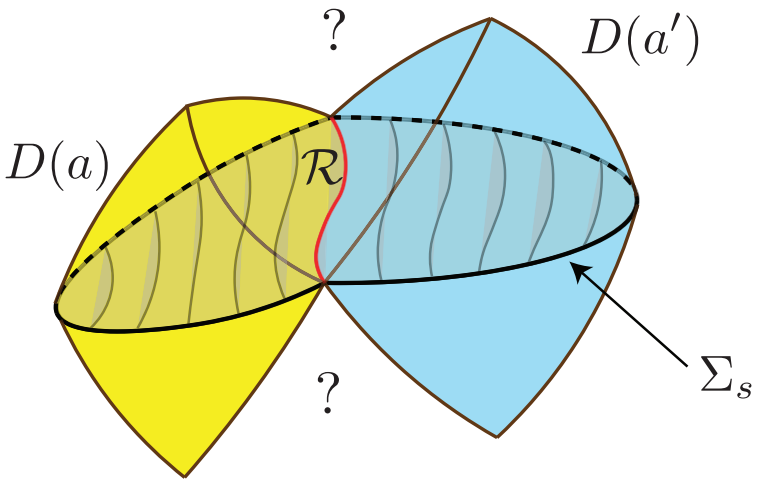

Figure 7.2: The kink-transformed spacetime $\mathcal{M}_s$ is generated by the Cauchy evolution of the kinked slice Σ_s. This reproduces the left and right entanglement wedges $D(a)$ and $D(a')$ of the original spacetime $\mathcal{M}$. The future and past of the extremal surface $\mathcal{R}$ are in general not related to the original spacetime.

This observation applies equally to any other vector field ξ^μ in the normal bundle to $\mathcal{R}$, if ξ^μ has a smooth extension into $D(a')$ and $D(a)$ in $\mathcal{M}$. The norm of ξ^μ and its inner products with $(t^\mu_{\Sigma_s})_L$ and $(t^\mu_{\Sigma_s})_R$ are unchanged by the kink transform. Hence, in $\mathcal{M}_s$, the left and right limits of ξ^μ to $\mathcal{R}$ will satisfy

$$\xi^\mu_R = (\Lambda_{2\pi s})^\mu_\nu \xi^\nu_L \ . \tag{7.3.22}$$

Now let $\Xi \supset \mathcal{R}$ be another Cauchy slice of $D(\Sigma)$. Since Ξ contains $\mathcal{R}$, its timelike normal vector field ξ^μ (at $\mathcal{R}$) lies in the normal bundle to $\mathcal{R}$. We have shown that Eq. (7.3.21) is equivalent to the kink transform of Σ; that Eq. (7.3.22) is equivalent to the kink transform of Ξ; and that Eq. (7.3.21) is equivalent to Eq. (7.3.22). Hence the kink transform of Σ is equivalent to the kink transform of Ξ. In other words, the spacetime resulting from a kink transform about $\mathcal{R}$ does not depend on which Cauchy surface containing $\mathcal{R}$ we apply the kink transform to.

The kink transform (with $s \neq 0$) always generates physically inequivalent initial data. However $\mathcal{M}_s$ need not differ from $\mathcal{M}$. They will be the same if and only if Σ_s is an initial data set in $\mathcal{M}$. There is an interesting special case where this holds for all values of s. Namely, suppose $\mathcal{M}$ has a Killing vector field that reduces to a boost in the normal bundle to $\mathcal{R}$. Then $\Sigma_s \subset \mathcal{M}$ (as a full initial data set), for all s. For example, the kink transform maps straight to kinked slices in the Rindler or maximally extended Schwarzschild spacetimes (see Fig. 7.3).

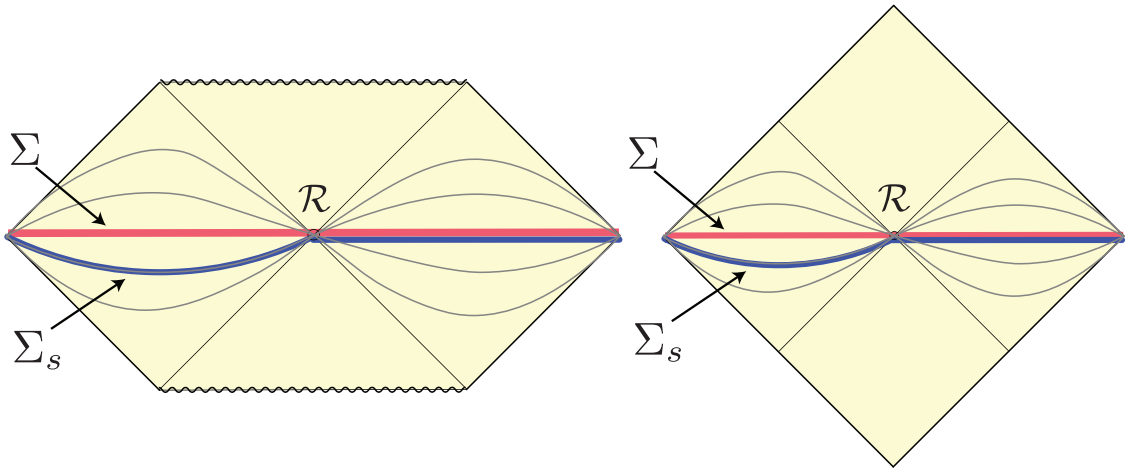

Figure 7.3: Straight slices Σ (red) in a maximally extended Schwarzschild (left) and Rindler (right) spacetime get mapped to kinked slices Σ_s (blue) under the kink transform about $\mathcal{R}$.

We can also consider the kink transform of matter fields on a fixed background spacetime with the above symmetry. Geometrically, $\mathcal{M} = \mathcal{M}_s$ for all s, but the matter fields will differ in $\mathcal{M}_s$ by a one-sided action of the Killing vector field. For example, let $\mathcal{M}$ be Minkowski space, with two balls at rest at $x = \pm 1$, $y = z = 0$ (see Fig. 7.4); and let $\mathcal{R}$ given by $x = t = 0$. In the spacetime $\mathcal{M}_s$ obtained by a kink transform, the two balls will approach with velocity $\tanh 2\pi s$ and so will collide. The right and left Rindler wedge, $D(a)$ and $D(a')$, are separately preserved; the collision happens in the past or future of $\mathcal{R}$.

7.4 Bulk Kink Transform = Boundary CC Flow

In this section, we will argue that the kink transform is the bulk dual of boundary CC flow. We will show that the kink transform satisfies two nontrivial necessary conditions. First, in Sec. 7.4, we show that the left and right bulk region are the subregion duals to the left and right boundary region, respectively. In Sec. 7.4 we show that the bulk kink transform leads to precisely the stress tensor shock at the boundary generated by boundary CC flow, Eq. (7.2.47). (In Sec. 7.5 we will show that the kink transform predicts additional shocks in the CC flowed state, which have not been derived previously purely from QFT methods.)

Matching Left and Right Reduced States

The entanglement wedge of a boundary region A in a (pure or mixed) state ρ_A,

$$\mathrm{EW}(\rho_A) = D[a(\rho_A)] \tag{7.4.1}$$

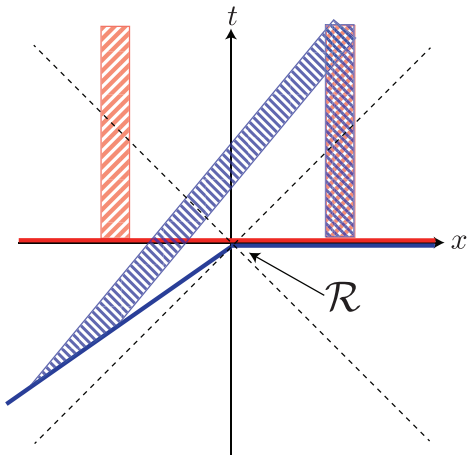

Figure 7.4: On a fixed background with boost symmetry, the kink transform changes the initial data of the matter fields. In this example, $\mathcal{M}$ is Minkowski space with two balls relatively at rest (red).The kink transform is still Minkowski space, but the balls collide in the future of $\mathcal{R}$ (blue).

is the domain of dependence of a bulk achronal region a satisfying the following properties [31, 32, 5, 6]:

1. The topological boundary of a (in the unphysical spacetime that includes the conformal boundary of AdS) is given by $\partial a = A \cup \mathcal{R}$.

2. $S_{\text{gen}}(a)$ is stationary under small deformations of $\mathcal{R}$.

3. Among all regions that satisfy the previous criteria, $\text{EW}(\rho_A)$ is the one with the smallest $S_{\text{gen}}(a)$.

We neglect end-of-the-world branes in this discussion [151, 152]. The generalized entropy is given by

$$S_{\text{gen}} = \frac{\text{Area}(\mathcal{R})}{4G\hbar} + S(a) + \dots \ , \tag{7.4.2}$$

where $S(a)$ is the von Neumann entropy of the region a and the dots indicate subleading geometric terms. The entanglement wedge is also referred to as the Wheeler-DeWitt patch of A.

There is significant evidence [35, 131] that $\text{EW}(\rho_A)$ represents the entire bulk dual to the boundary region A. That is, all bulk operators in $\text{EW}(\rho_A)$ have a representation in the algebra of operators $\mathcal{A}$ associated with A; and all simple correlation functions in A can be computed from the bulk. In other words, the entanglement wedge appears to be the answer [6] to the question [153, 154, 33, 37] of "subregion duality." A bulk surface $\mathcal{R}$ is

called quantum extremal (with respect to A in the state ρ) if it satisfies the first two criteria, and quantum RT if it satisfies all three. When the von Neumann entropy term in Eq. (7.4.2) is neglected, $\mathcal{R}$ is called an extremal or RT surface, respectively. This will be the case everywhere in this paper except in Sec. 7.6.

We now specialize to the setting in which CC flow was considered in Sec. 7.2. Recall that the pure boundary state $|\psi(\mathcal{C})\rangle$ is given on a boundary slice $\mathcal{C}$ corresponding to $t = 0$ in standard Minkowski coordinates; and that we regard $\mathcal{C}$ as the disjoint union of the left region A_0' ($x < 0$), with reduced state $\rho^{\psi}_{A_0}$; the cut ∂A_0 ($x = 0$); and the right region A_0 ($x > 0$), with reduced state $\rho^{\psi}_{A_0'}$. Let a_0' and a_0 be arbitrary Cauchy surfaces of the associated entanglement wedges $\mathrm{EW}(\rho^{\psi}_{A_0'})$ and $\mathrm{EW}(\rho^{\psi}_{A_0})$.

The entanglement wedges of non-overlapping regions are always disjoint, so

$$\mathrm{EW}(\rho^{\psi}_{A_0'}) \cap \mathrm{EW}(\rho^{\psi}_{A_0}) = \varnothing \ . \tag{7.4.3}$$

For the bipartition of a pure boundary state ψ, entanglement wedge complementarity holds:

$$a[|\psi(\mathcal{C})\rangle] = a_0' \cup \mathcal{R} \cup a_0 \ , \tag{7.4.4}$$

where $a[|\psi(\mathcal{C})\rangle]$ is a Cauchy surface of $\mathrm{EW}(|\psi(\mathcal{C})\rangle)$. In particular, the left and right entanglement wedge share the same HRT surface $\mathcal{R}$.

Crucially, the classical initial data on $a[|\psi(\mathcal{C})\rangle]$ is almost completely determined by the data on a_0' and a_0; however the data on $\mathcal{R}$ are not contained in a_0' nor in a_0. In the semiclassical regime, the quantum state on $a[|\psi(\mathcal{C})\rangle]$ also includes global information (through its entanglement structure) that neither subregion contains on its own. Hence in general

$$\mathrm{EW}(|\psi(\mathcal{C})\rangle) = D\left[\mathrm{EW}(\rho^{\psi}_{A_0'}) \cup \mathcal{R} \cup \mathrm{EW}(\rho^{\psi}_{A_0})\right] \tag{7.4.5}$$

is a proper superset of $\mathrm{EW}(\rho^{\psi}_{A_0'}) \cup \mathrm{EW}(\rho^{\psi}_{A_0})$ that also includes some of the past and future of $\mathcal{R}$.

Now consider the CC-flowed state on the precursor slice $|\psi_s(\mathcal{C}_s)\rangle$. By Eqs. (7.2.37) and (7.2.38), we have

$$\mathrm{EW}(\rho^{\psi_s}_{A_s'}) = \mathrm{EW}(\rho^{\psi}_{A_0'}) = D(a_0') \ , \tag{7.4.6}$$

$$\mathrm{EW}(\rho^{\psi_s}_{A_0}) = \mathrm{EW}(\rho^{\psi}_{A_0}) = D(a_0) \ , \tag{7.4.7}$$

Since $|\psi_s\rangle$ is again a pure state, $\mathrm{EW}[|\psi_s(\mathcal{C}_s)\rangle] = D\left(a[|\psi_s(\mathcal{C}_s)\rangle]\right)$ where

$$a[|\psi_s(\mathcal{C}_s)\rangle] = a_0' \cup \mathcal{R} \cup a_0 \ . \tag{7.4.8}$$

We see that this initial data slice has the same intrinsic geometry as that of the original bulk dual. Indeed, by the remarks following Eq. (7.4.4), the full classical initial data for the bulk dual to $|\psi_s\rangle$ will be identical on $a_0' \cup a_0$ and can only differ from the initial data for the original bulk at $\mathcal{R}$.

We pause here to note that a kink transform of $a[|\psi(\mathcal{C})\rangle]$ centered on $\mathcal{R}$ satisfies this necessary condition and hence becomes a candidate for $a[|\psi_s(\mathcal{C}_s)\rangle]$. However, this does not yet constrain the value of s. In order to go further, we would now like to show that a kink transform of $a[|\psi(\mathcal{C})\rangle]$ with parameter s yields a bulk slice whose boundary is geometrically the precursor slice $\mathcal{C}_s$.

The bulk metric takes the asymptotic form [155]:[5]

$$ds^2 = \frac{1}{z^2}\left[dz^2 + \eta_{AB}dx^A dx^B + O(z^d)\right] \ , \tag{7.4.9}$$

where η_{AB} is the metric of Minkowski space. Consider a stationary bulk surface $\mathcal{R}$ anchored on the boundary cut $u = v = 0$. At leading order, $\mathcal{R}$ will reside at $u = v = 0$ in the asymptotic bulk, in the above metric [29]. (The first subleading term, which appears at order z^d, will be crucial in our derivation of the boundary stress tensor shock in Sec. 7.4.)

Let Σ be a bulk surface that contains $\mathcal{R}$ and satisfies $t = 0 + O(z^d)$ in the metric of Eq. (7.4.9). Since the initial data on each side of $\mathcal{R}$ are separately preserved (see Sec. 7.3), Eq. (7.3.21) dictates that the kink transform Σ_s of Σ satisfies $t = 0$ $(x > 0)$ and $t = x \tanh 2\pi s$ $(x < 0)$, again up to corrections of order z^d. The corrections all vanish at $z = 0$, where Σ is bounded by $\mathcal{C}$ and Σ_s is bounded by $\mathcal{C}_s$ (see Eq. (7.2.32)). Recall also that the kink transform is slice-independent. Thus we have established that the kink transform of any Cauchy surface $a[|\psi(\mathcal{C})\rangle]$, by s along $\mathcal{R}$, yields a Cauchy surface bounded by the precursor slice $\mathcal{C}_s$.

The above arguments establish that

$$\mathrm{EW}[|\psi_s(\mathcal{C}_s)\rangle] = D\left(a[|\psi_s(\mathcal{C}_s)\rangle]\right) \ , \tag{7.4.10}$$

where $a[|\psi_s(\mathcal{C}_s)\rangle]$ is given by Eq. (7.4.8). In words, the bulk dual of the CC-flowed boundary state is the Cauchy development of the kink-transform of a Cauchy slice containing the HRT surface $\mathcal{R}$. Note that the classical initial data on this Cauchy surface is fully determined by the initial data on a_0' and a_0 inherited from the bulk dual of $|\psi(\mathcal{C})\rangle$, combined with the distributional geometric initial data consisting of the extrinsic curvature shock at $\mathcal{R}$. The full spacetime geometry will differ from $\mathrm{EW}[|\psi(\mathcal{C})\rangle]$ because of the different gluing at $\mathcal{R}$.

Matching Bulk and Boundary Shocks

In Sec. 7.3, we gave a prescription for generating bulk geometries in AdS by inserting a kink on the Cauchy surface, at the HRT surface. With the standard holographic dictionary, the resulting geometry manifestly yields the correct behavior of one-sided boundary observables under CC flow. This was shown in the previous subsection.

Another characteristic aspect of the CC flowed state $|\psi_s\rangle$ is the presence of a stress tensor shock at the cut (Sec. 7.2), proportional to shape derivatives of the von Neumann entropy; see Eq. (7.2.47). We will now verify that this shock is reproduced by the kink transform

[5] We set $\ell_{\mathrm{AdS}} = 1$.

in the bulk, upon applying the AdS/CFT dictionary. Notably, the shock is not localized to either wedge. Verifying kink/CC duality for this observable furnishes an independent, nontrivial check of our proposal.

We will now keep the first subleading term in the Fefferman-Graham expansion of the asymptotic bulk metric [155, 29]:

$$ds^2 = \frac{1}{z^2}\left(dz^2 + g_{AB}(x,z)dx^A dx^B\right) \ , \tag{7.4.11}$$

$$g_{AB}(x,z) = \eta_{AB} + z^d \frac{16\pi G}{d}\langle T_{AB}\rangle + o(z^d) \ , \tag{7.4.12}$$

where indices $A, B, \ldots$ correspond to directions along $z =$ const. surfaces.

The location of the RT surface $\mathcal{R}$ in the bulk can be described by a collection of $(d-1)$ embedding functions

$$X^\mu(y,z) = (z, X^A(y,z)) \ , \tag{7.4.13}$$

where (y,z) are intrinsic coordinates on $\mathcal{R}$. The expansion in z takes the simple form

$$X^A(y,z)) = z^d X^A_{(d)} + o(z^d) \ , \tag{7.4.14}$$

because the boundary anchor is the flat cut $u = v = 0$ of the Rindler horizon [29]. Stationarity of $\mathcal{R}$ can be shown to imply [29]

$$X^A_{(d)} = -\frac{4G}{d}\left.\frac{\delta S}{\delta X^A}\right|_{\mathcal{R}} \ . \tag{7.4.15}$$

We consider a bulk Cauchy slice $\Sigma \supset \mathcal{R}$ for which $\partial\Sigma$ corresponds to the $t = 0$ slice on the boundary. Since the subleading terms in Eqs. (7.4.12) and (7.4.14) start at z^d, we are free to choose Σ so that it is given by

$$t = z^d \varsigma(x) + o(z^d) \ , \tag{7.4.16}$$

Recall that the vector fields t^μ and x^μ are defined to be orthogonal to $\mathcal{R}$, and respectively orthogonal and tangent to Σ_s. In FG coordinates one finds:

$$t^A = z\left(t^A_{(0)} + z^d t^A_{(d)} + o(z^d)\right) \ , \tag{7.4.17}$$

$$x^A = z\left(x^A_{(0)} + z^d x^A_{(d)} + o(z^d)\right) \ , \tag{7.4.18}$$

$$t^z = z\left(z^{d-1} t^z_{(d-1)} + o(z^{d-1})\right) \ , \tag{7.4.19}$$

$$x^z = z\left(z^{d-1} x^z_{(d-1)} + o(z^{d-1})\right) \ . . \tag{7.4.20}$$

The overall factor of z is due to normalization. Note that $t^\mu_{(0)}$ is a coordinate vector field but in general, t^μ is not. Individual coordinate components of vectors and tensors are defined by contractions with $t^\mu_{(0)}$ and $x^\mu_{(0)}$ respectively, for example $t^t \equiv t_\mu t^\mu_{(0)}$.

We now consider a contraction of the extrinsic curvature tensor on Σ,

$$(K_\Sigma)_{ab}x^b = P_a^\mu x^\nu \nabla_{(\mu}t_{\nu)} \ . \tag{7.4.21}$$

We would like to further project the a index onto the z direction. Deep in the bulk the z direction does not lie entirely in Σ. However, note that $g_{\mu z}t^\mu \to 0$ in the limit $z \to 0$ due to Eq. (7.4.19). Therefore, at leading order in z, the z direction does lie entirely in Σ; moreover, $P_z^\mu \to \delta_z^\mu$ as $z \to 0$. We will only be interested in evaluating Eq. (7.4.21) at leading order in z so we may freely set $a = z$, which yields:[6]

$$\begin{aligned}(K_\Sigma)_{z\nu}x^\nu - x^\nu \partial_{(z}t_{\nu)} &= x^\nu t_\gamma \, \Gamma^\gamma_{\nu z} &(7.4.22)\\ &= z^2\Gamma_{txz} + zx^t\Gamma_{ttz} + zt^x\Gamma_{xxz} + x^z t^z \Gamma_{zzz} + o(z^{d-1}) &(7.4.23)\\ &= \frac{(d-2)}{2}z^{d-1}\frac{16\pi G}{d}\langle T_{tx}\rangle - z^{-3}t^z x^z - z^{d-1}(t^x_{(d)} - x^t_{(d)}) + o(z^{d-1}) \ . &(7.4.24)\end{aligned}$$

The condition $x_\mu t^\mu = 0$ implies that

$$z^{d-1}\frac{16\pi G}{d}\langle T_{tx}\rangle + z^{-3}x^z t^z + z^{d-1}(t^x_{(d)} - x^t_{(d)}) + o(z^{d-1}) = 0 \ . \tag{7.4.25}$$

Hence we find

$$(K_\Sigma)_{z\nu}x^\nu - x^\nu \partial_{(z}t_{\nu)} = z^{d-1}\, 8\pi G \, \langle T_{tx}\rangle + o(z^{d-1}) \ . \tag{7.4.26}$$

We now apply the kink transform to Σ (viewed as an initial data set). This yields a new initial data set on a slice Σ_s in a new spacetime $\mathcal{M}_s$. We again expand in Fefferman-Graham coordinates:

$$ds^2 = \frac{1}{z^2}\left(dz^2 + \tilde{g}_{AB}(\tilde{x}, z)d\tilde{x}^A d\tilde{x}^B\right) \ , \tag{7.4.27}$$

$$\tilde{g}_{AB}(x, z) = \tilde{\eta}_{AB} + z^d \frac{16\pi G}{d}\langle \tilde{T}_{AB}\rangle + o(z^d) \ . \tag{7.4.28}$$

Here $\tilde{\eta}_{AB}$ is still Minkowski space; any change in the bulk geometry will be encoded in the subleading term.

The notation $\tilde{\eta}_{AB}$ indicates that we will be using the specific coordinates in which the metric of d-dimensional Minkowski space takes the nonstandard form given by Eq. (7.2.41). This has the advantage that the *coordinate* form of all vectors, tensors, and embedding equations in $D(a') \cup \mathcal{R} \cup D(a)$ will be unchanged by the kink transform, if we use standard Cartesian coordinates before the transform and the tilde coordinates afterwards.

For example, the invariance of the left and right bulk domains of dependence under the kink transform implies that Σ_s is given by

$$\tilde{t} = z^d \varsigma(\tilde{x}) + o(z^d) \ , \tag{7.4.29}$$

[6] In $d > 2$ the terms involving $x^z t^z$ will be higher order, by Eqs. (7.4.19) and (7.4.20), and need not be included. Since they cancel out either way, we include them here to avoid an explicit case distinction.

with the *same* ς that appeared in Eq. (7.4.16). (In fact, this extends to at all orders in z.) As already shown in the previous subsection, $\partial\Sigma_s$ lies at $\tilde{t}=0$, $z=0$.

As another example, the coordinate components of the unit normal vector to Σ_s in $\mathcal{M}_s$, $\tilde{t}^\mu$, will be the same as the components of the normal vector to Σ in $\mathcal{M}$, t^μ, and therefore

$$\partial_{(z}t_{\nu)}\big|_{\mathcal{M}} = \partial_{(z}\tilde{t}_{\nu)}\big|_{\mathcal{M}_s} \ . \tag{7.4.30}$$

Below we will use the convention that any quantity with a tilde is evaluated in $\mathcal{M}_s$, in the coordinates of Eq. (7.4.28). Any quantity without a tilde is evaluated in $\mathcal{M}$, in the coordinates of Eq. (7.4.12). The only exception is the extrinsic curvature tensor, where the corresponding distinction is indicated by the subscript Σ_s or Σ, for consistency with Sec. 7.3.

We now consider the extrinsic curvature of Σ_s. A calculation analogous to the derivation of Eq. (7.4.26) implies

$$(K_{\Sigma_s})_{z\nu}\tilde{x}^\nu - \tilde{x}^\nu\partial_{(z}\tilde{t}_{\nu)} = z^{d-1}\,8\pi G\,\langle\tilde{T}_{\tilde{t}\tilde{x}}\rangle + o(z^{d-1}) \ . \tag{7.4.31}$$

From Eqs. (7.4.26) and (7.4.30) we find

$$z^{d-1}\langle\tilde{T}_{\tilde{t}\tilde{x}}\rangle = z^{d-1}\langle T_{tx}\rangle + \frac{1}{8\pi G}\left[(K_{\Sigma_s})_{z\nu} - (K_\Sigma)_{z\nu}\right]x^\nu + o(z^{d-1}) \ , \tag{7.4.32}$$

$$= z^{d-1}\langle T_{tx}\rangle - \frac{\sinh(2\pi s)}{8\pi G}\delta(\mathcal{R})x_z + o(z^{d-1}) \ . \tag{7.4.33}$$

In the first equality, we used the fact that x^μ and $\tilde{x}^\mu$ can be identified as vector fields, and the extrinsic curvature tensors can be compared, in the submanifold $\Sigma=\Sigma_s$. The second equality follows from the definition of the kink transform, Eq. (7.3.4).

By Eq. (7.4.18), $\delta(\mathcal{R}) = \delta(z^{-1}\tilde{x}) = z\delta(\tilde{x})$. The condition $x_\mu\partial_z X^\mu = 0$ yields

$$x_z = -d\,z^{d-2}\widetilde{X}_{(d)} + o(z^{d-2}) \ , \tag{7.4.34}$$

where $\widetilde{X}_{(d)}$ is the $A=\tilde{x}$ component of $X^A_{(d)}$. Taking $z\to 0$ and using Eq. (7.4.15), we thus find

$$\langle\tilde{T}_{\tilde{t}\tilde{x}}\rangle = \langle T_{tx}\rangle + \frac{1}{2\pi}\sinh(2\pi s)\frac{\delta S}{\delta\widetilde{X}}\Big|_{\widetilde{X}=0}\delta(\tilde{x}) \ , \tag{7.4.35}$$

which agrees precisely with Eq. (7.2.47).

Note that this derivation applies to any boosted coordinate system $(\check{t},\check{x})$ as well. Linear combinations of Eq. (7.4.35) with its boosted version reproduces both the $T_{\tilde{u}\tilde{u}}$ shock of Eq. (7.2.46) and the $T_{\tilde{v}\tilde{v}}$ shock of Eq. (7.2.45) holographically.

7.5 Predictions

Having found nontrivial evidence for kink transform/CC flow duality, we now change our viewpoint and assume the duality. In this section, we will derive a *novel* property of CC flow

from the kink transform: a shock in the $\langle T_{xx}\rangle$ component of the stress tensor in the CC flowed state. We do not yet know of a way to derive this directly in the quantum field theory, so this result demonstrates the utility of the kink transform in extracting nontrivial properties of CC flow. We further argue that $\langle T_{xx}\rangle$ and $\langle T_{tx}\rangle$ constitute all of the independent, nonzero stress tensor shocks in the CC flowed state.

Our holographic derivation only depends on near boundary behavior, and the value of the shock takes a universal form similar to Eq. (7.4.35). Thus, we expect that the properties we find in holographic CC flow hold in non-holographic QFTs as well.

To derive the $\langle T_{xx}\rangle$ shock, we use the Gauss-Codazzi relation [156]

$$P_a^\mu P_b^\nu P_c^\alpha P_d^\beta R_{\mu\nu\alpha\beta} = K_{ac}K_{bd} - K_{bc}K_{ad} + r_{abcd} \ , \tag{7.5.1}$$

where r_{abcd} is the intrinsic Riemann tensor of Σ. It is important to note that this relation is purely intrinsic to Σ. Since $\Sigma = \Sigma_s$ as submanifolds, we can not only evaluate Eq. (7.5.1) in both $\mathcal{M}$ and $\mathcal{M}_s$ but also meaningfully subtract the two. We emphasize that the following calculation is only nontrivial in $d > 2$ (in $d = 2$, the Gauss-Codazzi relation is trivial). We comment on $d = 2$ at the end.

First we evaluate Eq. (7.5.1) in $\mathcal{M}$. We will only be interested in evaluating it to leading order in z in the Fefferman-Graham expansion. As argued in Sec. 7.4, when working at leading order we can freely set $a = c = z$. We then compute the following at leading order in z:

$$R_{zxzx} = K_{zz}K_{xx} - (K_{xz})^2 + r_{zxzx} \ . \tag{7.5.2}$$

We start by computing K_{zz}. First we note that $\Gamma^\alpha_{zz} t_\alpha = 0$ identically. Therefore,

$$K_{zz} = \partial_z t_z = 4G(d-2)z^{d-3}\frac{\delta S}{\delta T}\Big|_{\mathcal{R}} + o(z^{d-3}) \ . \tag{7.5.3}$$

We have made use of

$$t_z = 4Gz^{d-2}\frac{\delta S}{\delta T}\Big|_{\mathcal{R}} + o(z^{d-2}) \ , \tag{7.5.4}$$

which follows from $t_\mu \partial_z X^\mu = 0$.

Next we compute

$$R_{zxzx} = \partial_x \Gamma^x_{zz} - \partial_z \Gamma^x_{xz} + \Gamma^x_{x\mu}\Gamma^\mu_{zz} - \Gamma^x_{z\mu}\Gamma^\mu_{xz} \ . \tag{7.5.5}$$

One finds

$$\partial_z \Gamma^x_{xz} = \frac{1}{2}(d-2)(d-1)z^{d-2}\frac{16\pi G}{d}\langle T_{xx}\rangle + o(z^{d-2}) \ , \tag{7.5.6}$$

$$\Gamma^x_{xz}\Gamma^z_{zz} = -\frac{1}{2}(d-2)z^{d-2}\frac{16\pi G}{d}\langle T_{xx}\rangle + o(z^{d-2}) \ , \tag{7.5.7}$$

with all other terms either subleading in z or identically vanishing, and hence

$$R_{zxzx} = -8\pi G(d-2)z^{d-2}\langle T_{xx}\rangle + o(z^{d-2}) \ . \tag{7.5.8}$$

Putting all this together, we have

$$-8\pi G(d-2)z^{d-2}\langle T_{xx}\rangle = 4Gz^{d-3}\frac{\delta S}{\delta\widetilde{T}}\Big|_{\mathcal{R}} K_{xx} - (K_{xz})^2 + r_{zxzx} + o(z^{d-2}) \ . \tag{7.5.9}$$

The analogous relation evaluated in $\mathcal{M}_s$ reads

$$-8\pi G(d-2)z^{d-2}\langle \tilde{T}_{\tilde{x}\tilde{x}}\rangle = 4Gz^{d-3}\frac{\delta S}{\delta\widetilde{T}}\Big|_{\mathcal{R}} \tilde{K}_{\tilde{x}\tilde{x}} - (\tilde{K}_{\tilde{x}z})^2 + \tilde{r}_{z\tilde{x}z\tilde{x}} + o(z^{d-2}) \ , \tag{7.5.10}$$

where we have made use of Eq. (7.4.30) to set $K_{zz} = \tilde{K}_{zz}$. We can now subtract these two relations. First note that $\tilde{r}_{abcd} = r_{abcd}$ since it is purely intrinsic to Σ. Next, recall from the definition of the kink transform Eq. (7.3.4) that

$$\tilde{K}_{\tilde{x}\tilde{x}} - K_{xx} = -z\sinh(2\pi s)\delta(\tilde{x}) \ . \tag{7.5.11}$$

Lastly, it is easy to check that $K_{xz} \sim o(z^{d-2})$ hence its contribution to Eq. (7.5.9) is subleading, and similarly for Eq. (7.5.10). Thus, subtracting Eq. (7.5.10) from Eq. (7.5.9) yields

$$\langle \tilde{T}_{\tilde{x}\tilde{x}}\rangle - \langle T_{xx}\rangle = \frac{1}{2\pi}\sinh(2\pi s)\frac{\delta S}{\delta\widetilde{T}}\Big|_{\widetilde{X}=0}\delta(\tilde{x}) \ . \tag{7.5.12}$$

The above calculation only works in $d > 2$. In $d = 2$, since the boundary theory is a CFT, tracelessness of the boundary stress tensor further implies that $\langle T_{tx}\rangle$ is the *only* independent component of the stress tensor shock so there is no need for a calculation analogous to the one above. We expect that this argument is robust under relevant deformations of the CFT since the shock is highly localized and should universally depend only on the UV fixed point.

Together with the $\langle T_{tx}\rangle$ shock we reproduced in the previous section, and using Lorentz invariance of the boundary, this result determines the transformation of the the stress tensor contracted with any pair of linear combinations of t and x, such as $\langle T_{tt}\rangle$. This linear space contains all of the independent nonvanishing components of the shock. To see this, note that

$$x^\nu\left(\nabla_\nu\tilde{y}_\mu - \nabla_\nu y_\mu\right) = 0 \ , \tag{7.5.13}$$

$$y^\nu\left(\nabla_\nu\tilde{y}_\mu - \nabla_\nu y_\nu\right) = 0 \ , \tag{7.5.14}$$

$$y^\nu\left(\nabla_\nu\tilde{t}_\mu - \nabla_\nu t_\mu\right) = 0 \ , \tag{7.5.15}$$

where $y^\mu = P^\mu_i y^i$ for any vector field y^i in the tangent bundle of $\mathcal{R}$. Eqs. (7.5.13) and (7.5.14) follow trivially from the fact that the prescription Eq. (7.3.22) only introduces a discontinuity in vector fields in the normal bundle of $\mathcal{R}$, while Eq. (7.5.15) simply follows from Eq. (7.3.4). Evaluating the $\mu = z$ components in the same way as in Sec. 7.4, we find,

$$\langle \tilde{T}_{\tilde{\mu}\tilde{y}}\rangle - \langle T_{\mu y}\rangle = 0 \ . \tag{7.5.16}$$

For $s \to \infty$, the shocks derived in the previous two sections agree with those found to be required for the existence of certain coarse grained bulk states in Ref. [139]. In that work, the cut was allowed to be a wiggly or flat cut of a bifurcate horizon such as a Rindler horizon, and the state could belong to any quantum field theory. Interpolation of these results suggests that the shocks we have derived here generalize to the case of CC flow for a wiggly cut of the Rindler horizon, in general QFTs with a conformal fixed point.

7.6 Discussion

Relation to JLMS and One Sided Modular Flow

The bulk dual of one-sided modular flow [141, 142] resembles the kink transform. CC flow yields a well defined state, however, whereas a one sided modular flowed state is singular in QFT. Correspondingly, the kink transform defined here yields a smooth bulk solution whereas the version implicitly defined in Ref. [142] results in a singular spacetime (see also Ref. [18], footnote 4). We will now explain this in detail.

Consider a boundary region A_0 with reduced state ρ_{A_0}, dual to a semi-classical state ρ_a in the bulk entanglement wedge a associated to A_0 as seen in Fig. 7.5. We denote by $K_A = -\log \rho_A$ and $K_a = -\log \rho_a$ the boundary and bulk modular Hamiltonians, respectively. The JLMS result [7] states that

$$\hat{K}_{A_0} = \frac{\hat{A}[\mathcal{R}]}{4G} + \hat{K}_a \ , \tag{7.6.1}$$

where $\hat{A}[\mathcal{R}]$ is the area operator that formally evaluates the area of the quantum extremal surface $\mathcal{R}$ [6].

Suppose now that A_0 has a nonempty boundary ∂A_0. Then there is an interesting asymmetry in Eq. (7.6.1). The one-sided boundary modular operator appearing on the left hand side is well-defined only with a UV cutoff. On the other hand, at least the leading (area) term in the bulk modular operator on the right hand side has a well-defined action. Let us discuss each side in turn.

In Einstein gravity, the area operator $\hat{A}$ is the generator of one-sided boosts. To see this, let us restrict the gravitational phase space to the bulk region $D(\Sigma_\epsilon)$. There exists a (non-unique) vector field ξ^a in $D(a') \cup \mathcal{R}$ such that ξ^a generates an infinitesimal one-sided boost at $\mathcal{R}$ [157, 158]. This boost can be quantified by a parameter s in the normal bundle to $\mathcal{R}$, as described in Sec. 7.3. The area functional $A[\mathcal{R}]/4G$ is the Noether charge at $\mathcal{R}$ associated to ξ^a, given by the expectation value of the area operator in the semi-classical bulk state:

$$A[\mathcal{R}] = \langle \hat{A}[\mathcal{R}] \rangle \ . \tag{7.6.2}$$

Each point in the gravitational phase space can be specified by the metric in $D(a')$, the metric in $D(a)$, and the boost angle s at $\mathcal{R}$ with which the two domains of dependence are

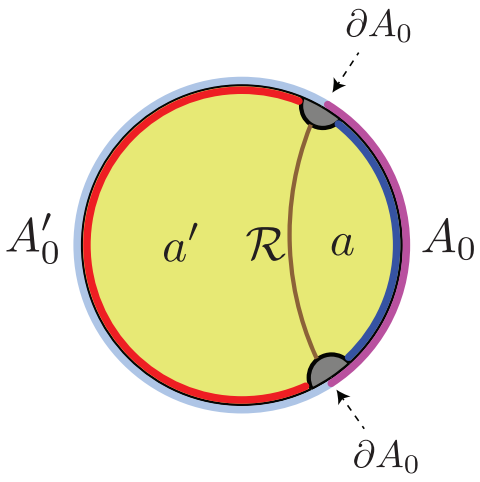

Figure 7.5: A boundary subregion A_0 (pink) has a quantum extremal surface denoted $\mathcal{R}$ (brown) and an entanglement wedge denoted a. The complementary region A_0' (light blue) has the entanglement wedge a'. CC flow generates valid states, but one-sided modular flow is only defined with a UV cutoff. For example, one can consider regulated subregions $A^{(\epsilon)}$ (deep blue) and $A'^{(\epsilon)}$ (red). In the bulk, this amounts to excising an infrared region (gray) from the joint entanglement wedge (yellow).

glued together [142, 158, 159, 144]. The action of

$$\langle e^{2\pi i s \hat{A}[\mathcal{R}]/4G} \rangle \tag{7.6.3}$$

on points in the gravitational phase space is to simply shift the conjugate variable, *i.e.*, the relative boost angle between the left and right domains of dependence, by s. Note that the metrics in the left and right domains of dependence are unchanged since the area functional acts purely on the phase space data at $\mathcal{R}$. This is the classical analogue of the statement that the area operator is in the center of the algebras of the domains of dependence [9]. Comparing with Sec. 7.3, we see that this action is equivalent to the kink transform of Σ_ϵ about $\mathcal{R}$ by s. We stress that this action is well-defined even if $\mathcal{R}$ extends all the way out to the conformal boundary, *i.e.*, in the far ultra-violet from the boundary perspective.

We turn to the right hand side of Eq. (7.6.1), still assuming that A_0 has a nonempty boundary ∂A. Since the algebra of a QFT subregion A_0 is a Type-III$_1$ von Neumann algebra, the Hilbert space does not factorize across ∂A_0 [146]. A reduced density matrix ρ_{A_0}, and hence $\hat{K}_{A_0}$, do not exist. Physically, the action of $\hat{K}_{A_0}$ on a fixed boundary time slice would break the vacuum entanglement of arbitrarily short wavelength modes across ∂A_0; this would create infinite energy.

Therefore, any discussion of $\hat{K}_{A_0}$ requires introducing a UV regulator. Consider the regulated subregions $A_0^{(\epsilon)}$ and $A_0'^{(\epsilon)}$ shown in Fig. 7.5. The split property in algebraic QFT [146,

160] guarantees the existence of a (non-unique) Type-I von Neumann algebra $\mathcal{N}$ nested between the algebras of subregion $A_0^{(\epsilon)}$ and the complementary algebra of $A_0'^{(\epsilon)}$, *i.e.*,

$$\mathcal{A}_0^{(\epsilon)} \subset \mathcal{N} \subset \left(\mathcal{A}_0'^{(\epsilon)}\right)' \ . \tag{7.6.4}$$

With this prescription, one can define a regulated version of the reduced density matrix ρ_A by using the Type-I factor $\mathcal{N}$ [161]. It has been suggested that there exists an $\mathcal{N}$ consistent with the geometric cutoff shown in Fig. 7.5 [162, 160]: the quantum extremal surface $\mathcal{R}$ in the bulk is regulated by a cutoff brane B demarcating the entanglement wedge of the subregion $A_0^{(\epsilon)} \cup A_0'^{(\epsilon)}$. The regulated area operator $\hat{A}[\mathcal{R}]/4G$ is well defined once boundary conditions on B are specified.

Let us now specialize to the case for which we have conjectured kink transform/CC-flow duality: the boundary slice $\mathcal{C} = A_0 \cup A_0'$ is a Cauchy surface of Minkowski space, and ∂A_0 is the flat cut $u = v = 0$ of the Rindler horizon. We have just argued that the kink transformation is generated by the area operator through Eq. (7.6.3). By Eq. (7.6.1), the boundary dual of this action should be one-sided modular flow, not CC flow. By Eqs. (7.2.4) and (7.2.6), these are manifestly different operations. Indeed, unlike one-sided modular flow, Connes cocycle flow yields a well-defined boundary state for all s, without any UV divergence at the cut ∂A_0: $|\psi(\mathcal{C})\rangle \to |\psi_s(\mathcal{C})\rangle$.

In fact there is no contradiction. For both modular flow and CC flow on the boundary, a bulk-dual Cauchy surface Σ_s is generated by the kink transform. The difference is in how Σ_s is glued back to the boundary.

For modular flow, Σ_s is glued back to the original slice $\mathcal{C}$. Generically, this would violate the asymptotically AdS boundary conditions, necessitating a regulator such as the excision of the grey asymptotic region in Fig. 7.5 and interpolation by a brane B. The boundary dual is an appropriately regulated modular flowed state with energy concentrated near the cut ∂A_0. This construction is possible even if ∂A_0 is not a flat plane, but the regulator is ambiguous and cannot be removed.[7]

For CC flow, Σ_s is glued to the precursor slice $\mathcal{C}_s$ as discussed in Sec. 7.3. This yields $|\psi_s(\mathcal{C}_s)\rangle$. Time evolution on the boundary yields $|\psi_s(\mathcal{C})\rangle$, the CC-flowed state on the original slice $\mathcal{C}$.

On the boundary, we can use the one-sided modular operator in two ways. As a map between states on $\mathcal{C}$ [7, 132] it requires a UV regulator. As a map that takes a state on $\mathcal{C}$ to a state on the precursor slice $\mathcal{C}_s$, $|\psi(\mathcal{C})\rangle \to |\psi_s(\mathcal{C}_s)\rangle$, it is equivalent to CC flow on $\mathcal{C}$ by Eqs. (7.2.35) and (7.2.36). This is a more natural choice due to its UV-finiteness. But it is available only if the vacuum modular operator for cut ∂A is geometric, so that the precursor slice is well-defined.

[7] There is evidence that a code subspace can be defined with an appropriate regulator such that one-sided modular flow keeps the state within the code subspace [144, 143, 145].

Quantum Corrections

It is natural to include semiclassical bulk corrections to all orders in G to our proposal. The natural guess would be to perform the kink transform operation about the quantum extremal surface along with a CC flow for the bulk state. In general, it is difficult to describe this procedure within EFT. In states far from the vacuum, the background spacetime changes under the kink transform, and it is unclear how to map states from one spacetime to another. However, we will find some evidence that suggests that the bulk operation relating the two states is a generalized version of CC flow in curved spacetime.

To see this, note that the quantum extremal surface $\mathcal{R}$ satisfies the equations

$$\mathcal{B}^{(t)}_{\mathcal{R}} + 4G\hbar \frac{\delta S}{\delta T} = 0 \ , \tag{7.6.5}$$

$$\mathcal{B}^{(x)}_{\mathcal{R}} + 4G\hbar \frac{\delta S}{\delta X} = 0 \ , \tag{7.6.6}$$

where $(\mathcal{B}^{(t)}_{\mathcal{R}})$ and $(\mathcal{B}^{(x)}_{\mathcal{R}})$ denote the trace of the extrinsic curvature (expansion) in the two normal directions to $\mathcal{R}$, *i.e.*, t^μ and x^μ respectively. Similarly $\frac{\delta S}{\delta T}$ and $\frac{\delta S}{\delta X}$ are the entropy variations in the t^μ and x^μ directions respectively.

The classical kink transform involves an extrinsic curvature shock at the classical RT surface. As shown in Sec. 7.3, extremality of the surface ensures that the constraint equations continue to be satisfied after the kink transform in this case. However, the quantum extremal surface has non-vanishing expansion, the constraint equations are not automatically satisfied when an extrinsic curvature shock is added at the quantum RT surface.

More precisely, the left hand side of the constraint equations on a slice Σ are modified by the kink transform by

$$\Delta\left(r_\Sigma - (K_\Sigma)_{ab}(K_\Sigma)^{ab} + (K_\Sigma)^2\right) = 8G\hbar \sinh(2\pi s) \frac{\delta S}{\delta T} \delta(X) \ , \tag{7.6.7}$$

$$\Delta\left(D_a(K_\Sigma)^a_x - D_x K_\Sigma\right) = 4G\hbar \sinh\left(2\pi s\right) \frac{\delta S}{\delta X} \delta(X) \ , \tag{7.6.8}$$

$$\Delta\left(D_a(K_\Sigma)^a_i - D_i K_\Sigma\right) = 0 \ , \tag{7.6.9}$$

where Δ represents the difference in the constraint equations between the original spacetime $\mathcal{M}$ and the kink transformed spacetime $\mathcal{M}_s$, and we have used Eqs. (7.6.5) and (7.3.4). These are essentially the analogs of Eqs. (7.3.13) and (7.3.14), and we have simplified the notation slightly.

