University of California
Santa Barbara

Quantum Gravity as a Holographic Theory: Lessons from the Gravitational Path Integral

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Sergio Hernández-Cuenca

Committee in charge:

Professor Gary Horowitz, Chair
Professor Xi Dong
Professor David Stuart

September 2022

The Dissertation of Sergio Hernández-Cuenca is approved.

Professor Xi Dong

Professor David Stuart

Professor Gary Horowitz, Committee Chair

August 2022

Quantum Gravity as a Holographic Theory:

Lessons from the Gravitational Path Integral

by

Sergio Hernández-Cuenca

To my family

Acknowledgements

Everyone warns you the journey through grad school is an arduous path. Only when you are already in it do you realize there is no path there to begin with. I am grateful to many for helping me find my way through this odyssey and accompanying me all along.

It has been a true privilege to have Gary Horowitz as my advisor, who has been a genuinely caring mentor committed to my success through grad school. I am indebted to Gary for his kind guidance into the bewildering world of academic research; for his steadfast support as I discovered my interests and encouragement to follow my curiosity wherever it took me; for giving me the confidence to think freely by listening to my ideas without prejudice; for always sharing his deep insight with pedagogy, never resorting to the misnomer "obviously"; for easing the burden of impostor syndrome with his humility, empowering me to evolve from a student into a researcher; ultimately, for his human value while embodying the researcher I aspire to become.

I owe much of my work throughout grad school to Veronika Hubeny, Mukund Rangamani, and Max Rota, who appreciated and welcomed the ideas of 2$^{\text{nd}}$-year me. In their unparalleled kindness I found a completely judgement-free environment where I have been able to benefit from countless discussions. Beside their fruitful collaboration, I have also been fortunate to count on Mukund and Veronika as mentors who have time and again provided me with precious advice (some even thought I studied at UC Davis!).

As I was halfway through grad school, there came the pandemic tsunami that hindered any form of spontaneous collaboration. Nonetheless, in spite of the difficulties of working virtually, Netta Engelhardt saw value in me as a collaborator and got me involved in the fascinating realm of wormholes. I am extremely grateful to Netta for inspiring me with her sharp ingenuity, which sets an example of impeccable balance between creativity and rigor; for instilling me with her highly contagious passion for research and, above all, for

putting her trust in my potential as a researcher.

My education as a grad student has benefited beyond words from interactions with many other researchers. Besides the above, I would like to thank each and every one of my other collaborators for their fundamental role in shaping me as a researcher: Chris Akers, David Avis, Ning Bao, Ven Chandrasekaran, Newton Cheng, Matteo Fadel, Sebastian Fischetti, Åsmund Folkestad, Temple He, Frederic Jia, Pratik Rath, Vincent Su, Gabriel Treviño, and Diandian Wang. Special thanks go to David, for guiding me into communicating physics to mathematicians; to Ning, for his exceptional outside-the-box thinking; and to Sebastian, for his inspiring dedication and insatiable work ethic. My incursion into research thus far has also benefited immensely from the wisdom I have been breathing at the 6^{th}-floor lounge in Broida, crucial to which have been all the discussions with professors David Berenstein, Xi Dong, James Hartle, Gary, Don Marolf, David Morrison, Mark Srednicki, and all the students and postdocs at UCSB and KITP. I would also like to thank David Stuart for his helpful guidance as a member of my committee; and Jennifer Farrar, for always being on top of every bureaucratic nuance.

I have been fortunate to share the grad school experience with many good friends. First and foremost, a warm thank you to Aleksei, Joe, Noah, Seamus, Will, Sarah, Josh, and Demi, for making Santa Barbara feel like home. A big thank you also to other physics student colleagues: Brianna, Diandian, Gabriel, Wayne, Zhencheng, Yaodong, Zhitao, Aaron, Mark, Brad, Remi, Seth, Chai, Brent, Xiaochuan, Tianji, Navin... for filling my time at UCSB with memorable experiences. And further thanks to an open-ended list of colleagues and friends who I have been bumping into in random countries and who have made attending conferences all the more productive and fun: Julio, Bogdan, Marija, Sebastian, Åsmund, Jon, Chris, Pratik, Ven, Vincent, Misha, Adam, Sean, Martin, Max, Masamichi, Dongsheng...

But none of this would have ever happened had the seeds of my academic journey not been planted. That a student from a foreign public school could be admitted to a world-leading UK university? That was unheard of to me... at least until my maths teacher back then, Alfonso Camaño Liceras, thought me deserving of dreaming big. Then came Sergio Montañez Naz, not the typical high-school teacher: a string theorist enthusiastically giving voluntary courses on quantum mechanics and relativity in his free time. I cannot thank Alfonso and Sergio enough for inspiring me to be where I am today.

I keep the fondest of memories of my years at Imperial, from where I proudly still wear my graduation ring. I am thankful to all the lecturers and instructors that formed me as a physics student, especially to my undergraduate advisor, James Roy Taylor, who accompanied and supported me through those memorable years. My experience as an undergrad was also enriched by those who facilitated my first incursions into research. It is a pleasure to thank Juan León, Antonello Scardicchio, and Toby Wiseman for generously dedicating their time to mentoring a then wannabe theorist with much to learn but little to give back. In particular to Toby, who I gratefully blame for getting me into the black hole information problem which eventually landed me at UCSB.

I cannot leave unmentioned my quintessential learning experience as a masters student at DAMTP in Cambridge. My PhD would not have been nearly as prolific had it not been for the brutally instructive experience that Part III was. I am indebted to all the instructors I had there for their superluminal lecturing skills (particularly Jorge Santos, who enthusiastically encouraged me to work with Gary), and to all my brilliant peer students who I learned so much from in conversations, if not by mere osmosis.

There happens to be a world outside the academic bubble though. My most heartfelt words of gratitude belong to my family. To my parents, Mari Luz Cuenca Alcalá and Carlos Hernández Barco, my staunchest advocates, for their unconditional love, support, and

faith in me; for granting me the opportunity to aspire to anything and everything; and for their readiness to move heaven and earth for my well-being. To Francisco Hernández Cuenca, my older brother with the younger spirit, for inadvertently rejuvenating me in every visit home; for always having my back and being my sporadically much-needed confidant; and for always celebrating my accomplishments as his own. And to my girlfriend and best friend, Laura Vidal Guerrero, for her patience, understanding, and unique mastery in lifting me up whenever I am down (a tough, full-time, multi-time-zone job); for always being by my side across parallels and meridians; and for her daily dose of heart-warming smiles.

And since family is more than blood, a shout-out to every one of my Vragabundos: Adri, Diego, Flow, Pedro, Helen, Luisal, María, and Rodri. What a rare group of high-school friends where physicists and mathematicians constitute a majority (blame the two teachers above). Your unbreakable and timeless friendship is an irreplaceable treasure.

Last but not least, thanks to Dr. Kayla Rosen, for keeping me afloat whenever the sea got rough; and to Prozac, for being so vitally supportive over the years. Thanks also to Mark Sherwin for his very kind support at a time I needed to reach out.

This dissertation has been supported by "la Caixa" Foundation (ID 100010434) under fellowship LCF/BQ/AA17/11610002, the National Science Foundation under grants PHY-1801805 and PHY-2107939, and by the University of California, Santa Barbara, through a Regents in Physics Fellowship, a Joseph and Dorothy Polchinski Graduate Student Fellowship, and a Len DeBenedictis Graduate Fellowship.

Curriculum Vitæ

Sergio Hernández-Cuenca

Education

2022	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2020	M.A. in Physics, University of California, Santa Barbara.
2017	M.A.St. in Applied Mathematics, University of Cambridge.
2016	B.Sc. in Physics with Theoretical Physics, Imperial College London.

Publications

1. *"A Tale of Two Saddles"*
   V. Chandrasekaran, N. Engelhardt, S. Fischetti, and S. Hernández-Cuenca
   arXiv:2207.09472 [hep-th] [1]
2. *"The holographic entropy cone from marginal independence"*
   S. Hernández-Cuenca, V. E. Hubeny, and M. Rota
   arXiv:2204.00075 [hep-th] [2]
3. *"The Symmetrized Holographic Entropy Cone"*
   M. Fadel and S. Hernández-Cuenca
   Phys. Rev. D **105** (2022) 086008 [arXiv:2112.03862 [quant-ph]] [3]
4. *"Topological Link Models of Multipartite Entanglement"*
   N. Bao, N. Cheng, S. Hernández-Cuenca, and V. P. Su
   Quantum **6** (2022) 741 [arXiv:2109.01150 [quant-ph]] [4]
5. *"Quantum Extremal Surfaces and the Holographic Entropy Cone"*
   C. Akers, S. Hernández-Cuenca, and P. Rath
   JHEP **11** (2021) 177 [arXiv:2108.07280 [hep-th]] [5]
6. *"Conformal Rigidity from Focusing"*
   Å. Folkestad and S. Hernández-Cuenca
   Class. Quant. Grav. **38** (2021) 215005 [arXiv:2106.09037 [gr-qc]] [6]
7. *"Boundary Causality Violating Metrics in Holography"*
   S. Hernández-Cuenca, G. T. Horowitz, G. Treviño, and D. Wang
   Phys. Rev. Lett. **127** (2021) 081603 [arXiv:2103.05014 [hep-th]] [7]
8. *"On the foundations and extremal structure of the holographic entropy cone"*
   D. Avis and S. Hernández-Cuenca
   (to appear in Discrete Applied Mathematics) arXiv:2102.07535 [math.CO] [8]

9. *"A Gap Between the Hypergraph and Stabilizer Entropy Cones"*
N. Bao, N. Cheng, S. Hernández-Cuenca, and V. P. Su
arXiv:2006.16292 [quant-ph] [9]

10. *"Bulk reconstruction of metrics with a compact space asymptotically"*
S. Hernández-Cuenca and G. T. Horowitz
JHEP **08** (2020) 108 [arXiv:2003.08409 [hep-th]] [10]

11. *"The Quantum Entropy Cone of Hypergraphs"*
N. Bao, N. Cheng, S. Hernández-Cuenca, and V. P. Su
SciPost Phys. **9** (2020) 067 [arXiv:2002.05317 [quant-ph]] [11]

12. *"The quantum marginal independence problem"*
S. Hernández-Cuenca, V. E. Hubeny, M. Rangamani, and M. Rota
arXiv:1912.01041 [quant-ph] [12]

13. *"The Holographic Entropy Cone for Five Regions"*
S. Hernández-Cuenca
Phys. Rev. D **100** (2019) 026004 [arXiv:1903.09148 [hep-th]] [13]

Abstract

Quantum Gravity as a Holographic Theory:
Lessons from the Gravitational Path Integral

by

Sergio Hernández-Cuenca

Holographic dualities like AdS/CFT provide a non-perturbative definition of a bulk theory of quantum gravity in terms of a lower-dimensional boundary quantum field theory. Elucidating quantum gravity thus becomes the challenge of deciphering the dictionary between bulk and boundary physics, and using it to translate basic properties of the latter into lessons about the former. A crucial intermediary between these two realms is the gravitational path integral, which defines the bulk theory in terms of an integral over spacetimes subject to conditions from the boundary theory. Though shallowly understood and rather formal, this piece of technology has thus far been able to provide deep insights into quantum gravity. This thesis is organized in three parts, each focusing on a different basic property in quantum theory and its consequences for quantum gravity through the gravitational path integral: entanglement, causality, and factorization.

Part I addresses the emergence of spacetime from entanglement, with a focus on understanding constraints on the entanglement structure of quantum states to posses classical geometries as holographic bulk duals. These constraints can be expressed as linear inequalities and used to define the holographic entropy cone (HEC). A systematic study of the HEC is accomplished by reformulating the holographic computation of von Neumann entropies as a graph-theoretic one, thereby recasting a complicated problem in differential geometry as a purely combinatorial one. This allows to prove important prop-

erties of the HEC, devise proof methods and algorithms for constructing it, derive precise relations to other polytope structures, and ultimately pursue a top-down understanding of the HEC from the universal quantum inequality of subadditivity. This part concludes with an exploration of how the machinery involved in the study of the HEC may also be generalized to settings where the von Neumann entropy receives contributions from bulk quantum fields, a regime where the graph-theoretic apparatus has to be upgraded.

Part II presents an alternative perspective on spacetime emergence, both in classical and quantum regimes, based on causality. Starting at a classical level, we explain how the conformal bulk geometry can be reconstructed by encoding its causal structure in data accessible from the boundary field theory. Through the use of field theory correlators, we propose a method for obtaining the full-dimensional bulk geometry up to a conformal factor. This generalizes the approach to bulk metric reconstruction based on light-cone cuts to a prescription which allows for recovering even those dimensions which become compact asymptotically. Moving away from the classical limit, we then resolve a known puzzle that arises from a tension between the bulk and boundary causal structures when the bulk theory is understood as a genuine gravitational path integral over spacetimes.

Finally, part III delves into the consequences of the lack of factorization that occurs in holography when wormholes are included in the gravitational path integral. In particular, we study generating functionals in quantum gravity and propose a recipe for their computation which accounts for the contribution of such connected topologies. This allows to differentiate between quenched and annealed quantities in quantum gravity, a distinction which may be used as a consistency test-ground for foundational aspects of the gravitational path integral regarding summing over topologies.

Contents

Chapter 1

Introduction

Quantum gravity holds the promise of elucidating profound questions about the universe. What happens when gravity and quantum phenomena strongly interact? Although we have astonishingly good descriptions of each separately, combining these has proven to be a paramount challenge already at a conceptual level. A key milestone in our quest for a theory of quantum gravity has been the discovery and our progressively sharper understanding of the framework of holographic dualities.