For the constraint equations to be solved, the kink transform would have to generate the same change on the right hand side of the constraints. It would thus have to induce an additional stress tensor shock of the form

$$\Delta T_{TT} = \frac{1}{2\pi} \sinh(2\pi s) \frac{\delta S}{\delta T} \delta(X) \ , \tag{7.6.10}$$

$$\Delta T_{TX} = \frac{1}{2\pi} \sinh(2\pi s) \frac{\delta S}{\delta X} \delta(X) \ . \tag{7.6.11}$$

Formally, these conditions agree precisely with the properties of CC flow discussed in Sec. 7.2. Thus, we might expect a generalized bulk CC flow to result in shocks of precisely this form.

In fact, the existence of semiclassical states satisfying the above equations was conjectured in [139]; the fact that CC flow generates such states in the non-gravitational limit was interpreted as non-trivial evidence in support of the conjecture. Thus, we expect a kink transform at the quantum extremal surface with a suitable modification of the state to provide the bulk dual of CC flow to all orders in G.

At a more speculative level, we can also discuss the bulk dual of CC flow in certain special states called fixed area states, which serve as a natural basis for modular flow [143, 144, 145]. These are approximate eigenstates of the area operator and are therefore unlike smooth semiclassical states which are analogous to coherent states. The Lorentzian bulk dual of such states potentially involves superpositions over geometries [163].

However, by construction, the reduced density matrix is maximally mixed at leading order in G. Thus, the state $|\psi\rangle$ is unaffected by one sided modular flow, and the only effect of CC flow is that we describe the state on a kinematically related slice $\mathcal{C}_s$. Thus, the dual description must be invariant under CC flow up to a diffeomorphism.

In such states, one could apply the semiclassical prescription using Eq. (7.6.1). As discussed above, the action of the area operators results in a diffeomorphism of the geometric description, if it exists. From Eq. 7.6.1, the remaining action of the boundary CC flow is to simply induce a bulk CC flow.

Beyond Flat Cuts

Kink transform/CC flow duality can be generalized to other choices of boundary subsystems, so long as a precursor slice can be defined. The precursor slice is generated by acting on the original slice with the vacuum modular Hamiltonian; this is well-defined only if this action is geometric. In Sec. 7.2, we ensured this by taking the boundary to be Minkowski space and choosing a planar cut. Precursor slices also exist in any conformally related choice, such as a spherical cut.

But there are other settings where the vacuum modular Hamiltonian acts geometrically. This includes multiple asymptotically AdS boundaries where the boundary manifold has a time translation symmetry. For example, consider a two-sided black hole geometry $\mathcal{M}$ with a compact RT surface $\mathcal{R}$ as seen in Fig. 7.6. The boundary manifold is of the form $\mathcal{C} \times \mathbb{R}$, where the first factor corresponds to the spatial geometry and the second corresponds to the time direction. The boundary Hilbert spaces factorizes; each boundary algebra is a Type-I factor. Thus, the version of CC flow defined in terms of density matrices in Eq. (7.2.6) becomes rigorous in this situation. A natural choice of vacuum state is the thermofield double [164, 165]. The reduced state on each side is thermal, $\rho_{A_0} \sim \exp(-\beta H)$. Thus the modular Hamiltonian is proportional to the ordinary Hamiltonian on each boundary. This generates time translations and so is geometric.

Now, in any such geometry $\mathcal{M}$ one can pick a Cauchy slice Σ that ends on boundary time slices on both sides and contains $\mathcal{R}$. In obvious analogy with Sec. 7.3, we conjecture

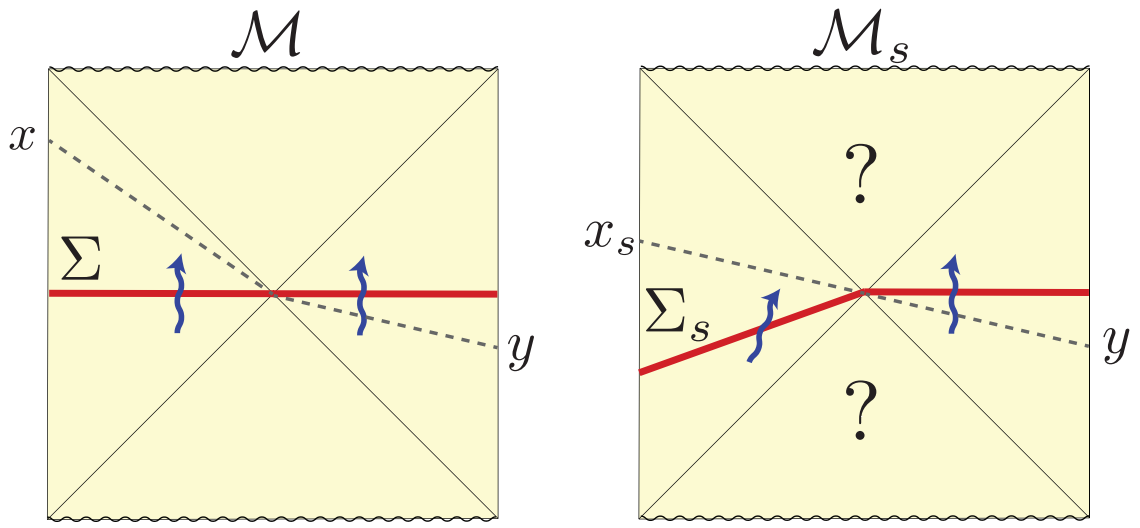

Figure 7.6: An arbitrary spacetime $\mathcal{M}$ with two asymptotic boundaries is transformed to a physically different spacetime $\mathcal{M}_s$ by performing a kink transform on the Cauchy slice Σ. A piecewise geodesic (dashed gray line) in $\mathcal{M}$ connecting x and y with boost angle $2\pi s$ at $\mathcal{R}$ becomes a geodesic between x_s and y in $\mathcal{M}_s$.

that the domain of dependence of the kink transformed slice Σ_s in a modified geometry $\mathcal{M}_s$ is dual to the boundary state:

$$|\psi_s(\mathcal{C}_s)\rangle_{LR} = \rho_L^{-is} |\psi(\mathcal{C})\rangle_{LR} \ , \tag{7.6.12}$$

where we have used the notation of Eq. (7.2.36).

In such a situation, it is again manifest that the Wheeler-DeWitt patches dual to either side are preserved by arguments similar to those made in Sec. 7.4. However, since there is no portion of $\mathcal{R}$ that reaches the asymptotic boundary, there is no analog of the shock matching done in Sec. 7.4. Notably, since $\partial A = \varnothing$ in this case, there is no subtlety regarding boundary conditions for JLMS and thus, one-sided modular flow makes sense without any regulator. Thus, our construction is simply kinematically related to the construction in [142].

An interesting situation arises for wiggly cuts of the Rindler horizon, *i.e.*, $u = 0$ and $v = V(y)$. The modular Hamiltonian acts locally, but only when restricted to the null plane [125]. Its action becomes non-local when extended to the rest of the domain of dependence. The properties of CC flow described in Sections 7.2-7.2 all hold for this choice of cut. This constrains one-sided operator expectation values on the null plane, subregion entropies for cuts entirely to one side of $V(y)$, and even the T_{vv} shock at the cut. Interestingly, all of them are matched by the kink transform, by the arguments given in Sec. 7.3. Even the expected stress tensor shock can still be derived, by taking a null limit of our derivation as described

in Appendix A.15. One might then guess that the kink transform is also dual to CC flow for arbitrary wiggly cuts.

Even in the vacuum, however, the kink transform across a wiggly cut results in a boundary slice that cannot be embedded in Minkowski space, due to the absence of a boost symmetry that preserves the entangling surface. Thus, the kink transform would have to be modified to work for wiggly cuts. The transformation of boundary observables off the null plane is quite complicated for wiggly cut CC flow. Thus, we also expect that regions of the entanglement wedge probed by such observables should be drastically modified, unlike the case where the entangling surface is a flat Rindler cut.

However, the wiggly-cut boundary transformation remains simple for observables restricted to the null plane. Thus one could try to formulate a version of the kink transform on Cauchy slices anchored to the null plane on the boundary and the RT surface in the bulk. Perhaps a non-trivial transformation of the entanglement wedge arises from the need to ensure that the kink transformed initial data be compatible with corner conditions at the junction where the slice meets the asymptotic boundary [166]. We leave this question to future work.

Other Probes of CC Flow

In Sec. 7.4, we provided evidence for kink transform/CC flow duality. The preservation of the left and right entanglement wedges under the kink transform ensures that all one-sided correlation functions transform as required. It would be interesting to consider two sided correlation functions. However, these do not change universally and are difficult to compute in general. In the bulk, this is manifested by the fact that the future and past wedges do not change simply and need to be solved for.

However, because of the shared role of the kink transform, we can take advantage of the modular toolkit for one-sided modular flow [142]. Let $|\widetilde{\psi}_s\rangle = \rho_A^{-i\,s}\,|\psi\rangle$ be a family of states generated by one-sided modular flow as discussed in Sec. 7.6. Then certain two sided correlation functions $\langle\widetilde{\psi}_s|\,O(x)\,O(y)\,|\widetilde{\psi}_s\rangle$ can be computed as follows.

Suppose $O(x)$ is an operator dual to a "heavy" bulk field with mass m such that $1/\ell_{\rm AdS} \ll m \ll 1/\ell_{\rm P}, 1/\ell_{\rm s}$. Correlation functions for such an operator can then be computed using the geodesic approximation,

$$\langle O(x)O(y)\rangle \approx \exp(-mL)\ , \tag{7.6.13}$$

where L is the length of the bulk geodesic connecting boundary points x and y. Now consider boundary points x and y such that there is a piecewise bulk geodesic of length $L(x,y)$ joining them in the spacetime dual to the state $|\psi\rangle$.

This kinked geodesic is required to pass through the RT surface of subregion A with a specific boost angle $2\pi s$ as seen in Fig. 7.6. (This is a fine-tuned condition on the set of points x, y.) Since single sided modular flow behaves locally as a boost at the RT surface, it straightens out the kinked geodesic such that it is now a true geodesic in the spacetime dual

to the state $|\widetilde{\psi}_s\rangle$. Thus, we have

$$\langle\widetilde{\psi}_s|\, O(x)\, O(y)\, |\widetilde{\psi}_s\rangle \approx \exp(-mL(x,y)) \ . \tag{7.6.14}$$

As discussed in Sec. 7.2, the CC flowed state can equivalently be thought of as the single sided modular flowed state $|\widetilde{\psi}_s\rangle$ on a kinematically transformed slice $\mathcal{C}_s$. Thus, the above rules can still be used to compute two sided correlation functions in the CC flowed state $|\psi_s\rangle = u_s\,|\psi\rangle$ as

$$\langle\psi_s|\, O(x_s)O(y)\, |\psi_s\rangle \approx \exp(-mL(x,y)) \ , \tag{7.6.15}$$

where x_s is the point related to x by the vacuum modular flow transformation.

We also note that the shock matching performed in Sec. 7.4 was a near boundary calculation. However, a bulk shock exists everywhere on the RT surface. One could solve for the position of the RT surface to further subleading orders and relate the bulk shock to the boundary stress tensor. This would yield a sequence of relations that the stress tensor must satisfy in order to be dual to the kink transform. In general these relations may be highly theory-dependent, but it would be interesting to see if some follow directly from CC flow or make interesting universal predictions for CC flow in holographic theories.

Higher Curvature Corrections

In Sec. 7.4, we argued that the bulk kink transform in a theory of Einstein gravity satisfies properties consistent with the boundary CC flow. However, this result is robust to the addition of higher curvature corrections in the bulk theory. The preservation of the two entanglement wedges, *i.e.*, Eq. (7.2.38), is a geometric fact that remains unchanged.[8]

Further, the matching of the stress tensor shock crucially depended on two ingredients. Firstly, Eq. (7.4.12), the holographic dictionary between the boundary stress tensor and the bulk metric perturbation and secondly, Eq. (7.4.15), the relation between the boundary entropy variation and the shape of the RT surface. Both of these relations are modified once higher curvature corrections are included [51, 92]. However, it follows generally from dimension counting arguments that

$$g^{(d)}_{ij} = \eta_1 \frac{16\pi G}{d} \langle T_{ij}\rangle \ , \tag{7.6.16}$$

$$X^A_{(d)} = -\eta_2 \frac{4G}{d}\, \frac{\delta S}{\delta X^A}\bigg|_{\mathcal{R}} \ , \tag{7.6.17}$$

where η_1 and η_2 are constants that depend on the higher curvature couplings. Using the first law of entanglement, it can be shown that in fact $\eta_1 = \eta_2$ [51, 92]. Hence, the boundary stress tensor shock obtained from the kink transform is robust to higher curvature corrections.

[8]Here we assume that the initial value formulation of Einstein gravity can be perturbatively adjusted to include higher curvature corrections despite the fact that a non-perturbative classical analysis of higher curvature theories is often problematic due to the Ostrogradsky instability [167].

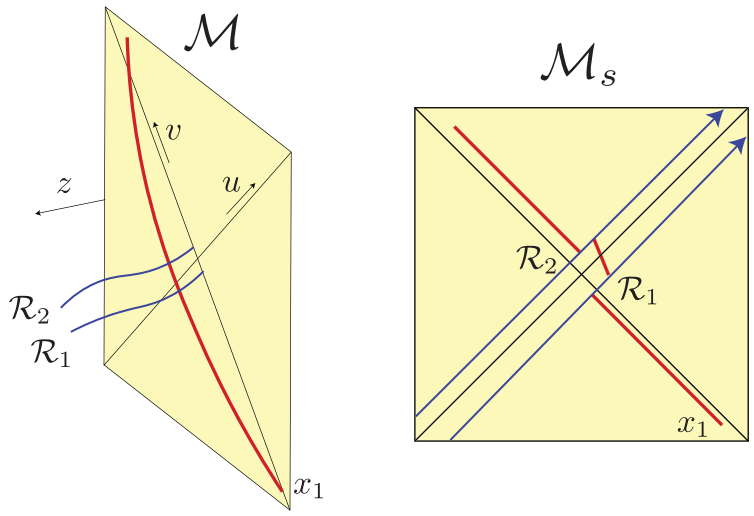

Figure 7.7: Holographic proofs. *Left:* Boundary causality is respected by the red curve that goes through the bulk in a spacetime $\mathcal{M}$; this is used in proving the ANEC. The RT surfaces $\mathcal{R}_1$ and $\mathcal{R}_2$ must be spacelike separated; this is used in proving the QNEC. *Right:* In the kink transformed spacetime $\mathcal{M}_s$ as $s \to \infty$, the QNEC follows from causality of the red curve, which only gets contributions from the Weyl shocks (blue) at $\mathcal{R}_1$ and $\mathcal{R}_2$, and the metric perturbation in the region between them.

Holographic proof of QNEC

A recent proof of the QNEC from the ANEC [18] considers CC flow for a subregion A on the null plane $u = 0$ with entangling surfaces $v = V_1(y)$ and $v = V_2(y)$ surrounding a given point p. From the transformation properties of the stress tensor under CC flow described in Sec. 7.2, $T_{vv} \to 0$ as $s \to \infty$. In addition, there are stress tensor shocks at ∂A, as described in Sec. 7.2, of weight $f(s)\frac{\delta S}{\delta V(y)}\Big|_{\psi,\partial A}$. In the limit $V_1(y) \to V_2(y)$, computing the ANEC in the CC flowed state, one obtains contributions from the stress tensor $T_{vv}(p)$ in subregion A, and a contribution proportional to $\frac{\delta^2 S}{\delta V(y_1)\delta V(y_2)}\Big|_{\psi,p}$ from the shocks. Positivity of the averaged null energy in the CC flowed state then implies the QNEC in the original state.

Prior to the QFT proofs, both the ANEC and QNEC had been proved holographically [168, 29]. The guiding principle behind both of these proofs was the fact that consistency of the holographic duality requires bulk causality to respect boundary causality as we demonstrate in Fig. 7.7. In the case of the ANEC, one considers an infinitely long curve connecting points on past null infinity to future null infinity through the bulk and demands that it respect boundary causality [168]. In the proof of the QNEC, one requires that curves joining the RT surfaces of subregions $v < V_1(y)$ and $v > V_2(y)$, denoted $\mathcal{R}_1$ and $\mathcal{R}_2$, respect bound-

ary causality [29, 137]. There are two contributions to the lightcone tilt of this bulk curve coming from the metric perturbation h_{vv} in the near boundary geometry, and the shape of the RT surface $X^\mu(y,z)$. By the holographic dictionary, these contributions can be related to the boundary stress tensor T_{vv} and the entropy variations $\frac{\delta S}{\delta V}$ as discussed in Sec. 7.4.

Now performing the kink transform removes the contribution coming from the shape of the RT surface and puts it into a time advance/delay coming from shocks in the bulk Weyl tensor that we compute in Appendix A.15. Considering the extended curve from past to future null infinity, we see that whether or not it respects boundary causality is determined entirely by the region between the entangling surfaces $\mathcal{R}_1$ and $\mathcal{R}_2$ since the bulk solution approaches the vacuum everywhere else in the limit $s \to \infty$. Requiring causality of the ANEC curve then results in the QNEC, making the connection to the boundary proof manifest.

Chapter 8

Quantum Information Bound on the Energy

8.1 Introduction

Semiclassical General Relativity allows for quantum matter while keeping the gravitational field classical, by coupling the metric to the expectation value of the stress tensor:

$$G_{ab} = 8\pi G\langle T_{ab}\rangle \ . \tag{8.1.1}$$

Since $\langle T_{ab}\rangle$ receives quantum contributions proportional to $\hbar$, this approximation can be organized as a perturbative expansion in $G\hbar$ and solved iteratively. This approach has proven to be quite useful, leading to the discovery of black hole thermodynamics and the associated information paradox.

Numerous theorems in General Relativity rely on the Null Energy Condition (NEC), which states that

$$T_{ab}k^a k^b \geq 0 \tag{8.1.2}$$

at every point in the spacetime, where k^a is any null vector. The NEC underlies the area theorems for event horizons [169] and for future holographic screens [170, 171], the focussing theorem [172], and Penrose's singularity theorem [173]. In other theorems, the stress tensor is assumed to obey even stronger conditions, which are nevertheless satisfied by known classical matter and radiation.

However, in relativistic quantum field theories (QFTs) such as the Standard Model, there are states in which $\langle T_{ab}\rangle$ violates the NEC in some regions of spacetime. Hence, none of the classical theorems mentioned above apply at the semiclassical level. The evaporation of a black hole, for example, is accompanied by violations of all of the above theorems. This is possible because the NEC is violated in the vicinity of the horizon.

Remarkably, there is considerable evidence that all of the above theorems admit a conjectural semiclassical extension. The key step to obtaining a viable proposal is to replace the

area of surfaces with their generalized entropy. Thus the area theorem becomes the Generalized Second Law (GSL) for event horizons [174, 175, 176] and for Q-screens [177]. The focussing theorem becomes the Quantum Focussing Conjecture (QFC) [178]; and Penrose's singularity theorem becomes Wall's Quantum Singularity Theorem [179].

Though these are conjectural statements about the semiclassical limit of quantum gravity, they can have interesting nongravitational limits. Some of these limit statements were already known, but others came as completely new and nontrivial results in QFT. The main example is the Quantum Null Energy Condition [178], which has since been rigorously proven within QFT, using a variety of methods [180, 181, 182]. Thus, the study of semiclassical gravity has had considerable impact in a seemingly unrelated arena.

The present work is inspired by these developments. We will study an important conjecture in classical General Relativity, the Penrose inequality [20]. This a relation between the area of certain marginally trapped surfaces μ in the spacetime and the total mass defined at spatial infinity [183]:

$$m \geq \sqrt{\frac{A[\mu]}{16\pi G^2}} \ . \tag{8.1.3}$$

The conjecture can be thought of as a generalization of the positive mass theorem [184]. For either statement, it is clearly essential that matter with negative energy be excluded. This can be implemented by assuming the dominant energy condition (DEC), that for any timelike future-directed vector t^a, $-T^a_{\ b}t^b$ is timelike and future-directed.

The Penrose inequality has not been proven and thus is not a theorem. But no counterexample to the conjecture is known. We will review the classical Penrose inequality in Sec. 8.2, where we provide both the reasoning motivating it, and a more careful formulation.

Since quantum matter can violate the NEC it can also violate the DEC, threatening the validity of the Penrose inequality. It is not immediately obvious that Eq. (8.1.3) fails, since the stress tensor in QFT cannot be dialed arbitrarily.

In fact, we find in Sec. 8.3 that Eq. (8.1.3) continues to be satisfied in a simple example of black hole formation and evaporation. However, we then provide an explicit counterexample to the classical Penrose inequality, by exploiting the thermal nature of the vacuum state near the horizon. When the thermal state is depleted, the vicinity of the horizon can contribute significant negative energy. This cancels an order one fraction of the black hole's mass, leading to a substantial violation of Eq. (8.1.3).

We are thus motivated, in Sec. 8.4, to propose a quantum-corrected version of the Penrose inequality. We introduce the relevant concepts of generalized entropy, quantum expansion, and quantum (marginally) trapped surfaces. We draw some lessons from the failure of the classical Penrose inequality in the semiclassical setting, and we formulate a Quantum Penrose Inequality (QPI).

In Sec. 8.5, we provide evidence for our proposal. We consider several interesting examples that could challenge the QPI, and we show that our proposal survives these tests. In Sec. 8.6, we discuss a number of alternative formulations of the QPI. We show why they are either

excluded or not ideal. In Sec. 8.7, we discuss the formulation of a QPI in asymptotically Anti-de Sitter spacetimes. This helps us identify subtleties that also affect the original QPI.

In Sec. 8.8, we discuss the classical and the nongravitational limits of the QPI.

Penrose's motivation in proposing Eq. (8.1.3) was as a test of the weak Cosmic Censorship Conjecture (CCC). In Sec. 8.9, we review this connection and the status of the CCC. We speculate that the Quantum Penrose Inequality could inform the formulation of a "quantum" CCC that accommodates the known, physically sensible violations of the classical CCC.

In Appendix A.16, we compute the expansion of outgoing null rays and the positions of classical and quantum marginally trapped surfaces for an evaporating Schwarzschild black hole. In Appendix A.17, we present a perturbative construction of Q-screens [177], which plays a role in our discussion of the Quantum Penrose Inequality in Anti-de Sitter spacetimes.

A brief summary of the main results of our investigation has appeared elsewhere [185].

8.2 Classical Penrose Inequality

In this section we describe the (classical) Penrose inequality; see Ref. [186] for a broader review.

Formulation

We formulate the classical Penrose inequality as follows:

Let m be the total mass of an asymptotically flat spacetime. Let μ be a trapped surface that has minimal area among all surfaces that enclose it, on some Cauchy surface that contains μ. Then

$$m \geq \sqrt{\frac{A[\mu]}{16\pi G^2}} \ . \tag{8.2.1}$$

Next, we provide detailed definitions and explanations of the terms appearing in this formulation.

Let (M, g_{ab}) be a connected Lorentzian spacetime with metric. Let μ be a codimension $1+1$ compact spacelike submanifold (a "surface"). [1] Let $\theta_\pm$ be the expansion of the future-directed light-rays emanating orthogonally from μ to either side. If $\theta_+ \leq 0$ and $\theta_- \leq 0$ then μ is called *trapped*. If $\theta_+ = 0$ and $\theta_- \leq 0$ then μ is *marginally trapped*.

Now let (M, g_{ab}) be in addition asymptotically flat. Note that we do not require μ to be connected; for example in a spacetime where multiple black holes are forming, μ could be the union of connected marginally trapped surfaces inside some or all of them.

Suppose that the surface μ has an "outer wedge" O_W that contains a single asymptotic region. By this we mean that μ forms the only boundary of any Cauchy surface of a globally hyperbolic region of space O_W that (in the "unphysical spacetime" or Penrose diagram)

[1] In the remainder of this paper we will specialize to 3+1 dimensional spacetime, so that μ will be a 2-dimensional surface. Generalization to higher dimensions is trivial.

contains a single copy of spatial infinity, i_0. This will be the case for trapped surfaces in a spacetime with a single asymptotic region. In the case of "two-sided" black hole solutions, it will hold if μ is homologous[2] to a horizon (with either choice of side), but not if μ is contractible. We will be interested in bounding the mass at spatial infinity [183] from below.

Finally, we assume that there exists a Cauchy surface Σ of O_W on which μ is the minimal area surface homologous to large spheres near i_0 (or in the AdS case, homologous to the boundary sphere) [187]. The purpose of this set of assumptions will become clear as we turn to presenting a heuristic argument that the Penrose Inequality should hold for μ.

Heuristic Argument

The Penrose Inequality was originally intended as a test of cosmic censorship, which guarantees that an asymptotically flat spacetime with regular initial conditions will be strongly asymptotically predictable [172]. If this latter property holds, then a compelling argument can be given that the Penrose inequality must hold; thus, any regular initial data set that violates the Penrose inequality would likely exclude cosmic censorship. We now present the argument.

Roughly speaking, strong asymptotic predictability establishes the existence of $\tilde{V}$, a globally hyperbolic open subset of M that contains any black hole horizons and their exterior, $\tilde{V} \supset J^-(\bar{\mathcal{I}}^+)$. (See Ref. [172] for more details.) The *black hole region* is $B \equiv M - J^-(\mathcal{I}^+)$. The black hole *event horizon* is its boundary $\dot{B}$.

Suppose that

$$R_{ab}k^a k^b \geq 0 \ , \tag{8.2.2}$$

as would be the case if Einstein's equations hold with matter satisfying the Null Energy Condition. Then any trapped or marginally trapped surface μ must lie in the black hole region:

$$\mu \subset B \ . \tag{8.2.3}$$

For a proof, see Propositions 12.2.2 in Ref. [172]. The key technical assumption is that M be strongly asymptotically predictable.[3]

Let $H = \dot{B} \cup \Sigma$ be the slice of the black hole event horizon (possibly with multiple disconnected components), on the Cauchy surface Σ of O_W. Since μ has minimal area on Σ, it follows that the horizon must be at least as large:[4]

$$A[H] \geq A[\mu] \ . \tag{8.2.4}$$

[2]Two cycles (closed submanifolds which are not boundaries of any other submanifolds) are said to be homologous, or equivalently, belong to the same homology class, if they can be continuously deformed into each other.

[3]The same property, $\nu \subset B$, follows from Proposition 12.2.3 in Ref. Wald for another class of surfaces called *outer trapped*. These would form an alternate starting point from which the classical and quantum Penrose conjectures could be developed along the same lines as we do here for trapped surfaces.

[4]Instead of assuming that μ has minimal area on *some* Cauchy slice of O_W, an alternative way of handling this issue is to replace $A[\mu]$ with the minimal area of all surfaces enclosing μ on a *given* initial Cauchy slice [186]. Verifying this assumption does not require knowledge of more than the initial slice.

The Null Curvature Condition, Eq. (8.2.2), and strong asymptotic predictability imply that the area of the event horizon cannot decrease with time [169]. Let $H' = \dot{B} \cup \Sigma'$, where Σ' is a Cauchy surface to the future of Σ. Then

$$A[H'] \geq A[H] \ . \tag{8.2.5}$$

Physically, it is reasonable to assume that regular initial data will eventually settle down to a Kerr black hole. (In four dimensions, this follows from the assumption of late-time stationarity, by the Israel-Hawking-Carter theorems [188].) Letting H' be a slice of the horizon at this late time, the formula for the area of a Kerr black hole implies that

$$16\pi G^2 m_{\text{Kerr}}^2 \geq A[H'] \ . \tag{8.2.6}$$

The spacetime will not be exactly Kerr, however. One expects that massive fields will have fallen into the black hole, but there may be massless fields that propagate to future null infinity. Because this radiation becomes dilute and well separated from the black hole, gravitational binding energy will be negligible. Hence the ADM mass, m, will be given by the sum

$$m = m_{\text{Kerr}} + m_{\text{rad}} \geq m_{\text{Kerr}} \ . \tag{8.2.7}$$

Combining the previous four inequalities, we obtain the Penrose conjecture, Eq. (8.2.1).

We would like to add a second, somewhat independent heuristic argument for Eq. (8.2.1). A *future holographic screen* is a hypersurface foliated by marginally trapped surfaces called *leaves* [189, 190, 170]. Assuming the Null Energy Condition, the area of the leaves increases monotonically along this foliation [170, 171]. In the spherically symmetric case, the screen eventually asymptotes to the event horizon (from the interior), so its final area will be equal to the late time event horizon area. Thus the screen area theorem implies the Penrose inequality in this case. More generally, given a marginally trapped surface μ, a future holographic screen can be constructed at least in a neighborhood. The Penrose inequality would follow from the stronger assumption that there exists a future holographic screen that interpolates from μ to the late-time event horizon, as in the spherical case.

8.3 Violation by Quantum Effects

In this section, we will show that there is a need for a quantum generalization of the classical Penrose inequality (CPI). We will construct an explicit counterexample that is based on a Boulware-like state outside a Schwarzschild black hole. It violates the CPI by a substantial, classical amount.

This will be a counterexample to the CPI in the same sense as black hole evaporation is a counterexample to Hawking's area theorem: we identify a physically allowed state in which a key assumption of the classical statement, the Null Energy Condition, does not hold, and we verify that the conclusion fails as well.

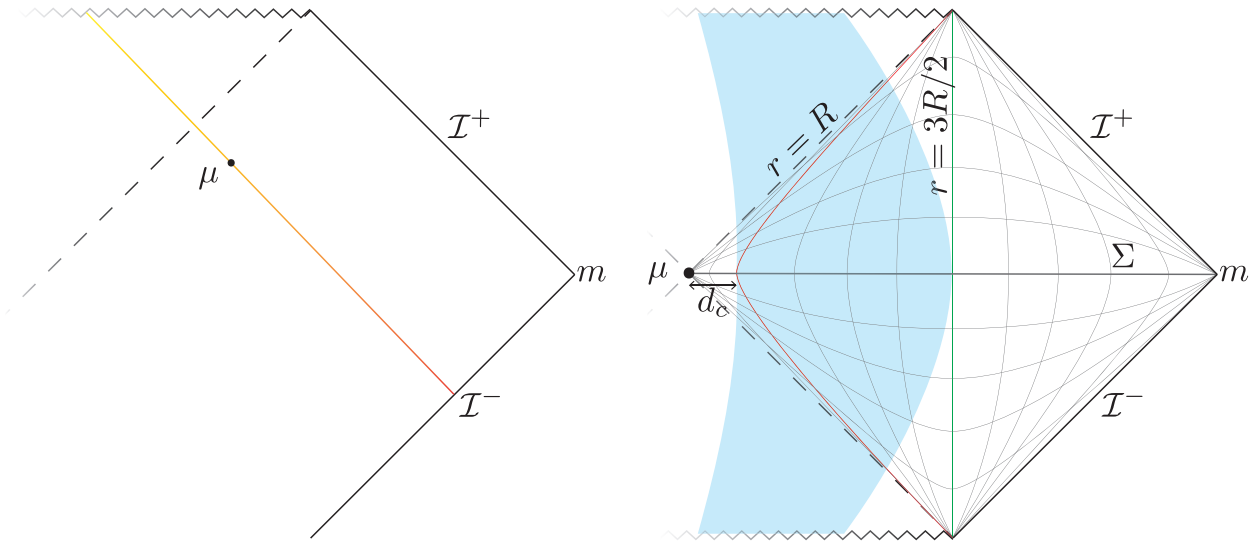

Figure 8.1: Left: A null shell collapsing in asymptotically flat spacetime. The classically marginally trapped surface μ is slightly outside of the event horizon due to the evaporation. It is not clear that this example violates the CPI. Right: initial data that violates the classical Penrose inequality. Here μ is the bifurcation surface of the Schwarzschild (Kruskal) solution. Inside a proper distance d_c, the state is the Hartle-Hawking vacuum. Outside of d_c,it becomes the Boulware vacuum, which has negative energy in the near-horizon zone (blue strip). This lowers the mass at infinity by an $O(1)$ fraction compared to a classical black hole.

However, before we turn to our counterexample, it is worth noting that no obvious violation of the CPI arises in the "normal" formation and evaporation of a black hole in the Unruh state. This is interesting, because in this setting the Null Energy Condition is already violated, and other theorems like the area theorem or the focussing theorem do fail. In order to have full control and exclude transient effects, let us consider the collapse of a null shell of mass m; see Fig. 8.1. Then by causality, there are no corrections to the classical solution on the shell and to its past, where the spacetime is a portion of Minkowski space. In particular, the marginally trapped surface on the shell will have the same area as in the classical case, and the CPI will be saturated. (The fact that the event horizon is inside of this surface is irrelevant.) At later times, we expect the apparent horizon area to decrease. Since the mass at infinity does not change during evaporation, the CPI will remain satisfied.

We do not claim that the CPI will hold for all black holes formed from collapse; and even in the above example, its validity may rely on idealizations, such as treating the collapsing null shell as infinitely thin and stable. But we would like to exhibit a situation where the CPI is definitely violated; in order to do this, we will consider a somewhat more artificial (but certainly valid) quantum state.

To demonstrate a violation of the classical PI by quantum effects, we now consider a Boulware-like state [191] of a massless scalar field, on one side of a maximally extended Schwarzschild black hole, at the time-symmetric slice; see Fig. 8.1. The Boulware vacuum is analogous to the Rindler vacuum. It corresponds to vanishing occupation number of the modes with support strictly outside the event horizon. This will contribute some negative energy outside of the black hole, in the near-horizon region $R < r < 3R/2$. Far from the black hole, the stress tensor vanishes in the Boulware vacuum.

Note that the classical Penrose inequality, applied to the bifurcation surface, is classically saturated. (That is, it is saturated if the stress tensor vanishes everywhere outside the black hole.) Thus, any net negative energy in the exterior will lead to a violation of Eq. (8.2.1).

The local stress tensor diverges in the Boulware vacuum as the horizon is approached [191, 192]. We regulate this divergence by building wavepackets with support strictly outside of a sphere H_c at proper distance $d_c > 0$ from the horizon (in this case, from the bifurcation surface). For full control of the semiclassical expansion, we choose

$$l_P \ll d_c \ll R \ . \tag{8.3.1}$$

Roughly speaking, this yields a Hartle-Hawking-like state (vanishing stress tensor) inside of H_c, and a Boulware-like state outside of H_c.

Integration of the QFT stress tensor computed in Ref. [192], outside the regulator sphere H_c, yields a QFT contribution to the energy at infinity of order $-(l_P/d_c)^2 M$, where $M = R/2G$ is the mass of the black hole [185]. Here we will go further; instead of naively gluing to QFT states across a surface (which is does not generally yield an allowed QFT state), we consider junction effects at H_c. Positivity of the energy for infalling observers requires some positive energy near H_c, which we wish to estimate and show to be negligible.

For this purpose it will be useful to analyze the problem mode by mode. This will allow us to distinguish between two cutoffs that we can freely choose: the angular momentum of the included QFT modes, and d_c. Establishing a small hierarchy between these cutoffs will give us a control parameter $1/n_{\rm node} \ll 1$, by which the positive energy at H_c is suppressed at infinity, relative to the negative contribution.

We will focus on the most relevant modes in the near-horizon zone, which have occupation number of order one in the thermal ensemble corresponding to the Hartle-Hawking state. These modes have the property that any wavepacket constructed from them has characteristic wavelength comparable to its distance from the horizon. Moreover, increasing the occupation number of the mode by 1 increases the energy at infinity by $\hbar/R$.

This set of modes includes s-waves as well as modes with nonzero angular momentum. Here we will use $\ell = 0, 1, \ldots$ for the angular momentum quantum number. The number of modes in the thermal atmosphere can be estimated from the number of nodes in a strictly outgoing Rindler mode in an interval beginning at proper distance d_c from the horizon and ending at a distance R (for the spherical modes, which we approximate as propagating freely) or $R/(\ell+1)$ (for the modes with angular momentum, which we approximate as being

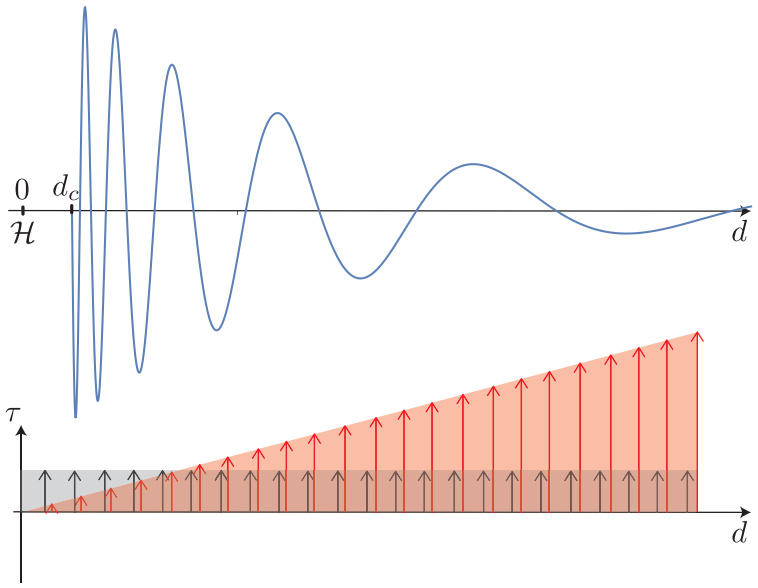

Figure 8.2: A typical wavepacket mode in the thermal atmosphere of the black hole, regulated to have support outside a sphere a proper distance d_c outside of the horizon. The classical Penrose inequality is violated in a Boulware-like state in which such modes have zero occupation number and negative energy. In a local inertial frame (black Killing vector field, ∂_τ, where τ is proper time), a large fraction of their energy is concentrated near the cutoff d_c. The total energy must appear positive in this frame; this can be satisfied by adding a comparable amount of positive energy inside of d_c. To an asymptotic observer (red Killing vector field, ∂_t), the negative energy is spread evenly over the mode, due to the greater redshift near the horizon. Thus the positive energy beyond the cutoff has a negligible effect on the ADM mass.

reflected by an angular momentum barrier). See Fig. 8.2. Hence there are

$$n_\ell = (2\ell+1)\log\left(\frac{R/(\ell+1)}{d_c}\right) \tag{8.3.2}$$

linearly independent modes with angular momentum ℓ.

In the Hartle-Hawking state, these modes are all thermally excited with $O(1)$ occupation numbers; this corresponds to vanishing stress tensor near the horizon. In the Boulware-like state, the modes are unoccupied. This corresponds to a negative stress tensor; it contributes an energy at infinity of order $-\hbar/R$, per mode. We choose a cutoff $\ell_{\max}$ on the angular momentum such that the angular momentum barrier is somewhat outside the short distance cutoff d_c:

$$\log\log\left(\frac{R/(\ell_{\max}+1)}{d_c}\right) \sim O(1)\ , \tag{8.3.3}$$

where the second log enforces a small hierarchy whose purpose will become clear below. From the previous two equations, the total number of unoccupied modes is

$$n_{\rm total} \equiv \sum_{\ell=0}^{\ell_{\rm max}} n_\ell \sim \frac{R^2}{d_c^2} \ . \tag{8.3.4}$$

Thus the total energy at infinity of the quantum field will be

$$E_{\rm neg} \sim -\frac{\hbar}{R} n_{\rm total} \sim -\alpha M \ , \tag{8.3.5}$$

where

$$\alpha = \frac{l_P^2}{d_c^2} \ . \tag{8.3.6}$$

The presence of a substantial amount of negative energy outside the black hole may seem suspect. However, we note that our construction cannot achieve vanishing or negative total ADM mass. Since the black hole contributes M, the total mass is $(1-\alpha)M$. Making this negative would require taking $d_c \lesssim l_P$, in conflict with Eq. (8.3.1), and so would take us outside of the semi-classical expansion. Moreover, our result is consistent with positive total matter energy in an appropriate neighborhood of the horizon. This is important since the spacetime can be treated as approximately flat on a distance scale $d_c \ll d_{\rm flat} \ll R$.

To see this, we note that the wavepackets we study have approximately constant Killing energy per cycle, where a cycle denotes the portion of a wavepacket between two nodes. See Fig. 8.2. The local proper wavelength of a given mode grows as the distance from the horizon, but this is precisely cancelled by the decreasing redshift. Thus from the viewpoint of infinity, each cycle of each mode contributes an ADM energy (per occupation number) of $\hbar/(R n_{\rm node})$, where

$$n_{\rm node}(\ell) \sim \log\left(\frac{R/(\ell+1)}{d_c}\right) \tag{8.3.7}$$

is the number of nodes or cycles in the wavepacket.

In a local inertial frame, on the other hand, there is no redshift effect. Yet, the proper wavelength grows exponentially away from the horizon, roughly doubling with every cycle. Thus an $O(1)$ fraction of the local energy of a mode is contained in the first phase cycle. In the Boulware-like state, this is the negative energy that must be cancelled. To have positive energy in the local frame, it suffices to have compensating positive energy just for this first cycle. The positive energy can be localized, for example, just below d_c.

This positive energy will partially cancel the negative ADM energy of the quantum state, Eq. (8.3.5). But because all cycles of the wavepacket contribute equally to the Killing energy, the correction is parametrically small, of order $|E_{\rm neg}|/n_{\rm node} \ll |E_{\rm neg}|$. In practice, $n_{\rm node}$ of order a few suffices, so we will not update Eq. (8.3.6). The purpose of the second log in Eq. (8.3.3) was to chose the angular momentum cutoff $\ell_{\rm max}$ so as to achieve $n_{\rm node} \sim$ a few, for all modes involved in the construction.

Finally, we note that the location and area of the marginally trapped surface do not receive large enough corrections to rescue the classical Penrose inequality. The bifurcation surface remains marginally trapped when we pass from the classical treatment to the Hartle-Hawking state, since the stress tensor vanishes there. Our construction keeps the Hartle-Hawking state near the bifurcation surface, up to corrections that can be suppressed arbitrarily by dialing $n_{\text{node}} \gg 1$.

To summarize, one can reduce the mass at infinity from M (in the Unruh state) to $(1-\alpha)M$ in the Boulware-like state. Since we require that $l_P \ll d_c$ for control, this correction is parametrically small, $\alpha \ll 1$. But since the Penrose inequality is saturated classically for a Schwarzschild black hole, our example violates it.

Moreover, the violation is substantial in the sense that it is not $O(\hbar)$ but $O(1)$. The contribution from each mode is $O(\hbar)$; but the number of available modes in the thermal atmosphere, at fixed control parameter l_P/d_c, is $n_{\text{total}} \sim O(\hbar^{-1})$. Thus, the negative energy of the quantum fields can cancel off an $O(1)$ fraction of the black hole's classical mass.

8.4 Quantum Penrose Inequality

In this section, we will formulate the Quantum Penrose Inequality (QPI). In Sec. 8.4, we review various concepts necessary for the quantum generalization of classical statements involving area and null expansion. In Sec. 8.4, we draw some conclusions from the failure of the classical Penrose inequality. In Sec. 8.4, we formulate our proposal for the QPI.

Generalized Entropy and Quantum Expansion

We begin by introducing the notion of generalized entropy and its main properties. We then use the generalized entropy to define certain quantum generalizations of various geometric quantities, necessary for formulating the Quantum Penrose Inequality; see [178] for more details.

The *generalized entropy* S_{gen}, was first introduced by Bekenstein [174, 175] as the total entropy of a system consisting of a black hole and its exterior on a given time slice. The definition can be extended to apply not only to the horizon of a black hole, but to any Cauchy-splitting surface σ:

$$S_{\text{gen}} \equiv \frac{A[\sigma]}{4G\hbar} + S_{\text{out}} + \dots \ , \tag{8.4.1}$$

where $A[\sigma]$ is the area of σ, and

$$S_{\text{out}} = -\text{Tr}\rho_{\text{out}} \log \rho_{\text{out}} \tag{8.4.2}$$

is the von Neumann entropy of the state of the quantum fields, restricted to one side of σ:

$$\rho_{\text{out}} = \text{Tr}_{\overline{\text{out}}}\, \rho \ . \tag{8.4.3}$$

Here, the state ρ is the global quantum state, and the trace is over the complement region, which we define as $\overline{\text{out}}$.

The von Neumann entropy S_{out} quantifies the amount of entanglement in the vacuum across σ, and as such, has divergences coming from short-distance entanglement. The leading divergence is given by A/ϵ^2, where ϵ is a short-distance cutoff. However, we can think of the geometric term in Eq. (8.4.1) as a counterterm. The dots indicate the presence of subleading divergences in S_{out} which come with their own geometric counterterms. It is expected that the divergences coming from the renormalization of G and from short-distance entanglement will cancel out [178], so as to keep S_{gen} a finite and well-defined quantity.

One can interpret S_{gen} in two distinct ways. Following the original motivation, one can view the area-term as a (large) "correction" to the entropy of quantum fields. Alternatively, we can define a quantum-corrected area of the surface σ:

$$A_Q[\sigma] \equiv A[\sigma] + 4G\hbar S_{\text{out}} + \dots \ , \tag{8.4.4}$$

in a semiclassical expansion in $G\hbar$. Hence, one can use the notion of generalized entropy to incorporate quantum effects into certain geometrical objects that derive from the area of surfaces.

One example is the notion of *quantum expansion*. Recall, the classical expansion of a surface σ at a point $y \in \sigma$ is defined as the trace of the null extrinsic curvature at y. Equivalently, one can define the classical expansion as a functional derivative,

$$\theta[\sigma; y] = \frac{1}{\sqrt{h(y)}} \frac{\delta A[V]}{\delta V(y)} \ , \tag{8.4.5}$$

where h represents the area element of the metric restricted to σ, inserted to ensure that the functional derivative is taken per unit proper area, not coordinate area. The function $V(y)$ is used to specify the affine location of σ and nearby surfaces along a congruence of null geodesics orthogonal to σ. The above definition of the classical expansion is needlessly complicated, in that it invokes the entire surface σ, even though θ depends only on its local extrinsic curvature at y. However, this definition naturally generalizes to the quantum expansion, Θ, which does depend on all of σ:

$$\Theta[\sigma; y] \equiv \frac{4G\hbar}{\sqrt{h(y)}} \frac{\delta S_{\text{gen}}[V]}{\delta V(y)} \ . \tag{8.4.6}$$

As in the classical case, we can use the notion of expansion to define certain types of surfaces (see Sec. 8.2). Let $\Theta_\pm$ be the quantum expansion of the future-directed light-rays orthogonal to a surface μ_Q. (As before, we take the + label to refer to the direction of spatial infinity.) If $\Theta_+ \leq 0$ ($\Theta_+ = 0$) and $\Theta_- \leq 0$, then we call μ_Q a *quantum (marginally) trapped surface*.

Quantum trapped surfaces, in the semiclassical setting, have some of the properties obeyed by trapped surfaces in the classical setting. For example, trapped surfaces cannot

lie outside the black hole, assuming weak cosmic censorship and the Null Energy Condition. When the NEC is violated, they can; however, quantum trapped surfaces must still lie inside or on the horizon [179] (still assuming weak cosmic censorship). This will prove to be important for our formulation of the quantum Penrose inequality.

A *quantum future holographic screen*, or Q-screen, is a hypersurface foliated by quantum marginally trapped surfaces. Assuming the quantum focussing conjecture [178], Q-screens obey a Generalized Second Law [177].

Lessons From the Counterexample

The failure of the classical PI in the presence of quantum matter (Sec. 8.3) illustrates the need for a Quantum Penrose Inequality. It also motivates some of the choices we will make below.

Let us distinguish two different time-scales: the time for the negative energy of the Boulware-like state to enter the black hole, and the evaporation time. The former is of order the scrambling time $\Delta t_s \sim R \log(R/l_P)$. The latter is much greater, of order $R^3/G\hbar$.

On the shorter time-scale, the process results in an outcome very similar to that invoked in motivating the classical Penrose inequality: a Kerr black hole with area $A_{\rm late}$ and no further evolution. That is, we neglect evaporation since it occurs on a much greater timescale; and by construction, no matter that will ever enter the black hole. Thus, the mass should obey $16\pi G^2 m^2 \geq A_{\rm late}$.

The key difference to the classical case is that the "late" area need not be greater than the area of trapped surfaces at earlier times; indeed our counterexample shows that it will not be. However, we know that the Generalized Second Law (GSL) takes the place of the area theorem in this setting. Thus, we expect that the generalized entropy of earlier quantum trapped surfaces should be less than $A_{\rm late}/4G\hbar$. And so, the generalized entropy of quantum trapped surfaces should replace the area of trapped surfaces when we replace the classical by a Quantum Penrose Inequality.

This argument is based on the GSL for the event horizon, and so involves an intermediate step where one argues that the generalized entropy of a quantum marginally trapped surfaces inside the black hole will not be greater than that of the event horizon. To avoid this step, we can generalize the second heuristic argument for the classical Penrose inequality, which was based on the area theorem for future holographic screens. Q-screens obey a GSL that interpolates directly between different marginally quantum trapped surfaces. If a suitable Q-screen connects μ_Q to the late-time event horizon, this establishes a Quantum Penrose Inequality. Of course this is far from a trivial assumption; our goal here was only to gain some intuition.

In the above heuristic arguments, it was important that the late-time generalized entropy should be given just by $A_{\rm late}$, i.e., that no entropy remains outside of the black hole. However, this will not be the case in general examples. This will motivate our choice, below, that the generalized entropy entering the Quantum Penrose Inequality should be evaluated on slices that remain inside the black hole. We will discuss this important issue further in Sec. 8.6.

Formulation

We will now obtain a Quantum Penrose Inequality from the classical PI, in three steps. First, we replace the area with generalized entropy in Eq. (8.2.1):

$$A \to 4G\hbar\, S_{\rm gen} \equiv A + 4G\hbar\, S_{\rm out} \ . \tag{8.4.7}$$

Thus we propose an inequality of the form

$$m \geq \sqrt{\frac{\hbar S_{\rm gen}}{4\pi G}} \ . \tag{8.4.8}$$

Secondly, we must specify the surfaces to which the inequality can be applied. In the classical case, a surface μ has to be trapped for the Penrose inequality to apply, corresponding to criteria satisfied by the classical expansion. For the QPI, it is natural to apply the same criteria to the quantum expansion:

$$\theta \to \Theta \ . \tag{8.4.9}$$

Thus in Eq. (8.4.8), $S_{\rm gen}$ is the generalized entropy of any surface μ_Q that is quantum trapped. We expect that the most interesting bounds will obtain when μ_Q is quantum marginally trapped, and we will only consider this case in all examples below.

Next, we must specify on which achronal hypersurface the generalized entropy appearing in Eq. (8.4.8) should be computed. As we will explain in Sec. 8.6, this *cannot* be chosen to be a Cauchy surface of the outer wedge. Instead, we will propose that this hypersurface should be entirely contained in the "black hole region" $B \equiv M - J^-(\mathcal{I}^+)$, i.e., inside or on the horizon.

More precisely, we require that $S_{\rm gen}$ should be evaluated on the "future portion" of the boundary of the outer wedge,

$$L(\mu_Q) \equiv \dot{O}_W(\mu_Q) - I^-(O_W(\mu_Q)) \ . \tag{8.4.10}$$

See Fig. 8.3. L is generated by the congruence of future-directed outgoing null geodesics orthogonal to μ_Q [172, 193]. Their initial quantum expansion is $\Theta_+ = 0$ by construction, so assuming the QFC [178], $\Theta_+ \leq 0$ everywhere on L. Hence L will be a (quantum) lightsheet of μ_Q. Assuming an appropriate version of weak cosmic censorship, L will terminate on the singularity inside the black hole. (Strictly, in order to remain in the semi-classical regime, one should terminate L slightly earlier, resulting in a second area term that can be made small by approaching the singularity.)

Note that the surface μ_Q must be quantum trapped with respect to L; it need not be quantum trapped with respect to any other hypersurface, such as a Cauchy surface of $O_W(\mu_Q)$. To find a suitable μ_Q, consider a null hypersurface N inside the black hole, for example the boundary of the future of an event q inside the black hole; see Fig. 8.3. Typically the area of N will increase near q and later decrease towards the singularity. Hence the area will have a maximum on some cut of N, and the generalized entropy of cuts of N (computed

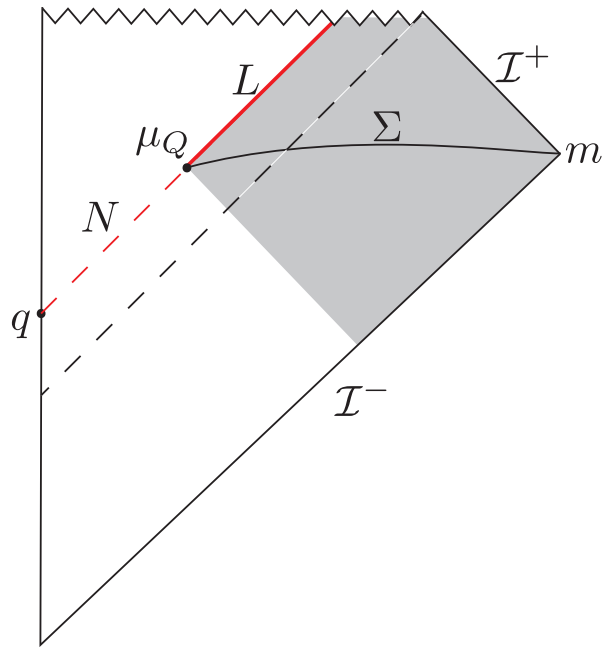

Figure 8.3: The Quantum Penrose Inequality bounds the mass at infinity in terms of the generalized entropy of a quantum marginally trapped surface μ_Q. The generalized entropy must be evaluated on the lightsheet L (red line), *not* on a Cauchy surface Σ of the outer wedge $O_W[\mu_Q]$ (shaded region).

with respect to the future of the cuts on N) will have a maximum on some nearby cut. This cut will be a suitable quantum marginally trapped surface μ_Q, and later cuts will also be quantum trapped.

Finally, we must impose a requirement analogous to the minimum area condition imposed on μ in the classical case. This condition demanded that there exist a Cauchy surface of O_W on which no surface enclosing μ_Q has area less than μ_Q. Here, we will instead consider the generalized entropy of any surface ν enclosing μ_Q, computed on the boundary of the future of the outer wedge of ν. For the QPI to apply to a quantum trapped surface μ_Q, we demand that there exist a Cauchy surface of $O_W[\mu_Q]$ on which no enclosing surface ν satisfies $S_{\rm gen}[\dot{O}_W(\nu) - I^-(O_W(\nu))] < S_{\rm gen}[L(\mu_Q)]$.

To summarize, we propose that the mass at spatial infinity of an asymptotically flat spacetime satisfies the Quantum Penrose Inequality

$$m \geq \sqrt{\frac{\hbar S_{\rm gen}[L(\mu_Q)]}{4\pi G}} \,, \tag{8.4.11}$$

where $S_{\rm gen}$ is computed on the future-outgoing lightsheet of μ_Q, and μ_Q is any quantum trapped surface homologous to spatial infinity that has minimal generalized entropy on some Cauchy surface of its outer wedge, in the sense described above.

We close by discussing a subtlety that introduces a small uncertainty in the formulation of the QPI. In Eq. (8.4.11), we used the classical functional relation between the area and mass of Schwarzschild black holes; we merely replaced the area with the generalized entropy. In fact, there will be a field-content-dependent quantum correction to the functional relation itself. However, this correction is small compared to the difference between our QPI and the classical Penrose inequality.

This is easier to discuss in asymptotically Anti-de Sitter (AdS) space, where the Schwarzschild black hole can be in thermal equilibrium. Therefore, we will revisit the issue in more detail in Sec. 8.7. In general, the black hole exterior will have nonzero energy density in equilibrium. This is a kind of Casimir energy associated with the potential well provided by the near horizon zone. It contributes to the total mass at infinity; but since it stays outside the black hole, it will not contribute to $S_{\rm gen}[L(\mu_Q)]$.

By dimensional analysis, one expects each field theory degree of freedom to contribute an amount of order $\hbar/R$ to this Casimir energy. In Eq. (8.4.11), this is equivalent to changing the area or generalized entropy by $O(c)$, where c is the number of matter quantum fields. For large black holes in AdS, it is possible to determine this correction and include it in the QPI (see Sec. 8.7). In general, however, we are presently unable to determine it.

Since $S_{\rm gen}$ is $O(\hbar^{-1})$ and c is $O(1)$, the undetermined Casimir term in Eq. (8.4.11) is subleading. But naively, it is comparable to the refinement we introduced in passing from the classical Penrose inequality to the QPI. However, the Casimir correction cannot be enhanced by factors proportional to $\hbar^{-1}$. Thus it is much smaller than the violations of the classical Penrose inequality that were exhibited in Sec. 8.3. Because of the $\hbar^{-1}$ enhancement, Eq. (8.2.1) can be violated by a *classical* amount through quantum effects. Correspondingly, a successful QPI cannot be a small modification of the classical Penrose inequality. Indeed, it is not: as we shall demonstrate in the next section, the counterexample to Eq. (8.2.1) is evaded by Eq. (8.4.11). In this and many other interesting examples, the Casimir correction is small compared to the difference between Eq. (8.2.1) and Eq. (8.4.11).

8.5 Evidence for the Quantum Penrose Inequality

We will now analyze the validity of our proposal in a number of examples. In the process, we will gain some intuition about the key quantity that appears in it: $S_{\rm gen}[L]$, the generalized entropy of the future-outgoing lightsheet L of a quantum marginally trapped surface μ_Q.

Black Hole in the Unruh State

As a first example, consider a black hole formed from collapse of a null shell; see Fig. 8.4. This is the example we analyzed in the context of the classical Penrose inequality, at the beginning of Sec. 8.3. We showed there that the CPI is saturated, since the area of the classically marginally trapped surface μ immediately after the collapse satisfies

$$16\pi G^2 m^2 = A[\mu] \ . \tag{8.5.1}$$

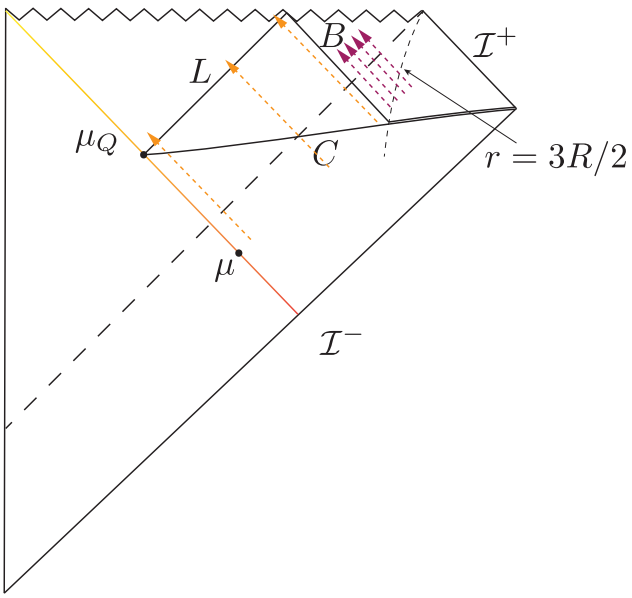

Figure 8.4: Black hole formed from the collapse of a null shell (orange line). The classically marginally trapped surface μ lies a Planckian distance outside of the event horizon. The quantum marginally trapped surface μ_Q lies a Planckian distance inside the horizon. The lightsheet $L(\mu_Q)$ captures $\sim \log(R/l_P)$ infalling Hawking modes (orange dashed lines); in the Unruh states these modes are unoccupied and so contribute negative entropy on L, compared to the Hartle-Hawking state. L ends at the singularity and does not encounter any later infalling modes (purple dashed lines). The entropy on L can also be computed using the mutual information, $S_L = S_C - S_B + I(L:B)$.

Here we are interested in a quantum marginally trapped surface with largest generalized entropy, for which the QPI provides the greatest lower bound on the mass. The area of (quantum) trapped surfaces decreases along with the event horizon, and the contribution from the entropy term is approximately time-independent. Hence we will again choose the earliest possible surface μ_Q, right after the collapse.

The quantum marginally trapped surface μ_Q must lie inside the event horizon [179], whereas μ lies outside. Therefore

$$A[\mu_Q] < A[\mu] \ . \tag{8.5.2}$$

We now turn to estimating $S_{\rm gen}[L]$. Strictly, $S_{\rm gen}[L]$ should be computed from the quantum state on a global Cauchy surface Σ that contains L. One would first compute the (divergent) field theory entropy $S[L]$ by tracing over the complement of L on Σ. One would then add the gravitational counterterms whose leading contribution is $A[\mu_Q]$. Locally, in a vacuum state, one expects $S_{\rm gen} \approx A[\mu_Q]/4G\hbar$, where G is the "infrared" value of Newton's constant that would be observed at large distances.

However, the state on L is not a standard vacuum state. L nearly coincides with the black hole horizon for a time $t \ll \Delta t_s$, where Δt_s is the scrambling time. The vacuum state on the horizon is the Hartle-Hawking state, which contains ingoing radiation. The ingoing radiation on L is entangled with modes on the other side of L. This contribution must be canceled by the counterterm so as to obtain $S_{\rm gen} \approx A[\mu_Q]/4G\hbar$ in the Hartle-Hawking state.

The actual state we consider here is the Unruh state, which does not have this ingoing radiation. As a result, the lightsheet will contain less entropy than in the vacuum state. Thus

$$S_{\rm gen}[L] < \frac{A[\mu_Q]}{4G\hbar} \ . \tag{8.5.3}$$

Combined with Eqs. (8.5.1) and (8.5.2) this establishes that the QPI is satisfied (and not saturated) in this example.

We would like to go further and estimate the "gap" by which the QPI fails to be saturated in this example,

$$\Delta \equiv \frac{4\pi G}{\hbar} m^2 - S_{\rm gen}[L] \ . \tag{8.5.4}$$

We will be interested only in the order of magnitude of this gap and so will make a number of approximations. We refer to Sec. 8.3 for notation and conventions.

First, we will assume that the higher angular momentum modes, $\ell > 0$, in the near-horizon zone completely reflect off of the angular momentum barrier and so will behave as if they were in the Hartle-Hawking state. In this approximation, the Unruh state differs only through the spherical ($\ell = 0$) modes, which we treat as having no angular momentum barrier at all. We also assume that the ingoing and outgoing s-waves do not interact.

A Planck sized, radially outgoing wavepacket starting a Planck distance from the horizon will be redshifted in such a way that its proper distance from the horizon remains comparable to its proper wavelength, while it propagates in the near horizon zone, $r \lesssim 3R/2$. Thus, the number of independent ingoing s-wave modes captured by L is of order $\log(R/l_P)$, as shown in Fig. 8.4. In other words, L "sees" what enters the black hole in the first scrambling time after infalling geodesics that would have crossed μ_Q (see also Appendix A.16).

Every such mode would contribute $O(1)$ entropy in the Hartle-Hawking state but is pure in the Unruh state (since it is in the ground state). The missing entropy, and the gap to saturating the QPI, is thus

$$\Delta \sim \log \frac{R}{l_P} \ . \tag{8.5.5}$$

The entropy on null surfaces can have surprising and counterintuitive properties [194]. As a check on the above arguments, we now verify this result by evaluating $S_{\rm gen}[L]$ using an alternative method, in which von Neumann entropies are evaluated only on spacelike hypersurfaces.[5]

[5]We thank Aron Wall for suggesting this approach.

The mutual information of any two systems is defined in terms of the von Neumann entropies of the individual and joint systems as follows:

$$I(L:B) \equiv S_L + S_B - S_{LB} \ . \tag{8.5.6}$$

Here we consider the lightsheet L and the partial Cauchy surface B shown in Fig. 8.4. We take B to be null until it meets the end of the near horizon zone, $r = 3R/2$, and to coincide approximately with a constant t hypersurface outside of this radius. To stay in the semiclassical regime, one can terminate L slightly before the singularity. We can choose this terminal surface to have area cl_P^2, where $1 \ll c \ll \log(R/l_P)$. The second inequality ensures that its contribution will be subleading to our result.

Note that the joint system LB is equivalent by unitary evolution to the purely spacelike Cauchy surface C. We can thus evaluate the von Neumann entropy on L as

$$S_L = S_C - S_B + I(L:B) \ . \tag{8.5.7}$$

Moreover, L and C have the same boundary, μ_Q, whereas B has a boundary of negligible area. It follows that

$$S_{\rm gen}[L] = S_{\rm gen}[C] - S_B + I(L:B) \ . \tag{8.5.8}$$

We chose μ_Q to be just after black hole formation, so there will be no outgoing Hawking radiation present on C. In the Unruh state, the ingoing spherical modes in the near-horizon zone are unoccupied, which reduces the entropy by $\log(R/l_P)$ compared to the Hartle-Hawking value. Hence

$$S_{\rm gen}[C] - \frac{A[\mu_Q]}{4G\hbar} \sim \log\frac{R}{l_P} \ . \tag{8.5.9}$$

In our approximation, B captures the same outgoing modes as C, but none of the ingoing modes that cross L, so $S_B = 0$. There is no data on L that is entangled with data on B, so $I(L:B) = 0$. Hence Eq. (8.5.7) implies $S_{\rm gen}[L] = S_{\rm gen}[C]$ in our example. Since $16\pi G^2 m^2 = A[\mu] = A[\mu_Q] + O(l_P^2)$, we recover Eq. (8.5.5).

Note that the Planck length enters Eq. (8.5.5) through the position of the quantum marginally trapped surface μ_Q, which is a proper distance of order l_P inside of the event horizon (or of μ). It would appear, therefore, that Δ could be minimized if one could arrange for μ_Q to lie a distance comparable to R inside the horizon. However, this requires a large perturbation of the black hole, to which the current analysis does not apply. We will revisit this question in Sec. 8.5.

Near-Saturation of the QPI

In the previous subsection, we found that in a newly formed Schwarzschild black hole with no exterior matter, the QPI will be satisfied but not quite saturated, with a gap of $\Delta \sim \log(R/l_P)$. The gap is only logarithmic, but it still becomes arbitrarily large for large black holes. Here we show that the logarithmic gap can be eliminated. Thus, the QPI can be saturated up to a fixed gap of order a Planck area, which we do not have full control over.

The simplest way to accomplish this is to time-reverse the state of the semiclassical fields on the partial Cauchy surface C shown in Fig. 8.4. In our approximation, this will not affect the $\ell > 0$ modes, but it will put the spherical waves in a time-reversed Unruh state. That is, the outgoing modes will be unoccupied and the ingoing modes will be occupied, reversing the situation considered in the previous subsection. Crucially, this modification will not change the mass m at infinity, so we still have

$$16\pi G^2 m^2 = A[\mu] = A[\mu_Q] + O(l_P^2) \ . \tag{8.5.10}$$

Because of the restriction to semiclassical modes, there is a cutoff near μ_Q at least of order l_P. Thus, while the initial conditions we now impose are somewhat unnatural, they will persist only for one scrambling time $\Delta t_s \sim R \log(R/l_P)$. After this time, the black hole will begin to evaporate. In particular, unlike the full Boulware state, there is no singularity at the horizon. Note also that this state differs from the one we considered in Sec. 8.3 in that the $\ell > 0$ modes are not in the Boulware vacuum.

The lightsheet L is sensitive only to the ingoing part of the radiation, so its generalized entropy will be the same as it would be in the Hartle-Hawking state:

$$S_{\rm gen}[L] = \frac{A[\mu_Q]}{4G\hbar} \ . \tag{8.5.11}$$

Thus we find that the QPI is nearly saturated:

$$\Delta \equiv \frac{4\pi G}{\hbar} m^2 - S_{\rm gen}[L] \sim O(1) \ . \tag{8.5.12}$$

Perturbative Regime: QPI from the GSL

Next, we will consider the more general case where matter enters into the black hole after its formation. We consider the same formation process as above. We will again focus on μ_Q right after formation so as to obtain the tightest bound. But now we will allow for a nontrivial quantum state outside of the black hole. This could be an ordinary matter system carrying some thermodynamic entropy. It could also be a quantum state with negative energy, such as the Boulware-like state that we considered in Sec. 8.3 as a counterexample to the CPI.

The future-outgoing lightsheet L of μ_Q will only receive matter that falls into the black hole within the first scrambling time after μ_Q; see Fig. 8.4. To be precise, consider a family of radially infalling geodesics that are initially at rest at some large radius $r \gg R$. The geodesics are all at the same angle but shifted in time. It is easy to check that the geodesic that passes through μ_Q and the last geodesic that reaches L are separated at large radius by a time of order $\Delta t_s \sim R \log(R/l_P)$. Any matter that falls in later will hit the singularity before reaching Σ. This statement does not depend on the initial radius, and it also holds also for ingoing null geodesics; see Appendix A.16.

In the following subsection, we will consider the effects of matter that falls in after the first scrambling time and so does not reach L. However, now we will focus on matter that

can be registered on L. By the above argument, we can take this matter to reside within the near-horizon zone, $R < r < 3R/2$, on the partial Cauchy surface C. Let H be the portion of the event horizon to the future of C, and let $S_{\rm gen}[H]$ be its generalized entropy.

We begin by making a simplifying assumption that will be relaxed below, that all of the matter that falls across the horizon will also cross L (as opposed to passing through the portion of B inside the black hole). The quantum marginally trapped surface μ_Q and the boundary of H have approximately the same area, so there is a simple relationship between the entropy on H and L:

$$S_{\rm gen}[L] = S_{\rm gen}[H] - \Delta S[H_{\rm late}] + \mathcal{O}(1), \tag{8.5.13}$$

where $H_{\rm late}$ is the portion of the horizon above a sufficiently late Cauchy slice, when the black hole has relaxed to equilibrium, but early enough that negligible Hawking radiation has been produced.

We have assumed a state in which there is negligible mutual information between L and $H_{\rm late}$. For example, if the black hole simply evaporates with no further matter falling in, $\Delta S[H_{\rm late}]$ is the (negative) renormalized entropy that exists on the horizon in the Unruh state (due to the missing infalling modes when to compared the Hartle-Hawking state).

From

$$S_{\rm gen}[H_{\rm late}] - \Delta S[H_{\rm late}] = \frac{A_{\rm late}}{4G\hbar} \tag{8.5.14}$$

and Eq. (8.5.13), the QPI follows:

$$S_{\rm gen}[L] = S_{\rm gen}[H] - \Delta S[H_{\rm late}] \leq S_{\rm gen}[H_{\rm late}] - \Delta S[H_{\rm late}] = \frac{A_{\rm late}}{4G\hbar} \leq \frac{4\pi G}{\hbar} m^2 \ . \tag{8.5.15}$$

The first inequality in this sequence is the GSL for event horizons. Note that we have ignored the $\mathcal{O}(1)$ additive uncertainty in Eq. (8.5.13) in light of the discussion at the end of Sec. 8.4.

This argument establishes the QPI for a large class of examples, including the Boulware-like state that served as a counterexample to the classical Penrose inequality in Sec. 8.3. In this case, $A_{\rm late}$ (which sets the mass) will be significantly smaller than the area of the trapped surface μ. Here we use the quantum trapped surface μ_Q, but its area is almost the same as that of μ. What saves the QPI is the contribution of the entropy on L, which is negative in this example. Specifically, the GSL guarantees that the lower bound, $S_{\rm gen}[L]$, is smaller than the area of μ_Q by a sufficient amount for the QPI to hold.

In the case where positive entropy registers on H and L, our QPI is stronger than the classical Penrose inequality. The lightsheet "knows" that more matter will enter the the black hole after μ_Q, and the GSL "knows" that this will result in an area increase. Effectively, this larger area becomes the lower bound on the mass.

Failed Counterexample: Negative Energy That Misses the Lightsheet

In the previous subsection, we considered the case where all matter outside the quantum trapped surface μ_Q crosses its lightsheet L. Here we generalize to discuss matter for which

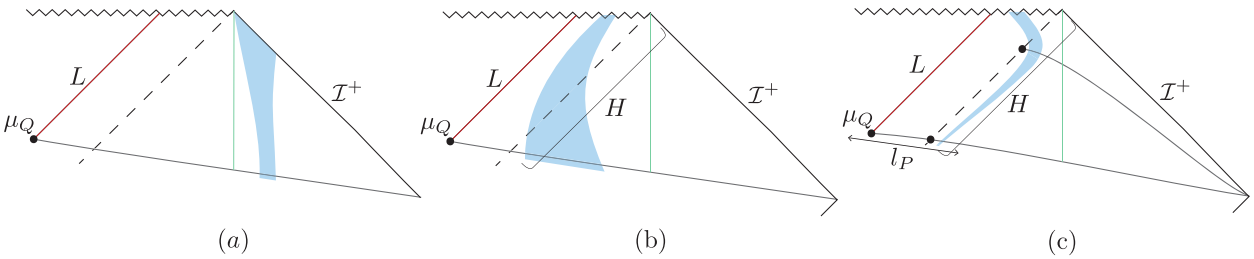

Figure 8.5: The QPI is threatened by any negative energy (blue worldvolume) that fails to register on the lightsheet L. We analyze three possibilities but find that none of them leads to a violation of the QPI. (a) Negative energy outside of the near horizon zone (vertical green line). (b) Negative energy that enters the black hole soon after μ_Q but evades L by accelerating outward. (c) Negative energy that remains near the black hole for more than a scrambling time.

this does not happen. In this case, we cannot use the GSL for the event horizon to constrain the relation between $S_{\text{gen}}[L]$ and the mass at infinity. However, we will give some plausibility arguments for the validity of the QPI.

In the previous subsections, we argued that the QPI will hold true if all matter outside of μ_Q passes through L. We can think of the present situation as a complication where we add matter that does not satisfy this property. Since this cannot affect $S[L]$, the only way that the QPI can now be violated is if the matter we added contributes negative mass at infinity. We will now argue that this is impossible in the semiclassical regime.

Matter outside of μ_Q can fail to register on L for any of the following three reasons (see Fig. 8.5):

1. The matter never enters the black hole.

2. The matter enters the black hole during the first scrambling time after C but escapes through the portion of B inside the black hole.

3. The matter enters the black hole later than a scrambling time after C.

In the first case, the matter can be approximately treated as isolated from the black hole. But the total mass of isolated systems is positive, so distant systems can never cause violations of the QPI. (This does not rule out regions with negative energy, but it implies that sufficient positive energy must be present nearby.)

In the second case, the matter system can be initially near the black hole and so could have regions of negative energy density (as in the example of Sec. 8.3). However, in order to miss L, it would have to accelerate outwards after crossing the horizon. This requires positive energy. We will not attempt to demonstrate here that this always results in a net

positive mass contribution; our goal is only to note that the QPI is not obviously violated in this setup. This question merits further study.

In the third case, we again must choose the matter system to be close to the horizon if we wish to give it negative energy. For example, the Boulware-like state of Sec. 8.3 would qualify. However, by assumption this state would have to be present more than one scrambling time after C. Moreover, the modes for which it is possible to obtain net negative energy are those that make up the thermal atmosphere of the black hole; these modes evolve exponentially close to the horizon under backward time evolution. Thus the state on C would contain transplanckian energy density (similar to a firewall). The initial state would not be a semiclassical state. This argument is robust and rules out an entire class of what naively seemed like promising counterexamples. We view this as nontrivial evidence in favor of our proposal.

8.6 Alternative Proposals

In this section we consider various alternative conjectures for the QPI. In Sec. 8.6 we give counterexamples to proposals that might otherwise seem natural. In Sec. 8.6 we discuss modifications of our proposal that appear viable, and we explain why we are not currently advocating for them.

Nonviable Alternatives

We will now discuss several alternative conjectures for a QPI that we considered in the process of this work. Our goal is to explain our choice in Sec. 8.4, and to illustrate that the problem is rather constrained. This proves neither that our formulation is unique, nor that it is correct. But we will see that it is remarkably difficult to find any alternative statement of the QPI that is not immediately ruled out.

Cauchy surfaces that reach spatial infinity First, we explain why we do not allow $\Sigma[\mu_Q]$ to reach outside the black hole. This prohibition is motivated by the asymptotically flat case, to which we will specialize for now. Let Σ_∞ be a Cauchy surface of $O_W[\mu_Q]$, in violation of our requirements. An example is the black slice in the Fig. 8.6. Let $S_{\text{gen}}[\Sigma_\infty(\mu_Q)]$ be the generalized entropy evaluated on Σ_∞. The alternative QPI thus would take the form

$$m \overset{?}{\geq} \sqrt{\frac{\hbar}{4\pi G} S_{\text{gen}}[\Sigma_\infty(\mu_Q)]}\ . \tag{8.6.1}$$

But it is easy to find a counterexample to Eq. (8.6.1): an arbitrary amount of matter entropy can be placed in regions far from the black hole, at arbitrarily little cost in mass. We now discuss this in detail.

Consider a dilute gas of N photon wave packets, each of characteristic size λ. Each photon occupies a region of volume λ^3, so the photons can be dilute if they occupy a region

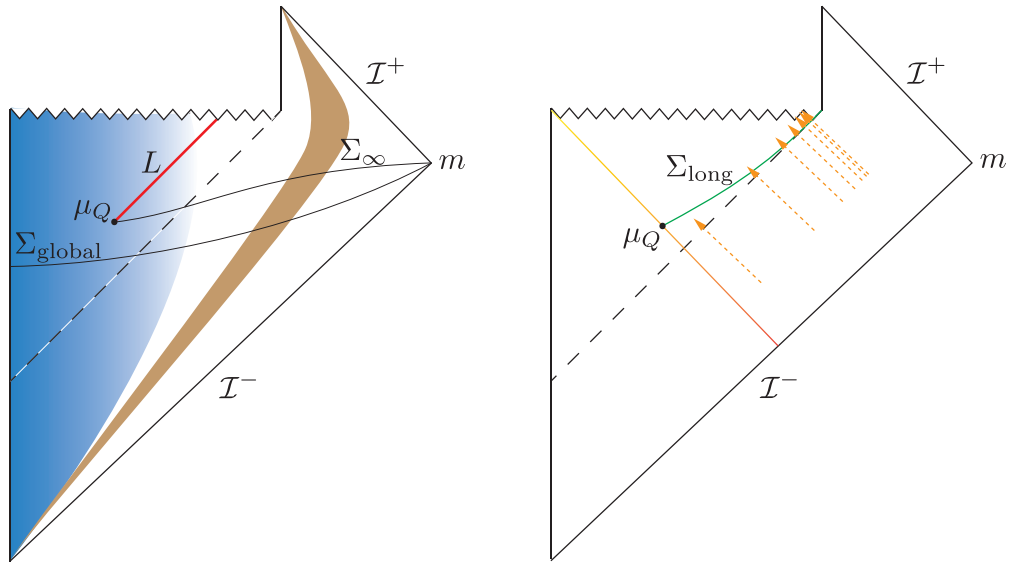

Figure 8.6: Left: the generalized entropy on the slice Σ_∞ can be dominated by distant soft particles (brown) and so does not yield a viable lower bound on the mass. The global Cauchy surface Σ_{global} plays a role in an alternative proposal discussed in the main text. Right: the long slice Σ_{long} captures all of the missing infalling Hawking modes.

of volume $N\lambda^3$. We can take each photon to be in a mixed state (say, of polarizations), and in a product state with respect to the rest of the universe. Then the gas contributes of order N to the generalized entropy on Σ.

We take the gas to be very far from the black hole or any other matter, so that gravitational binding energy to other objects is negligible. The gravitational binding energy of the photon cloud itself will be negligible if $NG\hbar/\lambda \ll N^{1/3}\lambda$, so we shall take $\lambda \gg N^{1/3} l_P$, where $l_P \equiv (G\hbar)^{1/2}$ is the Planck length. Then the gas of photons contributes a mass of order $N\hbar/\lambda$ to the ADM mass. This mass contribution can be taken to be arbitrarily small by taking $\lambda \to \infty$ at fixed N without violating any of the previous assumptions.

We are still free to choose N to take any value we like. Thus we have found a family of initial data with bounded m but unbounded $S_{\text{gen}}[\mu_Q] \approx c_1 + c_2 N$, where c_1 and c_2 are independent of N. For large enough N, this leads to a violation of Eq. (8.6.1).

Area of marginally quantum trapped surfaces A second alternative conjecture would be to use only the area of μ_Q, not its generalized entropy:

$$m \overset{?}{\geq} \sqrt{\frac{A[\mu_Q]}{16\pi G^2}} \,. \tag{8.6.2}$$

That is, one would conjecture that Eq. (8.2.1) holds if A is taken to be the area of a quantum trapped surface. This possibility is attractive because the entropy of distant soft radiation would never contribute to the lower bound in the first place.

However, Eq. (8.6.2) is ruled out (among other reasons) by the Boulware-like counterexample to the classical Penrose inequality. This is because the area of the bifurcation surface will receive only a correction that can be made parametrically small. This follows from the remarks concerning the classically marginally trapped surface at the end of Sec. 8.3. The same argument implies that the marginally quantum trapped surface area receives only a parametrically small correction, which cannot compete with the large decrease in mass.

Subtracting global entropy; interior generalized entropy Let us revisit the proposal of Sec. 8.6 and consider the generalized entropy $S_{\rm gen}[\Sigma_\infty(\mu_Q)]$ of a marginally trapped surface μ_Q, evaluated on a Cauchy surface that reaches outside of the black hole all the way to spatial infinity. This proposal suffered from the problem that distant soft modes can contribute unbounded entropy with bounded energy, so $S_{\rm gen}[\Sigma_\infty(\mu_Q)]$ is unrelated to any lower bound on the mass.

A natural idea is to subtract the von Neumann entropy on a global Cauchy surface (see Fig. 8.6):

$$m \stackrel{?}{\geq} \sqrt{\frac{\hbar(S_{\rm gen}[\Sigma_\infty(\mu_Q)] - S[\Sigma_{\rm global}]}{4\pi G}} \ . \tag{8.6.3}$$

If the distant soft modes have the same entropy in the global state as in the generalized entropy, then their dangerous contribution will cancel out.

However, this need not be the case. Consider a collapsing star that forms a Schwarzschild black hole of area A. The entropy of the star can be of order $S_{\rm star} \sim (A/G\hbar)^{3/4}$ or even $S_{\rm star} \sim A/G\hbar$ [195]. We can chose the global state to contain only distant soft radiation that purifies the star, so that $S[\Sigma_{\rm global}] = 0$ and

$$m = \sqrt{\frac{A[\mu_Q]}{16\pi G^2}} + \epsilon \ , \tag{8.6.4}$$

where ϵ can be arbitrarily small. But then

$$S_{\rm gen}[\Sigma_\infty(\mu_Q)] \approx \frac{A[\mu_Q]}{4G\hbar} + S_{\rm star} \ , \tag{8.6.5}$$

so that Eq. (8.6.3) is violated.

The violation in our example remains bounded, since $S_{\rm star}$ cannot exceed $A[\mu_Q]/4G\hbar$ by the GSL. One might consider absorbing this violation by adding a correction factor of $1/2$ to the right hand side of Eq. (8.6.3). But by considering initial data with a second asymptotic region, one can arrange $S[\Sigma_{\rm global}] = 0$ with unbounded $S_{\rm gen}[\Sigma_\infty(\mu_Q)]$ at fixed m, leading to unbounded violations.

A variation of this idea is to use the generalized entropy in the interior (not the exterior) of the surface μ_Q. It is easy to check that it fails for the same reasons.

Possible Modifications of the QPI

We will now discuss an alternative formulation of the QPI that we cannot currently rule out, and we comment on some of its properties that have led us to reject it as our main proposal.

The basic idea is to consider partial Cauchy surfaces other than L, still bounded by μ_Q and remaining inside the black hole. For example, we could assert that

$$m \geq \sqrt{\frac{\hbar S_{\text{gen}}[\Sigma]}{4\pi G}} \tag{8.6.6}$$

holds for any achronal hypersurface $\Sigma \subset B \cap O_W[\mu_Q]$ whose only boundary is μ_Q. This class includes the lightsheet L, so this conjecture would be strictly stronger than our main proposal. It is clear that the heuristic arguments in support of QPI in Sec. 8.5 also apply to this family of slices.

There are some clear downsides to this choice. The region B and therefore this family of slices are defined teleologically. Furthermore, it is not clear to us how one would formulate a minimality requirement in this case, analogous to the requirement that the classically trapped surface minimize the area on some Cauchy surface.

A variation would be to insist on a Cauchy surface that is as "long" as possible, i.e., which does not have any endpoint on the future singularity. Roughly, this means it ends on the future endpoints of the horizon generators, see Σ_{long} in Fig. 8.6. This proposal is weaker than the previous one and neither stronger nor weaker than our main proposal. We will now argue that for an evaporating black hole this results in a less stringent bound than the one obtained from L.

As discussed in Sec. 8.5, in the Unruh state there is negative entropy falling across the horizon, due to the missing ingoing modes compared to the Hartle-Hawking state. The long slice will capture this negative entropy through the entire process of evaporation. (Here we are assuming that the semiclassical expansion is valid until the black hole area is Planckian in size.) The generalized entropy on this slice is:

$$S_{\text{gen}}[\Sigma_{\text{long}}] = \frac{A[\mu_Q]}{4G\hbar} - \gamma \frac{A[\mu_Q]}{4G\hbar} \ , \tag{8.6.7}$$

where $\gamma \geq 1$ by the GSL, and the second term arises from the contribution of the missing ingoing modes on Σ.

It is difficult to compute γ exactly. If $\gamma > 1$, then S_{gen} will be negative. This renders (8.6) ill-defined. Negative S_{gen} is also conceptually in conflict with the interpretation of S_{gen} as an entropy in the fundamental theory of quantum gravity. This suggests that a careful computation will reveal that $\gamma = 1$, in which case Eq. (8.6) reduces back to the statement of the positivity of the ADM mass. Along with the downsides mentioned earlier, this conundrum shows that such long slices are not ideal for formulating the QPI.

8.7 Quantum Penrose Inequality in Anti-de Sitter Space

The classical Penrose inequality was motivated by the heuristic argument that a Schwarzschild black hole with no exterior matter should have the smallest possible mass for a given trapped surface area. In Eq. (8.2.1) we assumed a vanishing cosmological constant Λ. An analogous argument for asymptotically Anti-de Sitter spacetimes with curvature scale $L = (-\Lambda/3)^{1/2}$ yields the classical inequality

$$m \geq f_{\rm AdS}(A[\mu]) \ , \tag{8.7.1}$$

where

$$f_{\rm AdS}(A) \equiv \left(\frac{A}{16\pi G^2}\right)^{1/2} + \left(\frac{A}{16\pi G^2}\right)^{3/2} \frac{G^2}{L^2} \tag{8.7.2}$$

and μ is again a trapped surface satisfying an appropriate minimality condition (see Sec. 8.2).

Following our QPI proposal for asymptotically flat space, it would appear natural to propose the following QPI in asymptotically AdS spacetimes:

$$m \stackrel{?}{\geq} \left(\frac{\hbar S_{\rm gen}}{4\pi G}\right)^{1/2} + \left(\frac{\hbar S_{\rm gen}}{4\pi G}\right)^{3/2} \frac{G^2}{L^2} \ . \tag{8.7.3}$$

in asymptotically AdS spacetimes with curvature scale L. Here $S_{\rm gen}$ is defined with respect to slices defined in Sec. 8.4; see Fig. 8.7. However, due to $\mathcal{O}(1)$ subtleties discussed at the end of Sec. 8.4, it is not clear that Eq. (8.7.3) will hold exactly in the AdS Hartle-Hawking state (referred to as σ henceforth). The issue is the radiation mass outside of the black hole which could be negative, lowering the LHS of Eq. (8.7.3) to violation. As we will discuss here, in asymptotically AdS spacetimes one could fix this $\mathcal{O}(1)$ issue. Note that the quantum-corrected ADM mass in this state is

$$m = \left(\frac{A}{16\pi G}\right)^{1/2} + \left(\frac{A}{16\pi G}\right)^{3/2} \frac{G^2}{L^2} + m_{\rm rad} \ , \tag{8.7.4}$$

with

$$m_{\rm rad} = \int_{\Sigma_1} d\Sigma^\nu t^\mu \langle T_{\mu\nu}\rangle_\sigma \ , \tag{8.7.5}$$

where $\langle T_{\mu\nu}\rangle_\sigma$ is the renormalized stress tensor in σ, Σ_1 is a Cauchy slice stretching from the bifurcation surface to the boundary of AdS, and t^μ is the Killing field in Schwarzschild-AdS that is timelike at infinity. Also, note that the area term in Eq. (8.7.4) is not the quantum-corrected area. Furthermore, based on formulation in Sec. 8.4, $S_{\rm gen}$ in the σ is computed on the part of the horizon in the future of the bifurcation surface μ_Q; see Fig. 8.8. The quantum stress tensor $\langle T_{\mu\nu}\rangle$ has been computed in the Hartle-Hawking state in 2+1

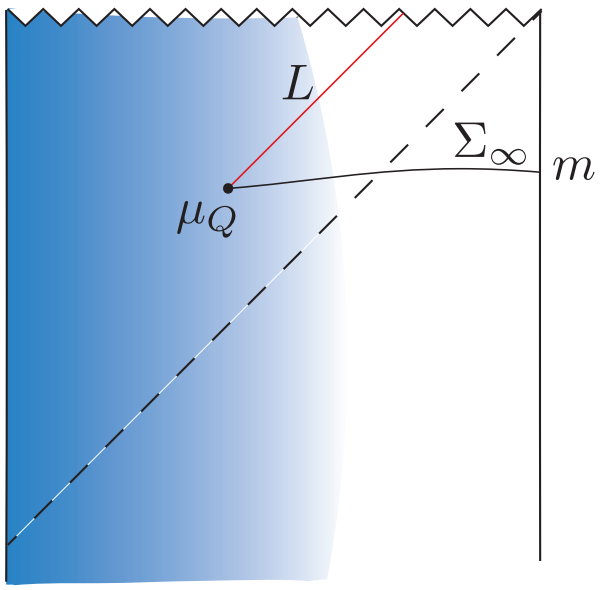

Figure 8.7: Different choices of slices anchored to the surface μ_Q on which one could compute S_{gen}. The red lightsheet L is defined analogously to the asymptotically flat case. Since distant soft modes do not exist for large black holes in AdS, one could also consider computing S_{gen} on the black slice Σ_∞ that ends on the asymptotic boundary.

dimensions with different choices of boundary conditions [196]. One finds that m_{rad} depends on the field content and the boundary conditions; moreover, m_{rad} does not have a definite sign [196]. Explicit calculations in 2+1 dimensions show that m_{rad} can be negative. We do not expect that the entropy of the matter on Σ_2 and the quantum corrections to the area term would compensate for this negative value of m_{rad} so as to uphold Eq. (8.7.3). Therefore, we expect that Eq. (8.7.3) can be violated in the Hartle-Hawking state. Furthermore, the non-universality of m_{rad} seems to suggest that the correct formulation of QPI for large AdS black holes must depend on various factors that m_{rad} depends on (e.g. the field content and the boundary conditions).