1.1 Holography

From the microscopic degrees of freedom of gravity to the macroscopic emergence of spacetime itself, the holographic principle has yielded deep insight into many aspects of quantum gravity [14, 15]. Holographic dualities provide a definition of quantum gravity in terms of an ordinary, non-gravitating quantum theory which encodes it in fewer dimensions – a hologram. Crucially, holography allows one to look at identical physics through two different lenses, one oftentimes providing a clearer image. This thesis fiddles with

various knobs of this holographic binocular to shed some light on the nature of quantum gravity.

The first and most studied realization of the holographic principle was formulated by Juan Maldacena in 1997 as a conjectured Anti-de Sitter / Conformal Field Theory (AdS/CFT) correspondence [16]. In particular, he was able to discern a detailed duality between aspects of type IIB string theory on an $\mathrm{AdS}_5 \times \mathbb{S}^5$ background (the bulk theory), and Yang-Mills theory with $\mathcal{N} = 4$ supersymmetry on its conformal boundary (the boundary theory) [16, 17, 18, 19, 20, 21, 22].[1] One of the basic ingredients of the original holographic dictionary involved the identification between fundamental parameters of each theory. On the boundary side, the Yang-Mills theory involves a gauge coupling constant g_{YM} and a choice of rank N of the gauge group, taken to be $SU(N)$. An important parameter derived from these is the 't Hooft coupling, $\lambda = g_{\mathrm{YM}} N^2$, which is crucial for establishing a connection between the planar diagrams that dominate Yang-Mills at large N and the sum over topologies for the perturbative strings in the bulk theory [23]. Indeed, a relation between the two expansions is made manifest by matching the string coupling g_s to the gauge coupling through $4\pi g_s = g_{\mathrm{YM}}^2$. The 't Hooft parameter also defines a ratio between two important scales in the bulk: the AdS radius ℓ_{AdS} and the string scale ℓ_s are related by $\ell_{\mathrm{AdS}}/\ell_s = \lambda^{1/4}$.

These simple relations alone already grant some interesting insight into the inner workings of the AdS/CFT correspondence as a strong/weak duality. The best understood corner of parameter space corresponds to the simultaneous limit of large N and λ, which leads to a strongly coupled gauge theory of many gauge fields on the boundary. In the bulk, however, this makes $\ell_s \to 0$ and the Newton constant $G_{\mathrm{N}} \sim 1/N^2 \to 0$, thereby

[1] For instance, at the level of symmetries, the boundary CFT has an $SO(4,2)$ conformal group and an $SU(4)$ R-symmetry, which precisely match the group of local isometries of AdS_5 and the global $SU(4) \simeq SO(6)$ symmetry of the $\mathbb{S}^5$ in the bulk, respectively.

reducing the theory to semiclassical (super)gravity. An intermediate regime of interest is the 't Hooft limit, where λ is kept finite as $N \to \infty$: the bulk remains under relative control as a perturbative theory of strings. In contrast, the regime of finite N and λ in which one may recover a tractable weakly coupled field theory on the boundary becomes a bewildering theory of strongly interacting strings of finite size in the bulk: a theory of quantum gravity in its full glory. We thus see AdS/CFT as a powerful strong/weak duality both ways, allowing one to gain insights into complicated strongly coupled theories in terms of weakly interacting ones that we can handle.

Given that the duality is well established for $N, \lambda \to \infty$, that CFT at finite N and λ is well-defined, and that we have no understanding of the bulk side in the latter regime, a conventional application of AdS/CFT is to take it as a definition of quantum gravity in the bulk.[2] Furthermore, although we have followed Maldacena's instance of AdS/CFT correspondence here, note that statements at the level of analysis above hold more generally in the realm of holographic gauge/gravity dualities. In general, N^2 encodes the number of degrees of freedom of the boundary theory (cf. the central charge of a CFT), and its inverse sets the strength of gravity in the bulk through the Newton constant. With holography at hand, the search for a theory of quantum gravity thus becomes the endeavour of first understanding the holographic dictionary when the bulk is semiclassical, and then building our knowledge up to finite N.

The most general statement of gauge/gravity duality at our disposal, irrespective of parameters or energy scales, is as an equivalence between bulk and boundary partition

[2]Were there to be found a non-perturbative string theory of quantum gravity independently of AdS/CFT, one could use it to test AdS/CFT as a conjecture at finite N and λ. Instead, the modern perspective is to understand whatever it is that AdS/CFT gives in the bulk as defining a theory of quantum gravity, thereby making the correspondence true by definition. This is reasonable precisely because the large N and λ limit is known to reduce to classical gravity.

functions,

$$\mathcal{Z}_{bulk} = \mathcal{Z}_{bdy}. \tag{1.1}$$

Several aspects that make this equivalence more precise have been left implicit; for instance, $\mathcal{Z}_{bulk}$ involves boundary conditions on bulk quantum fields, which on $\mathcal{Z}_{bdy}$ correspond to sources of operator deformations. The main ingredient we are interested in exploring in this thesis is the gravitational sector of $\mathcal{Z}_{bulk}$, which according to (1.1) must somehow emerge out of the purely quantum and non-gravitating theory that defines $\mathcal{Z}_{bdy}$. Understanding (1.1) as a path integral statement, a popular approach is to interpret the gravitational degrees of freedom in $\mathcal{Z}_{bulk}$ as being governed by a gravitational path integral over geometries.

Much progress in the field in recent years stems in one form or another from this idea. Schematically, the gravitational path integral may be written as

$$\mathcal{P}(B) = \int_{\partial M = B} \mathcal{D}g\, e^{-I[g]}, \tag{1.2}$$

which is formally understood as a path integral over the space of all metric manifolds (M, g) of all possible topologies subject to a fixed conformal structure at their asymptotic boundary $\partial M = B$. Here we have chosen to quote the Euclidean gravitational path integral, but an expression analogous to (1.2) may be written for the Lorentzian case.[3] On mathematical grounds, the gravitational path integral should be regarded as a rather formal object, given our poor understanding already of its measure. It is possible though to make sense of (1.2) in low-dimensional models of gravity (such as Jackiw-

[3]Whether or not the Euclidean and Lorentzian gravitational path integrals are equivalent in some precise sense as a Wick rotation of each other is the subject of much debate: the existence of richer topological manifolds in Riemannian geometry seems to leave the Lorentzian version at a disadvantage when it comes to computing certain observables.

Teitelboim (JT) [24]) or in a semiclassical limit where it can be reduced to a saddle-point approximation over solutions to the gravitational equations of motion. When properly interpreted, the gravitational path integral in (1.2) constitutes the pillar of our current understanding of holographic duality and the interplay between the geometric in the bulk and the quantum on the boundary. In this thesis, we use the gravitational path integral to draw lessons in the bulk from three basic properties of the boundary field theory: entanglement, causality, and factorization.

1.2 Entanglement

Most of what we know about quantum gravity finds its origins in the gedanken laboratory of black holes and the complexity of interpreting them as statistical mechanical objects. Even before Stephen Hawking discovered that black holes evaporate [25], some were already intrigued by the remarkable mathematical resemblance between the classical laws of black hole mechanics and the ordinary laws of thermodynamics [26]. However, there was still a clash between black holes and thermodynamics that made most skeptical about this similarity being physical: as originally emphasised by John A. Wheeler, throwing an entropic object into a black hole would lower the entropy of the universe outside and violate the 2^{nd} law of thermodynamics [27]. This motivated Jacob Bekenstein in 1973 to generalize the 2^{nd} law by proposing that, for black holes to be genuinely thermodynamic, some multiple of their area should be understood as its entropy [28]. With this interpretation and using the 2^{nd} law of black hole mechanics, he thus argued that when an entropic objects falls into a black hole and its area increases, so would do its entropy such that the total entropy of the universe never decreases [29].

Hawking's imminent discovery that black holes radiate at a nonzero temperature

provided the cornerstone for interpreting the analogies between black hole mechanics and thermodynamics as physically meaningful. This included Bekenstein's conjecture, which with Hawking's result led to a quantitative connection between entropy and area for black holes. The upshot was the Bekenstein-Hawking entropy,

$$S_{\text{BH}} = \frac{A}{4}, \tag{1.3}$$

a formula giving the entropy of a black hole in terms of the area A of its horizon in Planck units. String theory has provided successful evidence for the statistical interpretation of S_{BH}, suggesting that (1.3) may hold the answer to deep questions about the microscopic structure of quantum gravity.[4] More recently, with the advent of holography, it has even been possible to elevate our understanding of the Bekenstein-Hawking entropy from a thermodynamic to a quantum mechanical von Neumann entropy.

At the heart of this statement lies the Ryu-Takayanagi (RT) prescription [32], whose ingenuity was to generalize the relationship between areas and entropies in quantum gravity beyond black holes. In particular, they proposed that, for holographic theories, the von Neumann entropy S_{A} of a spatial boundary region A gets geometrized in the bulk in terms of the area of a specific bulk surface. Explicitly [32, 33, 34],

$$S_A = \min_{\mathcal{A}} \frac{\text{area } \mathcal{A}}{4}, \tag{1.4}$$

where the minimization is performed over all bulk codimension-2 surfaces $\mathcal{A}$ homologous to A relative to ∂A, i.e., such that $\partial A = \partial \mathcal{A}$ in the conformally compactified spacetime. Taking A to be a single boundary of a 2-sided black hole, (1.4) precisely reproduces

[4] The counting of states of long strings has been able to reproduced the characteristic area scaling of (1.3) [30] and even its precise numerical value in certain supersymmetric cases [31].

(1.3) with the RT surface $\mathcal{A}$ corresponding to the bifurcation surface. Given that such a geometry in AdS is dual to two entangled copies of a CFT in the thermofield-double state, we see that the RT formula grants S_{BH} the interpretation of an entropy of entanglement between the two boundary theories.

Though purely conjectural at its inception, a first-principles derivation of the RT formula has by now been understood in detail: it follows from the gravitational path integral when used to perform a bulk computation of a replica trick [35], a standard method for obtaining von Neumann entropies in ordinary quantum field theory. As a result, the RT prescription is a well-established and powerful entry in the holographic dictionary. Serving as a beacon for understanding more details about the microscopic structure of spacetime, (1.4) has been generalized to general time-dependent spacetimes [36], and to perturbative quantum gravity to all orders in the bulk Planck constant [37, 38]. Altogether, the RT formula and its generalizations provide a strong connection between bulk geometry and boundary correlations, and thereby embody the modern slogan that spacetime emerges from entanglement [39].

Importantly, that the RT formula should reproduce the results of the von Neumann entropy for arbitrary partitions of a quantum state establishes a necessary condition for it to be holographic, i.e., for it to admit a classical geometry as a bulk dual. Put differently, one is led to wonder if for any entanglement structure there may possibly exist a geometrization of it such that (1.4) applies. Very interestingly, the answer is no: a remarkable finding from the study of holographic entanglement is that, regardless of the theory, there exist strong constraints holographic states satisfy which general quantum states do not [40, 41].

These constraints following from the RT formula take the form of inequalities on the von Neumann entropy of subsystems of a quantum state. More precisely, for any positive

integer n, the constraints for arbitrary mixed quantum states on n parties can be written as homogeneous linear inequalities on all possible $2^n - 1$ nonempty subsystems. When combined, the collection of all such n-party holographic entropy inequalities can be used to define a polytope in the space of all subsystem entropies. The formalization of this idea was carried out in [42], where it was shown that the resulting object parameterizing the space of allowed entropies for holographic states is a rational polyhedral cone coined the holographic entropy cone (HEC).

Constructing the HEC is tantamount to the formidable challenge of characterizing relations among areas of surfaces in Riemannian manifolds subject to relative homology conditions. Remarkably, this problem at the interface between quantum gravity and information theory turns out to be amenable to systematic study. This is accomplished via a graph-theoretic reformulation of the RT formula pioneered by [42], who proved that the HEC could be equivalently defined as the space of entropies that can be realized by the weights of minimum cuts on weighted graphs. Despite this breakthrough, the general structure of the HEC for a general number of parties remains largely unknown, and has only been solved explicitly for up to 5 parties [13].

Part I of this thesis is devoted to the study of the HEC. It begins in Chapter 2, which rigorously establishes a graph-theoretic and computational framework for the study of the HEC, providing proofs of the most basic and structurally important results about this object. These are then used to develop two systematic approaches for finding the facets and extreme rays of the HEC, which are illustrated by algorithmically constructing it for 5 parties.

Having laid out the mathematical foundations of the HEC bottom-up, Chapter 3 then engages in the pursuit of a top-down derivation of it from a physical viewpoint. In particular, we argue that the HEC can be reconstructed from more fundamental data de-

termined solely by the subadditivity of quantum entropy, i.e., the property which states that the entropy of a system is no larger than the sum of the subsystem entropies for any bipartition of it. We formulate certain conjectures about graph models of holographic entanglement, for which we provide strong evidence, and rigorously prove that they all imply that such a reconstruction is possible. These conjectures (except only for the weakest) further imply that the necessary data is remarkably simple: all one needs to know to reconstruct the HEC is the holographic extreme rays of a cone whose representation in terms of inequalities is straightforward and simply involves subadditivity.