Here we propose a way to introduce this dependence into a Quantum Penrose Inequality for asymptotically Anti-de Sitter spacetimes. Let f^q_{AdS} be a function such that in the Hartle-Hawking state,

$$m = f^q_{\text{AdS}}(S_{\text{gen}}[\Sigma_2]) \ , \tag{8.7.6}$$

where m is the quantum-corrected ADM mass and S_{gen} is associated to the future portion Σ_2 of the horizon; see Fig. 8.8. Now, we propose

$$m \geq f^q_{\text{AdS}}(S_{\text{gen}}[L]) \ , \tag{8.7.7}$$

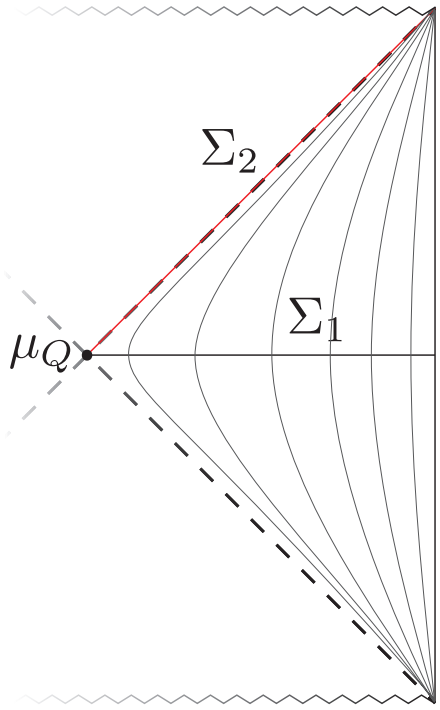

Figure 8.8: The Hartle-Hawking state is essential for our definition of $f^q_{\rm AdS}$ via $m = f^q_{\rm AdS}(S_{\rm gen})$. Here m is the ADM mass including the quantum corrections associated with the radiation mass. $m_{\rm rad}$ is computed on the black slice Σ_1 with respect to the time-like Killing field t^μ whose orbits are shown in the figure. $S_{\rm gen}$ is computed on the red null slice Σ_2 on the horizon that ends on the bifurcation surface μ_Q.

for any marginally trapped surface μ_Q in an asymptotically AdS spacetime with a large AdS black hole. A heuristic argument for Eq. (8.7.7) is as follows: First, the above inequality will follow from the classical Penrose inequality unless we are in a state perturbatively close to Kerr-AdS. In that limit, it can be shown (see Appendix A.17) that given any quantum marginally trapped surface, there exists a Q-screen that approaches the horizon of the Kerr-AdS at late times and has the quantum marginally trapped surface as a leaf. As discussed in Sec. 8.4, Q-screens are known to satisfy a generalized second law [177]. The QPI would then follow from:

$$\begin{aligned} &S_{\rm gen}|_{\rm early} \leq S_{\rm gen}|_{\rm Kerr-AdS} \\ &\qquad \implies f^q_{\rm AdS}\left(S_{\rm gen}|_{\rm early}\right) \leq f^q_{\rm AdS}\left(S_{\rm gen}|_{\rm Kerr-AdS}\right) \leq f^q_{\rm AdS}\left(S_{\rm gen}|_{\sigma}\right) = m \ , \end{aligned} \tag{8.7.8}$$

where the first inequality on the second line follows from the generalized second law of Q-screens and the second inequality follows from assuming $f^q_{\rm AdS}$ is a monotonic function.

In general we could have states where the AdS black hole is not large enough to reach stable thermal equilibrium with the asymptotic boundary of the spacetime, so a few words

about the case of these small AdS black holes are in order. For such black holes, we cannot define the function $f^q_{\rm AdS}$ as above. Our proposal would then follow more closely our proposal for asymptotically flat spacetimes, where we formulate our conjecture using the function f appearing in the classical Penrose inequality for AdS.

$$m \geq f_{\rm AdS}(4G\hbar S_{\rm gen}[L]) \ , \tag{8.7.9}$$

where $f_{\rm AdS}$ is defined in Eq. (8.7.2). The phase transition for (in)stability of AdS black holes happen around ADM mass L/G, so our proposal changes for mass above and below the phase transition point. The exact value of mass associated to a phase transition depends on the choice of boundary conditions and the field content.

An important difference between QPI for large AdS black holes and flat space black holes is the absence of the challenge associated with soft modes. As discussed in Sec. 8.6, in asymptotically flat space, one can add entropy far away from the black hole at negligible cost to the ADM mass. This prevents any formulation of QPI where the generalized entropy is computed on partial Cauchy slices approaching spatial infinity in asymptotically flat spacetimes.

However, in asymptotically AdS spacetimes and in the presence of a large black hole, excitations require considerable energy to remain outside of the black hole, so the arguments of Sec. 8.6 do not go through and matter entropy outside of the black hole has an energy cost. Therefore, in the presence of large AdS black holes the slice on which $S_{\rm gen}$ is evaluated could end on the asymptotic boundary of AdS. This possibility was discussed in the context of AdS/CFT in [187]. To define the function $f^q_{\rm AdS}$ in this version of QPI, we need to consider the Hartle-Hawking state and the generalized entropy on the spatial slice Σ_1 of Fig. 8.8,

$$m = f^q_{\rm AdS}(S_{\rm gen}[\Sigma_1]) \ . \tag{8.7.10}$$

The quantum extremal surface prescription [197] equates $S_{\rm gen}[\Sigma_1]$ with the von Neumann entropy of the dual CFT in the thermofield double state. Therefore, this definition of the function $f^q_{\rm AdS}$ has a very natural interpretation from the CFT perspective

$$\langle H \rangle_{TFD} = f^q_{\rm AdS}\left(S_{\rm CFT}[TFD]\right) \ , \tag{8.7.11}$$

where $\langle H \rangle_{TFD}$ is the expectation value of the CFT Hamiltonian in the thermofield double state.

8.8 Classical and Non-gravitational Limits

In this section we discuss two interesting limits of the QPI: the classical limit, $\hbar \to 0$; and the non-gravitational limit, $G \to 0$.

In the $\hbar \to 0$ of QPI, we recover the classical Penrose inequality. This is easy to see. The amount of matter entropy on L is $\mathcal{O}((G\hbar)^0)$, and therefore

$$\lim_{\hbar \to 0} 4G\hbar S_{gen}[L] = A[\mu_Q] \ . \tag{8.8.1}$$

Furthermore, the surface μ_Q is perturbatively close to a (classically) marginally trapped surface such that their area difference is due to quantum corrections and therefore of order $G\hbar$ and can be neglected. Lastly, any $\hbar$ corrections to the function f can trivially be ignored in the $\hbar \to 0$ limit. We therefore have the desired implication:

$$f^q\left(4G\hbar S_{gen}[\mu_Q]\right) \leq m \ \overset{\hbar \to 0}{\Longrightarrow} \ f^c\left(A[\mu]\right) \leq m \ . \tag{8.8.2}$$

We turn to the $G \to 0$ limit of the QPI. This is of interest because some semiclassical conjectures yield nontrivial and novel implications about QFT in this limit. For example, the QNEC was first discovered by taking the $G \to 0$ limit of the QFC in a particular setting [178]. In order to sidestep the small "Casimir uncertainty" discussed in Sec. 8.4, we will consider the QPI in AdS. We further restrict to two complementary scenarios.

First, consider a perturbation to a Hartle-Hawking state such that in a finite amount of time the state settles back down to a Hartle-Hawking state (with a different temperature). In this case, Eq. (8.7.8) shows that the QPI is equivalent to the GSL. The non-gravitational limit of the GSL the monotonicity of relative entropy. This is a nontrivial but well-known statement in quantum information theory, which applies in particular in QFT.

The second scenario is when the perturbation does not relax to equilibrium. This means that the excitation that takes the state away from the Hartle-Hawking state remains outside of the black hole. Therefore such excitations do not change the generalized entropy on L or the geometry of the event horizon. Let δm be the change in the ADM mass caused by this perturbation. Since the QPI is saturated in the Hartle-Hawking state, it reduces to

$$\delta m \geq 0 \tag{8.8.3}$$

in the $G \to 0$ limit. Here $\delta m = \int_{\Sigma_1} d\Sigma^\mu \xi^\nu T_{\mu\nu}$ (see Fig. 8.8), and t^ν is the timelike Killing vector field outside the black hole. This makes physical sense: if the field excitations are isolated from the black hole, they need to satisfy their own positive energy condition.

8.9 Cosmic Censorship Conjecture

In this section, we consider the current status of the cosmic censorship conjecture (CCC) and its relation to the Penrose inequality. We argue that there is a need for a quantum generalization of the CCC, and we suggest that the proposed Quantum Penrose Inequality may inform the formulation of a quantum CCC.

The formation of singularities in gravitational collapse is guaranteed by classical [173] and quantum [179] singularity theorems. However, it is not clear that the formation of a singularity implies the formation of a black hole.

The weak CCC asserts that singularities (regions of arbitrarily high curvature) will not be visible to a distant observer.[6] A precise statement of the conjecture can be formulated as follows [172]: *Let (Σ, $h_{\mu\nu}$, $K_{\mu\nu}$) be an asymptotically flat initial data set for Einstein's equation with (Σ, $h_{\mu\nu}$) a complete Riemannian manifold. Let the matter sources be such that $T_{\mu\nu}$ satisfies the dominant energy condition and the coupled Einstein-matter field equations are of the form $\Box\phi(x) = F(x, \phi, \nabla_\mu\phi)$, where F is a smooth function of its variables. In addition let the initial data for the matter fields on Σ satisfy appropriate asymptotic falloff conditions at spatial infinity. Then the maximal Cauchy evolution of these initial data is an asymptotically flat, strongly asymptotically predictable spacetime.*

The CCC has not been proven. Indeed, there are a number of known "mild" violations that we will discuss shortly. The (classical) Penrose inequality is only a necessary condition for the CCC, as explained in Sec. 8.2. Even this weaker statement has not been proven; but as a quantitative relation between mass and area, it has been extensively explored. The fact that no counterexample has been found can be viewed as indirect evidence that some version of the CCC may indeed hold.

Let us now discuss the mild violations mentioned in the previous paragraph. A black string in 4+1 dimensions suffers from the Gregory-Laflamme instability [198]. Further evolution causes the string to become arbitrarily thin in some regions [199, 200] and so arbitrarly high curvatures become visible to a distant observer.

In 3+1 dimensions, there exist fine-tuned initial data sets such that the solution exhibits a self-similar behavior near the threshold of formation of a black hole [201, 202, 203, 204]. At the threshold, a naked singularity forms. In some examples, the naked singularity propagates out to $\mathcal{I}^+$.

In the above two examples, the initial data satisfy the dominant energy condition, as required by the CCC. The black string is not asymptotically flat, but one expects that it can be truncated at a sufficiently great length so that local evolution far from the ends still leads to a naked singularity.

Let us now add a third example, which is physically relevant but does not obey the dominant energy condition: a black hole that evaporates completely. In this case, treating the spacetime as a classical manifold, a naked singularity is inevitable [205, 206].

Only the last example explicitly involves quantum effects. But it points to a resolution of all three violations: clearly, it makes no sense to treat the spacetime as a classical manifold near the endpoint of evaporation (i.e., arbitrarily close to the naked singularity). When the curvature formally exceeds the Planck curvature, the semiclassical expansion breaks down, and a classical geometric description of the spacetime need not exist.

But this observation also applies to the other known examples of CCC violation. One would expect a black string to pinch off before it becomes thinner than a Planck length. Similarly, one would expect quantum effects to smooth out the fine-tuned initial data, or at least the singularities they lead to.

[6] We will consider only the weak CCC here. The strong form of the CCC states, roughly, that no observer can see a singularity. In all cases, one assumes regular initial data.

Naively, all three examples violate the spirit of the CCC: starting from a highly classical regime, evolution produces an outcome in which quantum gravity is required to maintain predictability. But in an important sense, the violation is "small" in each case. The energies involved are likely no greater than the Planck mass, and we can "guess" a plausible future evolution without having a full quantum gravity theory. For example, the Planck-sized black hole will probably decay into a few more particles, and the black string will simply pinch off.

It would be of interest to formulate a quantum version of the CCC which accounts for these physically reasonable phenomena, i.e., one that is not formally violated by them.[7] We expect the Quantum Penrose Inequality to play a role analogous to the classical one: as a necessary condition for the quantum CCC, and thus as a useful test. Perhaps more importantly, the Quantum Penrose Inequality may be of some use in identifying the correct formulation of a quantum CCC in the first place.

[7] A specific proposal will be studied in forthcoming work.

Appendix A

Appendix

A.1 Notation and Definitions

Basic Notation

Notation for basic bulk and boundary quantities

- Bulk indices are $\mu, \nu, \ldots$.
- Boundary indices are $i, j, \ldots$. Then $\mu = (z, i)$.
- We assume a Fefferman–Graham form for the metric: $ds^2 = \frac{L^2}{z^2}(dz^2 + \bar{g}_{ij}dx^i dx^j)$.
- The expansion for $\bar{g}_{ij}(x, z)$ at fixed x is

$$\bar{g}_{ij} = g_{ij}^{(0)} + z^2 g_{ij}^{(2)} + z^4 g_{ij}^{(4)} + \cdots + z^d \log z g_{ij}^{(d,\log)} + z^d g_{ij}^{(d)} + \cdots . \tag{A.1.1}$$

  The coefficients $g_{ij}^{(n)}$ for $n < d$ and $g_{ij}^{(d,\log)}$ are determined in terms of $g_{ij}^{(0)}$, while $g_{ij}^{(d)}$ is state-dependent and contains the energy-momentum tensor of the CFT. If d is even, then $g_{ij}^{(d,\log)} = 0$. To avoid clutter we will often write $g_{ij}^{(0)}$ simply as g_{ij}. Unless otherwise indicated, i, j indices are raised and lowered by $g_{ij}^{(0)}$.

- We use $\mathcal{R}$, $\mathcal{R}_{\mu\nu}$, $\mathcal{R}_{\mu\nu\rho\sigma}$ to denote bulk curvature tensors, and R, R_{ij}, R_{ijmn} to denote boundary curvature tensors.

Notation for extremal surface and entangling surface quantities

- Extremal surface indices are $\alpha, \beta, \ldots$.
- Boundary indices are $a, b, \ldots$. Then $\alpha = (z, a)$.

- The extremal surface is parameterized by functions $\bar{X}^\mu(z, y^a)$. We choose a gauge such that $X^z = z$, and expand the remaining coordinates as

$$\bar{X}^i = X^i_{(0)} + z^2 X^i_{(2)} + z^4 X^i_{(4)} + \cdots + z^d \log z X^i_{(d,\log)} + z^d X^i_{(d)} + \cdots . \tag{A.1.2}$$

  The coefficients $X^i_{(n)}$ for $n < d$ and $X^i_{(d,\log)}$ are determined in terms of $X^i_{(0)}$ and $g^{(0)}_{ij}$, while $X^i_{(d)}$ is state-dependent and is related to the renormalized entropy of the CFT region.

- The extremal surface induced metric will be denoted $\bar{h}_{\alpha\beta}$ and gauge-fixed so that $\bar{h}_{za} = 0$.

- The entangling surface induced metric will be denoted h_{ab}.

- Note that we will often want to expand bulk quantities in z at fixed y instead of fixed x. For instance, the bulk metric at fixed y is

$$\begin{aligned}\bar{g}_{ij}(y, z) &= \bar{g}_{ij}(\bar{X}(z, y), z) = \bar{g}_{ij}(X_{(0)}(y) + z^2 X_{(2)}(y) + \cdots , z) \\ &= g^{(0)}_{ij} + z^2 \left(g^{(2)}_{ij} + X^m_{(2)} \partial_m g^{(0)}_{ij} \right) + \cdots \end{aligned} \tag{A.1.3}$$

  Similar remarks apply for things like Christoffel symbols. The prescription is to always compute the given quantity as a function of x first, the plug in $\bar{X}(y, z)$ and expand in a Taylor series.

Intrinsic and Extrinsic Geometry

Now will introduce several geometric quantities, and their notations, which we will need. First, we define a basis of surface tangent vectors by

$$e^i_a = \partial_a X^i. \tag{A.1.4}$$

We will also make use of the convention that ambient tensors which are not inherently defined on the surface but are written with surface indices (a, b, etc.) are defined by contracting with e^i_a. For instance:

$$g^{(2)}_{aj} = e^i_a g^{(2)}_{ij}. \tag{A.1.5}$$

We can form the surface projector by contracting the surface indices on two copies of e^i_a:

$$P^{ij} = h^{ab} e^i_a e^j_b = e^i_a e^{ja}. \tag{A.1.6}$$

We introduces a surface covariant derivative D_a that acts as the covariant derivative on both surface and ambient indices. So it is compatible with both metrics:

$$D_a h_{bc} = 0 = D_a g_{ij}. \tag{A.1.7}$$

Note also that when acting on objects with only ambient indices, we have the relationship

$$D_a V^{ij\cdots}_{pq\cdots} = e^m_a \nabla_m V^{ij\cdots}_{pq\cdots}, \tag{A.1.8}$$

where ∇_i is the ambient covariant derivative compatible with g_{ij}.

The extrinsic curvature is computed by taking the D_a derivative of a surface basis vector:

$$K^i_{ab} = -D_a e^i_b = -\partial_a e^i_b + \gamma^c_{ab} e^i_b - \Gamma^i_{ab}. \tag{A.1.9}$$

Note the overall sign we have chosen. Here γ^c_{ab} is the Christoffel symbol of the metric h_{ab}, and the lower indices on the Γ symbol were contracted with two basis tangent vectors to turn them into surface indices. Note that K^i_{ab} is symmetric in its lower indices. It is an exercise to check that it is normal to the surface in its upper index:

$$e_{ic} K^i_{ab} = 0. \tag{A.1.10}$$

The trace of the extrinsic curvature is denoted by K^i:

$$K^i = h^{ab} K^i_{ab}. \tag{A.1.11}$$

Below we will introduce the null basis of normal vectors k^i and l^i. Then we can define expansion $\theta_{(k)}$ ($\theta_{(l)}$) and shear $\sigma^{(k)}_{ab}$ ($\sigma^{(l)}_{ab}$) as the trace and traceless parts of $k_i K^i_{ab}$ ($l_i K^i_{ab}$), respectively.

There are a couple of important formulas involving the extrinsic curvature. First is the Codazzi Equation, which can be computed from the commutator of covariant derivatives:

$$\begin{aligned} D_c K^i_{ab} - D_b K^i_{ac} &= (D_b D_c - D_c D_b) e^i_a \\ &= R^i{}_{abc} - r^d{}_{abc} e^i_d. \end{aligned} \tag{A.1.12}$$

Here $R^i{}_{abc}$ is the ambient curvature (appropriately contracted with surface basis vectors), while $r^d{}_{abc}$ is the surface curvature. We can take traces of this equation to get others. Another useful thing to do is contract this equation with e^i_d and differentiate by parts, which yields the Gauss–Codazzi equation:

$$K_{cdi} K^i_{ab} - K_{bdi} K^i_{ac} = R_{dabc} - r_{dabc}. \tag{A.1.13}$$

Various traces of this equation are also useful.

Null Normals k and l

A primary object in our analysis is the bull vector k^i, which is orthogonal to the entangling surface and gives the direction of the surface deformation. It will be convenient to also introduce the null normal l^i, which is defined so that $l_i k^i = +1$. This choice of sign is different from the one that is usually made in these sorts of analysis, but it is necessary to

avoid a proliferation of minus signs. With this convention, the projector onto the normal space of the surface is

$$N^{ij} \equiv g^{ij} - P^{ij} = k^i l^j + k^j l^i = 2k^{(i} l^{j)}. \tag{A.1.14}$$

As we did with the tangent vectors e^i_a, we will introduce a shorthand notation to denote contraction with k^i or l^i: any tensor with k or l index means it has been contracted with k^i or l^i. As such we will avoid using the letters k and l as dummy indices. For instance.

$$R_{kl} \equiv k^i l^j R_{ij}. \tag{A.1.15}$$

Another quantity associated with k^i and l^i is the normal connection w^a, defined through

$$w_a \equiv l_i D_a k^i. \tag{A.1.16}$$

With this definition, the tangent derivative of k^i can be shown to be

$$D_a k^i = w_a k^i + K^k_{ab} e^{bi}, \tag{A.1.17}$$

which is a formula that is used repeatedly in our analysis.

At certain intermediate stages of our calculations it will be convenient to define extensions of k^i and l^i off of the entangling surface, so here we will define such an extension. Surface deformations in both the QNEC and QFC follow geodesics generated by k^i, so it makes sense to define k^i to satisfy the geodesic equation:

$$\nabla_k k^i = 0. \tag{A.1.18}$$

However, we will *not* define l^i by parallel transport along k^i. It is conceptually cleaner to maintain the orthogonality of l^i to the surface even as the surface is deformed along the geodesics generated by k^i. This means that l^i satisfies the equation

$$\nabla_k l^i = -w^a e^i_a. \tag{A.1.19}$$

These equations are enough to specify l^i and k^i on the null surface formed by the geodesics generated by k^i. To extend k^i and l^i off of this surface, we specify that they are both parallel-transported along l^i. In other words, the null surface generated by k^i forms the initial condition surface for the vector fields k^i and l^i which satisfy the differential equations

$$\nabla_l k^i = 0, \qquad \nabla_l l^i = 0\ . \tag{A.1.20}$$

This suffices to specify k^i an l^i completely in a neighborhood of the original entangling surface. Now that we have done that, we record the commutator of the two fields for future use:

$$[k, l]^i = \nabla_k l^i - \nabla_l k^i = -w^c e^i_c. \tag{A.1.21}$$

A.2 Surface Variations

Most of the technical parts of our analysis have to do with variations of surface quantities under the deformation $X^i \to X^i + \delta X^i$ of the surface embedding coordinates. Here δX^i should be interpreted a vector field defined on the surface. In principle it can include both normal and tangential components, but since tangential components do not actually correspond to physical deformations of the surface we will assume that δX^i is normal. The operator δ denotes the change in a quantity under the variation. In the case where $\delta X^i = \partial_\lambda X^i$, which is the case we are primarily interested in, δ can be identified with ∂_λ. With this in mind, we will always impose the geodesic equation on k^i whenever convenient. In terms of the notation we are introducing here, this is

$$\delta k^i = -\Gamma^i_{kk}. \tag{A.2.1}$$

To make contact with the main text, we will use the notation $k^i \equiv \delta X^i$, and assume that k^i is null since that is ultimately the case we care about. Some of the formulas we discuss below will not depend on the fact that k^i is null, but we will not make an attempt to distinguish them.

Ambient Quantities For ambient quantities, like curvature tensors, the variation δ can be interpreted straightforwardly as $k^i\partial_i$ with no other qualification. Thus we can freely use, for instance, the ambient covariant derivative ∇_k to simplify the calculations of these quantities. Note that δ itself is not the covariant derivative. As defined, δ is a coordinate dependent operator. This may be less-than-optimal from a geometric point of view, but it has the most conceptually straightforward interpretation in terms of the calculus of variations. In all of the variational formulas below, then, we will see explicit Christoffel symbols appear. Of course, ultimately these non-covariant terms must cancel out of physical quantities. That they do serves as a nice check on our algebra.

Tangent Vectors The most fundamental formula is that of the variation of the tangent vectors $e^i_a \equiv \partial_a X^i$. Directly from the definition, we have

$$\delta e^i_a = \partial_a k^i = D_a k^i - \Gamma^i_{ak} = w_a k^i + K^k_{ab} e^{bi} - \Gamma^i_{ak}. \tag{A.2.2}$$

This formula, together with the discussion of how ambient quantities transform, can be used together to compute the variations of many other quantities.

Intrinsic Geometry and Normal Vectors The intrinsic metric variation is easily computed from the above formula as

$$\delta h_{ab} = 2K^k_{ab}. \tag{A.2.3}$$

From here we can find the variation of the tangent projector, for instance:

$$\begin{aligned}\delta P^{ij} &= \delta h^{ab} e_a^i e_b^j + 2h^{ab} e_a^{(i} \partial_b k^{j)} \\ &= -2K_k^{ab} e_a^i e_b^j + 2h^{ab} e_a^{(i} D_b k^{j)} - 2h^{ab} e_a^{(i} \Gamma_{bk}^{j)} \\ &= 2w^a e_a^{(i} k^{j)} - 2h^{ab} e_a^{(i} \Gamma_{bk}^{j)}. \end{aligned} \tag{A.2.4}$$

Notice that the second line features a derivative of $k^i = \delta X^i$. In a context where we are taking functional derivatives, such as when computing equations of motion, this term would require integration by parts. We can write the last line covariantly as

$$\nabla_k P^{ij} = 2w^a e_a^{(i} k^{j)}. \tag{A.2.5}$$

Earlier we saw that l^i satisfied the equation $\nabla_k l^i = -w^a e_a^i$ as a result of keeping l^i orthogonal to the surface even as the surface is deformed. In the language of this section, this is seen by the following manipulation:

$$e_a^i \delta l_i = -l_i \partial_a k^i = -w_a - \Gamma_{ak}^l. \tag{A.2.6}$$

Again, note the derivative of k^i. It is easy to confirm that represents the only nonzero component of $\nabla_k l^i$.

The normal connection $w_a = l^i D_a k_i$ makes frequent appearances in our calculations, and we will need to know its variation. We can calculate that as follows:

$$\begin{aligned}\delta w_a &= \delta l^i D_a k_i + l^i \partial_a \delta k_i - l^i \delta \Gamma_{ji}^n e_a^j k_n - l^i \Gamma_{ji}^n \partial_a k^j k_n - l^i \Gamma_{ji}^n e_a^j \delta k_n \\ &= \nabla_k l^i D_a k_i + R_{klak} \\ &= -w^c K_{ac} + R_{klak}. \end{aligned} \tag{A.2.7}$$

Extrinsic Curvatures The simplest extrinsic curvature variation is that of the trace of the extrinsic curvature

$$\delta K^i = -K^m \Gamma_{mk}^i - D_a D^a k^i - R_{mkj}^i P^{mj} + \left(2D^a(K_{ad}^k) - D_d(K^k)\right) e^{di} - 2K_k^{ab} K_{ab}^i \tag{A.2.8}$$

Note that the combination $\delta K^i + K^k \Gamma_{km}^i k^m$ is covariant, so it makes sense to write

$$\nabla_k K^i = -D_a D^a k^i - R_{mkj}^i P^{mj} + \left(2D^a(K_{ad}^k) - D_d(K^k)\right) e^{di} - 2K_k^{ab} K_{ab}^i \tag{A.2.9}$$

This formula is noteworthy because of the first term, which features derivatives of $k^i = \delta X^i$. This is important because when K^i occurs inside of an integral and we want to compute the functional derivative then we have to first integrate by parts to move those derivatives off of k^i. This issue arises when computing Θ as in the QFC, for instance.

We can contract the previous formulas with l^i and k^i to produce other useful formulas. For instance, contracting with k^i leads to

$$\delta K^k = -K^{kab} K_{ab}^k - R_{kk}, \tag{A.2.10}$$

which is nothing but the Raychaudhuri equation.

The variation of the full extrinsic curvature K^i_{ab} is quite complicated, but we will not needed. However, its contraction with k^i will be useful and so we record it here:

$$k_i \delta K^i_{ab} = -K^j_{ab}\Gamma^m_{jn} k_m k^n - k_i D_a D_b k^i - R_{kakb}. \tag{A.2.11}$$

A.3 z-Expansions

Bulk Metric

We are focusing on bulk theories with gravitational Lagrangians

$$\mathcal{L} = \frac{1}{16\pi G_N}\left(\frac{d(d-1)}{\tilde{L}^2} + \mathcal{R} + \ell^2\lambda_1\mathcal{R}^2 + \ell^2\lambda_2\mathcal{R}^2_{\mu\nu} + \ell^2\lambda_{\rm GB}\mathcal{L}_{\rm GB}\right). \tag{A.3.1}$$

where $\mathcal{L}_{GB} = \mathcal{R}^2_{\mu\nu\rho\sigma} - 4\mathcal{R}^2_{\mu\nu} + \mathcal{R}^2$ is the Gauss-Bonnet Lagrangian, ℓ is the cutoff length scale of the bulk effective field theory, and the couplings λ_1, λ_2, and $\lambda_{\rm GB}$ are defined to be dimensionless. We have decided to include $\mathcal{L}_{GB}$ as part of our basis of interactions rather than $\mathcal{R}^2_{\mu\nu\rho\sigma}$ because of certain nice properties that the Gauss-Bonnet term has, but this is not important.

We recall that the Fefferman–Graham form of the metric is defined by

$$ds^2 = \frac{1}{z^2}(dz^2 + \bar{g}_{ij}dx^i dx^j), \tag{A.3.2}$$

where $\bar{g}_{ij}(x,z)$ is expanded as a series in z:

$$\bar{g}_{ij} = g^{(0)}_{ij} + z^2 g^{(2)}_{ij} + z^4 g^{(4)}_{ij} + \cdots + z^d \log z\, g^{(d,\log)}_{ij} + z^d g^{(d)}_{ij} + \cdots . \tag{A.3.3}$$

In principle, one would evaluate the equation of motion from the above Lagrangian using the Fefferman–Graham metric form as an ansatz to compute these coefficients. The results of this calculation are largely in the literature, and we quote them here. To save notational clutter, in this section we will set $g_{ij} = g^{(0)}_{ij}$.

The first nontrivial term in the metric expansion is independent of the higher-derivative couplings, and in fact is completely determined by symmetry [54]:

$$g^{(2)}_{ij} = -\frac{1}{d-2}\left(R_{ij} - \frac{1}{2(d-1)}Rg_{ij}\right). \tag{A.3.4}$$

The next term is also largely determined by symmetry, except for a pair of coefficients [54]. We are only interested in the kk-component of $g^{(4)}_{ij}$, and where one of the coefficients drops out. The result is

$$\begin{aligned} g^{(4)}_{kk} = \frac{1}{d-4}\Bigg[&\kappa C_{kijm}C_k{}^{ijm} + \frac{1}{8(d-1)}\nabla^2_k R - \frac{1}{4(d-2)}k^i k^j \Box R_{ij} \\ &- \frac{1}{2(d-2)}R^{ij}R_{kikj} + \frac{d-4}{2(d-2)^2}R_{ki}R^i_k + \frac{1}{(d-1)(d-2)^2}RR_{kk}\Bigg], \end{aligned} \tag{A.3.5}$$

where C_{ijmn} is the Weyl tensor and

$$\kappa = -\lambda_{GB}\frac{\ell^2}{L^2}\left(1 + O\left(\frac{\ell^2}{L^2}\right)\right). \tag{A.3.6}$$

In $d = 4$ we will need an expression for $g_{kk}^{(4,\log)}$ as well. One can check that this is obtainable from $g_{kk}^{(4)}$ by first multiplying by $4 - d$ and then setting $d \to 4$. We record the answer for future reference:

$$g_{kk}^{(4,\log)} = -\left[\kappa C_{kijm}C_k^{\ ijm} + \frac{1}{24}\nabla_k^2 R - \frac{1}{8}k^i k^j \Box R_{ij} - \frac{1}{4}R^{ij}R_{kikj} + \frac{1}{12}RR_{kk}\right]. \tag{A.3.7}$$

Extremal Surface Coordinates

The extremal surface position is determined by extremizing the generalized entropy functional [6, 49]:

$$S_{\text{gen}} = \frac{1}{4G_N}\int\sqrt{\bar{h}}\left[1 + 2\lambda_1\ell^2\mathcal{R} + \lambda_2\ell^2\left(\mathcal{R}_{\mu\nu}\mathcal{N}^{\mu\nu} - \frac{1}{2}\mathcal{K}_\mu\mathcal{K}^\mu\right) + 2\lambda_{\text{GB}}\ell^2\bar{r}\right] + S_{\text{bulk}}. \tag{A.3.8}$$

Here we are using $\mathcal{K}^i$ to denote the extrinsic curvature and $\bar{r}$ the intrinsic Ricci scalar of the surface.

The equation of motion comes from varying S_{gen} and is (ignoring the S_{bulk} term for simplicity)

$$\begin{aligned}0 = {}& \mathcal{K}^\mu\left[1 + 2\lambda_1\ell^2\mathcal{R} + \lambda_2\ell^2\left(\mathcal{R}_{\rho\nu}\mathcal{N}^{\rho\nu} - \frac{1}{2}\mathcal{K}_\rho\mathcal{K}^\rho\right) + 2\lambda_{\text{GB}}\ell^2\bar{r}\right] + 2\lambda_1\ell^2\nabla^\mu\mathcal{R} \\ & + \lambda_2\ell^2\Big(\mathcal{N}^{\rho\nu}\nabla^\mu\mathcal{R}_{\rho\nu} + 2\mathcal{P}^{\rho\nu}\nabla_\rho\mathcal{R}^\mu_\nu - 2\mathcal{R}^\mu_\rho\mathcal{K}^\rho + 2\mathcal{K}^{\mu\alpha\beta}\mathcal{R}_{\alpha\beta} + D_\alpha D^\alpha\mathcal{K}^\mu \\ & + \mathcal{K}^\rho\mathcal{R}_{\mu\sigma\rho\nu}\mathcal{P}^{\nu\sigma} + 2\mathcal{K}^{\mu\alpha\beta}\mathcal{K}_\nu\mathcal{K}^\nu_{\alpha\beta}\Big) - 4\lambda_{\text{GB}}\ell^2\bar{r}^{\alpha\beta}\mathcal{K}^\mu_{\alpha\beta}.\end{aligned} \tag{A.3.9}$$

This equation is very complicated, but since we are working in $d \leq 5$ dimensions we only need to solve perturbatively in z for $X^i_{(2)}$ and $X^i_{(4)}$[1]. Furthermore, $X^i_{(2)}$ is fully determined by symmetry to be [207]

$$X^i_{(2)} = \frac{1}{2(d-2)}D^a\partial_a X^i_{(0)} = -\frac{1}{2(d-2)}K^i, \tag{A.3.10}$$

where K^i denotes the extrinsic curvature of the $X^i_{(0)}$ surface, but we are leaving off the (0) in our notation to save space.

[1]It goes without saying that these formulas are only valid for $d > 2$ and $d > 4$, respectively.

The computation of $X^i_{(4)}$ is straightforward but tedious. We will only need to know $k_i X^i_{(4)}$ (where indices are being raised and lowered with $g^{(0)}_{ij}$), and the answer turns out to be

$$\begin{aligned}4(d-4)X^k_{(4)} &= 2X^k_{(2)}\left(P^{jm}g^{(2)}_{jm} - 4(X_{(2)})^2\right)\\ &\quad + K^k_{ab}g^{ab}_{(2)} + 4g^{(2)}_{km}X^m_{(2)} + 2X^{(2)}_j K^j_{ab}K^{kab} + k_i D_a D^a X^i_{(2)}\\ &\quad + k^j(\nabla_n g^{(2)}_{jm} - \frac{1}{2}\nabla_j g^{(2)}_{mn})P^{mn} + X^n_{(2)}R_{kmnj}P^{jm}\\ &\quad + 8\kappa\sigma^{ab}_{(k)}C_{kalb} - 2(d-4)\Gamma^k_{jm}X^j_{(2)}X^m_{(2)}. \end{aligned} \tag{A.3.11}$$

Here κ depends on $\lambda_{\rm GB}$ as in (A.3.6). Notice that the last term in this expression is the only source of noncovariant-ness. One can confirm that this noncovariant piece is required from the definition of $X^i_{(4)}$—despite its index, $X^i_{(4)}$ does not transform like a vector under boundary diffeomorphisms.

We also note that the terms in $X^k_{(4)}$ with covariant derivatives of $g^{(2)}_{ij}$ can be simplified using the extended k^i and l^i fields described §A.1 and the Bianchi identity:

$$k^j(\nabla_n g^{(2)}_{jm} - \frac{1}{2}\nabla_j g^{(2)}_{mn})P^{mn} = -\frac{1}{4(d-1)}\nabla_k R + \frac{1}{d-2}\nabla_l R_{kk}. \tag{A.3.12}$$

Finally, we record here the formula for $X^k_{(4,\log)}$ which is obtained from $X^k_{(4)}$ by multiplying by $4-d$ and sending $d \to 4$:

$$\begin{aligned}-4X^k_{(4,\log)} &= 2X^k_{(2)}\left(P^{jm}g^{(2)}_{jm} - 4(X_{(2)})^2\right)\\ &\quad + K^k_{ab}g^{ab}_{(2)} + 4g^{(2)}_{km}X^m_{(2)} + 2X^{(2)}_j K^j_{ab}K^{kab} + k_i D_a D^a X^i_{(2)}\\ &\quad + k^j(\nabla_n g^{(2)}_{jm} - \frac{1}{2}\nabla_j g^{(2)}_{mn})P^{mn} + X^n_{(2)}R_{kmnj}P^{jm}\\ &\quad + 8\kappa\sigma^{ab}_{(k)}C_{kalb}. \end{aligned} \tag{A.3.13}$$

We will not bother unpacking all of the definitions, but the main things to notice is that the noncovariant part disappears.

A.4 Details of the EWN Calculations

In this section we provide some insight into the algebra necessary to complete the calculations of the main text, primarily regarding the calculation of the subleading part of $(\delta\bar{X})^2$ in §3.2. The task is to simplify (3.2.13),

$$\begin{aligned}L^{-2}(\delta\bar{X})^2\big|_{z^2} &= 2k_i\delta X^i_{(4)} + 2g^{(2)}_{ij}k^i\delta X^j_{(2)} + g_{ij}\delta X^i_{(2)}\delta X^j_{(2)} + g^{(4)}_{ij}k^i k^j + X^m_{(4)}\partial_m g_{ij}k^i k^j\\ &\quad + 2X^m_{(2)}\partial_m g_{ij}k^i\delta X^j_{(2)} + X^m_{(2)}\partial_m g^{(2)}_{ij}k^i k^j + \frac{1}{2}X^m_{(2)}X^n_{(2)}\partial_m\partial_n g_{ij}k^i k^j. \end{aligned} \tag{A.4.1}$$

After some algebra, we can write this as

$$L^{-2}(\delta\bar{X})^2\big|_{z^2} = g^{(4)}_{kk} + 2\delta(X^k_{(4,\mathrm{cov})}) + 2g^{(2)}_{ik}\nabla_k X^i_{(2)} + \nabla_k X^{(2)}_j \nabla_k X^j_{(2)} - \frac{1}{d-2}(X^l_{(2)})\nabla_k R_{kk}. \tag{A.4.2}$$

Here we have defined

$$X^i_{(4,\mathrm{cov})} = X^i_{(4)} + \frac{1}{2}\Gamma^i_{lm}X^l_{(2)}X^m_{(2)}, \tag{A.4.3}$$

which transforms like a vector (unlike $X^i_{(4)}$). From here, the algebra leading to (3.2.14) is mostly straightforward, though tedious. The two main tasks which require further explanation are the simplification of one of the terms in $g^{(4)}_{kk}$ and one of the terms in $\delta X^k_{(4,\mathrm{cov})}$. We will explain those now.

$g^{(4)}_{kk}$ Simplification We recall the formula for $g^{(4)}_{kk}$ from (A.3.5):

$$\begin{aligned} g^{(4)}_{kk} = \frac{1}{d-4}\Bigg[&\kappa C_{kijm}C_k^{\ ijm} + \frac{1}{8(d-1)}\nabla_k^2 R - \frac{1}{4(d-2)}k^ik^j\Box R_{ij} \\ &-\frac{1}{2(d-2)}R^{ij}R_{kikj} + \frac{d-4}{2(d-2)^2}R_{ki}R^i_k + \frac{1}{(d-1)(d-2)^2}RR_{kk}\Bigg]. \end{aligned} \tag{A.4.4}$$

The main difficulty is with the term $k^ik^j\Box R_{ij}$. We will rewrite this term by making use of the geometric quantities introduced in the other appendices, and in particular we make use of the extended k and l field from §A.1. We first separate it into two terms:

$$k^ik^j\Box R_{ij} = k^ik^jN^{rs}\nabla_r\nabla_s R_{ij} + k^ik^jP^{rs}\nabla_r\nabla_s R_{ij}. \tag{A.4.5}$$

Now we compute each of these terms individually:

$$\begin{aligned} k^ik^jN^{rs}\nabla_r\nabla_s R_{ij} &= 2k^ik^jl^s\nabla_k\nabla_s R_{ij} + 2R_{kmlk}R^m_k \\ &= 2\nabla_k\nabla_l R_{kk} + 2w^ck^ik^jD_cR_{ij} + 2R_{kmlk}R^m_k \\ &= 2\nabla_k\nabla_l R_{kk} + 2w^cD_cR_{kk} - 4w^cw_cR_{kk} - 4w^cK^a_{ck}R_{ka} + 2R_{kmlk}R^m_k \\ &= 2\nabla_k\nabla_l R_{kk} + 2w^cD_cR_{kk} - 4w^cw_cR_{kk} + 2R_{kmlk}R^m_k. \end{aligned} \tag{A.4.6}$$

In the last line we assumed that $\sigma_{(k)} = 0$ and $\theta_{(k)} = 0$, which is the only case we will need to worry about. The other term is slightly messier, becoming

$$\begin{aligned} k^ik^jP^{rs}\nabla_r\nabla_s R_{ij} &= k^ik^je^{sc}D_c\nabla_s R_{ij} \\ &= D_c(k^ik^jD^cR_{ij}) - D_c(k^ik^je^{sc})\nabla_s R_{ij} \\ &= D_c(k^ik^jD^cR_{ij}) - 2w_cD^cR_{kk} + 4w_cw^cR_{kk} + 6w_cK^{ca}_kR_{ak} \\ &\quad - 2K^{ca}_kD_cR_{ka} + 2K^{ca}_kK^i_{ca}R_{ik} + 2K^{ca}_kK^{bk}_cR_{ab} + K^s\nabla_s R_{kk} \\ &= D_cD^cR_{kk} - 2D_c(w^cR_{kk}) - 2D_c(K^{cak}R_{ka}) - 2w_cD^cR_{kk} + 4w_cw^cR_{kk} + 6w_cK^{ca}_kR_{ak} \\ &\quad - 2K^{ca}_kD_cR_{ka} + 2K^{ca}_kK^i_{ca}R_{ik} + 2K^{ca}_kK^{bk}_cR_{ab} + K^s\nabla_s R_{kk} \\ &= D_cD^cR_{kk} - 2D_c(w^cR_{kk}) - 2D_c(K^{cak})R_{ka} - 2w_cD^cR_{kk} + 4w_cw^cR_{kk} + K^s\nabla_s R_{kk}. \end{aligned} \tag{A.4.7}$$

In the last line we again assumed that $\sigma_{(k)} = 0$ and $\theta_{(k)} = 0$. Putting the two terms together leads to some canellations:

$$\begin{aligned} k^i k^j \Box R_{ij} = {} & 2\nabla_k \nabla_l R_{kk} + 2R_{kmlk} R^m_k + D_c D^c R_{kk} - 2D_c(w^c R_{kk}) \\ & - 2(D_a \theta_{(k)} + R_{kcac}) R^a_k + K^s \nabla_s R_{kk}. \end{aligned} \tag{A.4.8}$$

$\delta X^k_{(4,\mathrm{cov})}$ **Simplification** The most difficult term in (A.3.11), which also gives the most interesting results, is

$$k_i D_a D^a X^i_{(2)} = -\frac{1}{2(d-2)} (D_a - w_a)^2 \theta_{(k)} + \frac{1}{2(d-2)} K_{ab} K^{abi} K_i. \tag{A.4.9}$$

The interesting part here is the first term, so we will take the rest of this section to discuss its variation. The underlying formula is (A.2.7),

$$\delta w_a = -w^c K_{ac} + R_{klak}. \tag{A.4.10}$$

From this we can compute the following related variations, assuming that $\theta_{(k)} = 0$ and $\sigma_{(k)} = 0$:

$$\delta(D^a w_a) = D^a R_{klak} + w^a \partial_a \theta_{(k)} - 3D_a(K^{ab}_k w_b) \tag{A.4.11}$$

$$\delta(w^a D_a \theta_{(k)}) = -3K^{ab}_k w_a D_b \theta_{(k)} + R_{klak} D^a \theta_{(k)} + w^a D_a \dot{\theta}_{(k)} \tag{A.4.12}$$

$$\delta(D^a D_a \theta_{(k)}) = D^a D_a \dot{\theta} - \partial_a \theta_{(k)} \partial^a \theta_{(k)} - 2P^{jm} R_{kjbm} D^b \theta_{(k)}. \tag{A.4.13}$$

Here $\dot{\theta}_{(k)} \equiv \delta\theta_{(k)}$ is given by the Raychaudhuri equation. We can combine these equations to get

$$\begin{aligned} \delta\left((D_a - w_a)^2 \theta_{(k)}\right) &= \delta\left(D^a D_a \theta_{(k)}\right) - 2\delta\left(w^a D_a \theta_{(k)}\right) - \delta\left((D_a w^a)\theta_{(k)}\right) + \delta\left(w_a w^a \theta_{(k)}\right) \\ &= -D^a D_a R_{kk} + 2w^a D_a R_{kk} + (D_a w^a) R_{kk} - w_a w^a R_{kk} \\ &\quad - \frac{d}{d-2} (D_a \theta_{(k)})^2 - 2R_{kb} D^b \theta_{(k)} - 2(D\sigma)^2. \end{aligned} \tag{A.4.14}$$

A.5 The $d = 4$ Case

As mentioned in the main text, many of our calculations are more complicated in even dimensions, though most of the end results are the same. The only nontrivial even dimension we study is $d = 4$, so in this section we record the formulas and special derivations necessary for understanding the $d = 4$ case. Some of these have been mentioned elsewhere already, but we repeat them here so that they are all in the same place.

Log Terms In $d = 4$ we get log terms in the extremal surface, the metric, and the EWN inequality. By looking at the structure of the extremal surface equation, it's easy to see that the log term in in the extremal surface is related to $X^i_{(4)}$ in $d \neq 4$ by first multipling by $4-d$ and then setting $d \to 4$. The result was recorded in (A.3.13), and we repeat it here:

$$\begin{aligned}-4X^k_{(4,\log)} &= 2X^k_{(2)}\left(P^{jm}g^{(2)}_{jm} - 4(X_{(2)})^2\right) \\ &+ K^k_{ab}g^{ab}_{(2)} + 4g^{(2)}_{km}X^m_{(2)} + 2X^{(2)}_j K^j_{ab}K^{kab} + k_i D_a D^a X^i_{(2)} \\ &+ k^j(\nabla_n g^{(2)}_{jm} - \frac{1}{2}\nabla_j g^{(2)}_{mn})P^{mn} + X^n_{(2)}R_{kmnj}P^{jm} \\ &+ 8\kappa\sigma^{ab}_{(k)}C_{kalb}. \end{aligned} \tag{A.5.1}$$

There is a similar story for $g^{(4,\log)}_{kk}$, which was recorded earlier in (A.3.7):

$$g^{(4,\log)}_{kk} = -\left[\kappa C_{kijm}C_k^{\ ijm} + \frac{1}{24}\nabla^2_k R - \frac{1}{8}k^i k^j \Box R_{ij} - \frac{1}{4}R^{ij}R_{kikj} + \frac{1}{12}RR_{kk}\right]. \tag{A.5.2}$$

From these two equations, it is easy to see that the log term in $(\delta\bar{X})^2$ has precisely the same form as the subleading EWN inequality (3.2.14) in $d \geq 5$, except we first multiply by $4-d$ and then set $d \to 4$. This results in

$$L^{-2}(\delta\bar{X})^2\big|_{z^2\log z, d=4} = -\frac{1}{4}(D_a\theta_{(k)} + R_{ka})^2 - \frac{1}{4}(D_a\sigma^{(k)}_{bc})^2. \tag{A.5.3}$$

Note that the Gauss-Bonnet term drops out completely due to special identities of the Weyl tensor valid in $d = 4$ [52]. The overall minus sign is important because $\log z$ should be regarded as negative.

QNEC in Einstein Gravity For simplicity we will only discuss the case of Einstein gravity for the QNEC in $d = 4$, so that the entropy functional is just given by the extremal surface area divided by $4G_N$. At order z^2, the norm of $\delta\bar{X}^\mu$ is formally the same as the expression in other dimensions:

$$L^{-2}(\delta\bar{X})^2\big|_{z^2} = g^{(4)}_{kk} + 2g^{(2)}_{ik}\nabla_k X^i_{(2)} + \nabla_k X^{(2)}_j \nabla_k X^j_{(2)} - \frac{1}{2}X^l_{(2)}\nabla_k R_{kk} + 2\delta(k_i X^i_{(4)\text{cov}}). \tag{A.5.4}$$

Now, though, $X^k_{(4)}$ and $g^{(4)}_{kk}$ are state-dependent and must be related to the entropy and energy-momentum, respectively.

We begin with the entropy. From the calculus of variations, we know that the variation of the extremal surface area is given by

$$\delta A = -\lim_{\epsilon\to 0}\frac{L^3}{\epsilon^3}\int\sqrt{h}\frac{1}{\sqrt{1 + g_{nm}\partial_z\bar{X}^n\partial_z\bar{X}^m}}g_{ij}\partial_z\bar{X}^i\delta X^j. \tag{A.5.5}$$

A few words about this formula are required. The $\bar{X}^\mu$ factors appearing here must be expanded in ϵ, but the terms without any (n) in their notation do *not* refer to (0), unlike elsewhere in this paper. The reason is that we have to do holographic renormalization carefully at this stage, and that means the boundary conditions are set at $z = \epsilon$. So when we expand out $\bar{X}^\mu$ we will find its coefficients determined by the usual formulas in terms of $X^i_{(0)}$. We need to then solve for $X^i_{(0)}$ in term of $X^i \equiv \bar{X}^i(z = \epsilon)$ re-express the result in terms of X^i alone. Since we are not in a high dimension this task is relatively easy. An intermediate result is

$$\frac{k^i}{L^3\sqrt{h}}\frac{\delta A}{\delta X^i}\bigg|_{\epsilon^0} = -2\, X^k_{(2)}\big|_{\epsilon^2} - 4\left(X^k_{(4)} - (X_{(2)})^2 X^k_{(2)}\right) - X^k_{(4,\log)}. \tag{A.5.6}$$

The notation on the first term refers to the order ϵ^2 part of $X^i_{(2)}$ that is generated when $X^i_{(2)}$ is written in terms of $\bar{X}^i(z = \epsilon)$. The result of that calculation is

$$\begin{aligned} -4\, X^k_{(2)}\big|_{\epsilon^2} &= 2X^{(2)}_j K^{jab} K^i_{ab} k_i + k_i D^b D_b X^i_{(2)} + K^m \Gamma^i_{ml} X^l_{(2)} k_i \\ &\quad + g^{ab}_{(2)} K^i_{ab} k_i + P^{kj} R^i_{jmk} X^m_{(2)} k_i + k^m \left(\nabla_j g^{(2)}_{mk} - \frac{1}{2}\nabla_m g^{(2)}_{jk}\right) P^{jk} \\ &= -4X^k_{(4,\log)} - 2X^k_{(2)}\left(P^{jm} g^{(2)}_{jm} - 4(X_{(2)})^2\right) - 4g^{(2)}_{km} X^m_{(2)} + K^m \Gamma^i_{ml} X^l_{(2)} k_i. \end{aligned} \tag{A.5.7}$$

We have dropped terms of higher order in ϵ. Thus we can write

$$\frac{k^i}{L^3\sqrt{h}}\frac{\delta A}{\delta X^i}\bigg|_{\epsilon^0} = -3X^k_{(\log)} - X^k_{(2)} P^{jm} g^{(2)}_{jm} + 8X^k_{(2)}(X_{(2)})^2 - 2g^{(2)}_{km} X^m_{(2)} - 4X^k_{(4)\mathrm{cov}}. \tag{A.5.8}$$

We will want to take one more variation of this formula so that we can extract $\delta X^k_{(4)\mathrm{cov}}$. We can get some help by demanding that the $z^2 \log z$ part of EWN be saturated, which states

$$g^{(\log)}_{kk} + 2\delta X^k_{\log} = 0. \tag{A.5.9}$$

Then we have

$$\delta\left(\frac{k^i}{L^3\sqrt{h}}\frac{\delta A}{\delta X^i}\bigg|_{\epsilon^0}\right) = \frac{3}{2} g^{(\log)}_{kk} - \delta(X^k_{(2)} P^{jm} g^{(2)}_{jm}) + 8\delta(X^k_{(2)}(X_{(2)})^2) - 2\delta(g^{(2)}_{km} X^m_{(2)}) - 4\delta X^k_{(4)\mathrm{cov}}. \tag{A.5.10}$$

Assuming that $\theta_{(k)} = \sigma_{(k)} = 0$, we can simplify this to

$$\delta\left(\frac{k^i}{L^3\sqrt{h}}\frac{\delta A}{\delta X^i}\bigg|_{\epsilon^0}\right) = \frac{3}{2} g^{(\log)}_{kk} - \frac{1}{4} R_{kk} P^{jm} g^{(2)}_{jm} - \frac{1}{4}\nabla_k(\theta_{(l)} R_{kk}) - \frac{1}{2} g^{(2)}_{kl} R_{kk} - 4\delta X^k_{(4)\mathrm{cov}}. \tag{A.5.11}$$

We can combine this with the holographic renormalization formula [72]

$$\begin{aligned} g^{(4)}_{kk} &= 4\pi G_N L^{-3} T_{kk} + \frac{1}{2}(g^2_{(2)})_{kk} - \frac{1}{4} g^{(2)}_{kk} g^{ij} g^{(2)}_{ij} - \frac{3}{4} g^{(\log)}_{kk} \\ &= 4\pi G_N L^{-3} T_{kk} + \frac{1}{8} R^i_k R_{ik} - \frac{1}{16} R_{kk} R - \frac{3}{4} g^{(\log)}_{kk} \end{aligned} \tag{A.5.12}$$

to get

$$L^{-2}(\delta \bar{X}^i)^2\big|_{z^2} = 4\pi G_N L^{-3} T_{kk} - \frac{1}{2}\delta\left(\frac{k^i}{L^3\sqrt{h}} \frac{\delta A}{\delta X^i}\bigg|_{\epsilon^0} \right). \tag{A.5.13}$$

After dividing by $4G_N$, we recognize the QNEC.

A.6 Connections to the ANEC

In A.6 we briefly review the connection between the relative entropy and the ANEC. Equation (4.1.2) then implies an interesting connection between the off-diagonal second variation of the entropy and the ANEC. In A.6 we analyze this result in more detail for holographic field theory states dual to perturbative bulk geometries.

ANEC and Relative Entropy

As in Section 4.2, the region $\mathcal{R}$ is a region whose boundary $\partial\mathcal{R}$ lies in the $u = 0$ plane. We also consider a one-parameter family of such regions, indexed by λ, with the convention that increasing λ makes the $\mathcal{R}$ smaller. In this section we will focus on a globally pure state reduced to these regions. The relative entropy (with respect to the vacuum) and its first two derivatives obey the following set of alternating inequalities:

$$S_{\text{rel}} \geq 0, \qquad \frac{dS_{\text{rel}}}{d\lambda} \leq 0, \qquad \frac{d^2 S_{\text{rel}}}{d\lambda^2} \geq 0. \tag{A.6.1}$$

The first two of these are general properties of relative entropy in quantum mechanics, known as the positivity and monotonicity of relative entropy, respectively. The third inequality is the QNEC together with strong subadditivity.

We can also consider the entropy $\bar{S}$ and relative entropy $\bar{S}_{\text{rel}}$ of the complement of $\mathcal{R}$, which we will denote by $\bar{\mathcal{R}}$. Since we specified that the global state is pure, we have $\bar{S} = S$. The set of inequalities obeyed by $\bar{S}_{\text{rel}}$ is

$$\bar{S}_{\text{rel}} \geq 0, \qquad \frac{d\bar{S}_{\text{rel}}}{d\lambda} \geq 0, \qquad \frac{d^2 \bar{S}_{\text{rel}}}{d\lambda^2} \geq 0. \tag{A.6.2}$$

From (4.2.6) and the analogous equation for $\bar{S}_{\text{rel}}$, together with the monotonicity of relative entropy inequalities, we can conclude

$$\frac{d\bar{S}_{\text{rel}}}{d\lambda} - \frac{dS_{\text{rel}}}{d\lambda} = 2\pi \int d^{d-2}y dv \;\; \langle T_{vv} \rangle \, \dot{V}(y) \geq 0. \tag{A.6.3}$$

This is the ANEC, and its connection to relative entropy was first pointed out in [208, 27].

The relation (A.6.3) has interesting implications. Note that the integral of T_{vv} is completely independent of λ. If we let $\lambda \to \infty$, it must be the case that $dS_{\text{rel}}/d\lambda \to 0$ or else positivity of relative entropy will be violated. Similarly, as $\lambda \to -\infty$ we must have $d\bar{S}_{\text{rel}}/d\lambda \to 0$. Then we can say

$$\int_{-\infty}^{\infty} d\lambda \, \frac{d^2 S_{\text{rel}}}{d\lambda^2} = \frac{dS_{\text{rel}}}{d\lambda}(\infty) - \frac{dS_{\text{rel}}}{d\lambda}(-\infty) = 2\pi \int d^{d-2}y dv \;\; \langle T_{vv} \rangle \, \dot{V}(y). \tag{A.6.4}$$

From the definition of relative entropy, this means that

$$\int_{-\infty}^{\infty} d\lambda \int d^{d-2}y \; S'' \dot{V}(y)^2 = - \int_{-\infty}^{\infty} d\lambda \int d^{d-2}y d^{d-2}y' \;\; \left(\frac{\delta^2 S}{\delta V(y) \delta V(y')} \right)_{\text{od}} \dot{V}(y) \dot{V}(y'). \tag{A.6.5}$$

So the diagonal and off-diagonal parts of the second variation entropy contribute equally when integrated over the entire one-parameter family of surface deformations. Since there are two y integrals on the RHS of (A.6.5), naïvely one might have thought that a limiting case for $\dot{V}(y)$ existed which caused the RHS of this equation to vanish while leaving the LHS finite, but this is not true. We will say more about the order-of-limits involved in the holographic context below. Applying the relation $S''_{vv} = 2\pi \langle T_{vv} \rangle$ we see that, after integration, the off-diagonal variations can be related back to the ANEC:

$$2\pi \int d^{d-2}y dv \;\; \langle T_{vv} \rangle \dot{V}(y) = - \int_{-\infty}^{\infty} d\lambda \int d^{d-2}y d^{d-2}y' \;\; \left(\frac{\delta^2 S}{\delta V(y) \delta V(y')} \right)_{\text{od}} \dot{V}(y) \dot{V}(y'). \tag{A.6.6}$$

This is a nontrivial consequence of (4.1.2). Note that $\delta^2 S^{\text{od}} / \delta V(y) \delta V(y') \leq 0$ by strong subadditivity [11].

ANEC in a Perturbative Bulk

In this section we will investigate (A.6.6) in AdS/CFT for perturbative bulk states. Once again, we will drop the contributions of S_{bulk} for simplicity. This amounts to considering coherent states in the bulk.

From (4.3.4), we can see that for perturbative classical bulk states the bulk boost energy completely accounts for the off-diagonal entropy variation. Then from (4.3.7) we get

$$\frac{\delta^2 S^{\text{od}}}{\delta V(y_1) \delta V(y_2)} = -2\pi \left(\frac{2^{d-2} \Gamma(\frac{d-1}{2})}{\pi^{\frac{d-1}{2}}} \right)^2 \int \frac{dz d^{d-2}y}{z^{d-1}} \, \langle T^{\text{bulk}}_{vv} \rangle \frac{z^{2d}}{(z^2 + (y-y_1)^2)^{d-1} (z^2 + (y-y_2)^2)^{d-1}} \tag{A.6.7}$$

As a consequence of (A.6.6) we then have the equation

$$\int d^{d-2}y dv \;\; \langle T_{vv} \rangle \dot{V}(y) = \int \frac{dv dz d^{d-2}y}{z^{d-1}} \;\; \langle T^{\text{bulk}}_{vv} \rangle \dot{\bar{V}}(y, z). \tag{A.6.8}$$

This is a nontrivial matching between the ANEC on the boundary and an associated ANEC in the bulk, made possible by the relationship between $\dot{V}$ and $\dot{\tilde{V}}$ that comes from solving the extremal surface equation:

$$\dot{\tilde{V}}(y,z) = \frac{2^{d-2}\Gamma(\frac{d-1}{2})}{\pi^{\frac{d-1}{2}}} \int d^{d-2}y' \frac{z^d}{(z^2+(y-y')^2)^{d-1}} \dot{V}(y'). \tag{A.6.9}$$

We can get some intuition for these equations by considering shockwave solutions in the bulk.

Shockwaves Consider a shockwave geometry in the bulk. The bulk stress tensor is [209]

$$\langle T^{\text{bulk}}_{vv} \rangle = E z_0^{d-1} \delta(v) \delta^{d-2}(y) \delta(z - z_0) \tag{A.6.10}$$

and the boundary stress tensor is

$$\langle T_{vv} \rangle = E \frac{2^{d-2}\Gamma\left(\frac{d-1}{2}\right) z_0^d}{\pi^{\frac{d-1}{2}}(z_0^2+y^2)^{d-1}} \delta(v) \tag{A.6.11}$$

The parameters z_0 and E characterize the solution. One can see directly that (A.6.8) holds.

It is also interesting to integrate over a finite range of the deformation parameter. As the range is extended to infinity we recover (A.6.8), but for finite amounts of deformation we can see how the diagonal and off-diagonal parts of the entropy compete. We take the undeformed surface at $\lambda = 0$ to be the flat plane $V(y) = 0$ and we place the shockwave at $v = v_0$. Then integrating over a range of deformations about zero we find on the boundary

$$\begin{aligned} \int_0^\lambda d\lambda' \int d^{d-2}y \, \langle T_{vv} \rangle \dot{V}(y)^2 &= \int d^{d-2}y \, E \frac{2^{d-2}\Gamma\left(\frac{d-1}{2}\right) z_0^d}{\pi^{\frac{d-1}{2}}(z_0^2+y^2)^{d-1}} \dot{V}(y) \Theta(\lambda \dot{V}(y=0) - v_0) \\ &= E \dot{\tilde{V}}(y=0, z=z_0) \Theta(\lambda \dot{V}(y=0) - v_0). \end{aligned} \tag{A.6.12}$$

As soon as the integration range crosses $v = v_0$, the total energy jumps from zero to the final answer. On the other hand, in the bulk we get

$$\int_0^\lambda d\lambda' \int \frac{dz d^{d-2}y}{z^{d-1}} \langle T^{\text{bulk}}_{vv} \rangle \dot{\tilde{V}}(y,z)^2 = E \dot{\tilde{V}}(y=0, z=z_0) \Theta\left(\lambda \dot{\tilde{V}}(y=0, z=z_0) - v_0\right). \tag{A.6.13}$$

This is a very similar answer, but now the jump does not occur until later: $\dot{\tilde{V}}(y=0, z=z_0)$ will always be less than $\dot{V}(y)$, which means λ has to get larger. How much larger? We can estimate it by looking at the example of a bump function deformation with $\dot{V}(y) = 1$ over a region of area $\mathcal{A} \ll z_0^{d-2}$ and zero elsewhere. Then the boundary energy will register at $\lambda = v_0$, while the bulk energy will register at

$$\lambda = \frac{\pi^{\frac{d-1}{2}}}{2^{d-2}\Gamma(\frac{d-1}{2})} \frac{z_0^{d-2}}{\mathcal{A}} v_0 \gg v_0 \ . \tag{A.6.14}$$

So for very narrow deformations, the off-diagonal contributions to the entropy can only be seen when integrated over a large range of the deformation parameter. From the boundary point of view, the parameter z_0 controls how diffuse the energy is in the y-directions. It is a measure of the nonlocality of the state. The off-diagonal entropy variations are sensitive to this nonlocality.

Note that the order of limits we have discovered here is worth repeating. If we take $\mathcal{A} \to 0$ before taking $\lambda \to \infty$ then our integration will only be sensitive to the diagonal entropy variation (i.e., the boundary stress tensor) and we will find apparent violations of (A.6.6). The reason is that there are important contributions to the off-diagonal entropy variations when $\lambda \sim z_0^{d-2}/\mathcal{A}$, where z_0 controls the level of nonlocality in the state.

Superpositions of Shockwaves At linear order in the bulk perturbations we can take superpositions of shockwaves. This allows us to create any bulk and boundary bulk stress tensor profile along the $u = 0$ plane, and in that sense represents the most general state for the purpose of this calculation. The bulk and boundary stress tensors would be

$$\langle T_{vv}^{\text{bulk}}(y, z, v)\rangle = z^{d-1}\rho(y, z, v) \tag{A.6.15}$$

and

$$\langle T_{vv}(y, v)\rangle = \frac{2^{d-2}\Gamma\left(\frac{d-1}{2}\right)}{\pi^{\frac{d-1}{2}}} \int d^{d-2}y' dz' \rho(y', z', v) \frac{(z')^d}{((z')^2 + (y - y')^2)^{d-1}} \tag{A.6.16}$$

The single shockwave is the special case $\rho = E\delta(v)\delta^{d-2}(y)\delta(z - z_0)$. We can repeat some of the calculations we did before, but qualitatively the results will be the same. The deformed bulk extremal surface always "lags behind" the deformed entangling surface in a way that depends on z and the width of the deformation, and as a result the bulk energy flux at finite deformation parameters will always be less than the boundary energy flux. Taking the deformation width to zero at finite deformation parameters will cause the bulk energy flux to drop to zero. It would be interesting to characterize this behavior directly in the field theory without the bulk picture.

A.7 Free and Weakly-Interacting Theories

Our conjectures (6.3.8) and (4.1.2) are only meant to apply to interacting theories. In this appendix we will explain how the null-null relation (4.1.2) is violated in free theories, and indicate how it might be fixed when interactions are included.

The Case of Free Scalars

The case of free scalar fields for entangling surfaces restricted to $u = 0$ was analyzed extensively in [16], and we will make use of that analysis here. As in Section 4.2 we have a

one-parameter family of regions indexed by λ. The deformation velocity $\dot{V}(y)$ is taken to be a unit step-function with support on a small region of area $\mathcal{A}$ in the y-directions. The crucial point is to focus attention on the pencil of the $u = 0$ plane that is the support of $\dot{V}(y)$. As λ varies, the entangling surface moves within this pencil but stays fixed outside of it.

The State and the Entropy For the purpose of constructing the state, we can model the full theory as a $1 + 1$-dimensional massless chiral boson living on the pencil, together with an auxiliary system consisting of the rest of the $u = 0$ plane. This is the formalism of null quantization, which is reviewed in [16].

There are two facts we're going to use to write down the sate $\rho(\lambda)$ on the pencil+auxiliary system. First, in the limit of small $\mathcal{A}$, the state on the pencil becomes approximately disentangled from the auxiliary system. The fully-disentangled part $\mathcal{A}^0$ part of the state looks like the vacuum, while the leading correction goes like $\mathcal{A}^{1/2}$ and consists of single-particle states on the pencil entangled with states of the auxiliary system. The second fact is that we can always translate our state in the pencil by an amount λ so that the entangling surface is at the origin and the operators which create the state are displaced by an amount λ from their original positions. A coordinate system where the entangling surface is fixed is preferable. Putting these facts together lets us write

$$\rho(\lambda) = \rho_{\text{vac}} \otimes \left(\sum_i e^{-2\pi K_i} |i\rangle\langle i| \right) + \mathcal{A}^{1/2} \sum_{i,j} \rho_{ij}^{(1/2)}(\lambda) \otimes \left(e^{-\pi(K_i+K_j)/2} |i\rangle\langle j| \right) + \cdots \tag{A.7.1}$$

The states $|i\rangle$ of the auxiliary system are merely those which diagonalize the $\mathcal{A}^0$ part of ρ, and the K_i are numbers specifying the eigenvalues.

As indicated above the state $\rho_{ij}^{(1/2)}(\lambda)$ should be interpreted as a state on the half-line $x > 0$. We can write this state in terms of a Euclidean path integral in the complex plane:

$$\rho_{ij}^{(1/2)}[\phi^-, \phi^+] = \int_{\phi(x^+)=\phi^+}^{\phi(x^-)=\phi^-} \mathcal{D}\phi \; \mathcal{O}_{ij}(\lambda) e^{-S_{\text{E}}}, \tag{A.7.2}$$

where $\phi(x^\pm)$ refers to boundary conditions just above/below the positive real axis. The insertion $\mathcal{O}_{ij}(\lambda)$ is a single-field insertion which specifies the state:

$$\mathcal{O}_{ij}(\lambda) = \int dz d\bar{z} \; \psi_{ij}(z, \bar{z}) \partial\phi(z - \lambda). \tag{A.7.3}$$

As in [16] we will normalize our field so that $\langle \partial\phi(z)\partial\phi(0)\rangle_{\text{vac}} = -1/z^2$ and $T_{vv} = (\partial\phi)^2/4\pi\mathcal{A}$. Then one can show that $Q \equiv S''_{vv} - 2\pi T_{vv}$ is given by

$$Q(\lambda) = -\frac{1}{2} \sum_{ij} \left| \int dx d\tau \; (z - \lambda)^{-2+i\alpha_{ij}} \psi_{ij}(x, \tau) \right|^2 \frac{\pi(1 + \alpha_{ij}^2)\alpha_{ij}}{\sinh \pi\alpha_{ij}} e^{2\pi\alpha_{ij}} \tag{A.7.4}$$

where $\alpha_{ij} = K_i - K_j$ and if $z = re^{i\theta}$ with $0 \leq \theta < 2\pi$ then

$$z^{i\alpha} = r^{i\alpha} e^{-\alpha\theta}. \tag{A.7.5}$$

The quantity Q is manifestly negative, as required by the QNEC, but it is not zero.

Recovering the ANEC In Appendix A.6 we showed how one can recover the ANEC by integrating the QNEC on a globally pure state. In the present context, we don't have any off-diagonal contributions to the entropy. Instead we have the function Q, and repeating the argument above would lead us to conclude

$$\int_{-\infty}^{\infty} d\lambda \; Q(\lambda) = -2\pi \int d\lambda \; \langle T_{vv}(\lambda) \rangle. \tag{A.7.6}$$

We can check this equation by integrating (A.7.4). Note that the assumption of global purity that was used in Appendix A.6 is crucial: the expectation value of $T_{vv}(\lambda)$ depends only on the part of the state proportional to $\mathcal{A}$, which we have not specified and in principle has many independent parameters. For a globally pure state there is a relationship between that part of the state and the $\mathcal{A}^{1/2}$ part of the state which we must exploit.

In the pencil+auxiliary model, the global Hilbert space consists of the full pencil plus a doubled auxiliary system. The doubling allows the auxiliary state to be purified. Let the global pure state by $|\Psi\rangle$. Then we have

$$|\Psi\rangle = |\text{vac}\rangle \otimes \left(\sum_i e^{-\pi K_i} |i\rangle \otimes |i\rangle \right) + \mathcal{A}^{1/2} \sum_{i,j} e^{-\pi \alpha_{ij}/2} |\Psi_{ij}\rangle \otimes |i\rangle \otimes |j\rangle + \cdots \tag{A.7.7}$$

Any subsequent terms will not affect the ANEC. The factor of $\exp(-\pi\alpha_{ij}/2)$ is purely for future convenience, and the $|\Psi_{ij}\rangle$ are not necessarily normalized. The expectation value of the ANEC operator in this state is given by

$$2\pi \int d\lambda \; \langle T_{vv}(\lambda) \rangle_\Psi = 2\pi \mathcal{A} \sum_{i,j} e^{-\pi\alpha_{ij}} \int d\lambda \, \langle \Psi_{ij}| \, T_{vv}(\lambda) \, |\Psi_{ij}\rangle \, . \tag{A.7.8}$$

We can make contact with our earlier formulas by computing the density matrix $|\Psi\rangle\langle\Psi|$ and tracing over the second copy of the auxiliary system. We find that

$$\rho_{ij}^{(1/2)} = \text{Tr}_{x<0} \left(|\Psi_{ij}\rangle\langle \text{vac}| + |\text{vac}\rangle\langle\Psi_{ji}| \right). \tag{A.7.9}$$

This lets us identify the part of $\mathcal{O}_{ij}$ in the lower half-plane as the operator which creates $|\Psi_{ij}\rangle$. Then, in our previous notation, we find

$$2\pi \int d\lambda \; \langle T_{vv}(\lambda) \rangle_\Psi = 4\pi i \sum_{i,j} e^{-\pi\alpha_{ij}} \int dx d\tau dx' d\tau' \frac{\psi_{ij}(x,\tau)\psi_{ij}(x',\tau')^*}{(z-w^*)^3} \Theta(-\tau)\Theta(-\tau'). \tag{A.7.10}$$

Our job now is to reproduce this by integrating (A.7.4) with respect to λ. The main identity we will need is

$$\int_{-\infty}^{\infty} \frac{d\lambda}{(z-\lambda)^{2-i\alpha_{ij}}(w^*-\lambda)^{2+i\alpha_{ij}}} = \frac{4ie^{-2\pi\alpha_{ij}}\sinh\pi\alpha_{ij}}{\alpha_{ij}(1+\alpha_{ij}^2)(w^*-z)^3}\left(e^{\pi\alpha_{ij}}\Theta(\tau)\Theta(\tau') - e^{-\pi\alpha_{ij}}\Theta(-\tau)\Theta(-\tau')\right). \tag{A.7.11}$$

Using this formula, the integral of (A.7.4) splits into two terms. We may combine them by exchanging i and j in the first term, leaving us with

$$\begin{aligned}\int d\lambda\, Q(\lambda) &= -2\pi i \sum_{ij}\int dx d\tau dx' d\tau'\, \frac{\psi_{ij}(x,\tau)\psi_{ij}(x',\tau')^*}{(w^*-z)^3}\left(e^{\pi\alpha_{ij}}\Theta(\tau)\Theta(\tau') - e^{-\pi\alpha_{ij}}\Theta(-\tau)\Theta(-\tau')\right)\\ &= -4\pi i\sum_{ij} e^{-\pi\alpha_{ij}}\int dx d\tau dx' d\tau'\, \frac{\psi_{ij}(x,\tau)\psi_{ij}(x',\tau')^*}{(z-w^*)^3}\Theta(-\tau)\Theta(-\tau').\end{aligned} \tag{A.7.12}$$

Coherent States For coherent states we obtain a correspondence between Q and T_{vv} without integrating over λ. This must be true because coherent states satisfy $S''_{vv}=0$, but it is reassuring to see it happen explicitly. In a coherent state of the original d-dimensional theory, the pencil and auxiliary system factorize and the pencil is in a $1+1$-dimensional coherent state. In other words, we have

$$\rho(\lambda)[\phi^-,\phi^+] = \left(\int_{\phi(x^+)=\phi^+}^{\phi(x^-)=\phi^-}\mathcal{D}\phi\; e^{-S_{\rm E}+\mathcal{A}^{1/2}\mathcal{O}(\lambda)}\right)\otimes\left(\sum_i e^{-2\pi K_i}|i\rangle\langle i|\right). \tag{A.7.13}$$

We can obtain Q for this state by taking the general equation (A.7.4) specializing to the case where $\psi_{ij}=\psi\delta_{ij}\exp(-\pi K_i)$. Making use of the normalization condition $\sum_i \exp(-2\pi K_i)=1$ we find the simple expression

$$Q_{\rm coherent}(\lambda) = -\frac{1}{2}\left|\int dx d\tau\, \frac{\psi(x,\tau)}{(z-\lambda)^2}\right|^2 = -\frac{1}{2\mathcal{A}}\langle\partial\phi(\lambda)\rangle^2_{\rm coherent}. \tag{A.7.14}$$

We recognize this as simply $-2\pi\langle T_{vv}\rangle_{\rm coherent}$, as expected.

Weakly Interacting Effective Field Theories

In the main text we provided evidence for that $S''_{vv}=2\pi\langle T_{vv}\rangle$ for interacting theories, but in the previous section we explained that for free theories $Q=S''_{vv}-2\pi\langle T_{vv}\rangle$ was nonzero, and in fact could be quite large. In this section we will show how we can transition from $S''_{vv}\neq 2\pi\langle T_{vv}\rangle$ to $S''_{vv}=2\pi\langle T_{vv}\rangle$ when a weak coupling is turned on.[2]

The essential point is that one should always consider the total variation $d^2S/d\lambda^2$ as the primary physical quantity. S''_{vv} is a derived quantity obtained by considering a limiting case

[2]We thank Thomas Faulkner for first pointing out the arguments we present in this section.

of arbitrarily thin deformations. However, a weakly-coupled effective field theory in the IR comes with a cutoff scale ϵ, and we cannot reliably compute $d^2S/d\lambda^2$ for deformations of width $\ell \lesssim \epsilon$. Now we will see how this can resolve the issue.

In the free theory, as we have explained above, the second functional derivative of the entropy has the form

$$\frac{\delta^2 S_{\text{free}}}{\delta V(y)\delta V(y')} = 2\pi\langle T_{vv}\rangle\delta^{(d-2)}(y-y') + Q\delta^{(d-2)}(y-y') + \left(\frac{\delta^2 S}{\delta V(y)\delta V(y')}\right)_{\text{od}}. \quad \text{(A.7.15)}$$

The function Q is related to the square of the expectation value of the field $\partial\phi$. This is especially obvious in the formula for the coherent state, (A.7.14), but the more general formula is essentially of the same form. In a free theory $(\partial\phi)^2$ has dimension d and is exactly of the right form to contribute to a δ-function. This fact was touched upon in the Introduction. When we turn on a weak coupling g, the dimension of ϕ will shift to $\Delta_\phi = (d-2)/2 + \gamma(g)$.[3] There will still be a term in the second variation of the entropy associated to $(\partial\phi)^2$, which we will call Q_g, but now it no longer comes with a δ-function:

$$\frac{\delta^2 S_g}{\delta V(y)\delta V(y')} = 2\pi\,\langle T_{vv}\rangle\,\delta^{(d-2)}(y-y') + Q_g f_g(y-y') + (\text{other off-diagonal terms})\,. \quad \text{(A.7.16)}$$

Here f_g is some function of mass dimension $d-2-2\gamma$ which limits to a δ-function as $g\to 0$, such as $f_g(y) \sim \gamma/y^{d-2-2\gamma}$. So the Q_g term has migrated from the δ-function to the off-diagonal part of the entropy variation.

Now consider integrating (A.7.16) twice against a deformation profile of width ℓ and unit height to get a total second derivative of the entropy. Suppose that ℓ is very small compared to the length scales of the state, but still large compared to the cutoff ϵ. Then we have

$$\frac{d^2 S_g}{d\lambda^2} = 2\pi\,\langle T_{vv}\rangle\,\ell^{d-2} + Q_g\ell^{d-2+2\gamma} + (\text{other smeared off-diagonal terms})\,. \quad \text{(A.7.17)}$$

We can write $Q_g \sim QM^{2\gamma}$, where M is a mass scale characterizing the state and Q is what we get in the $g\to 0$ limit. So at weak coupling, we can say that

$$Q_g\ell^{d-2+2\gamma} \sim Q\ell^{d-2}\,(1 + 2\gamma\log M\ell + \cdots)\,. \quad \text{(A.7.18)}$$

Thus we find that the answer for the weakly-coupled theory is approximately the same as for the free theory, as long as $\gamma\log M\ell \ll 1$. The smallest we can make ℓ is of order the cutoff ϵ, and the condition that $\gamma\log M\epsilon$ remain small is analogous to the problem of large logarithms in perturbation theory. The renormalization group is typically used to get around the problem of large logarithms, and it would be interesting to apply those same ideas to the present situation.

[3] We treat g and γ as fixed numbers that do not themselves depend on scale. A more complete treatment that incorporates the RG flow of the coupling would be interesting.

This argument hints that for general effective field theories S''_{vv} may not have a good operational meaning in terms of physical observables. The relevant condition for isolating the δ-function is that $(M\ell)^{2\gamma} \ll 1$ should be possible within the effective description. Clearly this can be done in an exact CFT with finite anomalous dimensions, but it should also be possible if the theory is approximately given by an interacting CFT over some large range of length scales. For instance, if an interacting CFT is weakly coupled to gravity and we consider states with energy M much less than the Planck scale then it should be possible to have $(M\ell)^{2\gamma} \ll 1$ while maintaining $\ell \gg \ell_{\text{Planck}}$.

Finally, a more precise version of the arguments given above can be given by interpreting the second functional derivative of the entropy as an OPE. We hope to use these techniques to find the exact form of f_g in future work [89].

A.8 Modified Ward identity

In this Appendix, we prove the following identity:

$$\int d^{d-2}y' \langle \Sigma_n^0 \hat{D}_+(y') \hat{D}_+(y) T_{--}(w, \bar{w}, 0)\rangle = -\partial_{\bar{w}} \langle \Sigma_n^0 \hat{D}_+(y) T_{--}(w, \bar{w}, 0)\rangle. \tag{A.8.1}$$

This is similar to the defect CFT ward identity of [99] except there is another insertion of the displacement operator. A priori it is not obvious that some form of the Ward identity carries through in the case where more than one operator is a defect operator. We will argue essentially that the second insertion of a $\hat{D}_+$ just comes along for the ride.

To show this, first we write the displacement operator as a stress tensor integrated around the defect:

$$\hat{D}_+(y) = i \oint d\bar{z}\; T_{++}(0, \bar{z}, y) \tag{A.8.2}$$

where we have suppressed the sum over replicas to avoid clutter. We will then argue that the following equality holds

$$\begin{aligned} & i \lim_{\varepsilon \to 0} \oint_{\varepsilon > |\bar{z}|} d\bar{z} \int_{|y-y'|>\epsilon} d^{d-2}y' \langle \Sigma_n^0 \hat{D}_+(y') T_{++}(0, \bar{z}, y) T_{--}(w, \bar{w}, 0)\rangle \\ & = \int d^{d-2}y' \langle \Sigma_n^0 \hat{D}_+(y') \hat{D}_+(y) T_{--}(w, \bar{w}, 0)\rangle \end{aligned} \tag{A.8.3}$$

for some appropriate $\varepsilon > 0$ that acts as the cutoff $|y' - y| > \varepsilon$.

To see this, simply note that we can replace $T_{++}(0, \bar{z}, y)$ by a sum over local defect operators at y using the bulk-defect OPE. The important point is that this OPE converges because the $\bar{z}$ contour is always inside of the sphere of size ε (by construction). We can take $|\bar{z}|$ to be arbitrarily small by making the size of the $\bar{z}$ contour as small as we like. The $\bar{z}$ integral outside now simply projects the sum onto the displacement operator since we only

consider the leading twist $d-2$ operators in the lightcone limit. Explicitly, we will be left with

$$\begin{aligned} i\lim_{\varepsilon\to 0}\oint_{\varepsilon>|\bar{z}|} d\bar{z}\int_{|y-y'|>\epsilon} d^{d-2}y'\langle\Sigma_n^0\hat{D}_+(y')T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)\rangle \\ =\lim_{\epsilon\to 0}\int_{|y-y'|>\epsilon} d^{d-2}y'\langle\Sigma_n^0\hat{D}_+(y')\hat{D}_+(y)T_{--}(w,\bar{w},0)\rangle. \end{aligned} \tag{A.8.4}$$

Note that perturbatively around $n=1$, the integral over $|y-y'|>\epsilon$ will miss the delta function contribution to the $\hat{D}_+\times\hat{D}_+$ OPE. Non-perturbatively away from $n=1$, however, there are no delta-function singularities in $|y-y'|$ present in the $\hat{D}_+\times\hat{D}_+$ OPE. In what follows, we must be careful to take $\epsilon\to 0$ *before* taking $n\to 1$.

Using this identity, we can view the displacement-displacement-bulk three point function as the contour integral of a displacement-bulk-bulk three point function. We can then use the regular displacement operator Ward identity on the latter three point function. This Ward identity follows from general diffeomorphism invariance [99]. To do this, define the deformation vector field

$$\xi(y')=f(y')\partial_+ \text{ with } f(y')=\Theta(|y'-y|-\varepsilon). \tag{A.8.5}$$

For this deformation, the Ward identity takes the form

$$\begin{aligned} &i\oint_{\varepsilon>|\bar{z}|} d\bar{z}\int_{|y-y'|>\epsilon} d^{d-2}y'\langle\Sigma_n^0\hat{D}_+(y')T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)\rangle \\ &=-f(0)\partial_{\bar{w}}\langle\Sigma_n^0\hat{D}_+(y)T_{--}(w,\bar{w},0)\rangle - i\oint d\bar{z}f(y)\partial_{\bar{z}}\langle\Sigma_n^0T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)\rangle \\ &-i\int_{\mathcal{M}_n} d^dx'\oint d\bar{z}\ \langle T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)T^{\mu\nu}(x')\partial_\mu\xi_\nu(x')\rangle \end{aligned} \tag{A.8.6}$$

where $\mathcal{M}_n$ is the full replica manifold.

The second term on the right hand side of the equality vanishes because $f(y)=0$. Since $f(0)=1$ by construction we just need to argue that the last term in (A.8.6) vanishes.

Arguing the last term vanishes

It is tempting at this stage to integrate by parts on the last term and conclude that this vanishes as one sends $\varepsilon\to 0$. Unfortunately, the last term in (A.8.6) can produce $1/\varepsilon$ enhancements due to T_{i+} operator coming ε close to T_{++}. Therefore one must take care to first do the x' integral and then take the $\varepsilon\to 0$ limit when evaluating this term.

To do so, note that

$$T^{\mu\nu}(x')\partial_\mu\xi_\nu(x')=\frac{1}{2}T_{i+}(x')\hat{n}^i\delta(|y'-y|-\varepsilon) \tag{A.8.7}$$

where $\hat{n}^i = (y'-y)^i/|y'-y|$. We then have the following

$$\int_{\mathcal{M}_n} d^d x' \oint d\bar{z}\ \langle T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)T^{\mu\nu}(x')\partial_\mu \xi_\nu(x')\rangle$$
$$= \frac{1}{2}\varepsilon^{d-3}\int \rho' d\rho' d\theta' \oint d\bar{z} \int d^{d-3}\vartheta'\ \hat{n}^i \langle T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)T_{i+}(|\vec{y}+\vec{\varepsilon}|,\vartheta'_{\vec{\varepsilon}},\rho' e^{-i\theta'},\rho' e^{-i\theta'})\rangle \tag{A.8.8}$$

where $|\vec{\varepsilon}| = \varepsilon$. In going to the second line we have done the coordinate transformation $x'^+ = \rho' e^{-i\theta'},\ x'^- = \rho' e^{i\theta'}$ because we are in the Euclidean section, and in going to the last line we have written y' in spherical coordinates on the defect. At this point we can safely send $w, \bar{w} \to 0$ so that T_{--} is simply fixed at the origin. Then, in particular, let us focus on

$$\int d\theta' \oint d\bar{z}\ \langle T_{++}(0,\bar{z},y)T_{--}(0)T_{i+}(|\vec{y}+\vec{\varepsilon}|,\vartheta'_{\vec{\varepsilon}},\rho' e^{-i\theta'},\rho' e^{-i\theta'})\rangle. \tag{A.8.9}$$

It is easy to see that this identically vanishes from the boost weights of the quantities involved. Specifically, T_{++} will yield a factor of $e^{2i\theta'}$, T_{i+} will yield a factor of $e^{i\theta'}$, T_{--} does not have a boost weight since it is fixed at the origin, and the measure $d\bar{z}$ will yield a factor of $e^{-i\theta'}$ so overall we will have $\int_0^{2\pi} d\theta' e^{i\theta'} = 0$. Therefore (A.8.8) is zero for any ε.

Thus, the identity in (A.8.6) becomes

$$i\lim_{\epsilon\to 0}\oint_{\varepsilon>|\bar{z}|} d\bar{z} \int_{|y-y'|>\epsilon} d^{d-2}y' \langle \Sigma_n^0 \hat{D}_+(y') T_{++}(0,\bar{z},y)T_{--}(w,\bar{w},0)\rangle$$
$$= -\partial_{\bar{w}}\langle \Sigma_n^0 \hat{D}_+(y) T_{--}(w,\bar{w},0)\rangle \tag{A.8.10}$$

which, using (A.8.3), proves (A.8.1).

A.9 Analytic Continuation of a Replica Three Point Function

In this section, we analytically continue a general $\mathbb{Z}_n$-symmetrized three point function of the form[4]

$$\mathcal{A}_n^{(3)} = n\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \mathrm{Tr}\left[e^{-2\pi n H}\mathcal{T}\mathcal{O}_a(0)\mathcal{O}_b(\tau_{ba}+2\pi j)\mathcal{O}_c(\tau_{ca}+2\pi k)\right] \tag{A.9.1}$$

where H is the vacuum modular Hamiltonian for the Rindler wedge and $\mathcal{T}$ denotes Euclidean time ordering with respect to this Hamiltonian.

[4]Note that we are writing this as a thermal three point function on $\mathbb{H}_{d-1}\times S_1$, which is related to the flat space replica answer via conformal transformation. For a review of the relevant conformal factors, which we suppress for convenience, see [115].

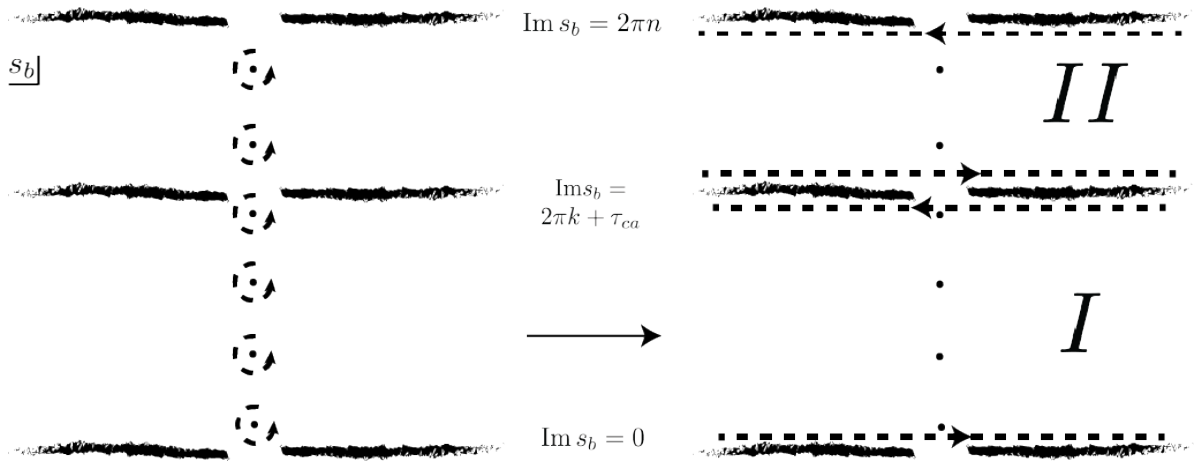

Figure A.1: The analytic structure of the integral in equation (A.9.2) represented in the s_b plane for fixed $s_k = i(2\pi k + \tau_{ca})$ for $n = 6$. The dots represent poles at $s_b = i(2\pi j + \tau_{ba})$ and the fuzzy lines denote light-cone branch cuts. The bottom and top branch cuts (which are identified by the KMS condition) arise from $\mathcal{O}_b$ becoming null separated from $\mathcal{O}_a$ and the middle branch cut arises from $\mathcal{O}_b$ becoming null separated from $\mathcal{O}_c$. Note that in this figure, $k = 3$ and $\tau_{ca} > \tau_{ba} > 0$. We start with the contour C_b represented by the dashed lines encircling the poles at $s_b = i(2\pi j + \tau_{ba})$ and unwrap so that it just picks up contributions from the branch-cuts. Region I corresponds to the ordering $\mathcal{O}_a\mathcal{O}_b\mathcal{O}_c$ whereas region II corresponds to $\mathcal{O}_a\mathcal{O}_c\mathcal{O}_b$.

.

Following [115], we begin by rewriting the the j-sum as as a contour integral

$$\frac{n}{2\pi i}\sum_{k=0}^{n-1}\oint_{C_b} ds_b \frac{\mathrm{Tr}\left[e^{-2\pi n H}\mathcal{T}\mathcal{O}_a(0)\mathcal{O}_b(-is_b)\mathcal{O}_c(2\pi k+\tau_{ca})\right]}{(e^{s_b - i\tau_{ba}}-1)} \tag{A.9.2}$$

where the contour C_b wraps the n poles at $s_b = i(2\pi j + \tau_{ba})$ for $j = 0, ..., n-1$. We will now unwrap the s_b contour integral in the complex plane, but will need to be careful as the analytic structure of the integrand in (A.9.2) is non-trivial as a function of s_b; the integrand has poles at $s_b = i(2\pi j + \tau_{ba})$ and light-cone branch cuts lying along the lines $\Im s_b = 0, 2\pi n$ and $\Im s_b = 2\pi k + \tau_{ca}$ for a fixed k. The first two branch cuts were discussed in [115]. The third (middle in the figure) branch cut arises from singularities due to $\mathcal{O}_b$ and $\mathcal{O}_c$ lying on the same light-cone.