However, when the entropy of quantum fields across bulk RT surfaces becomes comparable to their area, the RT formula is not valid anymore. Indeed, as alluded to previously, a generalization of it accounting for quantum corrections is required, and RT is superseded by Engelhardt-Wall's quantum extremal surface (QES) prescription for the generalized entropy [38]. The entropy inequalities that define the HEC are no longer satisfied once general quantum corrections are included by employing the QES formula. Nevertheless, the structure of the QES formula allows for a controlled study of how quantum contributions from bulk entropies interplay with HEC inequalities. In Chapter 4, we initiate an exploration of this problem by relating bulk entropy constraints to boundary entropy inequalities. In particular, we show that requiring the bulk entropies to satisfy the HEC inequalities implies that the boundary entropies also satisfy these. Further, we also show that requiring the bulk entropies to obey monogamy of mutual information implies the boundary entropies also obey this inequality. These results initiate the exciting program of characterizing holographic entanglement in quantum regimes.

1.3 Causality

Causality is a rather primitive property of the universe and how we perceive it. It is generally understood as an asymmetric relation between two events, a cause and an effect, where the former precedes and contributes to the occurrence of the latter. This fairly general metaphysical formulation can be formalized in different ways in the physical sciences, where the notion of causality is ubiquitous. Part II of this thesis is devoted to exploring how causality can be used to enhance our understanding of geometric emergence in holography not only semiclassically, but also when the bulk theory involves a genuine gravitational path integral over spacetimes.

All sensible implementations of this primitive physical property find common ground in Einstein's special theory of relativity through the idea that the speed of light in the vacuum sets a cosmic speed limit on how fast any physical interaction can propagate. Thereby so-called relativistic theories have causality built-in as the requirement that no superluminal signaling be allowed. This way, the trajectories followed by light rays, or null geodesics, take the distinguished role of delineating the light cones off of any point, which define the boundaries of the domain of influence of any physical events. Light cones thus endow spacetime with a causal structure which determines which events may or may not be causally connected, and a temporal (partial) order among them. In a precise sense, in fact, a causal structure itself completely defines a spacetime metric up to a scale function or conformal factor. This significant role that causality plays in building spacetime will be the central idea of Chapter 5. By confronting the causal structure of the boundary field theory with that of the bulk, we will then also ascertain some profound lessons about the gravitational path integral in Chapter 6.

As reviewed in Section 1.2, information theoretic measures have traditionally under-

pinned the idea that entanglement builds geometry. This has often been interpreted as the mechanism by which geometric properties of the gravitational bulk theory are encoded in the dual field theory, providing a crucial entry in the holographic dictionary. However, almost all work on metric reconstruction to date assumes settings where the bulk theory is just one dimension higher than the boundary theory. Realistic models of holography are not like this though: the bulk often involves extra internal dimensions which become compact asymptotically and trivialize on the conformal boundary.[5] In other words, the conformal boundary generally misses more than one bulk dimension, and it is then unclear how the entanglement entropy of the field theory can possibly encode all of these.

One may argue that those extra dimensions that become compact asymptotically simply deserve a different treatment than the extended ones where, rather than as geometric degrees of freedom, they are handled as matter fields resulting from a Kaluza-Klein dimensional reduction of spacetime. Unfortunately, such a distinction between compact and extended dimensions is not even well-defined in general: even when spacetime asymptotes to a product space, there may exist no unique choice of base space in the interior (cf. a warped geometry). Hence, while a distinction is available asymptotically, in the interior all bulk dimensions really are locally equivalent and should be recoverable from boundary data in the same way. The question thus remains whether entanglement alone can possibly be responsible for the emergence of the full-dimensional bulk geometry when the conformal boundary is just codimension 1.

At present, the answer is certainly no: RT surfaces and their generalizations are only well understood as codimension-2 bulk objects for codimension-1 boundaries.[6] We thus

[5]For instance, Maldacena's most notable instance of AdS/CFT correspondence [16] has $AdS_5 \times \mathbb{S}^5$ in the bulk. The compact space $\mathbb{S}^5$ of such a bulk spacetime shrinks away on the conformal boundary, which becomes just a 4-dimensional flat spacetime where the Yang-Mills theory is defined.

[6]In fact, for higher-codimension boundaries, it is not even clear what dimensionality the RT surfaces

pursue an alternative approach to the reconstruction of bulk metrics which is rooted in causality rather than entanglement. The relevant framework will be the light-cone cuts studied earlier by [45, 46]. Light-cone cuts are defined as the intersection of the light cone of bulk points with the conformal boundary. As [45, 46] originally showed, knowledge of the light-cone cuts can be used to completely reconstruct the metric of the bulk spacetime up to a conformal factor for all points in causal contact with the conformal boundary. Crucially, for most points in the causal wedge of the conformal boundary, these light-cone cuts are themselves obtainable intrinsically from within the boundary theory via a strategy which exploits the singularity structure of boundary correlation functions. Their results, however, implicitly assumed no extra bulk dimensions and thus only applied to codimension-1 boundaries.

In Chapter 5, we generalize the approach to bulk reconstruction using light-cone cuts and propose a prescription to obtain the full higher-dimensional metric of generic spacetimes up to an overall conformal factor. We first extend the definition of light-cone cuts to include information about the asymptotic compact dimensions, and show that the full conformal metric can be recovered from these extended cuts. We then give a prescription for obtaining these extended cuts from the dual field theory. The location of the usual cuts can still be obtained from bulk-point singularities of correlators, and the new information in the extended cut can be extracted by using appropriate combinations of operators dual to Kaluza-Klein modes of the higher-dimensional bulk fields.

The second part of this thesis concludes in Chapter 6 with an incursion into the wilderness of the gravitational path integral in a fully quantum regime beyond the semi-classical limit. A standard implementation of causality in quantum field theory at an algebraic level is as the requirement that local operators commute at spacelike separation.

should posses, or what asymptotic boundary conditions should be prescribed for them (see e.g. [43, 44]).

This is known as the condition of microcausality and ensures that causally-disconnected observables be unable to influence each other. A field theory is thus considered causal if its operator algebra is compatible with microcausality. Now, even for holographic theories which respect causality, the full bulk Lorentzian path integral includes metrics that violate the boundary causal structure. In other words, there generally exist off-shell geometries in the bulk which causally connect boundary points which would otherwise be spacelike with respect to the boundary causal structure alone. This leads to the following puzzle: for a causal field theory, the commutator of two local operators at spacelike-separated points on the boundary must vanish. However, if these points are causally related in a bulk metric, then the bulk calculation of the commutator will generically be nonzero. It would appear that the integral over all metrics of this commutator must vanish exactly for holography to hold. This is puzzling since it must also be true if the commutator is multiplied by any other operator. Upon a careful treatment of boundary conditions in holography and the role played by timefolds in their calculation, we show how the bulk path integral leads to a natural resolution of this puzzle.

1.4 Factorization

A rather trivial property of a theory involving subsystems that are completely decoupled is that its partition function factorizes into a product of partition functions associated to their independent degrees of freedom. If $Z(B)$ is the partition function of a quantum field theory on some manifold B, it thus follows that the partition function $Z(B^m)$ for m identical copies of the theory on disjoint copies of B is nothing but

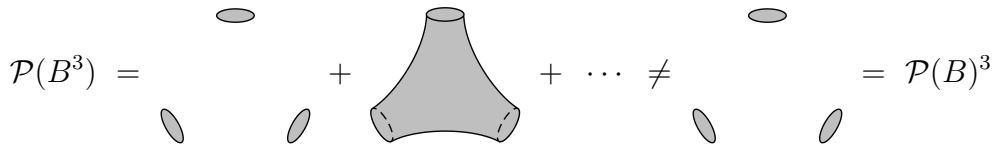
$$Z(B^m) = Z(B)^m. \tag{1.5}$$

In other words, the partition function factorizes.

The Euclidean gravitational path integral is generally understood as computing a thermal partition function, and treated as such in many computations in quantum gravity or in the very definition of holographic duality in (1.1). However, our current understanding of (1.2) is in stark conflict with the factorization property from (1.5), since generally

$$\mathcal{P}(B^m) \neq \mathcal{P}(B)^m. \tag{1.6}$$

The reason for this lack of factorization of the gravitational path integral is at the same time what makes it so interesting and puzzling: wormholes. Even when the boundary consists of multiple connected components, as is the case for B^m, there still exist connected wormhole geometries joining these in the bulk. For instance, for $m = 3$ and B a circle, one would pictorially have



where the ellipsis refers to partially-connected wormholes and we are ignoring higher topologies. Prima facie, the resolution of the factorization problem seems obvious: do not include such connected topologies in the gravitational path integral. Unfortunately, this would not be satisfactory either; one of the main lessons we have learned from recent explorations of the black hole information problem is that wormholes are crucial for restoring unitarity [47, 48, 49, 50]. In particular, wormholes contribute to the gravitational computation of the von Neumann entropy and are responsible for the turn-over of the Page curve that is required for unitarity of the process of of black hole evaporation.

Part III of this thesis delves into further consequences of the lack of factorization

of the gravitational path integral. As recently pointed out by [51], wormholes do not only arise in the replica trick for the von Neumann entropy, but also in the gravitational computation of other extensive quantities. This includes the arguably most fundamental object of a theory: the generating functional. Chapter 7 presents a novel semiclassical prescription for the computation of generating functionals in quantum gravity which accounts for the contribution of connected wormhole geometries. This prescription is then put to test in a model of JT gravity coupled to conformal matter, where we find wormhole saddles with lower action than the disconnected topologies. The stability of these saddles (and thus the resulting generating functional) turns out to be sensitive to an interesting subtlety in the definition of the Euclidean gravitational path integral: whether it is obtained as a deformation of a Lorentzian theory or independently defined.

1.5 Permissions and Attributions

1. The content of Chapter 2 and Appendix A is the result of a collaboration with David Avis [8].

2. The content of Chapter 3 and Appendix B is the result of a collaboration with Veronika Hubeny and Massimiliano Rota [2].

3. The content of Chapter 4 is the result of a collaboration with Chris Akers and Pratik Rath, and has previously appeared in the Journal of High Energy Physics [5]. It is reproduced here with the permission of the International School of Advanced Studies (SISSA): http://jhep.sissa.it/jhep/help/JHEP/CR_OA.pdf.

4. The content of Chapter 5 and Appendix C is the result of a collaboration with Gary Horowitz, and has previously appeared in the Journal of High Energy Physics [10].

It is reproduced here with the permission of the International School of Advanced Studies (SISSA): http://jhep.sissa.it/jhep/help/JHEP/CR_OA.pdf.

5. The content of Chapter 6 is the result of a collaboration with Gary Horowitz, Gabriel Treviño and Diandian Wang, and has previously appeared in Physical Review Letters [7]. It is reproduced here with the permission of the American Physical Society: http://journals.aps.org/copyrightFAQ.html#thesis.

6. The content of Chapter 7 and Appendix D is the result of a collaboration with Venkatesa Chandrasekaran, Netta Engelhardt and Sebastian Fischetti [1].

Part I

Entanglement

Chapter 2

Foundations and extremal structure of the holographic entropy cone

2.1 Introduction

Although the holographic entropy cone (HEC) can be defined and viewed entirely in a combinatorial setting, we first review its origin and importance in physics. Besides this short description, no physics background is assumed or needed for reading the rest of this chapter.

The tools of convex geometry have long been applied to systematically study entropy inequalities, from those obeyed by the Shannon entropy of random variables in classical probability distributions[52, 53, 54, 55], to the ones that the von Neumann entropy of marginals of density matrices of quantum states satisfy [56, 57, 58, 59]. As a measure of quantum entanglement, the study of the latter has proven to be of paramount importance to the development of the field of quantum information theory and, more generally, to the understanding of correlations in quantum physics [60].

Although the finding of universal inequalities obeyed by general quantum states has been elusive, significant progress has been made by the restriction of the domain of the entropy function to specific subclasses of quantum states of special relevance for which additional tools are at hand [61, 62, 63, 64, 12]. In the context of quantum gravity and holography [16, 18], one very important such class of quantum states are those which admit a semi-classical description in terms of a theory of gravity on a higher-dimensional spacetime. More specifically, in such cases, the Anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence asserts that a holographic state of the CFT, defined on a *boundary* spacetime M, has a gravitational *bulk* dual on a spacetime $\mathcal{M}$ with M as its conformal boundary, $\partial\mathcal{M} = M$. In the bulk, quantum entanglement of the CFT state acquires a geometric character which has been understood to play a fundamental role in the very emergence of spacetime itself [39]. These findings rely on the much celebrated Ryu-Takayanagi (RT) prescription [32, 34], according to which the von Neumann entropy $S(R)$ of a spatial boundary region $R \subset M$ is given holographically by

$$S(R) = \min_{\mathcal{R}\subset\mathcal{M}} \frac{\text{area}(\mathcal{R})}{4G\hbar}, \tag{2.1}$$

where G is Newton constant, $\hbar$ is Planck constant, and the minimization is over bulk hypersurfaces $\mathcal{R}$ in a time slice homologous to R relative to ∂R, i.e., subject to the condition $\partial\mathcal{R} = \partial R$. This geometric character that the von Neumann entropy acquires in the bulk turns out to place strong constraints on the allowed entanglement structures of holographic states. In a remarkable paper, [42] initiated a systematic exploration of these constraints with the objective of formalizing a set of conditions on quantum states to posses holographic duals. These were formulated as entropy inequalities satisfied by the RT formula, defining the facets of a polyhedral cone which was coined as the HEC.

More precisely, the HEC is a family of polyhedral cones H_n labelled by an integer $n \geq 1$, all related by projections from larger to smaller n. Their work laid the ground for the finding of new results about the HEC [13, 65, 66], and also lead to further generalizations and explorations of their methods [67, 68, 12, 11, 69, 9]. Most of these developments relied on two remarkable results of [42]: a proof of equivalence between holographic entropies obtained by the RT formula and minimum cuts on weighted graphs[1], and the invention of a combinatorial method to prove the validity of holographic entropy inequalities that we will review in Section 2.4. Crucially, their graph models allow for a complete study of the HEC from a purely combinatorial viewpoint without reference to the geometric RT formula or quantum physics.

In this chapter we build the foundations of the HEC from its graph-theoretic characterization in terms of minimum cuts in a complete graph, which is a natural generalization of the well-studied cone of cut functions. Our main focus is on the extremal structure of the HEC. At present no compact representation of either the extreme rays or the facets of H_n is known and a complete explicit description is only known up to $n = 5$, see [42, 13].

This chapter is structured as follows. Firstly, in Section 2.2 we give a formal definition of H_n and some basic structural results that will be needed throughout the chapter. These include new proofs that it is full-dimensional and polyhedral. In proving the latter result, using antichains in a lattice, we obtain a tighter bound on the size of the complete graph needed to realize all extreme rays of H_n. We then review some basic results on the H- and V-representations of cones and study H_2 relating it to the cone of cut functions. In Section 2.3 we discuss the extreme rays of H_n and describe some related cones that lead to methods to compute them. This gives a simple proof that the H_n is a rational

[1]Intuitively, the graph provides a discrete tessellation of the manifold which encodes sufficient information about its metric in the form of edge weights, with minimal surfaces and their areas becoming minimum cuts and their weights, respectively – see [42].

cone. It also allows us to give a description of H_3. Following that we give a general zero-lifting result for extreme rays. In Section 2.4 we describe valid inequalities and facets. We begin by reviewing the proof-by-contraction method that is used for proving validity of inequalities. In the proof we again use antichains, obtaining a reduction in the complexity of the original method. This is followed by a discussion of zero-lifting of valid inequalities and facets. In Section 2.5 we describe integer programs that can be used to test membership in H_n and prove non-validity of inequalities defined over it. Many of the results of the chapter are combined in Section 2.6, which describes two methods to derive complete facet and extreme-ray descriptions of H_n and illustrate these on computations of H_5. There are a lot of interesting open problems related to the HEC, and some of these are stated throughout the chapter and in the conclusion. Supplemental material, including input and output files, integer linear programs and C code for various functions mentioned, is available online[2].

2.2 Definitions and basic results

For any positive integers k and N, let $[k] = \{1, 2, \ldots, k\}$, and let K_N denote the undirected complete graph on the vertex set $[N]$. The edge set E_N consists of all edges $e = (i, j)$ between vertices $i, j \in [N]$ for every pair $1 \leq i < j \leq N$. A weight map $w : E_N \to \mathbb{R}_{\geq 0}$ is introduced to assign a nonnegative weight $w(e)$ to every $e \in E_N$. Any subset $W \subseteq [N]$ defines a *cut* $C(W)$ as the set of all edges (i, j) with $i \in W$ and $j \notin W$. Since both W and its complement define the same cut, we will normally consider cuts where $W \subseteq [N-1]$, and generally exclude the empty cut. We denote by $S_W = \|C(W)\|$ the total weight of the cut $C(W)$, which is the sum of the weights of all the edges in

[2] http://cgm.cs.mcgill.ca/~avis/doc/HEC/HEC.html

$C(W)$. Letting $n = N - 1$, consider the *S-vector* of length $2^n - 1$ with entries indexed by cardinality and then lexicographically by the non-empty subsets of $[n]$,

$$S = (S_1, S_2, \ldots, S_n, S_{12}, \ldots, S_{12\ldots n}), \tag{2.2}$$

where juxtaposition is a shorthand for the corresponding set of integers. The convex hull of the set of all S vectors for a given N forms a cone in $\mathbb{R}^{2^n-1}$. In fact, this cone is polyhedral, its facets are the subadditivity inequalities and the submodular inequalities are valid for it, as established independently by [70] and [71]. When the vector S is expressed as a function of W it is known as the *cut function*.

The HEC is a generalization of the cone defined by the cut function. We follow [42], but adapt its notation and terminology considerably. Instead of setting $n = N - 1$, we fix some integer $n \geq 2$ and consider K_N for all $N > n$. In any such graph, we call the vertices $[n]$ *terminals* (cf. boundary regions in holography). The vertex N is called the *purifying vertex* in the physics literature, but we will simply call it the *sink* here. Oftentimes, these will be combined into $[n; N] = [n] \cup \{N\}$ and collectively referred to as *extended terminals*. The other vertices, if any, are called *bulk* vertices (cf. the bulk spacetime).

Let I be a non-empty subset of terminals, i.e. $\varnothing \neq I \subseteq [n]$. We extend the definition of S above to this new setting. For any $N > n$ and weight map w defined on K_N, we introduce a construct which captures all the basic properties conveyed by the RT formula in (2.1). In particular, let

$$S_I = \min_{I = W \cap [n]} \|C(W)\|, \tag{2.3}$$

where the minimization is over all $W \subseteq [N-1]$. This says that S_I takes the minimum weight over all cuts which contain precisely the terminals in I and some (possibly empty)

subset of the bulk vertices. Note that when $n = N - 1$ we are minimizing over the single subset $W = I$ and the definition is equivalent to the one given earlier. In graph theory terms, S_I is just the capacity of the minimum-weight cut or *min-cut* in K_N separating I from $[n; N] \setminus I$. By the duality of cuts and flows, an equivalent definition is to let S_I be the value of the maximum flow between multiple-sources I and multiple-sinks $[n; N] \setminus I$ in K_N. The max flow problems are structurally different for each I but nevertheless give an efficient method of computing the S_I.

We form an S-vector from (2.3) of the form of (2.2) as we did previously, and say that w *realizes* S in K_N or, more compactly, that (S, w) is a *valid pair*.

Definition 2.1. *The holographic entropy cone on n terminals is defined as*

$$H_n = \{S \in \mathbb{R}^{2^n-1} \ : \ (S, w) \text{ is a valid pair for some } N \text{ and } w\}. \tag{2.4}$$

It follows from (2.3) that for any $\lambda > 0$, (S, w) is a valid pair for H_n if and only if $(\lambda S, \lambda w)$ is, so H_n is a cone. In fact it is full-dimensional. The proof employs S-vectors arising from K_{n+2} with all edges of zero weight except possibly edges $(i, n + 1)$ for $i \in [n; N]$. We call these *star graphs* and exhibit a family of $2^n - 1$ of them giving linearly independent S-vectors.

Proposition 2.1. *H_n is a cone of dimension $2^n - 1$.*

Proof. For each $\varnothing \neq J \subseteq [n]$, define a weighted star graph where the nonzero edge weights are[3]

$$w_i = 1, \quad \forall\ i \in J \qquad \text{and} \qquad w_N = \begin{cases} 1 & \text{if } |J| = 1, \\ |J| - 1 & \text{otherwise.} \end{cases} \tag{2.5}$$

[3]This class of star graphs were inspired by a construction of [68].

For every $n \geq 2$, their respective S-vectors S^J can be easily seen to be given by

$$S^J_I = |I \cap J| - \delta(I,J), \qquad \delta(I,J) = \begin{cases} 1 & \text{if } |J| \geq 2 \text{ and } I \supseteq J, \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

Using them as row vectors, we construct square matrices A^n. For example,

$$A^2 = \begin{array}{c} \begin{array}{ccc} 1 & 12 & 2 \end{array} \\ \left[\begin{array}{c|c|c} 1 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \end{array}\right] \end{array} \begin{array}{l} 1 \\ 12 \\ 2 \end{array}$$

$$A^3 = \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 12 & 13 & 23 & 123 & 3 \end{array} \\ \left[\begin{array}{cc|c|ccc|c} 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ \hline 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 2 & 2 & 1 \\ \hline 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right] \end{array} \begin{array}{l} 1 \\ 2 \\ 12 \\ 13 \\ 23 \\ 123 \\ 3 \end{array}$$

$$\cdots \quad A^{n+1} = \begin{array}{c} \begin{array}{ccc} 1 & I \cup \{n+1\} & \{n+1\} \end{array} \\ \left[\begin{array}{ccc} B^n & C^n & c \\ D^n & E^n & \vdots \\ d & \cdots & \cdot \end{array}\right] \end{array} \begin{array}{l} \text{J} \\ \text{J} \cup \{n+1\} \\ \{n+1\} \end{array} \tag{2.7}$$

where the general sketch partitions A^{n+1} into four square matrices B^n, C^n, D^n and E^n, of size $2^n - 1$, a final column c, and a final row d. Note that the rows and columns have been permuted from their usual ordering for subsets of $[n+1]$. Here, labels $\varnothing \neq I \subseteq [n]$ go first, then those of the form $I \cup \{n+1\}$, and $\{n+1\}$ last (cf. the block forms in (2.7)). We prove by induction on n that $\det(A^n) = (-1)^{n+1}$. This is immediate for $n = 2$. Matrix B^n in A^{n+1} is just A^n reordered as described above. Since we perform the same reordering for rows as for columns the determinant sign is unchanged, so by the induction hypothesis $\det(B^n) = (-1)^{n+1}$. As the rows of B^n and C^n are indexed by $J \not\ni n+1$, we have $C^n = B^n$. Additionally, one easily verifies that in c the first $2^n - 1$

entries are 0 and the rest are 1. Row d has the same pattern.

We now make a comparison between entries in column $I \not\ni n+1$ of D^n and column $I \cup \{n+1\}$ of E^n. Consider the diagonal elements of each. For row J, in D^n we have column $I = J \setminus \{n+1\}$ and so $S_I^J = |I| = |J| - 1$. In E^n the column label is also J and since $|J| \geq 2$ we have $\delta(I, J) = 1$ and so $S_I^J = |J| - 1$. Hence the diagonals are identical. Now consider the elements below them. For D^n each row index J contains $n+1$ but none of its column indices do, so $S_I^J = |I \cap J|$. In E^n the same applies but the intersection now includes $n+1$, so the corresponding entry is always bigger by one. These facts are illustrated by the coloured entries in A^3.

We now subtract the first $2^n - 1$ columns of A^{n+1} from the next $2^n - 1$ columns, then subtract c from each of these columns also, obtaining

$$A^{n+1} = \begin{bmatrix} B^n & B^n & \mathbf{0} \\ D^n & E^n & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & 1 \end{bmatrix} \quad \longrightarrow \quad \tilde{A}^{n+1} = \begin{bmatrix} B^n & \mathbf{0} & \mathbf{0} \\ D^n & E^n - D^n - \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}. \tag{2.8}$$

Here $\mathbf{0}$ and $\mathbf{1}$ respectively denote all-0 or all-1 matrices of suitable size. In the resulting $\tilde{A}^{n+1}$, notice that $E^n - D^n - \mathbf{1}$ is an upper triangular matrix with all diagonal elements equal to -1. Recalling that $\det(B^n) = (-1)^{n+1}$, we have $\det(A^{n+1}) = \det\left(\tilde{A}^{n+1}\right) = (-1)^{n+2}$, as desired. □

We will show in the following sections that H_n is also convex, polyhedral and rational. One important basic property the S-vectors do not possess is monotonicity, as can be seen by examples in Appendix A.1.2.

2.2.1 Polyhedrality of the HEC

In general, there may be more than one min-cut W for each $\varnothing \neq I \subseteq [n]$ achieving the minimum in (2.3). Among these, let W_I denote one which is minimal under set inclusion. We call W_I a *minimal min-cut* for I and have $S_I = \|C(W_I)\|$. The following basic theorem shows that these are unique and builds on results from Lemma 6 of [72], and Theorems 3.1 and 3.2 of [9].

Theorem 2.1. *For positive integers $n < N$, consider a weighted complete graph K_N with terminal set $[n]$. Let W_I and W_J be minimal min-cuts for $\varnothing \neq I, J \subseteq [n]$. Then:*

(a) Each $I \subseteq [n]$ has a unique minimal min-cut W_I.

(b) $$I \subseteq J \quad \Longleftrightarrow \quad W_I \subseteq W_J.$$

(c) $$I \cap J = \varnothing \quad \Longleftrightarrow \quad W_I \cap W_J = \varnothing.$$

(d) If $m = \left|\cup_{I \subseteq [n]} W_I\right|$, then all minimal min-cuts can be represented in a weighted K_{m+1}.

Proof.

(a) Suppose W and W' are minimal min-cuts for I. Submodularity of the cut function gives

$$\|C(W)\| + \|C(W')\| \geq \|C(W \cup W')\| + \|C(W \cap W')\|. \tag{2.9}$$

Clearly, $W \cup W'$ and $W \cap W'$ are cuts for I. Since W and W' are additionally min-cuts,

$$\|C(W \cup W')\| \geq \|C(W)\|, \qquad \|C(W \cap W')\| \geq \|C(W')\|, \tag{2.10}$$

thereby turning all inequalities above into equations. Hence $W \cap W'$ is a min-cut and, as an intersection of minimal ones, minimal as well. It must thus be the case that $W = W' = W_I$.

(b) First assume that $W_I \subseteq W_J$. Since W_I and W_J are cuts for I and J respectively, we have $W_I \cap [n] = I$ and $W_J \cap [n] = J$. As $W_I \subseteq W_J$, we have $W_I \cap [n] \subseteq W_J \cap [n]$. Hence $I \subseteq J$.

Now assume that $I \subseteq J$. Again, as min-cuts, $W_I \cap [n] = I$ and $W_J \cap [n] = J$, and therefore $(W_I \cap W_J) \cap [n] = I$ and $(W_I \cup W_J) \cap [n] = J$. This means $W_I \cap W_J$ and $W_I \cup W_J$ are respectively cuts for I and J. Then submodularity and minimality, applied to $W = W_J$ and $W' = W_I$ as in the proof of (a) above, imply $W_I \cap W_J = W_I$, which proves the claim.