We can unwrap the C_b contour now so that it hugs the branch cuts as in the right-hand

panel of Figure A.1. We will then be left with a sum of four Lorentzian integrals

$$\frac{n}{2\pi i}\sum_{k=0}^{n-1}\mathrm{Tr}\left[e^{-2\pi nH}\int_{-\infty}^{\infty}ds_b\times\right.$$
$$\frac{\mathcal{O}_a(0)\mathcal{O}_b(-is_b+\epsilon_j)\mathcal{O}_c(2\pi k+\tau_{ca})}{(e^{s_b-i\tau_{ba}}-1)}-\frac{\mathcal{O}_a(0)\mathcal{O}_b(-is_b+2\pi ik+\tau_{ca}-\epsilon)\mathcal{O}_c(2\pi k+\tau_{ca})}{(e^{s_b+2\pi ik+\tau_{ca}-i\epsilon-i\tau_{ba}}-1)}$$
$$\left.+\frac{\mathcal{O}_a(0)\mathcal{O}_c(2\pi k+\tau_{ca})\mathcal{O}_b(-is_b+2\pi k+\tau_{ca}+\epsilon)}{(e^{s_b+2\pi ik+\tau_{ca}+i\epsilon-i\tau_{ba}}-1)}-\frac{\mathcal{O}_a(0)\mathcal{O}_c(2\pi k+\tau_{ca})\mathcal{O}_b(-is_b+2\pi n-\epsilon)}{(e^{s_b+i2\pi n-i\epsilon-i\tau_{ba}}-1)}\right], \tag{A.9.3}$$

where we have set $2\pi k+\tau_{ca}=-is_c$ since the C_c contour still wraps the poles at these values.

We now need to make a choice about how to do the analytic continuation in n. The usual prescription, which was advocated for in [115], is to set $e^{2\pi in}=1$ in the last term of (A.9.3). We will follow this but also make one other choice. In the second and third terms in the integrand of (A.9.3) we make the choice to set $e^{2\pi ik}=1$ for all $k=0,...,n-1$.

Making this analytic continuation, we can now re-write the k-sum as a contour integral over s_c along some contour C_c. Unwrapping this s_c contour into the Lorentzian section, and after repeated use of the KMS condition to push operators back around the trace, we land on the relatively simple formula

$$\mathcal{A}_n^{(3)}=$$
$$\frac{-n}{4\pi^2}\int_{-\infty}^{\infty}ds_c ds_b\,\mathrm{Tr}\left[e^{-2\pi nH}\left(\frac{[[\mathcal{O}_a(0),\mathcal{O}_b(-is_b)],\mathcal{O}_c(-is_c)]}{(e^{s_b-i\tau_{ba}}-1)(e^{s_c-i\tau_{ca}}-1)}-\frac{[\mathcal{O}_a(0),[\mathcal{O}_b(-is_b-is_c),\mathcal{O}_c(-is_c)]]}{(e^{s_b+i\tau_{ca}-i\tau_{ba}}-1)(e^{s_c-i\tau_{ca}}-1)}\right)\right] \tag{A.9.4}$$

In deriving this formula, we have assumed $\tau_{ba}>0$ and $\tau_{ca}>0$ but we have not yet assumed any relationship between τ_{ba} and τ_{ca}. This formula is the full answer. One could stop here, but we will massage this formula into a slightly different form for future convenience. Instead of following [115] and applying ∂_n at this stage, which drops down powers of H, we will use a slightly different (although equivalent) technique.

We first focus on re-writing the two Lorentzian integrals in region I of Figure A.1 as one double integral.

Region I

Before re-writing the k-sum as a contour integral, the integrals in region I are[5]

$$\frac{n}{2\pi i}\sum_{k=0}^{n-1}\int_{-\infty}^{\infty}ds_b\left(\frac{\langle\mathcal{O}_a(0)\mathcal{O}_b(-is_b)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n}{(e^{s_b-i\tau_{ba}}-1)}-\frac{\langle\mathcal{O}_a(0)\mathcal{O}_b(-is_b+2\pi k+\tau_{ca}-\epsilon)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n}{(e^{s_b+i\tau_{ca}-i\tau_{ba}}-1)}\right) \tag{A.9.5}$$

[5]For ease of notation, we have switched to $\langle\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3\rangle_n=\mathrm{Tr}[e^{-2\pi nH}\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3]$.

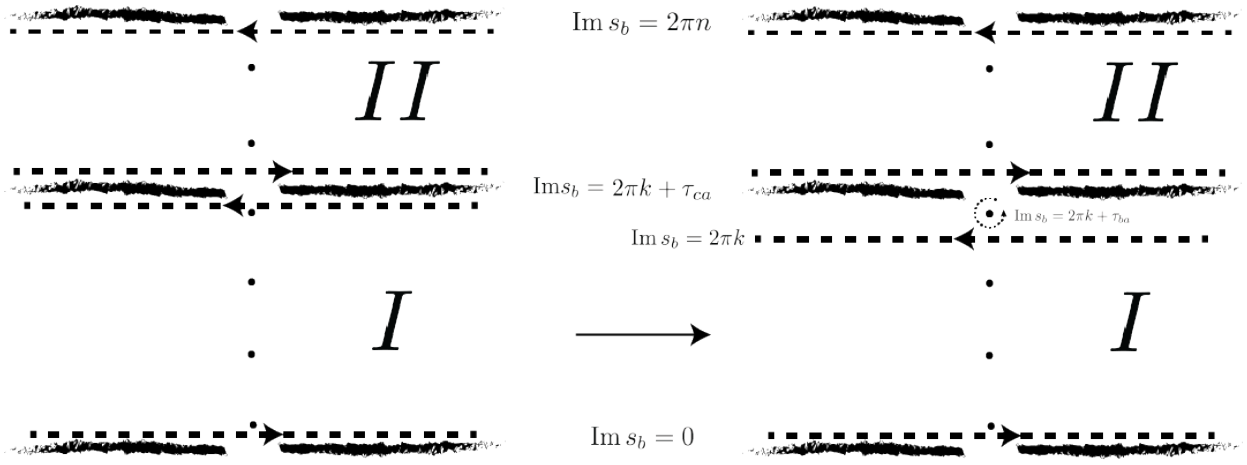

Figure A.2: This figure illustrates the contour shift $s_b \to s_b - i\tau_{ca}$ done at the cost of picking up the pole at $s = i(2\pi k + \tau_{ba})$ when $\tau_{cb} = \tau_{ca} - \tau_{ba} > 0$.

where as before we have set $e^{2\pi ik} = 1$ in the second term. The goal will be to make the denominators in these two terms the same so that we may combine their numerators. We will try to shift the s_b contour in the second term by an amount $-i\tau_{ca}$, making sure not to cross any poles or branch cuts. To make our lives easier, we will assume a fixed ordering of the operators. For now, we will pick $\tau_{ca} > \tau_{ba} > 0$. Note that any other ordering can be reached just by exchanging the a, b, c labels.

In this ordering, sending $s_b \to s_b - i\tau_{ca}$ crosses a pole at $\Im s_b = 2\pi k + \tau_{ba}$. This contour shift is illustrated in Figure A.2. After doing this shift, we get

$$\frac{n}{2\pi i}\sum_{k=0}^{n-1}\int_{-\infty}^{\infty} ds_b \left(\frac{\langle \mathcal{O}_a(0)\mathcal{O}_b(-is_b)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n - \langle \mathcal{O}_a(0)\mathcal{O}_b(-is_b+2\pi k)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n}{(e^{s_b - i\tau_{ba}}-1)}\right)$$
$$+ \theta(\tau_{cb}) \times (\text{terms with } j = k). \tag{A.9.6}$$

where we will mostly neglect the extra term coming from picking up the pole since it will not be important for most calculations we are interested in. We will refer to these terms as the "replica diagonal terms" since they arise from terms in the double sum over j, k in (A.9.1) where $j = k$.

The numerator for the first term in equation (A.9.6) then looks like the integral of a total derivative in some auxiliary parameter t_b which we write as

$$\frac{-n}{2\pi i}\sum_{k=0}^{n-1}\int_{-\infty}^{\infty} ds_b \int_0^{i2\pi k} dt_b \left(\frac{\frac{d}{dt_b}\langle \mathcal{O}_a(0)\mathcal{O}_b(-is_b - it_b)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n}{(e^{s_b - i\tau_{ba}}-1)}\right). \tag{A.9.7}$$

Since t_b shows up on equal footing with s_b in the numerator, we see we can re-write the derivative in t_b as one in s_b. Integrating by parts and dropping the boundary terms[6], we get

$$\frac{-n}{2\pi i}\sum_{k=0}^{n-1}\int_{-\infty}^{\infty}ds_b\int_0^{i2\pi k}dt_b\frac{\langle\mathcal{O}_a(0)\mathcal{O}_b(-is_b-it_b)\mathcal{O}_c(2\pi k+\tau_{ca})\rangle_n}{4\sinh^2((s_b-i\tau_{ba})/2)}. \tag{A.9.8}$$

We are now ready, as above, to turn the sum over k into a contour integral over some Lorentzian parameter s_c. We can then execute the same trick as before: we re-write two terms as the boundary terms of one integral in some new auxiliary parameter t_c. After all of this, the answer we find is the relatively simple result for region I

$$\begin{aligned}\text{region I} =&\frac{-n}{4\pi^2}\int_{-\infty}^{\infty}ds_cds_b\int_0^{i2\pi(n-1)}dt_c\int_0^{s_c+t_c}dt_b\frac{\langle\mathcal{O}_a(0)\mathcal{O}_b(-is_b-it_b)\mathcal{O}_c(-is_c-it_c+\tau_{ca})\rangle_n}{16\sinh^2((s_b-i\tau_{ba})/2)\sinh^2((s_c-i\epsilon)/2)}\\&+\theta(\tau_{cb})\times(\text{terms with } j=k).\end{aligned} \tag{A.9.9}$$

Note that the quadruple integral term is manifestly order $n-1$ because of the limits on the t_c integral.

Region II

In region II of Figure A.1, the calculations are exactly analogous, except now the ordering of the operators is different. We find that (up to terms that again come from picking up specific poles) the answer for region II is

$$\begin{aligned}&\text{region II} =\\&\frac{-n}{4\pi^2}\int_{-\infty}^{\infty}ds_cds_b\int_0^{i2\pi(n-1)}dt_c\int_{s_c+t_c+i2\pi}^{i2\pi n}dt_b\frac{\langle\mathcal{O}_a(0)\mathcal{O}_c(-is_c-it_c+\tau_{ca})\mathcal{O}_b(-is_b-it_b)\rangle_n}{16\sinh^2((s_b-i\tau_{ba})/2)\sinh^2((s_c-i\epsilon)/2)}\\&+\theta(\tau_{bc})\times(\text{terms with } j=k).\end{aligned} \tag{A.9.10}$$

Combining Regions I and II

Adding the Region I and Region II contributions, we get for the non-replica diagonal contributions to $\mathcal{A}_n^{(3)}$

$$\begin{aligned}&\frac{n}{4\pi^2}\int_{-\infty}^{\infty}ds_cds_b\int_0^{i2\pi(n-1)}dt_c\int_0^{s_c+t_c}dt_b\frac{\langle[\mathcal{O}_b(-is_b-it_b),\mathcal{O}_a(0)]\mathcal{O}_c(-is_c-it_c+\tau_{ca})\rangle_n}{16\sinh^2((s_b-i\tau_{ba})/2)\sinh^2((s_c-i\epsilon)/2)}\\+&\frac{n}{4\pi^2}\int_{-\infty}^{\infty}ds_cds_b\int_0^{i2\pi(n-1)}dt_c\int_{s_c+t_c}^{s_c+t_c+i2\pi(1-n)}dt_b\frac{\langle\mathcal{O}_b(-is_b-it_b)\mathcal{O}_a(0)\mathcal{O}_c(-is_c-it_c+\tau_{ca})\rangle_n}{16\sinh^2((s_b-i\tau_{ba})/2)\sinh^2((s_c-i\epsilon)/2)}\end{aligned} \tag{A.9.11}$$

[6]We will drop boundary terms at large Lorentzian time everywhere throughout this discussion, as we expect thermal correlators to fall off sufficiently quickly [115].

where we used the KMS condition to push $\mathcal{O}_b$ around to the left of $\mathcal{O}_a$ in (A.9.10). We then split the t_b contour in (A.9.10) into two pieces, one purely Lorentzian integral from $t_b = 0$ to $t_b = s_c + t_c$ and another purely Euclidean integral from $t_b = s_c + t_c$ to $t_b = s_c + t_c + 2\pi i(n-1)$. Again, this is the full answer for the replica three point function, $\mathcal{A}_n^{(3)}$, at all n excluding the replica diagonal terms.

From this we can compute the leading order in n correction to the three-point function (dropping the diagonal terms). Taking an n-derivative and setting $n \to 1$, the total correction is

$$\mathcal{A}_n^{(3)} \sim \frac{i(n-1)}{2\pi} \int_{-\infty}^{\infty} ds_c ds_b \int_0^{s_c} dt_b \frac{\langle [\mathcal{O}_b(-is_b - it_b), \mathcal{O}_a(0)] \mathcal{O}_c(-is_c + \tau_{ca}) \rangle_1}{16 \sinh^2((s_b - i\tau_{ba})/2) \sinh^2((s_c - i\epsilon)/2)} + (\text{replica diagonal terms}) + \mathcal{O}\left((n-1)^2\right). \tag{A.9.12}$$

Replica Diagonal Terms

For future reference, we now list the replica diagonal (or $j = k$) terms that we have suppressed. In the order we considered above, we have

$$\begin{aligned} & n\theta(\tau_{cb})\theta(\tau_{ba}) \sum_{k=0}^{n-1} \langle \mathcal{O}_a(0) \mathcal{O}_b(2\pi k + \tau_{ba}) \mathcal{O}_c(2\pi k + \tau_{ca}) \rangle_n \\ & = n\theta(\tau_{cb})\theta(\tau_{ba}) \Bigg(\langle \mathcal{O}_a(0) \mathcal{O}_b(\tau_{ba}) \mathcal{O}_c(\tau_{ca}) \rangle_n - \\ & \frac{1}{2\pi i} \int_{i2\pi}^{i2\pi n} dt_c \int_{-\infty}^{\infty} ds_c \frac{\langle \mathcal{O}_a(0) \mathcal{O}_b(-is_c - it_c - \tau_{cb}) \mathcal{O}_c(-is_c - it_c) \rangle_n}{4 \sinh^2((s_c - i\tau_{ca})/2)} \Bigg). \end{aligned} \tag{A.9.13}$$

Again, other orderings can be found just by swapping the a, b, c labels accordingly. Note that at $n = 1$, the integral term vanishes and the answer reduces to the angular ordered three-point function as expected.

A.10 Explicit Calculation of $c^{(2)}$

In this section, we compute the OPE coefficient of $\hat{T}_{++}$ in the $\hat{D}_+ \times \hat{D}_+$ OPE. This requires us to compute the twist defect three point function $\langle \Sigma_n^0 \hat{D}_+ \hat{D}_+ \hat{T}_{--} \rangle$. As described around equation (A.8.3), the appearence of a delta function in the $\hat{D}_+ \times \hat{D}_+$ OPE requires that the coefficient c_n for $\hat{T}_{--}$ must be at least of order $(n-1)^2$ near $n = 1$. We now show that this is indeed true. In the next section, we will explicitly compute the anomalous dimension of $\hat{T}_{--}$ and show that it behaves as $g_n \sim \gamma^{(1)}(n-1) + \mathcal{O}((n-1)^2)$. We will finally show that their ratio obeys the relation

$$c^{(2)}/\gamma^{(1)} = 2\pi/S_{d-3} \tag{A.10.1}$$

as required by the first law of entanglement entropy.

The three point function we are after, at integer n, takes the form

$$\begin{aligned}&\langle \Sigma_n^0 \hat{T}_{--}(y')\hat{D}_+(y)\hat{D}_+(y=0)\rangle \\ &= -\oint d\bar{z}\oint d\bar{w}\oint \frac{du}{2\pi i u}\,\langle \Sigma_n^0 T_{--}(u,\bar{u}=0,y')T_{++}(z=0,\bar{z},y)T_{++}(w=0,\bar{w},0)\rangle\end{aligned} \tag{A.10.2}$$

where it is understood that all the stress tensor operators should be $\mathbb{Z}_n$ symmetrized. Our goal is now to analytically continue this expression in n and then expand around $n=1$. We can turn to the previous section for this result, letting $\mathcal{O}_a = T_{++}(w=0,\bar{w},0)$, $\mathcal{O}_b = T_{++}(z=0,\bar{z},y)$ and $\mathcal{O}_c = T_{--}(u,\bar{u}=0,0)$.

Just as in Section 5.5, a major simplification occurs for this correlator; the two displacement operators are space-like separated from each other, so they commute even upon analytic continuation. Thus, any terms with commutators between $\mathcal{O}_a$ and $\mathcal{O}_b$ in the previous section drop out.

Furthermore, the so-called "replica diagonal" terms in the previous section will also vanish. This is because they do not contain enough s-integrals that produce necessary poles in $\bar{z}$ and $\bar{w}$. Thus, these terms vanish upon the contour integration over $\bar{z}$ and $\bar{w}$ in (A.10.2).

These considerations together with equation (A.9.11) of the previous section make it clear that the correlator in (A.10.2) vanishes up to order $(n-1)^2$. Indeed, the only surviving contribution is the second term in (A.9.11). Expanding that to second order while being careful to account for the spin of the stress tensors, we find

$$\begin{aligned}&\langle \Sigma_n^0 \hat{T}_{--}\hat{D}_+\hat{D}_+\rangle_n = \\ &\frac{-(n-1)^2}{2}\oint d\bar{z}d\bar{w}\frac{du}{2\pi i u}\int_0^\infty\int_0^\infty d\lambda_b d\lambda_c \lambda_b^2\lambda_c^2 \frac{\langle T_{++}(\bar{z}\lambda_b,y)T_{++}(\bar{w}\lambda_c)T_{--}(u,y')\rangle}{(\lambda_b-1-i\epsilon)^2(\lambda_c-1+i\epsilon)^2} + \mathcal{O}((n-1)^3).\end{aligned} \tag{A.10.3}$$

Rescaling $\lambda_b \to \lambda_b/\bar{z}$ and $\lambda \to \lambda_c/\bar{w}$, we can then expand the denominators in small $\bar{z},\bar{w}$ and perform the residue projections in $\bar{z},\bar{w}$ and u. The final answer is the simple result

$$\langle \Sigma_n^0 \hat{T}_{--}\hat{D}_+\hat{D}_+\rangle = 2\pi^2(n-1)^2\,\langle \mathcal{E}_+(y)\mathcal{E}_+(y=0)T_{--}(u=0,y')\rangle + \mathcal{O}((n-1)^3). \tag{A.10.4}$$

where $\mathcal{E}_+(y)$ is the half-averaged null energy operator

$$\mathcal{E}_+(y) = \int_0^\infty d\lambda\, T_{++}(z=0,\lambda,y) \tag{A.10.5}$$

We now set about computing this correlator. Expanding the stress tensor three point function in a general CFT into the free field basis, we have

$$\langle TTT\rangle = n_s\,\langle TTT\rangle_s + n_f\,\langle TTT\rangle_f + n_v\,\langle TTT\rangle_v \tag{A.10.6}$$

where n_s, n_f and n_v are charges characterizing the specific theory.

One can demonstrate that the only non-vanishing contribution from these three terms is from the scalar three point function. The way to see this is as follows. The fermion term can be computed by considering a putative free Dirac fermion theory with field ψ. The stress tensor looks like $T_{++} \sim \bar{\psi}\Gamma_+\partial_+\psi$. Then we can compute the $\langle TTT\rangle$ three point function via Wick contractions. There will always be at least one Wick contraction between operators in each T_{++}. The kinematics of these operators ensure that such a contraction vanishes because they are both on the same null plane.[7]

The same argument can be made for the vector fields. In fact, the *only* reason that the scalar contribution doesn't vanish is because of the presence of a total derivative term in the conformal stress tensor, namely $T_{++} \supset -\frac{d-2}{4(d-1)}\partial_+^2 :\phi^2:$. One can then show that the only non-vanishing term is

$$\langle \mathcal{E}_+(y)\mathcal{E}_+(0)T_{--}(y')\rangle = \frac{4n_s(d-2)}{(d-1)^3}\frac{1}{|y|^{d-2}|y'|^{2d}}. \tag{A.10.7}$$

Dividing by the two point function $\langle T_{++}(0)T_{--}(y')\rangle = \frac{c_T}{4|y'|^{2d}}$, we find

$$c^{(2)} = \frac{32\pi^2 n_s(d-2)}{c_T(d-1)^3}. \tag{A.10.8}$$

We now turn to computing the anomalous dimension $\gamma^{(1)}$ for the stress tensor operator $\hat{T}$ on the defect.

A.11 Explicit Calculation of $\gamma^{(1)}$

In this section, we will follow the steps laid out in [17] for computing the spectrum of defect operators and associated anomalous dimension induced by the bulk stress tensor. To do this, we must compute

$$n\sum_{j=0}^{n-1}\langle \Sigma_n^0 T_{--}(w,0,y)T_{++}(0,\bar{z},0)\rangle\,. \tag{A.11.1}$$

To leading order in $n-1$ this expression takes the form of a sum of two terms, a "modular energy" piece and a "relative entropy" piece

$$\begin{aligned}(\partial_n - 1)\left\langle \Sigma_n^0 \hat{T}_{--}\hat{T}_{++}\right\rangle|_{n=1} = (&-2\pi\left\langle H T_{--}(w,0,y)T_{++}(0,\bar{z},0)\right\rangle \\ &- \int_0^{-\infty} d\lambda \frac{\lambda^2}{(\lambda-1+i\epsilon)^2}\left\langle T_{--}(w,0,y)T_{++}(0,\bar{z}\lambda,0)\right\rangle\Big)\end{aligned} \tag{A.11.2}$$

[7] Actually these contractions will be proportional to a delta function $\delta^{d-2}(y)$ but we are assuming the three stress tensors sit at different y's.

We will try to extract the anomalous dimensions and spectra of operators by examining the two point function of the defect stress tensor. In this framework, the signal of an anomalous dimension is a logarithmic divergence. As explained in [17], the log needs to be cutoff by $\bar{z}w/y^2$ or $z\bar{w}/y^2$. In fact, there will be two such logarithms that will add to make the final answer single-valued on the Euclidean section.

We are thus tasked with looking for all of the terms containing log divergences in (A.11.2). Since the modular Hamiltonian is just a local integral of the stress tensor

$$H = \int d^{d-2}y' \int_0^\infty dx^+ x^+ T_{++}(x^- = 0, x^+, y') \tag{A.11.3}$$

then the first term on the r.h.s. of (A.11.2) is a stress tensor three point function. Following the method of the previous section, we can then break up (A.11.2) into the free field basis. This determines both terms on the r.h.s of (A.11.2) in terms of charges n_s, n_f and n_v. This allows us to instead compute the answer in a theory of free massless scalars, fermions and vectors. While this might seem like three times the work, it actually illuminates why g_n is only dependent on n_s. We start by examining the case of a free scalar and will see why the free fermion and free vector terms do not contribute to g_n.

Spectrum induced by free scalar

This spectrum of $\phi(z, \bar{z}, y)$ was analyzed in [98]. The authors found that the leading twist defect primaries are all twist one (in $d = 4$) and have dimension independent of n. As noted in Appendix C of that work, this can be understood in any dimension from the fact that ϕ is annihilated by the bulk Laplacian. This constraint - for defect primaries - enforces holomorphicity in $z, \bar{z}$ of the bulk-defect OPE which translates to a lack of anomalous dimensions. For free fermions and vectors, the same argument goes through since their two point functions are also annihilated by the Laplacian.

One might be confused because the anomalous dimension for scalar operators of dimension Δ was computed in [17] and found to be non-zero for operators of dimension $\Delta = \frac{d-2}{2}$. This discrepancy has to do with a subtlety related to the extra boundary term in the modular Hamiltonian for free scalars. This discrepancy is related to the choice of the stress tensor - the traceless, conformal stress tensor vs. the canonical stress tensor.

The authors of [98] worked with *canonical* free fields, for which the stress tensor is just $T_{++}^{\text{canonical}} = \partial_+\phi\partial_+\phi$. Indeed if one inserts the canonical stress tensor into the modular Hamiltonian in equation (3.20) of [17], then the anomalous dimension vanishes. On the other hand, if one uses the conformal stress tensor, $T_{++}^{\text{conformal}} = :\partial_+\phi\partial_+\phi: - \frac{(d-2)}{4(d-1)}\partial_+^2 :\phi^2:$, then anomalous dimension for ϕ is given by [17].

This discrepancy thus amounts to a choice of the stress tensor. Note that this is special to free scalars and does not exist for free fermions and vectors since there are no dimension $d-2$ scalar primaries in these CFTs. This proves that if one works with canonical free fields, there should be no anomalous dimension for the defect operators induced by the fundamental

fields ϕ, ψ and A_μ. This is enough to prove that the defect primary induced by the *canonical* bulk stress tensor must also have zero anomalous dimension since this is just formed by normal-ordered products of the defect primaries induced by the bulk fundamental fields.

Back to the stress tensor

The upshot is that we only need to worry about the terms in (A.11.2) proportional to n_s. Furthermore, we only need to worry about terms in the $\langle HTT\rangle$ term that involve the boundary term of the modular Hamiltonian. This reduces the expression down to the term

$$\langle HTT\rangle \supset -\frac{(d-2)}{4(d-1)}\int d^{d-2}y\,\langle :\phi^2: T_{++}(0,\bar{z},y)T_{--}(w,0,0)\rangle\,. \tag{A.11.4}$$

A simple calculation shows that the only contractions that give log divergences come from

$$\begin{aligned}\langle HTT\rangle &\supset \frac{n_s(d-2)^2}{4(d-1)^2}\int d^{d-2}y'\,\langle\phi(0,0,y')\phi(0,0,0)\rangle\,\langle\phi(0,0,y')\partial_{\bar{z}}^2\phi(0,\bar{z},0)T_{--}(0,0,y)\rangle \\ &= -\frac{n_s c_{\phi\phi}^3 d(d-2)^4}{16(d-1)^3}\int d^{d-2}y'\frac{1}{|y'|^{d-2}|y-y'|^{d-2}|y|^{d+2}}.\end{aligned} \tag{A.11.5}$$

This integral has two log divergences coming from $y'=0$ and $y'=y$, however they can be regulated by fixing $z,\bar{z}$ and $w,\bar{w}$ away from zero. The two singularities just add to make the final answer single valued under rotations by 2π about the defect as in [17]. We thus find

$$\langle HTT\rangle \supset -\,n_s\frac{c_{\phi\phi}^3 d(d-2)^4}{32(d-1)^3}S_{d-3}\log(w\bar{w}z\bar{z}/|y|^4)\frac{1}{|y|^{2d}} = -\frac{2n_s(d-2)}{(d-1)^3}S_{d-3}\log(w\bar{w}z\bar{z}/|y|^4)\frac{1}{|y|^{2d}}. \tag{A.11.6}$$

Dividing by $\langle T_{++}T_{--}\rangle$ gives

$$\gamma^{(1)} = \frac{16\pi n_s(d-2)}{c_T(d-1)^3}S_{d-3}. \tag{A.11.7}$$

Comparing with (A.10.8), we see that

$$\frac{c^{(2)}}{\gamma^{(1)}} = \frac{2\pi}{S_{d-3}} \tag{A.11.8}$$

as required by the first law of entanglement.

A.12 Calculating $\mathcal{F}_n$

At first glance, $\mathcal{F}_n$ seems difficult to calculate; we would like a method to compute this correlation function at leading order in $n-1$ without having to analytically continue a

$\mathbb{Z}_n$ symmetrized four point function. The method for analytic continuation is detailed in Appendix A.9.

As detailed in Appendix A.9, part of what makes the analytic continuation in n difficult is the analytic structure (branch cuts) due to various operators becoming null separated from each other in Lorentzian signature. One might naively worry that we have to track this for four operators in the four point function $\mathcal{F}_n$.

We will leverage the fact that the two stress tensors in $\hat{D}_+(y_1)$ and $\hat{D}_+(y_2)$ are in the lightcone limit with respect to the defect since

$$\hat{D}_+(y_1) = \lim_{|z|\to 0} i \oint d\bar{z} \sum_{j=0}^{n-1} T^{(j)}_{++}(z=0,\bar{z},y_1). \tag{A.12.1}$$

Thus, the stress tensors at y_1 and y_2 commute with each other even after a finite amount of boost. This means that these two operators do not see each other in the analytic continuation. In other words, the analytic structure for each of these operators is just that of a $\mathbb{Z}_n$ symmetrised *three* point function. This was computed in Appendix A.9.

We can thus jump straight to (A.9.12) but now with two $\mathcal{O}_b$ operators. The final replica four point function assuming $[\mathcal{O}_{b_1},\mathcal{O}_{b_2}]=0$ is given by[8]

$$\frac{(n-1)}{8\pi^2}\int_{-\infty}^{\infty} ds_c ds_{b_1} ds_{b_2} \int_0^{s_c} dt_{b_1} dt_{b_2} \frac{\langle [\mathcal{O}_{b_2}(-is_{b_2}-it_{b_2}),[\mathcal{O}_{b_1}(-is_{b_1}-it_{b_1}),\mathcal{O}_a(0)]]\mathcal{O}_c(-is_c+\tau_{ca})\rangle_1}{64\sinh^2((s_{b_1}-i\tau_{b_1a})\sinh^2((s_{b_2}-i\tau_{b_2a})/2)\sinh^2((s_c-i\epsilon)/2)} + \mathcal{O}((n-1)^2). \tag{A.12.2}$$

To make contact with $\mathcal{F}_n$, we assign

$$\begin{aligned}
\mathcal{O}_{b_1}(-is_1) &= \lim_{|z|\to 0} i\oint d\bar{z}\, e^{2s_1-2i\tau_{b_1a}} T_{++}(x^-=0,x^+=r_{\bar{z}}e^{s_1},y_1)\\
\mathcal{O}_{b_2}(-is_2) &= \lim_{|w|\to 0} i\oint d\bar{w}\, e^{2s_2-2i\tau_{b_2a}} T_{++}(x^-=0,x^+=r_{\bar{w}}e^{s_2},y_2)\\
\mathcal{O}_c(-is_c) &= \lim_{|u|\to 0} i\oint du\, e^{-2s_c+2i\tau_{ca}} T_{--}(x^-=-r_u e^{-s_c},x^+=0,y_4)\\
\mathcal{O}_a(0) &= \lim_{|v|\to 0} i\oint \frac{dv}{2\pi i}\, T_{--}(x^-=-r_v,x^+=0,y_3)
\end{aligned} \tag{A.12.3}$$

with $\bar{z},\bar{w} = r_{\bar{z},\bar{w}}e^{i\tau_{b_1,b_2}}$ and $u,v = r_{u,v}e^{-i\tau_{a,c}}$. The funny factors of $e^{2s-2i\tau}$ are to account for the spin of the stress tensor.

[8] We have dropped the so-called "replica diagonal" terms in (A.9.12) since they will drop out of the final answer after the residue projection in (A.12.1).

Shifting $s_{b_{1,2}} \to s_{b_{1,2}} - t_{b_{1,2}} - \log(r_{1,2})$ and moving to null coordinates $\lambda = e^s$, we find the expression

$$\begin{aligned}
&\mathcal{F}_n = \lim_{|z|,|w|,|u|,|v|\to 0} \oint d\bar{z}\, d\bar{w}\, du\, dv \times \\
&\frac{(n-1)}{8\pi^2} \int_{-\infty}^{\infty} ds_c \int_0^{\infty} \frac{d\lambda_{b_{1,2}}\, \lambda_{b_1}^2 \lambda_{b_2}^2}{\bar{z}^3 \bar{w}^3} \int_0^{s_c} dt_{b_1} dt_{b_2} e^{-s_c} e^{-t_{b_1}-t_{b_2}} e^{6i\tau_a} \times \\
&\frac{\langle [T_{++}(x^+ = \lambda_{b_1}), [T_{++}(x^+ = \lambda_{b_2}), T_{--}(x^- = -r_v)]] T_{--}(x^- = -r_u e^{-s_c - i\tau_{ca}}) \rangle_1}{\left(\frac{\lambda_{b_1} e^{i\tau_a}}{\bar{z} e^{t_{b_1}}} - 1\right)^2 \left(\frac{\lambda_{b_2} e^{i\tau_a}}{\bar{w} e^{t_{b_2}}} - 1\right)^2 (e^{s_c - i\epsilon} - 1)^2}.
\end{aligned} \tag{A.12.4}$$

The first line in (A.12.4) comes from the residue projections in the definitions of the displacement operators. Expanding the integrand at small $|\bar{z}|$ and $|\bar{w}|$, we can perform the residue integrals over $\bar{z}$ and $\bar{w}$ leaving us with

$$\begin{aligned}
&\mathcal{F}_n = \lim_{|u|,|v|\to 0} \oint du\, dv \times \\
&\frac{1-n}{2} \int_{-\infty}^{\infty} ds_c \int_0^{s_c} dt_{b_1} dt_{b_2} e^{-s_c + 2i\tau_a} e^{t_{b_1} + t_{b_2}} \frac{\langle [\mathcal{E}_+(y_1), [\mathcal{E}_+(y_2), T_{--}(x^- = -r_v)]] T_{--}(x^- = -u e^{-s_c + i\tau_a}) \rangle_1}{(e^{s_c - i\epsilon} - 1)^2}
\end{aligned} \tag{A.12.5}$$

where $\mathcal{E}_+(y_1)$ is a half-averaged null energy operator, $\int_0^{\infty} dx^+ T_{++}(x^+)$.

We can now do the t_{b_1} and t_{b_2} integrals which produce two factors of $e^{s_c} - 1$ precisely cancelling the denominator. Note that a similar cancellation occurred in equation (5.6.12). We can then replace commutators of half-averaged null energy operators with commutators of full averaged null energy operators. Using the fact that $\hat{\mathcal{E}}_+ |\Omega\rangle = 0$, we are left with the expression

$$\begin{aligned}
&\mathcal{F}_n = \lim_{|v|,|u|\to 0} \oint du dv \times \\
&\frac{(1-n)}{2} \int_{-\infty}^{\infty} ds_c\, e^{-s_c + 2i\tau_a} \left\langle T_{--}(x^- = -r_v, x^+ = 0, y_3) \hat{\mathcal{E}}_+(y_1) \hat{\mathcal{E}}_+(y_2) T_{--}(x^- = -u e^{-s_c + i\tau_a}, x^+ = 0, y_4) \right\rangle_1 .
\end{aligned} \tag{A.12.6}$$

Using boost invariance, we can also write this as

$$\mathcal{F}_n = 4\pi^2 (n-1) \int_{-\infty}^{\infty} ds_c\, e^{-s_c} \left\langle T_{--}(x^- = -1, x^+ = 0, y_3) \hat{\mathcal{E}}_+(y_1) \hat{\mathcal{E}}_+(y_2) T_{--}(x^- = -e^{-s_c}, x^+ = 0, y_4) \right\rangle_1 \tag{A.12.7}$$

where we have performed the projection over v, u.

This is precisely the formula we were after. From here, one can just insert the $\hat{\mathcal{E}}_+ \times \hat{\mathcal{E}}_+$ OPE as described in the main text.

A.13 Free Field Theories and Null Quantization

In this section we review the basics of null quantization (see [208, 28]). We then show that our computations in Section 5.6 can reproduce the results of [28]. In free (and super-renormalizable) quantum field theories, one can evolve the algebra of operators on some space-like slice up to the null plane $x_- = 0$ and quantize using the null generator $P_+ = \int d^{d-2}y\, dx^+\, T_{++}(x^+, y)$ as the Hamiltonian. One can show that for free scalar fields, the algebra on the null plane factorizes across each null-generator (or "pencil") of the $x^- = 0$ plane. For each pencil, the algebra $\mathcal{A}_{p_y}$ is just the algebra associated to a 1+1-d chiral CFT. Accordingly, the vacuum state factorizes as an infinite tensor product of $1+1$-d chiral CFT vacua:

$$|\Omega\rangle = \bigotimes_y |\Omega\rangle^{p_y} \tag{A.13.1}$$

where $|0\rangle_{p_y}$ is the vacuum for the chiral $1+1$-d CFT living on the pencil at transverse coordinate y.

Thus, if we trace out everything to the past of some (possibly wiggly) cut of the null plane defined by $x^+ = X^+(y)$, we will be left with an infinite product of reduced vacuum density matrices for a $1+1$-d CFT on the pencil

$$\sigma_{X^+(y)} = \bigotimes_y \sigma^{p_y}_{x^+ > X^+(y)}. \tag{A.13.2}$$

As discussed in [28], a general excited state on the null plane $|\Psi\rangle$ can also be expanded in the small transverse size of $\mathcal{A}$ of a given pencil. For any p_y, the full reduced density matrix above some cut of the null plane takes the form

$$\rho = \sigma^{p_y}_{X^+(y)} \otimes \rho^{(0)}_{\text{aux}} + \mathcal{A}^{1/2} \sum_{ij} \sigma^{p_y}_{X^+(y)} \int dr d\theta f_{ij}(r,\theta) \partial\phi(re^{i\theta}) \otimes E_{ij}(\theta) \tag{A.13.3}$$

where $\partial\phi$ is an operator acting on the pencil Hilbert space and $E_{ij}(\theta) = e^{\theta(K_i - K_j)} |i\rangle \langle j|$, with $|i\rangle$ eigenvectors for the auxiliary modular Hamiltonian, K_{aux}. Note that E_{ij} parameterizes our ignorance about the rest of the state on the null plane which is not necessarily the vacuum.

As a consistency check of (5.6.12), we now demonstrate agreement with the result of [28]. In null quantization, the delta function piece of the shape deformation corresponds to a shape deformation of the pencil while keeping the auxiliary system fixed. Note that the ansatz A.13.3 is analogous to the λ expansion in Section 5.6 even though we are now considering a general excited state

$$\rho = \sigma + \mathcal{A}^{1/2} \delta\rho + \mathcal{O}(\mathcal{A}). \tag{A.13.4}$$

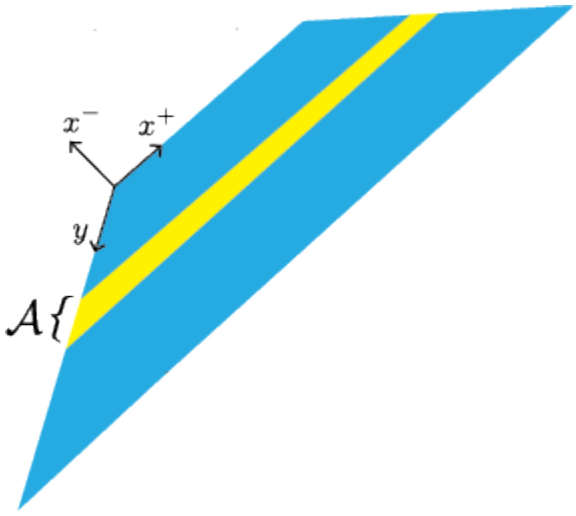

Figure A.3: The Hilbert space on a null hypersurface of a free (or superrenormalizable) quantum field theory factorizes across narrow pencils of width $\mathcal{A}$. One pencil is shown above in yellow. The neighboring pencils then can be thought of as an auxiliary system (shown in blue). In the vacuum, the state between the pencil and the auxiliary system factorizes, but in an excited state there could be nontrivial entanglement between the two systems.

We now just plug in our expression of $\delta\rho$ into (5.6.8) and find that the relative entropy second variation is

$$\frac{d^2}{dX^+(y)^2}S_{\text{rel}}(\rho|\rho_0) = \frac{1}{2}\sum_{ij}\int\int (drd\theta)_1(drd\theta)_2(f_{ij}(r,\theta))_1(f_{ji}(r,\theta))_2 \int ds\ e^s\langle(\partial\phi)_1\mathcal{E}_+\mathcal{E}_+(\partial\phi)_2(s)\rangle_{\text{p}}\langle E_{ij}(\theta_1)E_{ji}(\theta_2-is)\rangle_{\text{aux}}. \tag{A.13.5}$$

Now on the pencil, $\mathcal{E}_+$ is the translation generator so we can use the commutator $i[\mathcal{E}_+,\partial\phi] = \partial^2\phi$ and the fact that $\mathcal{E}_+|0\rangle = 0$ to get

$$\frac{d^2}{dX^+(y)^2}S_{\text{rel}}(\rho|\rho_0) = \frac{1}{2}\sum_{ij}\int\int (drd\theta)_1(drd\theta)_2(f_{ij}(r,\theta))_1(f_{ji}(r,\theta))_2 \int dse^s\langle(\partial^3\phi)_1(\partial\phi)_2(s)\rangle_{\text{p}}\langle E_{ij}(\theta_1)E_{ji}(\theta_2-is)\rangle_{\text{aux}}. \tag{A.13.6}$$

Using the chiral two-point function we have

$$\langle(\partial^3\phi)_1(\partial\phi)_2(s)\rangle_{\text{p}} = \frac{e^s}{(r_1e^{i\theta_1}-r_2e^{i\theta_2+s})^4}. \tag{A.13.7}$$

Moreover, the auxiliary correlator is given by

$$\langle E_{ij}(\theta_1)E_{ji}(\theta_2 - is)\rangle = e^{-2\pi K_i}e^{\nu_{ij}(\theta_1-\theta_2+is)},\ \nu_{ij} = K_i - K_j \tag{A.13.8}$$

We now shift the integration contour by $s \to s + i(\theta_1 - \theta_2) + i\pi + \log(r_1/r_2)$. Putting this all together we are left with evaluating

$$e^{-\pi(K_i+K_j)}e^{-2i(\theta_1+\theta_2)}\left(\frac{r_1}{r_2}\right)^{i\nu_{ij}}\frac{1}{(r_1r_2)^2}\int_{-\infty}^{\infty}ds\frac{e^{is\nu_{ij}}e^{2s}}{(1+e^s)^4}. \tag{A.13.9}$$

The θ integrals project us onto the $m = 2$ Fourier modes of f_{ij}, $f_{ij}^{(m=2)}(r)$, and we find the final answer

$$\frac{d^2}{dX^+(y)^2}S_{\rm rel}(\rho|\rho_0) = \frac{1}{2}\sum_{ij}|F_{ij}^{(2)}|^2e^{-\pi(K_i+K_j)}g(\nu_{ij}) \tag{A.13.10}$$

where

$$F_{ij}^{(m)} = \int\frac{dr}{r^m}r^{i\nu_{ij}}f_{ij}^{(m)}(r),\ g(\nu) = \frac{\pi\nu(1+\nu^2)}{\sinh(\pi\nu)}. \tag{A.13.11}$$

This is precisely the answer that was found by different methods in [28]. Note that the right hand side of (A.13.10) is manifestly positive as required by the QNEC.

A.14 Ant Conjecture and Properties of Energy Minimizing States

In Sec. 6.5, we showed that the Ceyhan-Faulkner construction proves our conjecture in the pure-QFT limit. The original purpose of the CF construction, however, was to prove Wall's "ant conjecture" [66] (and thus, the Quantum Null Energy Condition [16]). It is therefore of interest to ask how closely related our coarse-graining conjecture is to the ant conjecture on Killing horizons.It is easy to see that Eqs. (6.4.14) and (6.4.15) imply the ant conjecture. Conversely, we will show in this section that the ant conjecture implies Eqs. (6.4.14) and (6.4.15), but only in 1+1 dimensions.

In Appendix A.14, we will review the ant conjecture. In Appendices A.14 and A.14, we establish some general properties that energy-minimizing states must satisfy. We show that the minimum energy completion has vanishing stress tensor on the unconstrained half-space, with all of the remaining energy appearing as a shock immediately on the cut. We also show that for a pure minimum energy state, the von Neumann entropy of semi-infinite regions is constant so long as the region's boundary lies on the unconstrained side. In 1+1 dimensions, we can also show that the integrated left stress tensor vanishes. Thus the ant conjecture implies Eqs. (6.4.14) and (6.4.15), the key properties of the field theory limit of our coarse-graining conjecture. In higher dimensions, we are unable to establish this result.

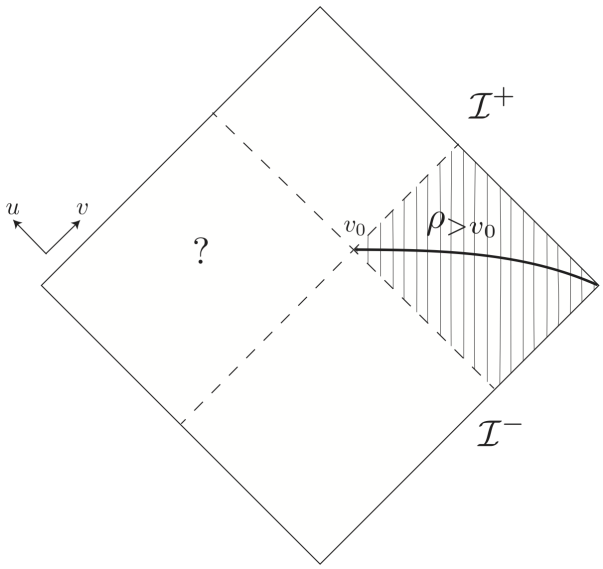

Figure A.4: The ant conjecture in 1+1 dimensions. A left-walking ant has access to all the information in the right wedge. It asks what is the least amount of additional energy it might still encounter to the left of v_0. The conjecture states that this is $\hbar S'/2\pi$, where S' is the right derivative of the von Neumann entropy of the reduced state on the right, evaluated at the cut. We show that this statement is equivalent to the nongravitational limit of our coarse-graining conjecture.

Ant Conjecture

Wall's "ant argument" for the Quantum Null Energy Condition in 1+1 dimensions invokes an ant that has walked left from $+\infty$ to v_0. (See Fig. A.4.) That is, given a global state ρ, the ant has knowledge only of the right half-space state $\rho_{>v_0}$. Pausing for rest, the ant contemplates how much energy it might still encounter in the remainder of its path, the interval $(-\infty, v_0]$. Because of global energy conditions, this amount is bounded from below.