(c) By the definitions, $W_I \cap W_J = \varnothing$ implies that $I \cap J = \varnothing$.

For the converse, suppose that $I \cap J = \varnothing$ and that there exists a vertex $x \in W_I \cap W_J$. Let a, b and c be the total weight of edges from x to, respectively, $W_I \setminus W_J$, $W_J \setminus W_I$ and $[N] \setminus (W_I \cup W_J)$. Since W_I is a min-cut, $a > b + c$, for otherwise we could remove x from W_I without increasing the weight of the cut. Similarly, by considering W_J, we have $b > a + c$. As edge weights are nonnegative, this gives the desired contradiction.

(d) Firstly, we renumber the vertices $n + 1, \ldots, N$ in K_N so that vertices $[m]$ cover all of the vertices in the union of the W_I. In K_{m+1} we will let $m + 1$ take the role of the sink N and adjust weights as follows. We leave the edge weights unchanged between edges with both endpoints in $[m]$. For $i \leq m$ we give edge $(i, m + 1)$ the weight corresponding to the sum of the weights of all edges (i, j)

with $j = m+1, \ldots, N$. It is easy to verify that the weights of the min-cuts W_I are preserved: if a smaller weight cut for a terminal set I existed in K_{m+1}, then it could be reproduced in the original K_N, a contradiction.

□

Unfortunately, part (b) above does not generalize to the intersection of three or more sets. A simple example is given by the K_5 star graph with unit weights for the 3 terminal edges and zero for the sink edge. In particular, the intersection of the three pairs of terminals is of course empty, but the intersection of their minimal min-cuts is not as it contains the bulk vertex.

Each S-vector is realized in K_N for some N, and we are interested in the smallest such N. More generally, for a given n, is there a smallest integer $m(n)$ such that all S-vectors on $[n]$ can be realized in $K_{m(n)}$? The answer is yes and this was proved by [42] (Lemma 6) who obtained $m(n) \leq 2^{2^n-1}$. A tighter bound can be obtained from Theorem 2.1 as follows.

Let $Bool_n$ denote the *Boolean lattice* of all subsets of $[n]$ ordered under inclusion. A family of subsets of $[n]$, $\mathcal{I} \in Bool_n$, is an *upper set* if for each $I \in \mathcal{I}$ and $J \subseteq [n]$ that contains I we have $J \in \mathcal{I}$. We call $\mathcal{I}$ *pairwise intersecting* if each pair of its constituent subsets has a non-empty intersection. If $\mathcal{I}$ is the empty set or consists of a singleton, we consider $\mathcal{I}$ to be pairwise intersecting. An *antichain* in $Bool_n$ is a collection of subsets of $[n]$ which are pairwise incomparable, i.e. none of them is contained in any of the others. Notice that the minimal elements of any upper set form an antichain and that two non-empty antichains pairwise-intersect if and only if their upper sets do. Let $M(n)$ denote the number of antichains in $Bool_n$ which are pairwise intersecting. We can use this value to bound $m(n)$ as follows:

Corollary 2.1. *For $n \geq 2$, every S-vector for n terminals can be realized in $K_{m(n)}$, where*

$$m(n) \leq M(n). \tag{2.11}$$

Proof. We first sketch the argument in [42] for their upper bound on $m(n)$. Suppose a given S-vector on n terminals can be realized in a weighted K_N, for some given N. For $I \subseteq [n]$, W_I partitions the vertex set $[N]$ of K_N into two subsets. If we intersect these by W_J, for some $I \neq J \subseteq [n]$, we get 4 subsets, some possibly empty. After repeating for all $2^n - 1$ non-empty subsets of $[n]$ we obtain a partition of $[N]$ into 2^{2^n-1} subsets, many of which may be empty. However, in each of the non-empty subsets, the vertices of K_N may be merged into a single vertex by combining edge weights (cf. Theorem 2.1(d)). This new complete graph has at most 2^{2^n-1} vertices and realizes the same min-cut weights as before, giving their result.

To improve this bound we use Theorem 2.1. For a set $W \subseteq [N]$, denote its complement by $W^{\complement} = [N] \setminus W$. Any atom in the partition just described is formed by splitting the non-empty subsets of $[n]$ into two disjoint, spanning families $\mathcal{I}$ and $\mathcal{J}$, and taking the intersection

$$\bigcap_{I \in \mathcal{I}} W_I \quad \cap \quad \bigcap_{J \in \mathcal{J}} W_J^{\complement}. \tag{2.12}$$

Suppose this intersection is non-empty. Theorem 2.1(c) implies that $\mathcal{I}$ is pairwise intersecting, for otherwise the left intersection in (2.12) is empty. In particular, this implies that both a subset and its complement cannot be in $\mathcal{I}$. In addition, one can show that $\mathcal{I}$ must either be empty or an upper set in $Bool_n$ as follows. If $\mathcal{I} = \varnothing$, then (2.12) is in fact never empty because it will always contain vertex N. As for $\mathcal{I} \neq \varnothing$, consider two subsets $I \subset K \subseteq [n]$ and suppose $I \in \mathcal{I}$ is non-empty. We have by Theorem 2.1(b) that $W_I \subseteq W_K$ and so $W_I \cap W_K^{\complement} = \varnothing$, implying that if $K \in \mathcal{J}$, then (2.12) is empty. As a

result, either $\mathcal{I} = \varnothing$ or $\mathcal{I}$ must be a pairwise intersecting upper set in $Bool_n$, with $\mathcal{J}$ containing all other non-empty subsets of $[n]$.

Because empty atoms from (2.12) do not contribute to min-cut weights, it follows that when considering S-vectors we need only be concerned with pairwise intersecting upper sets $\mathcal{I}$ in $Bool_n$ and $\mathcal{I} = \varnothing$. As described above, the upper sets can be equivalently enumerated as the number of pairwise intersecting antichains in $Bool_n$. Since $M(n)$ counts their number, we conclude that all S-vectors with n terminals can be realized in $K_{M(n)}$. □

We have the following reasonably tight asymptotic bounds on $M(n)$. Let $\bar{M}(n)$ be the total number of antichains in $Bool_n$. Then,

$$\binom{n}{\lfloor \frac{n}{2} \rfloor + 1} < \log_2 M(n) < \log_2 \bar{M}(n) \sim \binom{n}{\lfloor \frac{n}{2} \rfloor} \sim \frac{2^{n+1}}{\sqrt{2\pi n}}. \tag{2.13}$$

The asymptotic upper bound on $\bar{M}(n)$ is due to [73]. The lower bound can be obtained by considering all subsets of $[n]$ of size $\lfloor n/2 \rfloor + 1$. Each pair of such subsets intersects and none can properly contain another. So any collection of these subsets forms an intersecting antichain. While we do not know of tighter asymptotic bounds for $M(n)$, exact values are known [74][4] for $1 \leq n \leq 8$:

$$2, \quad 4, \quad 12, \quad 81, \quad 2646, \quad 1422564, \quad 229809982112, \quad 423295099074735261880. \tag{2.14}$$

However, it seems that $M(n)$ is a very poor upper bound on $m(n)$. For example, data for $1 \leq n \leq 4$ shows that $m(n) = 2, 3, 5, 6$ – see Section 2.5 for more details.

Problem 2.1. *Find tighter bounds on $m(n)$. In particular, does $\log_2 m(n)$ admit an*

[4] $M(n)$ is entry $n+1$ in Proposition 1.2 of [74].

upper bound that is polynomial in n?

Definition 2.1 suggests the following family of cones, which are useful in proving the polyhedrality of H_n. For any pair of integers n and N such that $2 \leq n+1 \leq N$, consider

$$H_{N,n} = \text{conv } \{S \in \mathbb{R}^{2^n-1} \ : \ \textit{nonnegative weighted } K_N \textit{ s.t. } S \textit{ satisfies (2.3)}\}. \quad (2.15)$$

This is a generalization of the cone defined by the cut functions, which corresponds to the specific case $n = N-1$. Without the convex hull operator in (2.15), $H_{N,n}$ would not be convex in general, as shown by example in Appendix A.2. An important property of $H_{N,n}$ is that it is naturally invariant under the action of the symmetric group Sym_n which permutes the elements of the set $[n]$. In fact, $H_{N,n}$ enjoys a larger symmetry group: it is symmetric under permutations of vertices in the extended terminal set $[n;N]$. The permutations of K_N under Sym_{n+1} yield S-vectors (2.2) which are related by the simple fact that $C(W) = C(W^{\complement})$. Similarly, in an undirected graph any min-cut is insensitive to the exchange of its sources and sinks. When talking about symmetries under Sym_{n+1}, it is thus convenient to identify $S_{[n;N]\setminus I} = S_I$.

It is easy to see that in general $H_{N,n} \subseteq H_{N+1,n}$, since an additional bulk vertex can always be added to K_N with all edges containing it of weight zero. We are now able to prove that H_n is a convex polyhedral cone:

Corollary 2.2. *For any $n \geq 1$, H_n is a convex, polyhedral cone given by $H_n = H_{m(n),n}$.*

Proof. Firstly, suppose $S \in H_n$. Then for some N and weight set w the pair (S, w) satisfies (2.3) and so $S \in H_{N,n} \subseteq H_n$. Conversely, suppose S is in the convex hull of extreme rays of H_n. Each of these rays can be realized in a weighted K_N for some $N \leq m(n)$. For a specific conical combination of extreme rays giving S, consider the

graph obtained by identifying all of their associated graphs at their terminal vertices, each with its edge weights multiplied by the coefficient in the associated conical combination. This graph clearly still realizes S,[5] and can be thought of as a K_N for some large but finite N with many zero-weight edges omitted. However, the bound on N in Corollary 2.1 guarantees that by merging vertices this graph can be reduced to one with $N \leq m(n)$, thus proving convexity. □

Of course, H_n inherits the Sym_{n+1} symmetry discussed above. As a result, when considering the extremal structure of H_n, we will only specify single representatives of symmetry orbits under Sym_{n+1}.

Almost nothing is known about the complexity of computational problems related to H_n.

Problem 2.2. *Given a vector $q \in \mathbb{Z}^{2^n-1}$, what is the complexity of deciding if $q \in H_n$? What is the complexity of deciding whether $qx \geq 0$ is satisfied for all $x \in H_n$?*

2.2.2 Representations of the HEC

A basic result of polyhedral geometry is that any polyhedral cone C can be represented by a non-redundant set of facet-defining inequalities, which we can write as $Ax \geq 0$ for a suitably dimensioned matrix A and variables x, and is called an *H-representation*. One can also represent C by a non-redundant list E of its extreme rays, such that $C = \text{conv}\{E\}$, which is called a *V-representation*. Both representations are unique up to row scaling. In this chapter we are concerned with computing the H- and V-

[5] This is basically the statement that any flow network problem with multiple source/sink vertices can be equivalently reformulated in terms of a single supersource/supersink vertex connected to each of the sources/sinks with edges of infinite capacity. In our case, however, it is preferable to simply merge together terminals of the same type into a single "superterminal", rather than having unnecessary infinite-weight edges.

representations of H_n. Normally, for cones (or polyhedra) arising in discrete optimization, we have available one or the other of the representations. However, this is not the case for H_n. To proceed we will initially try to find both *valid inequalities* for C, which are those satisfied by all rays in C, and to find *feasible rays* of C. A set of valid inequalities forms an *outer approximation* of H_n and a set of valid rays forms an *inner approximation*. The following well known basic result shows when such sets respectively constitute an H- and V-representation.

Proposition 2.2. *For a given polyhedral cone C, let $Ax \geq 0$ be a non-redundant set of valid inequalities and let E be a non-redundant set of feasible rays. Then:*

(a) An extreme ray of $Ax \geq 0$ is an extreme ray of C if it is feasible for C.

(b) A facet of $\mathrm{conv}\{E\}$ is a facet of C if it is valid for C.

(c) E is precisely the set of extreme rays of $Ax \geq 0$ if and only if they respectively constitute V- and H-representations of C.

Parts (a) and (b) lead to a kind of bootstrapping process which terminates once (c) can be applied. This is described in detail in Section 2.6, but to illustrate we now look at some small values of n.

For $n = 1$, the S-vectors are the nonnegative real numbers. There is one facet $S_1 \geq 0$ and one extreme ray with $S_1 = 1$, and Proposition 2.2(c) is readily verified. Notice that this simple $n = 1$ extreme ray can be represented in K_2 with the single edge $(1, 2)$ having weight 1. More generally, any K_N where two distinct singletons $i, j \in [n; N]$ share a unit-weight edge and all other weights are zero will be called an (i, j) *Bell pair*. The S-vector of such a Bell pair has nonzero $S_I = 1$ if and only if either $i \in I$ or $j \in I$, but not both.

For $n = 2$ we have $S = (S_1, S_2, S_{12})$ and a valid inequality $S_1 + S_2 \geq S_{12}$. This follows since the union of cuts for two terminals $i, j \in [n; N]$ is always a cut for the union of the two terminals $\{i, j\}$. Hence the latter's min-cut cannot have larger weight than the sum of the weights of the two other cuts. The full orbit of $S_1 + S_2 \geq S_{12}$ contains $S_1 + S_{12} \geq S_2$ and $S_{12} + S_2 \geq S_1$. All three are related by the S_3 symmetry of permutations of $[2; N]$ and the identification of $S_{[2;N]\setminus I} = S_I$ for all $I \subseteq [2; N]$. In the context of information theory, the first one is known as *subadditivity* (SA), while the latter two are called the Araki-Lieb inequalities.