Let $M(v_0)$ be the lowest energy of any global state that reduces to the same $\rho_{>v_0}$.[9] More precisely,

$$M(v_0) \equiv \inf_{\hat{\rho}} \left[\int_{-\infty}^{\infty} d\tilde{v} \, \langle T_{vv} \rangle \big|_{\hat{\rho}} \right] \ . \tag{A.14.1}$$

[9]We should point out two differences in our conventions compared to [66]. First, we have switched the side on which the state is held fixed, from left to right. Secondly, in [66], M was the infimum of the energy density integrated only over the complement of that fixed half-space, whereas here it is the infimum the global energy. This choice is more convenient as otherwise the presence of distributional sources at the cut v_0 would lead to ambiguities and require a more elaborate definition. In this respect, our conventions agree with [18].

The infimum is over all global states $\hat{\rho}$ that agree with ρ in the region $v > v_0$: $\mathrm{Tr}_{\leq v_0}\,\hat{\rho} = \rho_{>v_0}$. A strictly larger set of global states will agree with ρ on a smaller region, ρ_{v_1}, $v_1 > v_0$, so the infimum can only decrease with v:

$$\partial_v M(v) \leq 0 \ . \tag{A.14.2}$$

One can readily establish a lower bound on $M(v)$. The global energy appearing in the infimum can be written as $(\hbar/2\pi)(\partial_v \bar{K} - \partial_v K)$, by Eq. (6.4.10) and its left analogue. Moreover, Eq. (6.4.12) must hold for all states appearing in the infimum, so by adding $\partial_v K$ to it one finds that

$$M(v_0) \geq -\frac{\hbar}{2\pi}\partial_v S_{\mathrm{rel}}|_{v_0} \ , \tag{A.14.3}$$

where we have used Eq. (6.4.9). Note that the lower bound is determined solely by the input state ρ.

Wall conjectured [66] that this inequality is saturated:

$$M(v_0) = -\frac{\hbar}{2\pi}\partial_v S_{\mathrm{rel}}|_{v_0} \ . \tag{A.14.4}$$

This conjecture is equivalent to the existence of a sequence of states $\hat{\rho}^{(n)}$, all of which reduce to $\rho_{>v_0}$ on the right, such that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} d\tilde{v} \langle T_{vv}\rangle|_{\hat{\rho}^{(n)}} = -\frac{\hbar}{2\pi}\partial_v S_{\mathrm{rel}}|_{v_0} \ . \tag{A.14.5}$$

Here we will assume the conjecture to be true. We will be interested in certain universal properties of the states in this sequence that emerge in the limit as $n \to \infty$.

Properties of the minimum energy completion in 1+1 dimensions

For compactness of notation, we will ascribe any limiting properties of the states $\hat{\rho}^{(n)}$ as $n \to \infty$ to a "limit state" $\hat{\rho}^{\infty}$. We stress that such a state need not exist. Rather, $\hat{\rho}^{\infty}$ is shorthand for $\lim_{n\to\infty} \hat{\rho}^{(n)}$, where the limit should be moved outside of any maps of the state to other quantitities. Moreover, we indicate $\hat{\rho}^{\infty}$ as the argument of a map by the superscript ∞. For example,

$$S^{\infty}(v_0) = \lim_{n\to\infty} \left[-\mathrm{Tr}\,\hat{\rho}^{(n)}_{>v_0} \log \hat{\rho}^{(n)}_{>v_0}\right] \ . \tag{A.14.6}$$

By Eq. (A.14.4) and the discussion leading to Eq. (A.14.3), the state $\hat{\rho}^{\infty}$ must saturate both inequalities in Eq. (6.4.12):

$$\partial_v \bar{K}^{\infty}|_{v_0} = \partial_v \bar{S}^{\infty}|_{v_0} = \partial_v S^{\infty}|_{v_0} \ . \tag{A.14.7}$$

The first equality implies

$$\partial_v \bar{S}^{\infty}_{\mathrm{rel}}|_{v_0} = 0 \ . \tag{A.14.8}$$

Applying the left analogues of Eqs. (A.14.2) and (A.14.4) to $\bar{M}$ (with $\hat{\rho}^{\infty}$ as the input state!), we have

$$\partial_v^2 \bar{S}^{\infty}_{\text{rel}} \geq 0 \ , \tag{A.14.9}$$

for all v. The above two consequences of Wall's conjecture, combined with positivity and monotonicity of the left relative entropy,

$$\bar{S}^{\infty}_{\text{rel}} \ \geq \ 0 \ , \tag{A.14.10}$$

$$\partial_v \bar{S}^{\infty}_{\text{rel}} \ \geq \ 0 \ , \tag{A.14.11}$$

imply that

$$\partial_v \bar{S}^{\infty}_{\text{rel}} = 0 \text{ for all } v < v_0 \ . \tag{A.14.12}$$

This is a very strong condition and it intuitively suggests that for $v < v_0$ we have a vacuum-like state. In particular all local observable in the region between v and v_0 for $v < v_0$ need to register vacuum values otherwise we would have $S^{\infty}_{\text{rel}}(v_0) > S^{\infty}_{\text{rel}}(v)$. This is particular tells us that

$$\langle T_{vv}(v) \rangle|_{\hat{\rho}^{\infty}} = 0 \text{ for } v < v_0 \ . \tag{A.14.13}$$

The above equation combined with Eq. (A.14.12) implies

$$\partial_v^2 \bar{S}^{\infty} = 0 \implies \partial_v \bar{S}^{\infty} = \alpha \text{ for } v \leq v_0 \ , \tag{A.14.14}$$

In fact, in 1+1 CFTs we can argue that $\alpha = 0$ by invoking the strengthened version of the QNEC [210, 19] [10]:

$$\langle T_{vv} \rangle \geq \frac{\hbar}{2\pi} \partial_v^2 \bar{S} + \frac{6\hbar}{c} (\partial_v \bar{S})^2 \ . \tag{A.14.15}$$

Now, Eq. (A.14.12) implies that

$$\langle T_{vv} \rangle = \frac{\hbar}{2\pi} \partial_v^2 \bar{S} \text{ for } v < v_0 \ , \tag{A.14.16}$$

which together with Eq. (A.14.15) implies that $\partial_v \bar{S} = 0$. So, we conclude that for $v < v_0$,

$$\partial_v \bar{S}^{\infty} = 0 \text{ and } \partial_v \bar{S}^{\infty}_{\text{rel}} = 0 \implies \tag{A.14.17}$$

$$\lim_{\epsilon \to 0} \int_{-\infty}^{v_0 - \epsilon} d\tilde{v} \, \langle T_{vv} \rangle \big|_{\hat{\rho}^{(\infty)}} = 0 \ . \tag{A.14.18}$$

We also know that

$$\lim_{\epsilon \to 0} \left[\int_{-\infty}^{v_0 + \epsilon} d\tilde{v} \, \langle T_{vv} \rangle \big|_{\hat{\rho}^{\infty}} \right] = \frac{\hbar}{2\pi} \partial_v S \big|_{v_0} \ . \tag{A.14.19}$$

[10] We thank Aron Wall for suggesting the use of strengthened QNEC here

This along with Eq. (A.14.17) implies that the minimum energy state contains a shock (a delta function in energy density) at v_0, and vanishing energy to its left:

$$\langle T_{vv}\rangle = \left(\frac{\hbar}{2\pi}\partial_v S\big|_{v_0}\right)\delta(v-v_0) \text{ for } v \le v_0 \ . \tag{A.14.20}$$

If $\hat{\rho}^\infty$ is a pure state[11] this further implies that

$$\partial_v S = 0 \text{ for } v < v_0 \ . \tag{A.14.21}$$

In fact, we expect that $\hat{\rho}^\infty$ can always taken to be pure. The basic idea is that any density operator can be purified by a suitable auxiliary system. In general the auxiliary system has to be external, but we now argue it can be taken to be distant soft modes in the quantum field itself.

Suppose we had identified a sequence $\hat{\rho}^{(n)}$ that limits to a mixed $\hat{\rho}^\infty$. Finiteness of the energy requires that each state in the sequence looks like the vacuum in some sufficiently distant left region $v < v^{(n)}$ with $v^{(n)} < v_0$. We can take $v^{(n)} \to -\infty$ as $n \to \infty$. We can add a purification of the state $\hat{\rho}^{(n)}$ in soft wavepackets localized to the region $v < v^{(n)}$. This results in a new, pure state and we redefine $\hat{\rho}^{(n)}$ to be that state. Since we have not modified the state in the region $v > v_0$, it will still reduce to the given right state $\rho_{>v_0}$; and since the region $v < v^{(n)}$ is semi-infinite, we can take the purifying wave-packets to have arbitrarily small energy. In particular, we can take their contribution to the energy to vanish in the limit as $n \to \infty$.

Higher-dimensional case

The generalization of the above result to higher dimensions is straightforward. We can consider any Killing horizon $N = \mathbf{R} \times \mathcal{B}$, with $v \in \mathbf{R}$ an affine parameter along light-rays orthogonal to the $d-2$ dimensional spatial surface $\mathcal{B}$ with collective coordinates y.

The analogue of the 1+1 dimensional ant is now an army of ants that have walked along the null generators from $v = +\infty$ to the position $v = V(y)$, so that they know the state $\rho_{>V(y)}$. (See Fig. A.5.) The ants again ask about the minimum global energy consistent with this knowledge, $M[V(y)]$. This quantity can only decrease under deformations of $V(y)$ that are everywhere positive:

$$\frac{\delta M}{\delta V(y)} \le 0 \ . \tag{A.14.22}$$

The definition of M differs from the 1+1 case only through an additional transverse integral over $d^{d-2}y$. It can be shown [112, 68] that the modular Hamiltonian, too, is simply the sum of the local Rindler energies associated with the individual null generators, Eq. (6.4.6):

$$\Delta K(V_0(y)) = \frac{2\pi}{\hbar}\int d^{d-2}y \int_{V_0(y)}^{\infty} dv\,(v - V_0(y))T_{vv} \ , \tag{A.14.23}$$

[11] The conclusion would extend to mixed states under the assumption that $\Delta S(v)$ remains bounded from below for any v in the limit as $n \to \infty$. The status of this assumption is not clear to us, however.

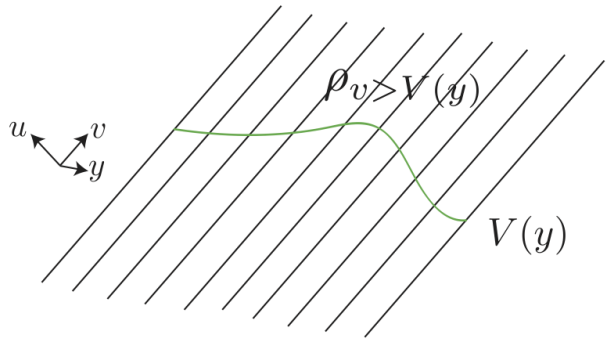

Figure A.5: A general cut of the Rindler horizon in $d > 2$. An army of ants marches down along the null direction towards the cut. Given the state above the cut, they ask what is the minimum energy still to come.

By the analogue of Eq. (6.4.12),

$$\frac{\delta \bar{K}}{\delta V(y)} \geq \frac{\delta \bar{S}}{\delta V(y)} \geq \frac{\delta S}{\delta V(y)} \ , \tag{A.14.24}$$

one finds

$$M \geq -\frac{\hbar}{2\pi} \frac{\delta S_{\rm rel}}{\delta V(y)} \ . \tag{A.14.25}$$

The ant conjecture again demands that this be an equality. That is, there exists a global state $\hat{\rho}^\infty$ that saturates Eq. (A.14.25) (or if not, saturation can at least be approached, in the limit of a sequence of global states). The same arguments as in the 1+1 dimensional case imply that $\hat{\rho}^\infty$ satisfies

$$\frac{\delta \bar{S}^\infty_{\rm rel}}{\delta V(y)} = 0 \text{ for all } v < V_0(y) \ . \tag{A.14.26}$$

Exactly as in the 1+1 case, the above condition implies

$$\langle T_{vv}(v) \rangle|_{\hat{\rho}^\infty} = 0 \quad \text{for} \quad v < V_0(y) \ , \tag{A.14.27}$$

$$\frac{\delta^2 \bar{S}^\infty}{\delta V(y_1) \delta V(y_2)} = 0 \implies \frac{\delta \bar{S}^\infty}{\delta V(y)} = \alpha \quad \text{for all } v < V_0(y) \ . \tag{A.14.28}$$

where α is some constant. As was discussed at the end of the previous section, we can take ρ^∞ to be a limit of pure states where we additionally have

$$\frac{\delta S^\infty}{\delta V(y)} = \alpha \quad \text{for all } v < V_0(y) \ , \tag{A.14.29}$$

At this point, it would be nice to argue that $\alpha = 0$ as in the 1+1 dimensional case, but we will leave this to future work. If we assume that $\alpha = 0$, then the purity of the global state implies

$$\frac{\delta \bar{S}^{\infty}}{\delta V(y)} = 0 \quad \text{for all } v < V_0(y) \ , \tag{A.14.30}$$

and together with Eq. (A.14.26) one obtains

$$\lim_{\epsilon \to 0} \left[\int_{-\infty}^{V_0(y)-\epsilon} d\tilde{v} \, \langle T_{vv} \rangle \big|_{\hat{\rho}^{\infty}} \right] = 0 \ , \tag{A.14.31}$$

for all y. Note that Eq. (A.14.31) does not otherwise follow from Eq. (A.14.27): because $\hat{\rho}^{\infty}$ is defined as a limit of a sequence, it would be possible for $\langle T_{vv} \rangle$ to approach zero while its integral approaches a finite value. Assuming the ant conjecture, that Eq. (A.14.25) is an equality, it follows that

$$\begin{aligned} \langle T_{vv}(v,y) \rangle \big|_{\hat{\rho}^{\infty}} &= \left(\frac{\hbar}{2\pi} \frac{\delta S}{\delta V(y)} \big|_{V_0} \right) \delta(v - V_0(y)) \\ &\quad \text{for } v \leq V_0(y) \ . \end{aligned} \tag{A.14.32}$$

To summarize, in 1+1 dimensions, the ant conjecture implies the key properties of the coarse-graining states we conjectured: Eqs. (6.4.14) and (6.4.15) hold on a Killing horizon. In greater than 1+1 dimensions, this implication obtains only with the unproven assumption that $\alpha = 0$ above.

A.15 Null Limit of the Kink Transform

In this appendix we apply the kink transform to a Cauchy slice Σ that has null segments. In the null limit we express the kink transform in terms of the null initial value problem. We then show that this leads to a shock in the Weyl tensor for $d > 2$. From this Weyl shock we extract the boundary stress tensor shock. This serves a two-fold purpose. The first is that it provides direct intuition for how the kink transform modifies the geometry. The second is that, as will be evident from the calculation below, the derivation of the stress tensor shock from the Weyl shock works even for wiggly cuts of the Rindler horizon on the boundary.[12]

Let N_k be a null segment of Σ in a neighborhood of $\mathcal{R}$ and let k^a be the null generator of N_k. We now allow the boundary anchor of $\mathcal{R}$ to be an arbitrary cut $V_0(y)$ of the Rindler horizon, as considered in Sec. 7.2. Lastly, denote by P^a_μ and P^i_μ the projectors onto N_k and

[12]The results of this section do not apply when $d = 2$, as the shear and the Weyl tensor vanish identically. However in $d = 2$ there is no distinction between flat and wiggly cuts on the boundary so we gain neither additional intuition nor generality compared with the analysis in Sec. 7.4.

cross-sections of N_k (including the RT surface $\mathcal{R}$), respectively. We can compose these to obtain the projector P_a^i.

By Eq. (7.3.4), when Σ is spacelike in a neighborhood of $\mathcal{R}$ the kink transform can be contracted as follows:

$$x^a(K_\Sigma)_{ab} \to x^a(K_\Sigma)_{ab} - \sinh(2\pi s) x_b \, \delta(\mathcal{R}) \ . \tag{A.15.1}$$

In the null limit both x^a and t^μ approach k^a. Therefore, the quantity in the LHS of Eq. (A.15.1) has the following null limit:

$$x^a(K_\Sigma)_{ab} \overset{\text{null}}{\to} k^a \nabla_a k_b \ . \tag{A.15.2}$$

The transformation of Eq. (A.15.1) then becomes

$$\kappa \to \kappa - \sinh(2\pi s)\delta(\lambda) \ , \tag{A.15.3}$$

where λ is a null parameter adapted to k^a and κ is the inaffinity defined by

$$k^b \nabla_b k^a = \kappa k^a \ . \tag{A.15.4}$$

We refer to this transformation as the *left stretch*, as it arises from a one-sided dilatation along N_k. This transformation was originally described in [139] in the context of black hole coarse-graining.

We now show that the left stretch generates a Weyl tensor shock at the RT surface. The shear of a null congruence is defined by

$$\sigma_{ij} = P_i^a P_j^b \nabla_{(a} k_{b)} \ . \tag{A.15.5}$$

It satisfies the evolution equation [156]

$$\mathcal{L}_k \sigma_{ij} = \kappa \sigma_{ij} + \sigma_i{}^k \sigma_{kj} - P_i^\mu P_j^\mu k^a k^b C_{a\mu b\nu} \ . \tag{A.15.6}$$

Now let λ be a parametrization of N_k adapted to k^a, with $\lambda = 0$ corresponding to $\mathcal{R}$. In terms of λ, the evolution equation can be written as

$$\partial_\lambda \sigma_{ij} = \kappa \sigma_{ij} + \sigma_i{}^k \sigma_{kj} - C_{\lambda i \lambda j} \ . \tag{A.15.7}$$

Consider now the new spacetime $\mathcal{M}_s$ generated by the left stretch. As in Sec. 7.4, we denote quantities in $\mathcal{M}_s$ with tildes. We can then write the evolution equation in $\mathcal{M}_s$,

$$\partial_{\tilde{\lambda}} \tilde{\sigma}_{ij} = \tilde{\kappa} \tilde{\sigma}_{ij} + \tilde{\sigma}_i{}^k \tilde{\sigma}_{kj} - \tilde{C}_{\lambda i \lambda j} \ . \tag{A.15.8}$$

Since k^a is tangent to N_k, and $(N_k)_s = N_k$ as submanifolds, we can identify k^a with $\tilde{k}^a$. Thus we can use the same parameter λ in both spacetimes. Since σ_{ij} is intrinsic to N_k, we

can identify σ_{ij} and $\tilde{\sigma}_{ij}$ for the same reason. Comparing Eqs. (A.15.7) and (A.15.8), and inserting Eq. (A.15.3), we find that there is a Weyl shock

$$\tilde{C}_{\lambda i\lambda j} = C_{\lambda i\lambda j} - \sinh(2\pi s)\sigma_{ij}\delta(\lambda) \ . \tag{A.15.9}$$

We now show that the Weyl shock Eq. (A.15.9) reproduces the near boundary shock Eq. (7.2.44), but now for wiggly cuts of the Rindler horizon. To do this, we evaluate both σ_{ij} and $C_{\lambda i\lambda j}$ in Fefferman-Graham coordinates to leading non-trivial order. The Fefferman-Graham coordinates for $\mathcal{M}$ and $\mathcal{M}_s$ are defined exactly as in Sec. 7.4, except we now use null coordinates (u, v) and $(\tilde{u}, \tilde{v})$ on the boundary as defined in Sec. 7.2. To start with, we note that $k_a\partial_z\bar{X}^a = 0$ since $\partial_z\bar{X}^a$ is tangent to the RT surface. Evaluating this at leading order yields the relation

$$k_z = -dz^{d-3}\mathcal{U}_{(d)} + \mathcal{O}(z^{d-4}) \ . \tag{A.15.10}$$

We recall that

$$\mathcal{U}_{(d)} = -\frac{4G}{d}\frac{\delta S}{\delta V}\Big|_{V_0} \ . \tag{A.15.11}$$

Moreover, the projector is given by

$$P_i^\mu = \partial_i\bar{X}^\mu \ . \tag{A.15.12}$$

From this definition, one can check that

$$P_i^z = \delta_i^z + \mathcal{O}(z^{d-1}) \ , \tag{A.15.13}$$
$$P_i^A = \mathcal{O}(z^{d-1}) \ . \tag{A.15.14}$$

Furthermore,

$$\begin{aligned} &\nabla_z k_A,\ \nabla_A k_z \sim \mathcal{O}(z^{-1}) \ , \\ &\nabla_A k_B \sim \mathcal{O}(1) \ , \\ &\nabla_z k_z = -d(d-2)\mathcal{U}_{(d)}z^{d-4} + \mathcal{O}(z^{d-5}) \ , \end{aligned} \tag{A.15.15}$$

where we have used that $k^A \sim \mathcal{O}(1)$. Hence to leading order we simply have

$$\sigma_{ij} = -d(d-2)\mathcal{U}_{(d)}z^{d-4}\delta_i^z\delta_j^z + \mathcal{O}(z^{d-5}) \ . \tag{A.15.16}$$

Finally, a straightforward but tedious calculation of the Weyl tensor yields

$$\tilde{C}_{\tilde{v}i\tilde{v}j} = C_{vivj} - 8\pi G(d-2)\left(\langle\tilde{T}_{\tilde{v}\tilde{v}}\rangle - \langle T_{vv}\rangle\right)z^{d-4}\delta_i^z\delta_j^z\delta(\tilde{v} - V_0) + \mathcal{O}(z^{d-5}) \ , \tag{A.15.17}$$

where we have used that $\lambda \to v, \tilde{v}$ as $z \to 0$ in $\mathcal{M}, \mathcal{M}_s$ respectively. Putting this together yields the desired shock for wiggly cuts of the Rindler horizon.

A.16 (Quantum) Trapped Surfaces in the Schwarzschild Geometry

Classical Solution and Semiclassical Corrections

The Schwarzschild metric is

$$ds^2 = -\left(1 - \frac{R}{r}\right) dt^2 + \frac{dr^2}{1 - R/r} + r^2 d\Omega^2 \ . \tag{A.16.1}$$

where $R = 2GM$ is the Schwarzschild radius. In ingoing Eddington-Finkelstein coordinates,

$$ds^2 = -\left(1 - \frac{R}{r}\right) dv^2 + 2dv\, dr + r^2 d\Omega^2 \ , \tag{A.16.2}$$

where

$$v = t + r_* \ , \quad r_* = r + R \log|\frac{r}{R} - 1| \ , \quad \frac{dr}{dr_*} = 1 - \frac{R}{r} \ . \tag{A.16.3}$$

Ingoing radial null congruences are at constant v, so $dv = 0$. Outgoing null congruences satisfy $dv = 2dr_*$, so

$$v = 2r_* + \text{ const} \ . \tag{A.16.4}$$

We are interested in their expansion,

$$\theta = \frac{dA/d\lambda}{A} \tag{A.16.5}$$

in terms of a convenient affine parameter, λ.

To find λ, first note that r is an affine parameter. This follows because $A = 4\pi r^2$, so

$$\theta = \frac{2}{r}\frac{dr}{d\lambda} \ ; \tag{A.16.6}$$

and Raychaudhuri's equation in the vacuum, for spherical symmetry, reduces to

$$\frac{d\theta}{d\lambda} + \frac{1}{2}\theta^2 = 0 \ . \tag{A.16.7}$$

This implies that $dr/d\lambda$ must be constant for any affine λ. We can take that constant to be 1 if we like, and choose another constant of integration so that $r = \lambda$.

However, this choice is not convenient for outgoing lightrays, because we are interested in radial null congruences near and on the event horizon,

$$|r - R| \ll R \ . \tag{A.16.8}$$

Intuitively, the radius r does not change much for these congruences, so small changes in r correspond to large motions along the congruence. On the horizon, r is degenerate, and inside the black hole, r runs towards the past.

To remedy this, let us consider the coordinate distance $c = r - R$ from the horizon. We will work in the near-horizon limit of Eq. (A.16.8), i.e., to first order in $c/R \ll 1$. For example, $r_* = R + R\log(|c|/R)$ in this approximation; and by Eq. (A.16.4), an outgoing congruence satisfies $v = 2R\log(|c|/R)+$ const. Inverting this, we find

$$c = c_0 e^{v/2R} \tag{A.16.9}$$

where c_0 is the coordinate distance from the horizon at $v = 0$. This is the quantity that vanishes on the horizon and goes negative inside, so we can define a nondegenerate, always future-directed parameter by choosing $\lambda = c/c_0$. This is affine since $\lambda = (r - R)/c_0$ and r is affine.

To summarize, we choose the affine parameter

$$\lambda = e^{v/2R} \tag{A.16.10}$$

on outgoing null geodesics near the horizon. By Eq. (A.16.6), the expansion of any such congruence is given by

$$\theta = \frac{2c_0}{R} \ , \tag{A.16.11}$$

where we again used $r - R \ll R$. All surfaces on the event horizon have $c_0 = 0$ and hence $\theta = 0$; they are marginally outer trapped. It is easy to check that these are the only such surfaces.

Any null vector tangent to the outgoing congruences must be proportional to $\partial_t + \partial_{r_*}$. Let k^a be the particular null vector associated to the affine parameter λ. From Eq. (A.16.10) we have

$$k = \frac{d}{d\lambda} = \frac{2R}{\lambda}\,\frac{d}{dv}\bigg|_{\text{cong}} = \frac{R}{\lambda}(\partial_t + \partial_{r_*}) \ , \tag{A.16.12}$$

For the second equality, we used that on the outgoing congruence $t = (v + \text{const})/2$, $r_* = (v - \text{const})/2$.

For all ingoing spherical congruences in the region covered by the ingoing Eddington-Finkelstein coordinates, $-r$ is a future-directed nondegenerate affine parameter. Thus Eq. (A.16.6) implies that their expansion, θ_l, is everywhere negative. This establishes that every spherical cut of the event horizon is marginally trapped, i.e., satisfies $\theta = 0$ and $\theta_l \le 0$.

To treat quantum matter as a small perturbation, we expand the Einstein equation, $G_{ab} = 8\pi G\langle T_{ab}\rangle$, in powers of $G\hbar$, to first order. (We drop the expectation value symbol below.) In this approximation, we can compute matter effects on the expansion of congruences by integrating the Raychaudhuri equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \varsigma^2 - 8\pi G T_{kk} \ . \tag{A.16.13}$$

Here $T_{kk} = T_{ab}k^a k^b$, and $k^a = (\frac{d}{d\lambda})^a$ is the affine tangent vector to the null congruence. The shear term vanishes for the spherical congruences we consider. In general, the θ^2 term will be $O((G\hbar)^0)$ and thus dominant.

However, here we will be interested in surfaces where classical and quantum effects compete. Such surfaces must have $\theta \sim O(G\hbar)$ classically. By Eq. (A.16.11) they are found in a neighborhood $|c| \leq O(G\hbar)$ of the event horizon. Hence $\theta^2 \sim O((G\hbar)^2)$ will be negligible in the region of interest, and Eq. (A.16.13) reduces to

$$\theta(\lambda) - \theta(\lambda_0) = -8\pi G \int_{\lambda_0}^{\lambda} T_{kk} \ . \tag{A.16.14}$$

Classically Trapped Surfaces During Evaporation

We will now compute the effect of the quantum stress tensor for the Unruh state [192] on the position of (marginally) trapped surfaces in the Schwarzschild geometry.

The renormalized stress tensor in the Unruh vacuum takes the form

$$\langle U| T_a^{\ b} |U\rangle_{\rm ren} \xrightarrow{r\to 2M} \frac{L}{4\pi R^2} \begin{pmatrix} f^{-1} & -1 \\ f^{-2} & -f^{-1} \end{pmatrix}, \tag{A.16.15}$$

where $f = (1 - R/r)$, $R = 2M$, a and b range over t and r, and

$$L \sim \frac{\hbar}{R^2} \tag{A.16.16}$$

is the luminosity of the black hole. Lowering indices we find

$$\langle U| T_{ab} |U\rangle_{\rm ren} \xrightarrow{r\to 2M} \frac{L}{4\pi R^2} \begin{pmatrix} -1 & -f^{-1} \\ -f^{-1} & -f^{-2} \end{pmatrix}, \tag{A.16.17}$$

Using

$$\partial_{r*} = \frac{dr}{dr*}\partial_r = \left(1 - \frac{R}{r}\right)\partial_r \ , \tag{A.16.18}$$

we can express the null vector k in (t, r) coordinates,

$$k = \frac{R}{\lambda}\left(\partial_t + \left(1 - \frac{R}{r}\right)\partial_r\right) = k^t\partial_t + k^r\partial_r \ . \tag{A.16.19}$$

and we obtain

$$\begin{aligned} \langle T_{\mu\nu}k^\mu k^\nu\rangle &= \langle T_{tt}k^t k^t\rangle + \langle T_{rr}k^r k^r\rangle + 2\langle T_{tr}k^t k^r\rangle \\ &= -\frac{L}{\pi\lambda^2} = -\frac{\hbar}{\pi R^2\lambda^2} \end{aligned} \tag{A.16.20}$$

Next we compute the change in the expansion induced by the above quantum stress tensor. We consider a black hole at the onset of evaporation, for which there is no Hawking radiation outside the near horizon zone yet. Thus we expect the geometry to revert to the

classical vacuum Schwarzschild solution far from the black hole. And so, to find the corrected expansion, we integrate backwards from $\lambda = \infty$ to find the shift:

$$\begin{aligned}\delta\theta \equiv \theta(\lambda) - \theta(\infty) &= -8\pi G \int_{\infty}^{\lambda} \langle T_{\mu\nu} k^\mu k^\nu \rangle d\lambda' = \\ &= 8\pi G \int_{\lambda_0}^{\lambda} \frac{\hbar}{\pi R^2 \lambda'^2} d\lambda' = -\frac{8G\hbar}{R^2\lambda} \ .\end{aligned} \tag{A.16.21}$$

To find the (classically) marginally trapped surfaces in the Unruh state, we solve

$$\theta^{(0)} + \delta\theta = 0 \ , \tag{A.16.22}$$

where $\theta^{(0)}$ is the uncorrected classical expansion given in Eq. A.16.11. Using $c = c_0\lambda$, we find that the classical marginally trapped surfaces are located at

$$c_{\rm MTS} \sim \frac{G\hbar}{R} \tag{A.16.23}$$

in the quantum-corrected geometry. Very near the horizon, we can treat the radial coordinate to be essentially R to zeroth order.

An alternative useful notion of distance is the proper radial distance from the horizon, ℓ, which satisfies

$$d\ell = \frac{dr}{\sqrt{1-\frac{R}{r}}} \simeq \sqrt{R}\frac{dr}{\sqrt{r-R}} \quad \to \quad \ell \simeq 2\sqrt{R(r-R)} \sim (Rc)^{1/2} \tag{A.16.24}$$

Since $G\hbar = l_p^2$, we see that the trapped surfaces are about a Planck length outside the horizon:

$$\ell_{\rm MTS} \sim \mathcal{O}(l_p). \tag{A.16.25}$$

Thus, the area of the classical marginally trapped surface is increased by the quantum correction, by

$$\Delta A_{\rm MTS} \sim G\hbar = l_P^2 \tag{A.16.26}$$

Quantum Trapped Surfaces During Evaporation

We still consider the quantum-corrected geometry in the Unruh state, so the classical expansion is given by

$$\theta = \theta^{(0)} + \delta\theta \sim \frac{c_0}{R} - \frac{G\hbar}{R^2\lambda} \ . \tag{A.16.27}$$

The generalized entropy is

$$S_{\rm gen} = \frac{A}{4G\hbar} + S \ , \tag{A.16.28}$$

where $S = -\mathrm{Tr}\,\rho \log \rho$ and ρ is the quantum state in the region exterior to the Cauchy-splitting sphere. The quantum expansion Θ is ($4G\hbar$ times) the rate of change of the generalized entropy, per unit area, under shape deformations. In the spherically symmetric case,

$$\Theta = \theta + \frac{4G\hbar}{A}\frac{dS}{d\lambda} \ , \tag{A.16.29}$$

Quantum marginally trapped surfaces are characterized by $\Theta = 0$.

The Generalized Second Law (GSL) states that any outgoing radial congruence on or outside the event horizon must satisfy $\Theta \geq 0$, so the quantum marginally trapped surfaces must lie inside the horizon [179]. By Eq. (A.16.27), $\theta < 0$ on and inside the horizon. We see from Eq. (A.16.29) that the GSL requires

$$\frac{4G\hbar}{A}\frac{dS}{d\lambda} = -\alpha\theta|_{\mathcal{H}} \ , \tag{A.16.30}$$

where $\mathcal{H}$ refers to the horizon. We take $\alpha - 1 \sim O(1)$, in line with Page's explicit calculation for an evaporating black hole in the Unruh state [211].

Combining these results and neglecting factors of order unity where appropriate, we find

$$\Theta = \theta - \alpha\theta|_{\mathcal{H}} = \frac{c}{R\lambda} - \frac{G\hbar}{R^2\lambda} + \alpha\frac{G\hbar}{R^2\lambda} \ . \tag{A.16.31}$$

Setting $\Theta = 0$ yields

$$\frac{c}{R\lambda} = -(\alpha - 1)\frac{G\hbar}{R^2\lambda} \quad \rightarrow \quad c \sim -\frac{G\hbar}{R}. \tag{A.16.32}$$

Using the proper area, we find

$$\Delta A_{\mathrm{QMTS}} \sim -l_P^2 \ . \tag{A.16.33}$$

Thus, the quantum marginally trapped surfaces are a proper distance of order the Planck length inside of the horizon.

We will now show that the "duration" of the lightsheet L of a quantum marginally trapped surface μ_Q is of order of scrambling time

$$\Delta t_s \sim R \log \frac{R}{l_P} \ . \tag{A.16.34}$$

This assumes that μ_Q is about one Planck length inside of the event horizon, as would be the case for an isolated, slowly evaporating black hole. Of course, the points on L are null or spacelike separated. What we mean by the "duration" of L is the amount of time, as measured at large radius r, for which it will be the case that matter falling in radially from this radius will cross L (see Fig. A.6).

We will approximate the infalling matter as ingoing radial null geodesics; the result would be the same for timelike geodesics starting at rest at large radius. Let the earliest geodesic crossing L be at $v = v_1$ in the Eddington-Finkelstein coordinates defined in Appendix A.16. It will meet L at μ_Q, whose radius satisfies $R - r_{\mu_Q} \sim l_P^2/R$. The last geodesic that meets

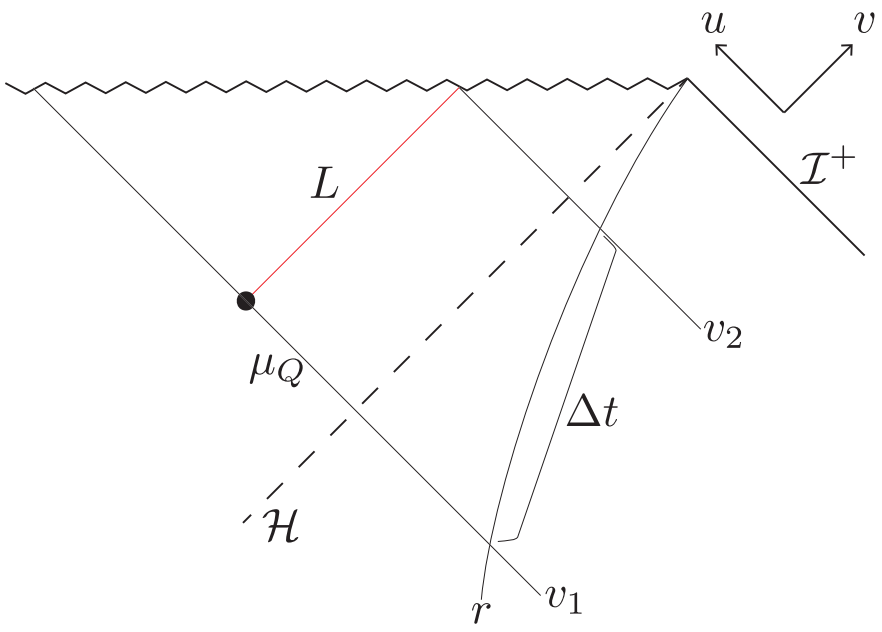

Figure A.6: The future outgoing lightsheet of μ_Q (top red line) is crossed by two ingoing radial null geodesics at v_1 (at μ_Q) and v_2 (at the singularity). Their Schwarzschild time difference at fixed r is the scrambling time, Δt_s.

L will do so where L hits the singularity, at $r = 0$. The lightsheet L is characterized by $u = const$, where u is the ingoing Eddington-Finkelstein coordinate, $u \equiv t - r_*$. Here r_* is the tortoise coordinate defined in Eq. A.16.3. Since r_* depends only on r, we have

$$\Delta t = t_2 - t_1 = r_*(r_{\mu_Q}) - r_*(0) = r_{\mu_Q} + R \log \frac{R}{l_P^2/R} \sim \Delta t_s \ . \tag{A.16.35}$$

A similar analysis demonstrates that the scrambling time is how long it takes a geodesic to propagate from about a Planck distance outside the horizon to the edge of the near-horizon zone, at $r = 3R/2$.

A.17 Perturbative Construction of Q-screens

Let μ_Q be a quantum marginally trapped surface near a perturbed Killing horizon that approaches the Hartle-Hawking state in the future. Then there exists a Q-screen that approaches the Killing horizon in the future and contains μ_Q as a leaf.

This fact is useful in sketching a heuristic argument for our conjectured QPI in asymptotically AdS spacetime, following Eq. (8.7.8). We will now demonstrate this claim by explicit construction.

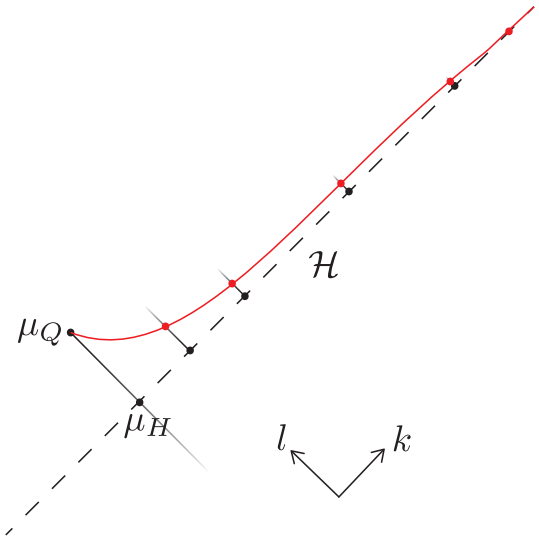

Figure A.7: A quantum marginally trapped surface μ_Q in the vicinity of a perturbed Killing horizon $\mathcal{H}$. We construct a Q-screen containing μ_Q that asymptotes to the Killing horizon at late times. We first fire a null plane towards $\mathcal{H}$ that intersects it on μ_H. We then foliate $\mathcal{H}$ starting from μ_H. At every leaf of this foliation, we fire null planes inwards and to the future. On each null plane, we find a quantum marginally trapped surface at an affine distance δU from $\mathcal{H}$. The Q-screen is the union of these quantum marginally trapped surfaces.

Consider an event horizon $\mathcal{H}$ which is a perturbation to a Killing horizon caused by matter excitations $T_{\mu\nu} \sim \mathcal{O}(\hbar)$ such that in the far future $\mathcal{H}$ settles down to a Killing horizon in the Hartle-Hawking state. Furthermore, assume that there exists a quantum marginally trapped surface near $\mathcal{H}$. It is known [179] that quantum marginally trapped surfaces are behind event horizons, so μ_Q will be a small distance in the inward direction l from $\mathcal{H}$. Given any co-dimension 2 surface in this spacetime, k and l respectively represent the outward and inward null vectors perpendicular to the surface. Let y parametrize the transverse position of the surface; see Fig. A.7.

For the construction of the Q-screen, we start by emanating a past outwards-directed null plane from μ_Q and mark its intersection with the horizon as μ_H. Now, we can pick a foliation of the horizon that starts from μ_H and continues towards the future of $\mathcal{H}$ such that it eventually approaches the preferred foliation of the Killing horizon. Mark the leaves of this foliation by λ such that $\lambda = 0$ is μ_H and λ grows along the future leaves. We construct the Q-screen by shooting null future-directed inward null planes from the leaves μ_H and on that null plane look for a quantum marginally trapped surface.

Suppose that a given leaf of our foliation of $\mathcal{H}$ (marked by λ) has a quantum expansion $\Theta_k(\lambda; y)$ at a given transverse position y. By the generalized second law, $\Theta_k \geq 0$. Then,

perturbatively we can find the location of a quantum marginally trapped surface as

$$\Theta_k(\lambda; y) + \delta U(\lambda; y)\,(\partial_l \Theta_k(\lambda; y)) = 0 \ , \tag{A.17.1}$$

where δU is the amount of affine parameter in the l direction we need to venture to find a quantum marginally trapped surface and $\Theta_k = \mathcal{O}(G\hbar)$.

We need to solve for a function $\delta U(y)$ and show that it approaches zero as we go towards higher values of λ. From the definition of quantum expansion it follows that

$$\partial_l \Theta_k = \partial_l \theta_k + 4G\hbar \partial_l \partial_k S_{\text{out}} \ . \tag{A.17.2}$$

The cross-focusing equation is

$$\partial_l \theta_k = -\frac{1}{2}\mathcal{R} - \theta_l \theta_k + \nabla.\chi + \chi^2 + 8\pi G\ T_{kl} \ , \tag{A.17.3}$$

where $\mathcal{R}$ is the intrinsic Ricci scalar of the leaf and χ is its twist [212]. From Eq. (A.17.1), we see that in order to solve for δU to first non-trivial order in $G\hbar$, we only need the leading order expression for $\partial_l \Theta_k$. The leading order term is

$$\partial_l \Theta_k = -\frac{1}{2}\mathcal{R}^{(0)} + \mathcal{O}(G\hbar) \ , \tag{A.17.4}$$

where $\mathcal{R}^{(0)}$ is the (y-independent) intrinsic Ricci scalar of the leaf on the unperturbed Killing horizon. For a 2-sphere $\mathcal{R}^{(0)} = 2$. Combining the above equations with (A.17.1), we can solve for δU to the first non-trivial order in $\mathcal{O}(G\hbar)$:

$$\delta U(y; \lambda) = \Theta_k(\lambda; y) \ . \tag{A.17.5}$$

Since by assumption $\mathcal{H}$ approaches a Killing horizon in the Hartle-Hawking state in the future, we have

$$\lim_{\lambda \to \infty} \Theta_k(\lambda; y) = 0 \implies \lim_{\lambda \to \infty} \delta U(\lambda; y) = 0 \ , \tag{A.17.6}$$

where the implication follows from Eq. (A.17.5). This means that the leaves of the Q-screen start at μ_Q and approach the late times of the event horizon, which is what we set out to show.

Bibliography

[1] Juan Martin Maldacena. "The Large N limit of superconformal field theories and supergravity". In: *Int. J. Theor. Phys.* 38 (1999), pp. 1113–1133. DOI: 10.1023/A:1026654312961. arXiv: hep-th/9711200.

[2] Edward Witten. "Anti-de Sitter space and holography". In: *Adv. Theor. Math. Phys.* 2 (1998), pp. 253–291. DOI: 10.4310/ATMP.1998.v2.n2.a2. arXiv: hep-th/9802150.

[3] Shinsei Ryu and Tadashi Takayanagi. "Aspects of Holographic Entanglement Entropy". In: *JHEP* 08 (2006), p. 045. DOI: 10.1088/1126-6708/2006/08/045. arXiv: hep-th/0605073 [hep-th].

[4] Aitor Lewkowycz and Juan Maldacena. "Generalized Gravitational Entropy". In: *JHEP* 08 (2013), p. 090. DOI: 10.1007/JHEP08(2013)090. arXiv: 1304.4926 [hep-th].

[5] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. "Quantum Corrections to Holographic Entanglement Entropy". In: *JHEP* 11 (2013), p. 074. DOI: 10.1007/JHEP11(2013)074. arXiv: 1307.2892 [hep-th].

[6] Netta Engelhardt and Aron C. Wall. "Quantum Extremal Surfaces: Holographic Entanglement Entropy Beyond the Classical Regime". In: *JHEP* 01 (2015), p. 073. DOI: 10.1007/JHEP01(2015)073. arXiv: 1408.3203 [hep-th].

[7] Daniel L. Jafferis et al. "Relative entropy equals bulk relative entropy". In: *JHEP* 06 (2016), p. 004. DOI: 10.1007/JHEP06(2016)004. arXiv: 1512.06431 [hep-th].

[8] Ahmed Almheiri, Xi Dong, and Daniel Harlow. "Bulk Locality and Quantum Error Correction in AdS/CFT". In: *JHEP* 04 (2015), p. 163. DOI: 10.1007/JHEP04(2015)163. arXiv: 1411.7041 [hep-th].

[9] Daniel Harlow. "The Ryu-Takayanagi Formula from Quantum Error Correction". In: (2016). arXiv: 1607.03901 [hep-th].

[10] Raphael Bousso et al. "Entropy on a null surface for interacting quantum field theories and the Bousso bound". In: (June 2014). eprint: 1406.4545. URL: https://arxiv.org/abs/1406.4545.

[11] Raphael Bousso et al. "A Quantum Focussing Conjecture". In: (2015). arXiv: 1506.02669 [hep-th].

[12] Geoffrey Penington. "Entanglement Wedge Reconstruction and the Information Paradox". In: (May 2019). arXiv: 1905.08255 [hep-th].

[13] Jacob D. Bekenstein. "Black Holes and Entropy". In: *Phys. Rev.* D7 (1973), pp. 2333–2346. DOI: 10.1103/PhysRevD.7.2333.

[14] Leonard Susskind and John Uglum. "Black Hole Entropy in Canonical Quantum Gravity and Superstring Theory". In: *Phys. Rev.* D50 (1994), pp. 2700–2711. DOI: 10.1103/PhysRevD.50.2700. arXiv: hep-th/9401070 [hep-th].

[15] Aron C. Wall. "A proof of the generalized second law for rapidly changing fields and arbitrary horizon slices". In: *Phys. Rev.* D85 (2012). [Erratum: Phys. Rev.D87,no.6,069904(2013)], p. 104049. DOI: 10.1103/PhysRevD.87.069904,10.1103/PhysRevD.85.104049. arXiv: 1105.3445 [gr-qc].