Subadditivity generalizes to a valid inequality for any disjoint, non-empty subsets of terminals I, J with $I \cup J \subset [n; N]$ to give

$$\text{SA:} \qquad S_I + S_J \geq S_{IJ}. \tag{2.16}$$

We exclude $I \cup J = [n; N]$ since in this case SA reduces to nonnegativity. Every inequality in the enlarged symmetry orbit Sym_{n+1} of (2.16) clearly remains valid. For any disjoint subsets of terminals $I, J \subseteq [n; N]$, those of Araki-Lieb type take the form $S_I + S_{I \cup J} \geq S_J$. The qualitative difference between SA and Araki-Lieb is that in the former the disjoint subsets $I, J \subseteq [n; N]$ do not contain the sink, whereas in the latter one of them does, and whenever $N \in I \subseteq [n; N]$ one identifies $S_I = S_{[n;N]\setminus I}$. For future convenience, we introduce

$$q_{I:J} S = S_I + S_J - S_{IJ}, \tag{2.17}$$

known as the *mutual information* in the physics community. This way, SA corresponds to the nonnegativity of the mutual information $q_{I:J} S \geq 0$.

Proceeding as suggested by Proposition 2.2(a), we can compute the extreme rays arising from (2.16) by the action of Sym_3, together with the 3 nonnegativity inequalities

(in fact, the latter are redundant and can be ignored). Doing so we obtain 3 extreme rays related by symmetry, of which one is $S = (1, 1, 0)$. This can be represented in K_3 by a $(1, 2)$ Bell pair. Obviously, this can also be obtained from the $(1, 2)$ Bell pair for $n = 1$ in K_2 by adding a new vertex with all edge weights to it zero. This is a process called a *zero-lifting* of extreme rays and which we discuss in detail in Section 2.3.2. So if we let E be the set of 3 output rays and $Ax \geq 0$ be the 3 SA inequalities we are again done by Proposition 2.2(c).

2.3 Extreme rays

In this section we introduce a lifting of $H_{N,n}$ to a cone for which we can explicitly write an H-representation. Computing the extreme rays of this cone and then making a projection allow us to compute a superset of the extreme rays of $H_{N,n}$. We also describe a zero-lifting operation, which allows known extreme rays for S-vectors defined on n terminals to generate extreme rays for those defined on $n + 1$ terminals.

2.3.1 A lifting of the HEC

For integers n and N such that $3 \leq n + 1 \leq N$, consider the following system of inequalities whose variables are the S-vector entries S_I and the edge weights $w(e)$ of K_N:

$$S_I \leq \|C(W)\|, \qquad \forall\, W \subseteq [N-1],\ \varnothing \neq I \subseteq [n] \ s.t.\ I = W \cap [n], \tag{2.18a}$$

$$w(e) \geq 0, \qquad \forall\, e \in E_N. \tag{2.18b}$$

Note that for each cut $C(W)$ there is precisely one I such that $I = W \cap [n]$. Since we are excluding the case $I = \varnothing$, this implies that (2.18a) contains $2^{N-1} - 2^{N-n-1}$ inequalities. There are an additional $N(N-1)/2$ nonnegative inequalities from (2.18b). Since each $\|C(W)\|$ is just a sum of weights $w(e)$ for every $e \in C(W)$, together with the $2^n - 1$ variables S_I, there are a total of $M_{N,n} = 2^n - 1 + N(N-1)/2$ variables involved in (2.18). Let (S, w) denote a vector of length $M_{N,n}$ representing these variables.

Definition 2.2. *The cone $P_{N,n}$ denotes the set of all (S, w) satisfying (2.18).*

Since $P_{N,n}$ is given explicitly by (2.18), it is a rational cone. Note that the variables S_I in $P_{N,n}$ are not bounded from below. One could add the inequalities $S_I \geq 0$ but this greatly increases the complexity of the cone, as explained below.

The cones $H_{N,n}$ and $P_{N,n}$ cannot be directly compared, since they are defined in different spaces. So we first define a lifting of $H_{N,n}$ by

$$H^+_{N,n} = \text{conv } \{(S, w) \in P_{N,n} \ : \ S \in H_{N,n}\}. \tag{2.19}$$

which is by definition a subset of $P_{N,n}$ and whose projection onto the S coordinates is $H_{N,n}$. We similarly define H^+_n. An example in Appendix A.2 shows that $H^+_{N,n}$ is in general non-convex without the convex hull operator. It is easy to see that $P_{N,n}$ contains $M_{N,n}$ *trivial* extreme rays. These are formed by setting either one $S_I = -1$ or one $w(e) = 1$, and all other variables zero.

The non-trivial extreme rays of $P_{N,n}$ include all extreme rays of $H_{N,n}$, as we now show:

Theorem 2.2.

(a) If (S, w) is a non-trivial extreme ray of $P_{N,n}$, then $S \in H_{N,n}$.

(b) If S is an extreme ray of $H_{N,n}$, then there is a weight vector w such that (S, w) is a non-trivial extreme ray of $P_{N,n}$.

Proof.

(a) Suppose that (S, w) defines a non-trivial extreme ray of $P_{N,n}$. Then there must be a set of at least $M_{N,n} - 1$ inequalities in (2.18) satisfied as equations whose solutions have the form $\lambda(S, w)$, with $\lambda \geq 0$. Each S_I must be present in at least one of these equations or else it could be increased independently of the others and the resulting vector (S', w) would still be a solution of the equations but not of that form. This in turn implies that S satisfies (2.3) for the given weight function w. Hence $S \in H_{N,n}$.

(b) Suppose S is an extreme ray of $H_{N,n}$. Since S satisfies (2.3), all of its values are nonnegative and there must exist a corresponding weight assignment $\bar{w}$. We now define a face F of $P_{N,n}$ by intersecting it with the hyperplanes

$$S_I = \|C(W)\|, \qquad \forall\ W \text{ achieving the minimum in (2.3)}, \tag{2.20a}$$

$$w(e) = 0, \qquad \forall\ e \in E_N \ s.t.\ \bar{w}(e) = 0. \tag{2.20b}$$

Each S_I appears in at least one equation. F is defined by a set of extreme rays of $P_{N,n}$ but none of these can be a trivial ray of the type $S_I = -1$ since $S_I = \|C(W)\| \geq 0$. There may be a trivial extreme ray of the type $w(e) = 1$ as long as $\bar{w}(e) \neq 0$ and the edge e does not appear in any of the cuts $C(W)$ in the system of hyperplanes. Suppose that there are s of these and write each of them as 1_e. Also, denote the non-trivial extreme rays of $P_{N,n}$ that lie on F by (S^i, w^i), with $i \in [t]$. From part (a) above, we have that $S^i \in H_{N,n}$ for all $i \in [t]$. Writing

(S, w) as a conical combination of the extreme rays that define F,

$$(S, w) = \sum_{i=1}^{t} \lambda_i (S^i, w^i) + \sum_{j=1}^{s} \mu_j 1_{e_j}, \qquad \lambda_i, \mu_j \geq 0, \tag{2.21}$$

we deduce that $S = \sum_{i=1}^{t} \lambda_i S^i$. Since S is an extreme ray of $H_{N,n}$ it follows that $S = S^i$ for all i for which $\lambda_i > 0$. For each such i, (S, w^i) is a non-trivial extreme ray of $P_{N,n}$.

□

It was initially hoped that non-trivial extreme rays of $P_{N,n}$ would project to extreme rays of $H_{N,n}$, but an example in Appendix A.2 show that this is not the case. Also, a direct projection of $P_{N,n}$ onto its S coordinates does not give $H_{N,n}$. For any given values of the S coordinates, a feasible solution to inequalities (2.18) can be obtained by choosing any suitably large w coordinates. Hence the projection is simply the whole space R^{2^n-1}. Nevertheless, we do obtain a method in principle for obtaining a complete description of H_n:

Corollary 2.3. *A V-description of H_n can be obtained by computing the extreme rays (S, w) of $P_{m(n),n}$, projecting to the S coordinates, and removing both the trivial and the redundant rays.*

Proof. By Theorem 2.2(b) every extreme ray $S \in H_{m(n),n}$ appears in some non-trivial extreme ray (S, w) of $P_{m(n),n}$. Projecting the latter to the S coordinates produces all extreme rays of $H_{m(n),n}$. Redundant rays can be removed by linear programming. Since $H_n = H_{m(n),n}$, the result follows by Corollary 2.2. □

Since $P_{m(n),n}$ is a rational cone, as remarked earlier, it follows that H_n is also. This fact, proven in Proposition 7 of [42] by a different technique, means extreme rays and

facets always admit integral representations. With our current upper bound on $m(n)$ the computation is impractical except for small values of n. It does not help here to include the extra inequalities $S_I \geq 0$ because this introduces a large number of new extreme rays (S, w) of $P_{N,n}$ which are not valid pairs. For example, while $P_{6,4}$ has 50 extreme rays of which 15 are trivial, adding nonnegativity yields 49915 extreme rays, all but 35 of which do not yield valid (S, w) pairs. A computationally lighter method to test whether a given S-vector is realizable is as follows:

Theorem 2.3. *If $\bar{S} \in H_{N,n}$, then there is a weight vector $\bar{w}$ such that $(\bar{S}, \bar{w})$ is a vertex of the polyhedron Q defined as the intersection of $P_{N,n}$ and the hyperplanes $S = \bar{S}$.*[6]

Proof. First we suppose that $\bar{S} \in H_{N,n}$. Then there exists a weight vector $\bar{w}$ for K_N that realizes $\bar{S}$. By construction $(\bar{S}, \bar{w})$ is contained in Q and so Q is non-empty. Since all components of S in Q are fixed, Q is a possibly unbounded polyhedron with vertices of the form $(\bar{S}, w)$ and possibly additional extreme rays of the form 1_{e_j} for edges e_j that do not appear in any minimum-weight cut W obeying $W \cap [n] = I$ and $\|C(W)\| = S_I$. We can write

$$(\bar{S}, \bar{w}) = \sum_{i=1}^{t} \lambda_i \, (\bar{S}, w^i), \qquad \sum_{i=1}^{t} \lambda_i = 1, \qquad \lambda_i \geq 0, \tag{2.22}$$

for a set of t vertices $(\bar{S}, w^i)$ of Q. For a cut W in K_N, let $\left\|\bar{C}(W)\right\|$ and $\|C^i(W)\|$ denote its weight using $\bar{w}$ and w^i, respectively. Since $\bar{w}$ is a realization of $\bar{S}$, for each $\varnothing \neq I \subseteq [n]$ there exists a min-cut W_I such that $\bar{S}_I = \bar{C}(W_I)$. Now, by (2.22) and the linearity of the cut weight function,

$$\left\|\bar{C}(W_I)\right\| = \sum_{i=1}^{t} \lambda_i \left\|C^i(W_I)\right\|, \qquad \sum_{i=1}^{t} \lambda_i = 1, \qquad \lambda_i \geq 0. \tag{2.23}$$

[6] The converse of this is false, as Q generally has many vertices $(\bar{S}, w)$ for which w does not realize $\bar{S}$.

Hence by (2.18a) we must also have $\bar{S}_I = \|C^i(W_I)\|$ for all $i \in [t]$. It follows that each $(\bar{S}, w^i)$ is a vertex of Q representing $\bar{S}$. □

We now show how these theorems can help in determining the extremal structure of H_3. For example, we can compute the extreme rays of $P_{5,3}$, project onto the S-coordinates, delete the trivial rays and remove redundancy, getting a V-representation of $H_{5,3}$. It contains 17 extreme rays in 7 dimensions. The H-representation of $H_{5,3}$ is easy to compute and contains 7 facets. One facet is new and the other 6 are SA inequalities:

$$S_i + S_j \geq S_{ij}, \qquad i \neq j \in [3; N], \tag{2.24}$$

where N is the sink and we recall that $S_{[3;N]\setminus I} = S_I$. Earlier we saw that for $n = 2$ there is one SA orbit of 3 facets: $S_1 + S_2 \geq S_{12}$, $S_1 + S_{12} \geq S_2$, and $S_{12} + S_2 \geq S_1$. Although the first one remains a facet of the form of (2.24) for $n = 3$, the other two do not, which may seem surprising. We return to this in Section 2.4.2 when discussing the lifting of facets (see Proposition 2.8).

The cone $H_{5,3}$ has one new facet which is inequality (2.28) below. If we can prove this is valid for H_3 then, by Proposition 2.2(c), we will have complete H- and V-representations. How to prove validity of an inequality is discussed in Section 2.4.