[16] Raphael Bousso et al. "Proof of the Quantum Null Energy Condition". In: *Phys. Rev. D* 93 (2016), p. 024017. eprint: 1509.02542. URL: https://arxiv.org/abs/1509.02542.

[17] Srivatsan Balakrishnan et al. "A General Proof of the Quantum Null Energy Condition". In: (June 2017). eprint: 1706.09432. URL: https://arxiv.org/abs/1706.09432.

[18] Fikret Ceyhan and Thomas Faulkner. "Recovering the QNEC from the ANEC". In: (2018). arXiv: 1812.04683 [hep-th].

[19] Jason Koeller and Stefan Leichenauer. "Holographic Proof of the Quantum Null Energy Condition". In: *Phys. Rev. D* 94 (2016), p. 024026. eprint: 1512.06109. URL: https://arxiv.org/abs/1512.06109.

[20] R." "Penrose. "Naked singularities". In: *Annals N. Y. Acad. Sci.* 224 (1973), pp. 125–134. DOI: 10.1111/j.1749-6632.1973.tb41447.x.

[21] J. J Bisognano and E. H. Wichmann. "On the Duality Condition for Quantum Fields". In: *J. Math. Phys.* 17 (1976), pp. 303–321. DOI: 10.1063/1.522898.

[22] Horacio Casini, Marina Huerta, and Robert C. Myers. "Towards a Derivation of Holographic Entanglement Entropy". In: *JHEP* 05 (2011), p. 036. DOI: 10.1007/JHEP05(2011)036. arXiv: 1102.0440 [hep-th].

[23] John Cardy and Erik Tonni. "Entanglement hamiltonians in two-dimensional conformal field theory". In: *J. Stat. Mech.* 1612.12 (2016), p. 123103. DOI: 10.1088/1742-5468/2016/12/123103. arXiv: 1608.01283 [cond-mat.stat-mech].

[24] Raphael Bousso et al. "Proof of a Quantum Bousso Bound". In: *Phys. Rev.* D90.4 (2014), p. 044002. DOI: 10.1103/PhysRevD.90.044002. arXiv: 1404.5635 [hep-th].

[25] Matthias Burkardt. "Light front quantization". In: *Adv. Nucl. Phys.* 23 (1996), pp. 1–74. DOI: 10.1007/0-306-47067-5_1. arXiv: hep-ph/9505259 [hep-ph].

[26] Geoffrey L. Sewell. "Quantum fields on manifolds: PCT and gravitationally induced thermal states". In: *Annals Phys.* 141 (1982), pp. 201–224. DOI: 10.1016/0003-4916(82)90285-8.

[27] Thomas Faulkner et al. "Modular Hamiltonians for Deformed Half-Spaces and the Averaged Null Energy Condition". In: (2016). arXiv: 1605.08072 [hep-th].

[28] Raphael Bousso et al. "Proof of the Quantum Null Energy Condition". In: (2015). arXiv: 1509.02542 [hep-th].

[29] Jason Koeller and Stefan Leichenauer. "Holographic Proof of the Quantum Null Energy Condition". In: (2015). arXiv: 1512.06109 [hep-th].

[30] Chris Akers et al. "Geometric Constraints from Subregion Duality Beyond the Classical Regime". In: (2016). arXiv: 1610.08968 [hep-th].

[31] Shinsei Ryu and Tadashi Takayanagi. "Holographic derivation of entanglement entropy from AdS/CFT". In: *Phys. Rev. Lett.* 96 (2006), p. 181602. DOI: 10.1103/PhysRevLett.96.181602. arXiv: hep-th/0603001 [hep-th].

[32] Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. "A Covariant holographic entanglement entropy proposal". In: *JHEP* 07 (2007), p. 062. DOI: 10.1088/1126-6708/2007/07/062. arXiv: 0705.0016 [hep-th].

[33] Bartlomiej Czech et al. "The Gravity Dual of a Density Matrix". In: *Class. Quant. Grav.* 29 (2012), p. 155009. DOI: 10.1088/0264-9381/29/15/155009. arXiv: 1204.1330 [hep-th].

[34] Matthew Headrick et al. "Causality & Holographic Entanglement Entropy". In: *JHEP* 12 (2014), p. 162. DOI: 10.1007/JHEP12(2014)162. arXiv: 1408.6300 [hep-th].

[35] Xi Dong, Daniel Harlow, and Aron C. Wall. "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality". In: *Phys. Rev. Lett.* 117.2 (2016), p. 021601. DOI: 10.1103/PhysRevLett.117.021601. arXiv: 1601.05416 [hep-th].

[36] Aron C. Wall. "Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy". In: *Class. Quant. Grav.* 31.22 (2014), p. 225007. DOI: 10.1088/0264-9381/31/22/225007. arXiv: 1211.3494 [hep-th].

[37] Veronika E. Hubeny and Mukund Rangamani. "Causal Holographic Information". In: *JHEP* 06 (2012), p. 114. DOI: 10.1007/JHEP06(2012)114. arXiv: 1204.1698 [hep-th].

[38] Jason Koeller et al. "In Progress". In: (2017).

[39] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. 10th. New York, NY, USA: Cambridge University Press, 2011. ISBN: 1107002176, 9781107002173.

[40] Roger Penrose. "Gravitational Collapse and Space-Time Singularities". In: *Phys. Rev. Lett.* 14 (1965), pp. 57–59. DOI: 10.1103/PhysRevLett.14.57.

[41] S. W. Hawking. "Gravitational Radiation from Colliding Black Holes". In: *Phys. Rev. Lett.* 26 (1971), pp. 1344–1346. DOI: 10.1103/PhysRevLett.26.1344.

[42] S. W. Hawking. "The Chronology Protection Conjecture". In: *Phys. Rev.* D46 (1992), pp. 603–611. DOI: 10.1103/PhysRevD.46.603.

[43] John L. Friedman, Kristin Schleich, and Donald M. Witt. "Topological Censorship". In: *Phys. Rev. Lett.* 71 (1993). [Erratum: Phys. Rev. Lett.75,1872(1995)], pp. 1486–1489. DOI: 10.1103/PhysRevLett.71.1486. arXiv: gr-qc/9305017 [gr-qc].

[44] Aron C. Wall. "Proving the Achronal Averaged Null Energy Condition from the Generalized Second Law". In: *Phys. Rev.* D81 (2010), p. 024038. DOI: 10.1103/PhysRevD.81.024038. arXiv: 0910.5751 [gr-qc].

[45] Aron C. Wall. "The Generalized Second Law Implies a Quantum Singularity Theorem". In: *Class. Quant. Grav.* 30 (2013). [Erratum: Class. Quant. Grav.30,199501(2013)], p. 165003. DOI: 10.1088/0264-9381/30/19/199501,10.1088/0264-9381/30/16/165003. arXiv: 1010.5513 [gr-qc].

[46] Raphael Bousso and Netta Engelhardt. "Generalized Second Law for Cosmology". In: *Phys. Rev.* D93.2 (2016), p. 024025. DOI: 10.1103/PhysRevD.93.024025. arXiv: 1510.02099 [hep-th].

[47] Chris Akers et al. "Geometric Constraints from Subregion Duality Beyond the Classical Regime". In: (Oct. 2016). eprint: 1610.08968. URL: https://arxiv.org/abs/1610.08968.

[48] Zicao Fu, Jason Koeller, and Donald Marolf. "Violating the Quantum Focusing Conjecture and Quantum Covariant Entropy Bound in $d \geq 5$ dimensions". In: (May 2017). eprint: 1705.03161. URL: https://arxiv.org/abs/1705.03161.

[49] Xi Dong and Aitor Lewkowycz. "Entropy, Extremality, Euclidean Variations, and the Equations of Motion". In: (May 2017). eprint: 1705.08453. URL: https://arxiv.org/abs/1705.08453.

[50] Sebastian de Haro, Sergey N. Solodukhin, and Kostas Skenderis. "Holographic Reconstruction of Space-Time and Renormalization in the AdS / CFT Correspondence". In: *Commun. Math. Phys.* 217 (2001), pp. 595–622. DOI: 10.1007/s002200100381. arXiv: hep-th/0002230 [hep-th].

[51] Thomas Faulkner et al. "Gravitation from Entanglement in Holographic CFTs". In: *JHEP* 03 (2014), p. 051. DOI: 10.1007/JHEP03(2014)051. arXiv: 1312.7856 [hep-th].

[52] Zicao Fu, Jason Koeller, and Donald Marolf. "The Quantum Null Energy Condition in Curved Space". In: (June 2017). eprint: 1706.01572. URL: https://arxiv.org/abs/1706.01572.

[53] Stefan Leichenauer. "The Quantum Focusing Conjecture Has Not Been Violated". In: (2017). arXiv: 1705.05469 [hep-th].

[54] C. Imbimbo et al. "Diffeomorphisms and Holographic Anomalies". In: *Class. Quant. Grav.* 17 (2000), pp. 1129–1138. DOI: 10.1088/0264-9381/17/5/322. arXiv: hep-th/9910267 [hep-th].

[55] C. Robin Graham and Edward Witten. "Conformal Anomaly of Submanifold Observables in AdS / CFT Correspondence". In: *Nucl. Phys.* B546 (1999), pp. 52–64. DOI: 10.1016/S0550-3213(99)00055-3. arXiv: hep-th/9901021 [hep-th].

[56] Xi Dong. "Holographic Entanglement Entropy for General Higher Derivative Gravity". In: *JHEP* 01 (2014), p. 044. DOI: 10.1007/JHEP01(2014)044. arXiv: 1310.5713 [hep-th].

[57] Metin Gurses, Tahsin Cagri Sisman, and Bayram Tekin. "New Exact Solutions of Quadratic Curvature Gravity". In: *Phys. Rev.* D86 (2012), p. 024009. DOI: 10.1103/PhysRevD.86.024009. arXiv: 1204.2215 [hep-th].

[58] Robert C. Myers, Razieh Pourhasan, and Michael Smolkin. "On Spacetime Entanglement". In: *JHEP* 06 (2013), p. 013. DOI: 10.1007/JHEP06(2013)013. arXiv: 1304.2030 [hep-th].

[59] Horacio Casini, Eduardo Teste, and Gonzalo Torroba. "The a-theorem and the Markov property of the CFT vacuum". In: (2017). arXiv: 1704.01870 [hep-th].

[60] Lisa Randall and Raman Sundrum. "An Alternative to compactification". In: *Phys. Rev. Lett.* 83 (1999), pp. 4690–4693. DOI: 10.1103/PhysRevLett.83.4690. arXiv: hep-th/9906064 [hep-th].

[61] Herman L. Verlinde. "Holography and compactification". In: *Nucl. Phys.* B580 (2000), pp. 264–274. DOI: 10.1016/S0550-3213(00)00224-8. arXiv: hep-th/9906182 [hep-th].

[62] Steven S. Gubser. "AdS / CFT and gravity". In: *Phys. Rev.* D63 (2001), p. 084017. DOI: 10.1103/PhysRevD.63.084017. arXiv: hep-th/9912001 [hep-th].

[63] H. Casini. "Relative entropy and the Bekenstein bound". In: *Class.Quant.Grav.* 25 (2008), p. 205021. eprint: 0804.2182. URL: https://arxiv.org/abs/0804.2182.

[64] Raphael Bousso et al. "Entropy on a null surface for interacting quantum field theories and the Bousso bound". In: *Phys. Rev. D* 91 (2015), p. 084030. eprint: 1406.4545. URL: https://arxiv.org/abs/1406.4545.

[65] Thomas Hartman, Sandipan Kundu, and Amirhossein Tajdini. "Averaged Null Energy Condition from Causality". In: (2016). arXiv: 1610.05308 [hep-th].

[66] Aron C. Wall. "A Lower Bound on the Energy Density in Classical and Quantum Field Theories". In: *Phys. Rev. Lett.* 118 (2017), p. 151601. eprint: 1701.03196. URL: https://arxiv.org/abs/1701.03196.

[67] Christian Ecker et al. "Saturation of the Quantum Null Energy Condition in Far-From-Equilibrium Systems". In: *Phys. Rev. D* 97.12 (2018), p. 126016. DOI: 10.1103/PhysRevD.97.126016. arXiv: 1710.09837 [hep-th].

[68] Horacio Casini, Eduardo Teste, and Gonzalo Torroba. "Modular Hamiltonians on the null plane and the Markov property of the vacuum state". In: (Mar. 2017). eprint: 1703.10656. URL: https://arxiv.org/abs/1703.10656.

[69] Chris Akers et al. "The Quantum Null Energy Condition, Entanglement Wedge Nesting, and Quantum Focusing". In: (June 2017). eprint: 1706.04183. URL: https://arxiv.org/abs/1706.04183.

[70] Ted Jacobson. "Thermodynamics of Spacetime: The Einstein Equation of State". In: *Phys.Rev.Lett.* 75 (1995), pp. 1260–1263. eprint: gr-qc/9504004. URL: https://arxiv.org/abs/gr-qc/9504004.

[71] Ling-Yan Hung, Robert C. Myers, and Michael Smolkin. "Some Calculable Contributions to Holographic Entanglement Entropy". In: *JHEP* 08 (2011), p. 039. DOI: 10.1007/JHEP08(2011)039. arXiv: 1105.6055 [hep-th].

[72] Sebastian de Haro, Sergey N. Solodukhin, and Kostas Skenderis. "Holographic reconstruction of space-time and renormalization in the AdS / CFT correspondence". In: *Commun. Math. Phys.* 217 (2001), pp. 595–622. DOI: 10.1007/s002200100381. arXiv: hep-th/0002230 [hep-th].

[73] Donald Marolf and Aron C. Wall. "State-Dependent Divergences in the Entanglement Entropy". In: *JHEP* 10 (2016), p. 109. eprint: 1607.01246. URL: https://arxiv.org/abs/1607.01246.

[74] Netta Engelhardt and Aron C. Wall. "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime". In: (Aug. 2014). eprint: 1408.3203. URL: https://arxiv.org/abs/1408.3203.

[75] Igor R. Klebanov and Edward Witten. "AdS / CFT correspondence and symmetry breaking". In: *Nucl. Phys.* B556 (1999), pp. 89–114. DOI: 10.1016/S0550-3213(99)00387-9. arXiv: hep-th/9905104 [hep-th].

[76] Vijay Balasubramanian and Per Kraus. "A Stress tensor for Anti-de Sitter gravity". In: *Commun. Math. Phys.* 208 (1999), pp. 413–428. DOI: 10.1007/s002200050764. arXiv: hep-th/9902121.

[77] Masahiro Nozaki et al. "Dynamics of Entanglement Entropy from Einstein Equation". In: (Apr. 2013). eprint: 1304.7100. URL: https://arxiv.org/abs/1304.7100.

[78] Thomas Faulkner, Robert G. Leigh, and Onkar Parrikar. "Shape Dependence of Entanglement Entropy in Conformal Field Theories". In: *JHEP* 1604 (2016), p. 088. eprint: 1511.05179. URL: https://arxiv.org/abs/1511.05179.

[79] Jyotirmoy Bhattacharya et al. "Entanglement Density and Gravitational Thermodynamics". In: *Phys. Rev.* D91.10 (2015), p. 106009. DOI: 10.1103/PhysRevD.91.106009. arXiv: 1412.5472 [hep-th].

[80] Mark Van Raamsdonk. "Building up spacetime with quantum entanglement". In: *Gen.Rel.Grav.42:2323-2329,2010; Int.J.Mod.Phys.D* 19 (2010), pp. 2429–2435. eprint: 1005.3035. URL: https://arxiv.org/abs/1005.3035.

[81] Juan Maldacena and Leonard Susskind. "Cool horizons for entangled black holes". In: (June 2013). eprint: 1306.0533. URL: https://arxiv.org/abs/1306.0533.

[82] Ted Jacobson. "Entanglement Equilibrium and the Einstein Equation". In: *Phys. Rev. Lett.* 116 (2016), p. 201101. eprint: 1505.04753. URL: https://arxiv.org/abs/1505.04753.

[83] Nima Lashkari, Michael B. McDermott, and Mark Van Raamsdonk. "Gravitational Dynamics from Entanglement 'Thermodynamics'". In: *JHEP* 04 (2014), p. 195. DOI: 10.1007/JHEP04(2014)195. arXiv: 1308.3716 [hep-th].

[84] Thomas Faulkner et al. "Gravitation from Entanglement in Holographic CFTs". In: (Dec. 2013). eprint: 1312.7856. URL: https://arxiv.org/abs/1312.7856.

[85] Brian Swingle and Mark Van Raamsdonk. "Universality of Gravity from Entanglement". In: (2014). arXiv: 1405.2933 [hep-th].

[86] Thomas Faulkner et al. "Nonlinear Gravity from Entanglement in Conformal Field Theories". In: *JHEP* 08 (2017), p. 057. DOI: 10.1007/JHEP08(2017)057. arXiv: 1705.03026 [hep-th].

[87] Stefan Leichenauer. "The Quantum Focusing Conjecture Has Not Been Violated". In: (May 2017). eprint: 1705.05469. URL: https://arxiv.org/abs/1705.05469.

[88] Chris Akers et al. "Quantum null energy condition, entanglement wedge nesting, and quantum focusing". In: *Phys. Rev. D* 101.2 (2020), p. 025011. DOI: 10.1103/PhysRevD.101.025011. arXiv: 1706.04183 [hep-th].

[89] Srivatsan Balakrishnan et al. In: *To Appear* (2018).

[90] Andrea Allais and Márk Mezei. "Some results on the shape dependence of entanglement and Rényi entropies". In: *Phys. Rev.* D91.4 (2015), p. 046002. DOI: 10.1103/PhysRevD.91.046002. arXiv: 1407.7249 [hep-th].

[91] Márk Mezei. "Entanglement entropy across a deformed sphere". In: *Phys. Rev. D* 91 (2015), p. 045038.

[92] Stefan Leichenauer, Adam Levine, and Arvin Shahbazi-Moghaddam. "Energy is Entanglement". In: (2018). arXiv: 1802.02584 [hep-th].

[93] Christoph Holzhey, Finn Larsen, and Frank Wilczek. "Geometric and renormalized entropy in conformal field theory". In: *Nucl. Phys.* B424 (1994), pp. 443–467. DOI: 10.1016/0550-3213(94)90402-2. arXiv: hep-th/9403108 [hep-th].

[94] Pasquale Calabrese and John L. Cardy. "Entanglement entropy and quantum field theory". In: *J. Stat. Mech.* 0406 (2004), P06002. DOI: 10.1088/1742-5468/2004/06/P06002. arXiv: hep-th/0405152 [hep-th].

[95] Nima Lashkari. "Modular Hamiltonian for Excited States in Conformal Field Theory". In: *Phys. Rev. Lett.* 117.4 (2016), p. 041601. DOI: 10.1103/PhysRevLett.117.041601. arXiv: 1508.03506 [hep-th].

[96] Gábor Sárosi and Tomonori Ugajin. "Relative entropy of excited states in two dimensional conformal field theories". In: *JHEP* 07 (2016), p. 114. DOI: 10.1007/JHEP07(2016)114. arXiv: 1603.03057 [hep-th].

[97] Paola Ruggiero and Pasquale Calabrese. "Relative Entanglement Entropies in 1+1-dimensional conformal field theories". In: *JHEP* 02 (2017), p. 039. DOI: 10.1007/JHEP02(2017)039. arXiv: 1612.00659 [hep-th].

[98] Lorenzo Bianchi et al. "Renyi entropy and conformal defects". In: *Journal of High Energy Physics* 2016.7 (2016). ISSN: 1029-8479. DOI: 10.1007/jhep07(2016)076. URL: http://dx.doi.org/10.1007/JHEP07(2016)076.

[99] Marco Billo et al. "Defects in conformal field theory". In: *Journal of High Energy Physics* 2016.4 (2016), pp. 1–56. ISSN: 1029-8479. DOI: 10.1007/jhep04(2016)091. URL: http://dx.doi.org/10.1007/JHEP04(2016)091.

[100] Ferdinando Gliozzi et al. "Boundary and Interface CFTs from the Conformal Bootstrap". In: *JHEP* 05 (2015), p. 036. DOI: 10.1007/JHEP05(2015)036. arXiv: 1502.07217 [hep-th].

[101] Davide Gaiotto, Dalimil Mazac, and Miguel F. Paulos. "Bootstrapping the 3d Ising twist defect". In: *JHEP* 03 (2014), p. 100. DOI: 10.1007/JHEP03(2014)100. arXiv: 1310.5078 [hep-th].

[102] M. Billó et al. "Line defects in the 3d Ising model". In: *JHEP* 07 (2013), p. 055. DOI: 10.1007/JHEP07(2013)055. arXiv: 1304.4110 [hep-th].

[103] Diego M Hofman and Juan Maldacena. "Conformal collider physics: energy and charge correlations". In: *Journal of High Energy Physics* 2008.05 (2008), pp. 012–012. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/05/012. URL: http://dx.doi.org/10.1088/1126-6708/2008/05/012.

[104] Matthew Headrick. "Entanglement Renyi entropies in holographic theories". In: *Phys. Rev.* D82 (2010), p. 126010. DOI: 10.1103/PhysRevD.82.126010. arXiv: 1006.0047 [hep-th].

[105] Pasquale Calabrese, John Cardy, and Erik Tonni. "Entanglement entropy of two disjoint intervals in conformal field theory II". In: *J. Stat. Mech.* 1101 (2011), P01021. DOI: 10.1088/1742-5468/2011/01/P01021. arXiv: 1011.5482 [hep-th].

[106] Cesar Agón and Thomas Faulkner. "Quantum Corrections to Holographic Mutual Information". In: *JHEP* 08 (2016), p. 118. DOI: 10.1007/JHEP08(2016)118. arXiv: 1511.07462 [hep-th].

[107] Raphael Bousso et al. "Entropy on a null surface for interacting quantum field theories and the Bousso bound". In: *Phys. Rev.* D91.8 (2015), p. 084030. DOI: 10.1103/PhysRevD.91.084030. arXiv: 1406.4545 [hep-th].

[108] Madalena Lemos et al. "Universality at large transverse spin in defect CFT". In: *JHEP* 09 (2018), p. 091. DOI: 10.1007/JHEP09(2018)091. arXiv: 1712.08185 [hep-th].

[109] Miguel S. Costa, Tobias Hansen, and João Penedones. "Bounds for OPE coefficients on the Regge trajectory". In: *Journal of High Energy Physics* 2017.10 (2017). ISSN: 1029-8479. DOI: 10.1007/jhep10(2017)197. URL: http://dx.doi.org/10.1007/JHEP10(2017)197.

[110] Petr Kravchuk and David Simmons-Duffin. "Light-ray operators in conformal field theory". In: *JHEP* 11 (2018), p. 102. DOI: 10.1007/JHEP11(2018)102. arXiv: 1805.00098 [hep-th].

[111] Murat Kologlu et al. "The light-ray OPE and conformal colliders". In: (2019). arXiv: 1905.01311 [hep-th].

[112] Jason Koeller et al. "Local Modular Hamiltonians from the Quantum Null Energy Condition". In: (Feb. 2017). eprint: 1702.00412. URL: https://arxiv.org/abs/1702.00412.

[113] Srivatsan Balakrishnan, Souvik Dutta, and Thomas Faulkner. "Gravitational dual of the Renyi twist displacement operator". In: *Phys. Rev.* D96.4 (2017), p. 046019. DOI: 10.1103/PhysRevD.96.046019. arXiv: 1607.06155 [hep-th].

[114] Thomas Faulkner et al. "Nonlinear gravity from entanglement in conformal field theories". In: *Journal of High Energy Physics* 2017.8 (2017). ISSN: 1029-8479. DOI: 10.1007/jhep08(2017)057. URL: http://dx.doi.org/10.1007/JHEP08(2017)057.

[115] Thomas Faulkner. "Bulk Emergence and the RG Flow of Entanglement Entropy". In: (Dec. 2014). eprint: 1412.5648. URL: https://arxiv.org/abs/1412.5648.

[116] Gábor Sárosi and Tomonori Ugajin. "Modular Hamiltonians of excited states, OPE blocks and emergent bulk fields". In: *JHEP* 01 (2018), p. 012. DOI: 10.1007/JHEP01(2018)012. arXiv: 1705.01486 [hep-th].

[117] Nima Lashkari, Hong Liu, and Srivatsan Rajagopal. "Perturbation Theory for the Logarithm of a Positive Operator". In: (2018). arXiv: 1811.05619 [hep-th].

[118] Netta Engelhardt and Aron C. Wall. "Decoding the Apparent Horizon: Coarse-Grained Holographic Entropy". In: *Phys. Rev. Lett.* 121.21 (2018), p. 211301. DOI: 10.1103/PhysRevLett.121.211301. arXiv: 1706.02038 [hep-th].

[119] Jacob D. Bekenstein. "Black Holes and Entropy". In: *Phys. Rev. D* 7 (8 1973), pp. 2333–2346. DOI: 10.1103/PhysRevD.7.2333. URL: https://link.aps.org/doi/10.1103/PhysRevD.7.2333.

[120] Netta Engelhardt and Aron C. Wall. "Coarse Graining Holographic Black Holes". In: (2018). arXiv: 1806.01281 [hep-th].

[121] Raphael Bousso and Netta Engelhardt. "New Area Law in General Relativity". In: *Phys. Rev. Lett.* 115.8 (2015), p. 081301. DOI: 10.1103/PhysRevLett.115.081301. arXiv: 1504.07627 [hep-th].

[122] Chris Akers et al. "Boundary of the future of a surface". In: *Phys. Rev.* D97.2 (2018), p. 024018. DOI: 10.1103/PhysRevD.97.024018. arXiv: 1711.06689 [hep-th].

[123] Netta Engelhardt and Sebastian Fischetti. "Surface Theory: the Classical, the Quantum, and the Holographic". In: (2019). arXiv: 1904.08423 [hep-th].

[124] J. J Bisognano and E. H. Wichmann. "On the Duality Condition for a Hermitian Scalar Field". In: *J. Math. Phys.* 16 (1975), pp. 985–1007. DOI: 10.1063/1.522605.

[125] Horacio Casini, Eduardo Teste, and Gonzalo Torroba. "Modular Hamiltonians on the null plane and the Markov property of the vacuum state". In: *J. Phys.* A50.36 (2017), p. 364001. DOI: 10.1088/1751-8121/aa7eaa. arXiv: 1703.10656 [hep-th].

[126] Aron C. Wall. "Lower Bound on the Energy Density in Classical and Quantum Field Theories". In: *Phys. Rev. Lett.* 118.15 (2017), p. 151601. DOI: 10.1103/PhysRevLett.118.151601. arXiv: 1701.03196 [hep-th].

[127] Edward Witten. "APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory". In: *Rev. Mod. Phys.* 90.4 (2018), p. 045003. DOI: 10.1103/RevModPhys.90.045003. arXiv: 1803.04993 [hep-th].

[128] H. Araki. "Relative Entropy of States of Von Neumann Algebras". In: *Publ. Res. Inst. Math. Sci. Kyoto* 1976 (1976), pp. 809–833.

[129] Alex Hamilton et al. "Holographic representation of local bulk operators". In: *Phys. Rev. D* 74 (2006), p. 066009. DOI: 10.1103/PhysRevD.74.066009. arXiv: hep-th/0606141.

[130] S.S. Gubser, Igor R. Klebanov, and Alexander M. Polyakov. "Gauge theory correlators from noncritical string theory". In: *Phys. Lett. B* 428 (1998), pp. 105–114. DOI: 10.1016/S0370-2693(98)00377-3. arXiv: hep-th/9802109.

[131] Patrick Hayden and Geoffrey Penington. "Learning the Alpha-bits of Black Holes". In: *JHEP* 12 (2019), p. 007. DOI: 10.1007/JHEP12(2019)007. arXiv: 1807.06041 [hep-th].

[132] Thomas Faulkner and Aitor Lewkowycz. "Bulk locality from modular flow". In: *JHEP* 07 (2017), p. 151. DOI: 10.1007/JHEP07(2017)151. arXiv: 1704.05464 [hep-th].

[133] Yiming Chen. "Pulling Out the Island with Modular Flow". In: *JHEP* 03 (2020), p. 033. DOI: 10.1007/JHEP03(2020)033. arXiv: 1912.02210 [hep-th].

[134] Jordan Cotler et al. "Entanglement Wedge Reconstruction via Universal Recovery Channels". In: *Phys. Rev. X* 9.3 (2019), p. 031011. DOI: 10.1103/PhysRevX.9.031011. arXiv: 1704.05839 [hep-th].

[135] Chi-Fang Chen, Geoffrey Penington, and Grant Salton. "Entanglement Wedge Reconstruction using the Petz Map". In: *JHEP* 01 (2020), p. 168. DOI: 10.1007/JHEP01(2020)168. arXiv: 1902.02844 [hep-th].

[136] Geoff Penington et al. "Replica wormholes and the black hole interior". In: (Nov. 2019). arXiv: 1911.11977 [hep-th].

[137] Srivatsan Balakrishnan et al. "A General Proof of the Quantum Null Energy Condition". In: *JHEP* 09 (2019), p. 020. DOI: 10.1007/JHEP09(2019)020. arXiv: 1706.09432 [hep-th].

[138] Nima Lashkari. "Constraining Quantum Fields using Modular Theory". In: *JHEP* 01 (2019), p. 059. DOI: 10.1007/JHEP01(2019)059. arXiv: 1810.09306 [hep-th].

[139] Raphael Bousso, Venkatesa Chandrasekaran, and Arvin Shahbazi-Moghaddam. "From black hole entropy to energy-minimizing states in QFT". In: *Phys. Rev.* D101.4 (2020), p. 046001. DOI: 10.1103/PhysRevD.101.046001. arXiv: 1906.05299 [hep-th].

[140] Jan De Boer and Lampros Lamprou. "Holographic Order from Modular Chaos". In: *JHEP* 06 (2020), p. 024. DOI: 10.1007/JHEP06(2020)024. arXiv: 1912.02810 [hep-th].

[141] Daniel L. Jafferis and S. Josephine Suh. "The Gravity Duals of Modular Hamiltonians". In: (2014). arXiv: 1412.8465 [hep-th].

[142] Thomas Faulkner, Min Li, and Huajia Wang. "A modular toolkit for bulk reconstruction". In: *JHEP* 04 (2019), p. 119. DOI: 10.1007/JHEP04(2019)119. arXiv: 1806.10560 [hep-th].

[143] Chris Akers and Pratik Rath. "Holographic Renyi Entropy from Quantum Error Correction". In: *JHEP* 05 (2019), p. 052. DOI: 10.1007/JHEP05(2019)052. arXiv: 1811.05171 [hep-th].

[144] Xi Dong, Daniel Harlow, and Donald Marolf. "Flat entanglement spectra in fixed-area states of quantum gravity". In: *JHEP* 10 (2019), p. 240. DOI: 10.1007/JHEP10(2019)240. arXiv: 1811.05382 [hep-th].

[145] Xi Dong and Donald Marolf. "One-loop universality of holographic codes". In: (2019). arXiv: 1910.06329 [hep-th].

[146] Edward Witten. "APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory". In: *Rev. Mod. Phys.* 90.4 (2018), p. 045003. DOI: 10.1103/RevModPhys.90.045003. arXiv: 1803.04993 [hep-th].

[147] Srivatsan Balakrishnan and Onkar Parrikar. "Modular Hamiltonians for Euclidean Path Integral States". In: (Jan. 2020). arXiv: 2002.00018 [hep-th].

[148] Jun-ichirou Koga. "Asymptotic symmetries on Killing horizons". In: *Phys. Rev. D* 64 (2001), p. 124012. DOI: 10.1103/PhysRevD.64.124012. arXiv: gr-qc/0107096.

[149] Abhay Ashtekar, Christopher Beetle, and Jerzy Lewandowski. "Geometry of generic isolated horizons". In: *Class. Quant. Grav.* 19 (2002), pp. 1195–1225. DOI: 10.1088/0264-9381/19/6/311. arXiv: gr-qc/0111067 [gr-qc].

[150] Robert M Wald. *General relativity*. University of Chicago press, 2010.

[151] Tadashi Takayanagi. "Holographic Dual of BCFT". In: *Phys. Rev. Lett.* 107 (2011), p. 101602. DOI: 10.1103/PhysRevLett.107.101602. arXiv: 1105.5165 [hep-th].

[152] Ioanna Kourkoulou and Juan Maldacena. "Pure states in the SYK model and nearly-AdS_2 gravity". In: (July 2017). arXiv: 1707.02325 [hep-th].

[153] Raphael Bousso, Stefan Leichenauer, and Vladimir Rosenhaus. "Light-sheets and AdS/CFT". In: *Phys. Rev.* D86 (2012), p. 046009. DOI: 10.1103/PhysRevD.86.046009. arXiv: 1203.6619 [hep-th].

[154] Raphael Bousso et al. "Null Geodesics, Local CFT Operators and AdS/CFT for Subregions". In: *Phys. Rev. D* 88 (2013), p. 064057. DOI: 10.1103/PhysRevD.88.064057. arXiv: 1209.4641 [hep-th].

[155] Charles Fefferman and C. Robin Graham. "The ambient metric". In: *Ann. Math. Stud.* 178 (2011), pp. 1–128. arXiv: 0710.0919 [math.DG].

[156] Eric Gourgoulhon. "3+1 formalism and bases of numerical relativity". In: (Mar. 2007). arXiv: gr-qc/0703035.

[157] Nima Lashkari et al. "Gravitational positive energy theorems from information inequalities". In: *Progress of Theoretical and Experimental Physics* 2016.12 (2016), p. 12C109. ISSN: 2050-3911. DOI: 10.1093/ptep/ptw139. URL: http://dx.doi.org/10.1093/ptep/ptw139.

[158] William Donnelly and Laurent Freidel. "Local subsystems in gauge theory and gravity". In: *Journal of High Energy Physics* 2016.9 (2016). ISSN: 1029-8479. DOI: 10.1007/jhep09(2016)102. URL: http://dx.doi.org/10.1007/JHEP09(2016)102.

[159] Steven Carlip and Claudio Teitelboim. "The off-shell black hole". In: *Classical and Quantum Gravity* 12.7 (1995), 1699–1704. ISSN: 1361-6382. DOI: 10.1088/0264-9381/12/7/011. URL: http://dx.doi.org/10.1088/0264-9381/12/7/011.

[160] Souvik Dutta and Thomas Faulkner. "A canonical purification for the entanglement wedge cross-section". In: (May 2019). arXiv: 1905.00577 [hep-th].

[161] Sergio Doplicher and Roberto Longo. "Standard and split inclusions of von Neumann algebras". In: *Invent. Math.* 75 (1984), pp. 493–536. DOI: 10.1007/BF01388641.

[162] Tadashi Takayanagi and Koji Umemoto. "Entanglement of purification through holographic duality". In: *Nature Phys.* 14.6 (2018), pp. 573–577. DOI: 10.1038/s41567-018-0075-2. arXiv: 1708.09393 [hep-th].

[163] Donald Marolf. "Microcanonical Path Integrals and the Holography of small Black Hole Interiors". In: *JHEP* 09 (2018), p. 114. DOI: 10.1007/JHEP09(2018)114. arXiv: 1808.00394 [hep-th].

[164] W. Israel. "Thermo field dynamics of black holes". In: *Phys. Lett. A* 57 (1976), pp. 107–110. DOI: 10.1016/0375-9601(76)90178-X.

[165] Juan Martin Maldacena. "Eternal black holes in anti-de Sitter". In: *JHEP* 04 (2003), p. 021. DOI: 10.1088/1126-6708/2003/04/021. arXiv: hep-th/0106112.

[166] Gary T. Horowitz and Diandian Wang. "Gravitational Corner Conditions in Holography". In: *JHEP* 01 (2020), p. 155. DOI: 10.1007/JHEP01(2020)155. arXiv: 1909.11703 [hep-th].

[167] Richard P. Woodard. "Ostrogradsky's theorem on Hamiltonian instability". In: *Scholarpedia* 10.8 (2015), p. 32243. DOI: 10.4249/scholarpedia.32243. arXiv: 1506.02210 [hep-th].

[168] William R. Kelly and Aron C. Wall. "Holographic Proof of the Averaged Null Energy Condition". In: *Phys. Rev.* D90.10 (2014). [Erratum: Phys. Rev.D91,no.6,069902(2015)], p. 106003. DOI: 10.1103/PhysRevD.90.106003,10.1103/PhysRevD.91.069902. arXiv: 1408.3566 [gr-qc].

[169] S. W. Hawking. "Gravitational radiation from colliding black holes". In: *Phys. Rev. Lett.* 26 (1971), pp. 1344–1346.

[170] Raphael Bousso and Netta Engelhardt. "New Area Law in General Relativity". In: *Phys. Rev. Lett.* 115.8 (2015), p. 081301. DOI: 10.1103/PhysRevLett.115.081301. arXiv: 1504.07627 [hep-th].

[171] Raphael Bousso and Netta Engelhardt. "Proof of a New Area Law in General Relativity". In: *Phys. Rev.* D92.4 (2015), p. 044031. DOI: 10.1103/PhysRevD.92.044031. arXiv: 1504.07660 [gr-qc].

[172] Robert M. Wald. *General Relativity*. Chicago: The University of Chicago Press, 1984.

[173] Roger Penrose. "Gravitational collapse and space-time singularities". In: *Phys. Rev. Lett.* 14 (1965), pp. 57–59. DOI: 10.1103/PhysRevLett.14.57.

[174] J. D. Bekenstein. "Black holes and the second law". In: *Nuovo Cim. Lett.* 4 (1972), pp. 737–740.

[175] Jacob D. Bekenstein. "Black Holes and Entropy". In: *Phys. Rev. D* 7 (1973), p. 2333.

[176] Jacob D. Bekenstein. "Generalized second law of thermodynamics in black hole physics". In: *Phys. Rev. D* 9 (1974), p. 3292.

[177] Raphael Bousso and Netta Engelhardt. "Generalized Second Law for Cosmology". In: *Phys. Rev.* D93.2 (2016), p. 024025. DOI: 10.1103/PhysRevD.93.024025. arXiv: 1510.02099 [hep-th].

[178] Raphael Bousso et al. "Quantum Focusing Conjecture". In: *Phys. Rev.* D93.6 (2016), p. 064044. DOI: 10.1103/PhysRevD.93.064044. arXiv: 1506.02669 [hep-th].

[179] Aron C. Wall. "The Generalized Second Law implies a Quantum Singularity Theorem". In: *Class.Quant.Grav.* 30 (2013), p. 165003. DOI: 10.1088/0264-9381/30/19/199501,10.1088/0264-9381/30/16/165003. arXiv: 1010.5513 [gr-qc].

[180] Raphael Bousso et al. "Proof of the Quantum Null Energy Condition". In: *Phys. Rev.* D93.2 (2016), p. 024017. DOI: 10.1103/PhysRevD.93.024017. arXiv: 1509.02542 [hep-th].

[181] Jason Koeller and Stefan Leichenauer. "Holographic Proof of the Quantum Null Energy Condition". In: (2015). arXiv: 1512.06109 [hep-th].

[182] Srivatsan Balakrishnan et al. "A General Proof of the Quantum Null Energy Condition". In: (2017). arXiv: 1706.09432 [hep-th].

[183] R. Arnowitt, S. Deser, and C. W. Misner. "The Dynamics of General Relativity". In: *Gravitation: an Introduction to Current Research*. Ed. by L. Witten. New York: Wiley, 1962, pp. 227–265.

[184] Richard Schoen and Shing Tung Yau. "Proof of the positive mass theorem. II". In: *Comm. Math. Phys.* 79.2 (1981), pp. 231–260. URL: https://projecteuclid.org:443/euclid.cmp/1103908964.

[185] Raphael Bousso, Arvin Shahbazi-Moghaddam, and Marija Tomasevic. "Quantum Penrose Inequality". In: (2019). arXiv: 1908.02755 [hep-th].

[186] Marc Mars. "Present status of the Penrose inequality". In: *Class. Quant. Grav.* 26 (2009), p. 193001. DOI: 10.1088/0264-9381/26/19/193001. arXiv: 0906.5566 [gr-qc].

[187] Netta Engelhardt and Gary T. Horowitz. "A Holographic Argument for the Penrose Inequality in AdS". In: *arXiv e-prints*, arXiv:1903.00555 (2019), arXiv:1903.00555. arXiv: 1903.00555 [hep-th].

[188] S. W. Hawking and G. F. R. Ellis. *The large scale stucture of space-time*. Cambridge, England: Cambridge University Press, 1973.

[189] Raphael Bousso. "A covariant entropy conjecture". In: *JHEP* 07 (1999), p. 004. eprint: hep-th/9905177.

[190] Raphael Bousso. "Holography in general space-times". In: *JHEP* 06 (1999), p. 028. eprint: hep-th/9906022.

[191] David G. Boulware. "Quantum field theory in Schwarzschild and Rindler spaces". In: *Phys. Rev. D* 11 (6 1975), pp. 1404–1423. DOI: 10.1103/PhysRevD.11.1404. URL: https://link.aps.org/doi/10.1103/PhysRevD.11.1404.

[192] P. Candelas. "Vacuum polarization in Schwarzschild spacetime". In: *Phys. Rev. D* 21 (8 1980), pp. 2185–2202. DOI: 10.1103/PhysRevD.21.2185. URL: https://link.aps.org/doi/10.1103/PhysRevD.21.2185.

[193] Chris Akers et al. "Boundary of the future of a surface". In: *Phys. Rev.* D97.2 (2018), p. 024018. DOI: 10.1103/PhysRevD.97.024018. arXiv: 1711.06689 [hep-th].

[194] Raphael Bousso et al. "Entropy on a null surface for interacting quantum field theories and the Bousso bound". In: *Phys.Rev.* D91.8 (2015), p. 084030. DOI: 10.1103/PhysRevD.91.084030. arXiv: 1406.4545 [hep-th].

[195] Raphael Bousso, Ben Freivogel, and Stefan Leichenauer. "Saturating the holographic entropy bound". In: *Phys. Rev.* D82 (2010), p. 084024. DOI: 10.1103/PhysRevD.82.084024. arXiv: 1003.3012 [hep-th].

[196] Alan R. Steif. "The Quantum stress tensor in the three-dimensional black hole". In: *Phys. Rev.* D49 (1994), pp. 585–589. DOI: 10.1103/PhysRevD.49.R585. arXiv: gr-qc/9308032 [gr-qc].

[197] Netta Engelhardt and Aron C. Wall. "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime". In: *JHEP* 1501 (2015), p. 073. DOI: 10.1007/JHEP01(2015)073. arXiv: 1408.3203 [hep-th].

[198] Ruth Gregory and Raymond Laflamme. "Black strings and p-branes are unstable". In: *PRL* 70.19 (1993), pp. 2837–2840. DOI: 10.1103/PhysRevLett.70.2837. arXiv: hep-th/9301052 [hep-th].

[199] Ruth Gregory. "The Gregory-Laflamme instability". In: *arXiv e-prints*, arXiv:1107.5821 (2011), arXiv:1107.5821. arXiv: 1107.5821 [gr-qc].

[200] Luis Lehner and Frans Pretorius. "Final State of Gregory-Laflamme Instability". In: *arXiv e-prints*, arXiv:1106.5184 (2011), arXiv:1106.5184. arXiv: 1106.5184 [gr-qc].

[201] Robert M. Wald. "Gravitational Collapse and Cosmic Censorship". In: gr-qc/9710068 (1997). arXiv: gr-qc/9710068 [gr-qc].

[202] Carsten Gundlach and José M. Martín-García. "Critical Phenomena in Gravitational Collapse". In: *Living Reviews in Relativity* 10.1, 5 (2007), p. 5. DOI: 10.12942/lrr-2007-5. arXiv: 0711.4620 [gr-qc].

[203] Matthew W. Choptuik. "Universality and scaling in gravitational collapse of a massless scalar field". In: *Phys. Rev. Lett.* 70 (1 1993), pp. 9–12. DOI: 10.1103/PhysRevLett.70.9. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.9.

[204] Demetrios Christodoulou. "Violation of cosmic censorship in the gravitational collapse of a dust cloud". In: *Commun. Math. Phys.* 93 (1984), pp. 171–195. DOI: 10.1007/BF01223743.

[205] Hideo Kodama. "INEVITABILITY OF A NAKED SINGULARITY ASSOCIATED WITH THE BLACK HOLE EVAPORATION". In: *Prog. Theor. Phys.* 62 (1979), p. 1434. DOI: 10.1143/PTP.62.1434.

[206] Robert M. Wald. "BLACK HOLES, SINGULARITIES AND PREDICTABILITY". In: (Jan. 1984).

[207] A. Schwimmer and S. Theisen. "Entanglement Entropy, Trace Anomalies and Holography". In: *Nucl. Phys.* B801 (2008), pp. 1–24. DOI: 10.1016/j.nuclphysb.2008.04.015. arXiv: 0802.1017 [hep-th].

[208] Aron C. Wall. "A Proof of the Generalized Second Law for Rapidly-Evolving Rindler Horizons". In: *Phys. Rev.* D82 (2010), p. 124019. DOI: 10.1103/PhysRevD.82.124019. arXiv: 1007.1493 [gr-qc].

[209] Nima Afkhami-Jeddi et al. "Shockwaves from the Operator Product Expansion". In: (Sept. 2017). eprint: 1709.03597. URL: https://arxiv.org/abs/1709.03597.

[210] Aron C. Wall. "Testing the Generalized Second Law in 1+1 Dimensional Conformal Vacua: an Argument for the Causal Horizon". In: *Phys. Rev.* D85 (2012), p. 024015. DOI: 10.1103/PhysRevD.85.024015. arXiv: 1105.3520 [gr-qc].

[211] D. N. Page. "Particle Emission Rates From a Black Hole. II. Massless Particles From a Rotating Hole". In: *Phys. Rev. D* 14 (1976), p. 3260.

[212] Raphael Bousso and Mudassir Moosa. "Dynamics and Observer-Dependence of Holographic Screens". In: *Phys. Rev.* D95.4 (2017), p. 046005. DOI: 10.1103/PhysRevD.95.046005. arXiv: 1611.04607 [hep-th].