2.3.2 Zero-lifting extreme rays

Extreme rays for $P_{N,n}$ remain extremal for larger values of N as the following result describes:

Proposition 2.3. *If (S, w) is an extreme ray of $P_{N,n}$, then, by adding suitably many new weight coordinates set to zero, it is an extreme ray (S, w') of $P_{N',n}$ for every $N' > N$. Hence $P_{N,n}$ is a projection of $P_{N',n}$.*

Proof. Let (S, w) define an extreme ray of $P_{N,n}$ and consider the base graph K_{N+1}. We extend the weight vector w by adding N new coordinates to get a vector w' of length $(N+1)N/2$. We set $w'(i,j) = w(i,j)$ for $1 \leq i < j < N$ and $w'(i, N+1) = w(i, N)$ for $1 \leq i < N$. All edges containing vertex N receive weight zero in w', and vertex $N+1$ is the new sink. Since N is now a bulk vertex, it may participate in a cut W, but since $\|C(W)\| = \|C(W \cup \{N\})\|$, it will never change its total weight. Hence $(S, w') \in P_{N+1,n}$. Since (S, w) defines an extreme ray of $P_{N,n}$ we can choose $M_{N,n} - 1$ inequalities in (2.18) which, when satisfied as equations, have solutions in $P_{N,n}$ of the form $\lambda(S, w)$ with $\lambda \geq 0$. To these equations we add the N equations $w(i, N) = 0$. The resulting system has solutions in $P_{N+1,n}$ of the form $\lambda(S, w')$ with $\lambda \geq 0$, and the extremality of the ray defined by (S, w') follows. □

Extreme rays can also be preserved under the addition of new terminal vertices as follows. Given $(S, w) \in P_{N,n}$, we define its *zero-lift* as the vector $(S', w') \in P_{N+1,n+1}$, where S' has dimension $2^{n+1} - 1$ and w' has dimension $(N+1)N/2$, by

$$S'_{\{n+1\}} = 0, \qquad S'_{I \cup \{n+1\}} = S'_I = S_I, \qquad \varnothing \neq I \subseteq [n], \tag{2.25a}$$

$$w'(i, n+1) = 0, \qquad i \in [n], \qquad w'(i,j) = w(i,j), \qquad 1 \leq i < j \leq N. \tag{2.25b}$$

Similarly, $S' \in H_{N+1,n+1}$ above defines the *zero-lift* of the given $S \in H_{N,n}$. A vector x that satisfies an inequality $qx \geq 0$ as an equation is called a *root* of that inequality.

Proposition 2.4. *If (S, w) is an extreme ray of $P_{N,n}$, then its zero-lift (S', w') is an extreme ray of $P_{N+1,n+1}$.*

Proof. Let (S, w) define an extreme ray of $P_{N,n}$. For each $\varnothing \neq I \subseteq [n]$ choose an inequality from (2.18a) for which it is a root. These are linearly independent inequalities

since the S_I coordinates define minus the identity matrix. To these inequalities add all those from (2.18b) for which (S, w) is a root. Since (S, w) is an extreme ray we have a linearly independent set of $2^n + N(N-1)/2 - 2$ such inequalities. Call this system L. We now add terminal $n+1$ and define S' and w' as above. It is easy to verify that $(S', w') \in P_{N+1,n+1}$. Note $P_{N+1,n+1}$ has $2^n + N$ more dimensions than $P_{N,n}$ and we will augment L by this many linearly independent inequalities. Firstly, for each $\varnothing \neq I \subseteq [n]$ the inequality previously chosen will also be satisfied as an equation when I is replaced by $I \cup \{n+1\}$. This gives an additional $2^n - 1$ inequalities which are linearly independent from the others in L since the new variables $S_{I\cup\{n+1\}}$ again form minus the identity matrix. To these we add N equations $w(i, n+1) = 0$ for $1 \leq i \leq n$ and the equation $S_{\{n+1\}} = \|C(\{n+1\})\|$. This gives L the required number of tight linearly independent constraints. By construction, their solution is $\lambda(S', w')$ with $\lambda \geq 0$, which proves that (S', w') defines an extreme ray of $P_{N+1,n+1}$. □

Proposition 2.5. *If S is an extreme ray of H_n, then its zero-lift S' is an extreme ray of H_{n+1}. Hence H_n is a projection of H_{n+1}.*

Proof. If S is an extreme ray of H_n, then it is a root of a linearly independent set of $2^n - 2$ of its facet inequalities. Clearly S' is also a root of these inequalities. Additionally, S' is a root of the following $2^n - 1$ instances of SA,

$$S'_I + S'_{\{n+1\}} \geq S'_{I\cup\{n+1\}}, \tag{2.26}$$

for all $\varnothing \neq I \subseteq [n]$, and also obeys $S'_{\{n+1\}} = 0$. This gives 2^n more equations which are linearly independent and are also independent of the former $2^n - 2$ because they independently involve the new, distinct variables $S'_{I\cup\{n+1\}}$ for every $I \subseteq [n]$. Since S' satisfies $2^{n+1} - 2$ independent valid inequalities as equations, it is an extreme ray of H_{n+1}

by Proposition 2.2(a). □

To illustrate the use of the results in this subsection we consider the case $n = 4$. If we compute $P_{6,4}$ and remove the trivial rays we have 35 extreme rays. When we project onto the 15 coordinates S_I and remove redundancy, there remain 20 extreme rays. Only 5 of these are new, whereas the other 15 come from zero-lifts of the two extreme-ray classes that define H_3. The convex hull of this set is bounded by 20 facets that are examples of what is called zero-lifting from the two facet classes for H_3. We will define this process and prove that it preserves validity and facets in Section 2.4.2. By this result and Proposition 2.2(c) we have obtained the H- and V-representations of H_4.

2.4 Valid inequalities and facets

As remarked in Section 2.2, there is no general explicit H-representation known for H_n, although it can in principle be computed by using the method of Corollary 2.3 and then converting the resulting V-representation into an H-representation. In Section 2.5 we give an integer linear program (ILP) for testing whether an inequality is valid over H_n or not. However, to prove validity would require solving an ILP whose size depends on $m(n)$ and so is impractical with current bounds. The main result of this section is a tractable method known as *proof by contraction* to prove inequalities valid for H_n. By exhibiting the required number of extreme rays it is then possible to prove they are facets.

A general inequality $qS \geq 0$ over H_n is specified by a vector $q \in \mathbb{R}^{2^n-1}$. Let $K \subseteq [n]$ be the subset of terminals appearing in it, i.e. $i \in K$ if and only if $q_I \neq 0$ for some $I \ni i$. If $|K| < n$, one can turn it into an inequality over $H_{|K|}$ by relabelling terminals $K \to [\,|K|]$, if necessary. We say that an inequality over H_n is in *canonical form* if $|K| = n$, and write

it canonically as

$$\sum_{l=1}^{L} \alpha_l S_{I_l} \geq \sum_{r=1}^{R} \beta_r S_{J_r}, \tag{2.27}$$

where L and R are respectively the number of positive and negative entries in the vector q, and for all $l \in [L]$ and $r \in [R]$, the coefficients $\alpha_l, \beta_r > 0$ and the sets $I_l, J_r \subseteq [n]$ are distinct and non-empty. Because H_n is a rational cone, the normalization of (2.27) of any inequality of interest is always set such that all coefficients $\alpha_l, \beta_r > 0$ are together coprime integers.

At the end of Section 2.2.2 we discussed the cases $n = 1, 2$. Recall that H_1 is 1-dimensional and corresponds to a nonnegative half-line. Its only facet is $S_1 \geq 0$, which trivially follows from nonnegativity of the weights in (2.15). For $n = 2$, we saw that the resulting 3-dimensional cone H_2 is a simplex bounded by the 3 facets in the symmetry orbit of the SA inequality (2.16). We discussed H_3 at the end of Sections 2.3.1 and 2.3.2. Apart from the SA orbit, containing 6 facet inequalities, an additional inequality was discovered to bound H_3 [40]:

$$S_{12} + S_{13} + S_{23} \geq S_1 + S_2 + S_3 + S_{123}. \tag{2.28}$$

This is known in physics as the *monogamy of mutual information* (MMI) due to its rewriting using (2.17) as $q_{1:23}S \geq q_{1:2}S + q_{1:3}S$. Since one can write submodularity as $q_{1:23}S \geq q_{1:2}S$, by nonnegativity of $q_{1:3}S \geq 0$ one sees that MMI is a strictly stronger inequality. The proof of the validity of (2.28) will be presented in Section 2.4.1 (see Table 2.1) as an example of a general combinatorial proof method for valid inequalities of H_n. This will show that H_3 has 7 facets and so is also a simplex. As we saw at the end of Section 2.3.2, no new inequalities arise for $n = 4$. However, for $n \geq 3$ note that MMI acquires a more general form which we show is valid for H_n: for disjoint non-empty

subsets $I, J, K \subseteq [n]$,

$$\text{MMI:} \qquad S_{IJ} + S_{IK} + S_{JK} \geq S_I + S_J + S_K + S_{IJK}, \tag{2.29}$$

which is valid for H_n, as we show in Section 2.4.2.

2.4.1 Proof by contraction

We now describe the proof-by-contraction method for proving validity of inequalities for H_n. Although our description is complete, we refer the reader to [42] and [11] for more details on its derivation. By making use of antichains, as in Corollary 2.1, we obtain a stronger result than previous ones. Our discussion will be exemplified with MMI as given in (2.28).

To set the stage, let $Q_m = \{0,1\}^m \subset \mathbb{R}^m$ denote the (vertices of) the unit m-cube and refer to $x \in Q_m$ as a *bitstring*. At times, it will also be useful to think of Q_m as an m-ary Boolean domain. Given some vector of positive entries $\gamma \in \mathbb{R}^m_+$, we can turn Q_m into a metric space with distance function d_γ by endowing it with a weighted Hamming norm $\|\cdot\|_\gamma$ via

$$d_\gamma(x, x') = \|x - x'\|_\gamma, \qquad \|x\|_\gamma = \sum_{k=1}^{m} \gamma_k |x_k|. \tag{2.30}$$

Consider now a candidate inequality in canonical form over H_n, written as in (2.27). Encode each side of it into $n+1$ *occurrence vectors*, $x^{(i)} \in Q_L$ and $y^{(i)} \in Q_R$ for $i \in [n; N]$, with entries

$$x^{(i)}_l = \delta(i \in I_l), \qquad y^{(i)}_r = \delta(i \in J_r), \tag{2.31}$$

where δ is a Boolean indicator function, i.e. it yields 1 or 0 depending on whether its argument is true or false, respectively. Clearly, the occurrence vectors for $i = N \notin [n]$

are all-0 vectors. For inequality (2.28), the $i \in [n]$ occurrence vectors are the following bitstrings:

$$\begin{aligned} x^{(1)} &= (1,1,0), \qquad & y^{(1)} &= (1,0,0,1), \\ x^{(2)} &= (1,0,1), \qquad & y^{(2)} &= (0,1,0,1), \\ x^{(3)} &= (0,1,1), \qquad & y^{(3)} &= (0,0,1,1). \end{aligned} \tag{2.32}$$

Bitstrings are a bookkeeping device for partitioning the vertex set $[N]$ of K_N into specific disjoint subsets suitable for studying a given candidate inequality. In particular, consider the minimal min-cuts W_{I_l} for $l \in [L]$ associated to every term on the left-hand side of (2.27). Each of the 2^L different bitstrings $x \in Q_L$ indexes a disjoint vertex subset $W(x) \subseteq [N]$ defined by

$$W(x) = \bigcap_{l=1}^{L} W_{I_l}^{x_l}, \qquad W^b = \begin{cases} W & \text{if } b = 1, \\ W^{\complement} & \text{if } b = 0. \end{cases} \tag{2.33}$$

The attentive reader will notice that these $W(x)$ sets would be precisely the atoms introduced in (2.12) when proving Corollary 2.1 if one were to iterate the intersection over all possible $2^n - 1$ non-empty subsets of $[n]$. In the current discussion, one need only consider the pertinent L subsets $I_l \subseteq [n]$ involved in the left-hand side of (2.27). The resulting $W(x)$ sets are again disjoint by construction, i.e. $W(x) \cap W(x') = \varnothing$ unless $x = x'$. The converse is certainly not true though, as expected from Corollary 2.1. In particular, $W(x)$ will be empty whenever the family of sets $\mathcal{I}(x) = \{I_l \subseteq [n] \ : \ x_l = 1\}$ is not a pairwise intersecting upper set in $\mathcal{I}_L = \{I_l \subseteq [n] \ : \ l \in [L]\}$. The relation between this statement and the one in Corollary 2.1 that refers to upper sets in $Bool_n$ is better understood in terms of their associated antichains. Namely, $W(x)$ will be empty whenever the minimal elements in $\mathcal{I}(x)$ are not a pairwise intersecting antichain in $\mathcal{I}_L$,

which holds if and only if the same is true in $Bool_n$. In other words, there is a non-trivial $W(x)$ precisely for every pairwise intersecting antichain in $Bool_n$ that is also so in $\mathcal{I}_L$, and thus the number of relevant bitstrings will be considerably smaller than $M(n)$.

The discussion above motivates introducing the subset $A_n(\mathcal{I}_L) \subseteq Q_{2^n-1}$ of all bitstrings $x \in A_n(\mathcal{I}_L)$ such that $\mathcal{I}(x)$ is a pairwise intersecting antichain in $\mathcal{I}_L$. Crucially, these suffice to characterize the minimal min-cuts for $I_L \in \mathcal{I}_L$ in any K_N, in the sense that these are all reconstructible via

$$W_{I_l} = \bigcup_{x:x_l=1} W(x), \tag{2.34}$$

where the union here, and in all that follows next, runs over all bitstrings $x \in A_n(\mathcal{I}_L)$ subject to the given conditions. Furthermore, one can use these vertex sets $W(x)$ to construct (not necessarily minimum) cuts for any subset of terminals $J \subseteq \bigcup_{l=1}^{L} I_l \subseteq [n]$. To see this, let $i \in [n]$ be any one of the terminals involved in the subsets $I_l \in \mathcal{I}_L$. We want to find which of the $W(x)$ sets the vertex i lands on. Since $i \in W_{I_l}$ if and only if $i \in I_l$ by the definition of a cut for I, it follows that $i \in W(x)$ if and only if $x_l = 1$ precisely when $I_l \ni i$ and $x_l = 0$ otherwise. In other words, the bitstring we are after is precisely the occurrence vector $x^{(i)} \in Q_L$ defined in (2.31), and we thus have $i \in W(x^{(i)})$.

Lemma 2.1. *Let $f : Q_m \to \{0,1\}$ be an m-ary Boolean function. Given some collection of terminal subsets $\mathcal{I}_L \ni I_l$ and the partitioning of $[N]$ defined in (2.33), construct the vertex set*

$$U^f = \bigcup_{x:f(x)=1} W(x). \tag{2.35}$$

Then, for any subset of terminals $J \subseteq \bigcup_{l=1}^{L} I_l \subseteq [n]$*, we have*

$$U^f \cap [n] = J \qquad \Longleftrightarrow \qquad f(x^{(i)}) = \delta(i \in J), \quad \forall\ i \in [n]. \tag{2.36}$$

Proof. A trivial rephrasing of $U^f \cap [n] = J$ is that, for $i \in [n]$, one has $i \in J$ if and only if $i \in U^f$. Since $i \in W(x^{(i)})$ for every $i \in [n]$ and all $W(x)$ are disjoint, it follows that $i \in U^f$ if and only if $W(x^{(i)}) \subseteq U^f$. Finally, since by construction $W(x^{(i)}) \subseteq U^f$ if and only if $f(x^{(i)}) = 1$, the desired result is obtained. □

The min-cut edges $C(W_I)$ can also be conveniently organized in terms of bitstrings via

$$E(x, x') = \{(i, j) \in E_N \ :\ i \in W(x) \text{ and } j \in W(x')\}. \tag{2.37}$$

Because the $W(x)$ vertex sets are disjoint, so are the $E(x, x')$ edge sets for any distinct pair of bitstrings $x, x' \in Q_L$. This leads to the following useful result for $C(W_I)$:

Lemma 2.2. *The edges of a min-cut* W_{I_l} *for some* $I_l \in \mathcal{I}_L$ *and their total weight are, respectively,*

$$C(W_{I_l}) = \bigcup_{x,x':x_l \neq x'_l} E(x, x'), \qquad \|C(W_{I_l})\| = \sum_{x,x'} |x_l - x'_l||E(x, x')|, \tag{2.38}$$

where the index sets are unordered pairs of bitstrings $x, x' \in A_n(\mathcal{I}_L)$*.*

Proof. By definition, an edge $(i, j) \in C(W_{I_l})$ if and only if $i \in W_{I_l}$ and $j \in W_{I_l}^{\complement}$. Using (2.34), one can write $W_{I_l} = \bigcup_{x:x_l=1} W(x)$ and, similarly, $W_{I_l}^{\complement} = \bigcup_{x:x_l=0} W(x)$. Hence $(i, j) \in E(x, x')$ is contained in $C(W_{I_l})$ if and only if x and x' differ in their l^{th} bit $x_l \neq x'_l$. It follows that $C(W_{I_l})$ can be constructed by joining all edge sets $E(x, x')$ with bitstrings $x, x' \in A_n(\mathcal{I}_L)$ such that $x_l \neq x'_l$, thereby proving the first equation in (2.38).

Furthermore, since all $E(x, x')$ are disjoint for distinct pairs of bitstrings, the total weight of their union reduces to the sum over the total weights of every $E(x, x')$ involved, which is precisely what the second equation computes. □

Given two metric spaces (M, d) and (M', d'), we call $f : M \to M'$ a *d-d' contraction map* if

$$d'(f(x), f(y)) \;\leq\; d(x, y), \qquad \forall\; x, y \in M. \tag{2.39}$$

The general proof method can now be stated:

Theorem 2.4. *Inequality (2.27) is valid for H_n if there exists a d_α-d_β contraction map*

$$f : A_n(\mathcal{I}_L) \to Q_R, \tag{2.40}$$

satisfying $f(x^{(i)}) = y^{(i)}$ for all $i \in [n; N]$.

Proof. Associate a cut U_{J_r} to each subsystem J_r that appears on the right-hand side of (2.27) by using the map f to pick which sets $W(x)$ to include in the definition of U_{J_r} as follows:

$$U_{J_r} = \bigcup_{x : f(x)_r = 1} W(x). \tag{2.41}$$

That this indeed obeys the cut condition $U_{J_r} \cap [n] = J_r$ is guaranteed by Lemma 2.1 and the fact that f is required to respect occurrence vectors, i.e. $f(x^{(i)}) = y^{(i)}$ for every $i \in [n; N]$.

$$\sum_{l=1}^{L} \alpha_l S_{I_l} = \sum_{l=1}^{L} \alpha_l \, \|C(W_{I_l})\| = \sum_{x,x'} |E(x, x')| \sum_{l=1}^{L} \alpha_l |x_l - x'_l| = \sum_{x,x'} |E(x, x')| \, d_\alpha(x, x'). \tag{2.42}$$

Similarly, for the U_{J_r} cuts one has

$$\sum_{r=1}^{R} \beta_r \|C(U_{J_r})\| = \sum_{x,x'} |E(x,x')|\, d_\beta(f(x), f(x')). \tag{2.43}$$

Therefore, by hypothesis, the contraction property of f implies

$$\sum_{l=1}^{L} \alpha_l \|C(W_{I_l})\| \;\geq\; \sum_{r=1}^{R} \beta_r \|C(U_{J_r})\|. \tag{2.44}$$

Because every set U_{J_r} is a cut for each J_r appearing on the right-hand side of (2.27), by minimality $\|C(U(J_r)\| \geq S(J_r)$ for every $r \in [R]$. Hence the right-hand side of (2.44) is no smaller than that of (2.27). Finally, since their respective left-hand sides are equal, validity of (2.27) follows. □

This theorem was proved in [42] (Theorem 8) in a somewhat weaker form. Whereas the domain of their contraction map is Q_L, enumerating all subsets of $\mathcal{I}_L \subseteq Bool_n$, we reduce this to $A_n(\mathcal{I}_L) \subseteq Q_L$, enumerating only the pairwise intersecting antichains in $\mathcal{I}_L$ (cf. Corollary 2.1). This leads to a reduction of the worst-case complexity of the search space.

As an example, a contraction map which proves validity of (2.28) is shown in Table 2.1.[7] One easily checks that occurrence vectors are respected, e.g. for $3 \in [n]$ we have $(0,1,1) \mapsto (0,0,1,1)$, which matches (2.32). Iterating through every pair of rows, one can also check that the contraction property holds. Here the vectors defining the distance function are $\alpha = (1,1,1)$ and $\beta = (1,1,1,1)$ for left and right, respectively. For instance, occurrence vectors 1 and 3 in the domain give $d_\alpha(x^{(1)}, x^{(3)}) = 1+0+1 = 2$, while their

[7] In fact, this contraction map which proves (2.28) is unique. This is generically not the case for larger-n facets, for which there usually exists many contraction maps compatible with the requirements of Theorem 2.4.

images $d_\beta(f(x^{(1)}), f(x^{(3)})) = 1 + 0 + 1 + 0 = 2 \leq d_\alpha(x^{(1)}, x^{(3)})$. Notice that the map f need not be injective nor surjective.

	S_{12}	S_{13}	S_{23}	S_1	S_2	S_3	S_{123}
0	0	0	0	0	0	0	0
	0	0	1	0	0	0	1
	0	1	0	0	0	0	1
3	0	1	1	0	0	1	1
	1	0	0	0	0	0	1
2	1	0	1	0	1	0	1
1	1	1	0	1	0	0	1
	1	1	1	0	0	0	1

Table 2.1: Representation of the contraction map which proves the MMI inequality (2.28). The left-most column labels the occurrence vectors shown in (2.32), including the one for $N \sim 0$. The top row labels bitstring entries, separating domain (left) from codomain (right). For the domain, S_{I_l} labels entries x_l, $l \in [L]$ for $x \in Q_L$ and, for the codomain, S_{J_r} labels entries y_r, $r \in [R]$ for $y \in Q_R$. Every row represents one entry of the map $f : x \mapsto y$ by listing all entries as $\{x, y\}$.

The proof of Theorem 2.4 is constructive and, as shown in [42], leads to an algorithm for finding a contraction map or showing none exists. The enumeration of all contraction maps is prohibitively expensive in all but very small cases. However, the authors developed a greedy technique for partial search which is successful in finding a map, when one exists. Indeed, we have found this method very powerful in proving new inequalities valid for H_6. For proving an inequality is invalid, the previously-mentioned ILP approach, which will be presented in Section 2.5, is also very effective.

Theorem 2.4 provides a robust sufficient condition for an inequality to be valid, but it is not a necessary one. For example, even after exhausting all of the possibilities given

by the theorem, it was not able to prove the validity of this inequality over H_5:

$$3S_{123} + 3S_{124} + S_{125} + S_{134} + 3S_{135} + S_{145} + S_{234} + S_{235} + S_{245} + S_{345} \geq$$
$$2S_{12} + 2S_{13} + S_{14} + S_{15} + S_{23} + 2S_{24} + 2S_{35} + S_{45} + 2S_{1234} + 2S_{1235} + S_{1245} + S_{1345}. \tag{2.45}$$

However, this inequality can be proved valid by expanding the codomain of f by replacing coefficients greater than one on the right-hand side by a sum of terms with unit coefficients. For instance, a term like $2S_I$ gets replaced by $S_I + S_I$, with the obvious generalization applied to larger coefficients. Theorem 2.4 still applies and this time the desired contraction map does exist, thereby proving validity of (2.45) By expanding the right-hand side there are more possible images for the contraction map, while the number of contraction conditions remains fixed. This may explain why this approach worked well here and in other cases we have tried for larger n.

This mild generalization of the proof technique of Theorem 2.4 has been remarkably successful in proving inequalities for $n = 6$, which motivates the following problem:

Problem 2.3. *In (2.27), if we replace terms $\beta_r S_{J_r}$ with $\beta_r \geq 2$ by $\sum_{i=1}^{\beta_r} S_{J_r}$ and accordingly adjust R to $\sum_{r=1}^{R} \beta_r$, does Theorem 2.4 provide a necessary condition for validity over H_n?*

2.4.2 Zero-lifting of valid inequalities and facets

Given an inequality $qS \geq 0$ over H_n, let $K \subseteq [n]$ be the subset of terminals appearing in it. Then consider a family of disjoint, non-empty subsets $\{I_i \subseteq [n+1]\}_{i \in K}$ (not necessarily spanning). The *zero-lifting* of the inequality given by this family is obtained by replacing each singleton $i \in I$ in every S_I in $qS \geq 0$ by its corresponding I_i. For example, the zero-lifting of $S_1 + S_2 \geq S_{12}$ from H_3 to H_4 corresponding to $I_1 = \{2, 3\}$

and $I_2 = \{1, 4\}$ yields $S_{23} + S_{14} \geq S_{1234}$. The zero-lift where $I_i = \{i\}$ for every $i \in K$ is called the *trivial zero-lift*.

Proposition 2.6. *If an inequality $qS \geq 0$ is valid for H_n, then any zero-lift $q'S' \geq 0$ is valid for H_{n+1}.*

Proof. Assume $qS \geq 0$ is valid for H_n. Proceed by contradiction by supposing q has a zero-lift q' such that $q'S' < 0$ for some $S' \in H_{n+1}$. Such S' must be realized by some weight map w applied to K_N for some N. In this K_N, contract each terminal set I_i to the vertex in I_i with the minimum label, combining parallel edges and summing their weights into a single edge, and deleting any loops. Let S be the realized S-vector in the new graph. We have $q'S' = qS \geq 0$, the desired contradiction. □

Before discussing lifting facets we need to recall some terminology from Proposition 2.1, in particular the weighted star graphs and the construction of matrix A^{n+1} in (2.7). We will make frequent use of the square matrix D^n of size $2^n - 1$, defined by

$$D^n_{I,J} = |I \cap J|, \qquad \varnothing \neq I \subseteq [n], \quad \varnothing \neq J \subseteq [n+1],\ n+1 \in J, \tag{2.46}$$

and use the notation D^n_j to refer to row J of D^n. An inequality $qS \geq 0$ in $\mathbb{R}^{2^n-1}$ is called *balanced* if $D_{\{j\}}q = 0$ for all $j \in [n]$. By definition, balance is invariant under any permutation of terminals $[n]$, but need not be so under permutations of the extended terminals $[n; N]$. For instance, $S_1 + S_2 \geq S_{12}$ is balanced but $S_1 + S_{12} \geq S_2$ is not. Balance is equivalent to a seemingly stronger condition:

Lemma 2.3. *An inequality $qS \geq 0$ in $\mathbb{R}^{2^n-1}$ is balanced if and only if $D^n q = 0$.*

Proof. Obviously, $D^n q = 0$ implies balance. For the converse, writing out row J of $D^n q$,

$$D^n_J q = \sum_{\varnothing \neq I \subseteq [n]} q_I |I \cap J| = \sum_{j=1}^{n} \delta^j_J \sum_{\varnothing \neq I \subseteq [n]} q_I \delta^j_I = \sum_{j=1}^{n} \delta^j_J \; D^n_{\{j\}} q, \tag{2.47}$$

where $\delta^i_I = \delta(i \in I)$ (cf. (2.31)) and we used $|I \cap J| = \sum_{k=1}^{n} \delta^k_I \delta^k_J$. So $D^n q = 0$ by balance. □

The next result relates the trivial zero-lifting of facets to the notion of balance.

Proposition 2.7. *If $qS \geq 0$ is a balanced facet of H_n, then its trivial zero-lift $q'S' \geq 0$ is a balanced facet of H_{n+1}.*

Proof. Suppose $qS \geq 0$ is a balanced facet of H_n. Then it is a valid inequality of H_{n+1} by Proposition 2.6. We adopt the notation of Proposition 2.1 and build a matrix A^{n+1} with the structure in (2.8), except it will now have only $2^{n+1} - 2$ rows. Let B^n consist of $2^n - 2$ linearly independent roots of $qS \geq 0$ as rows, so that the first $2^n - 2$ rows of A^{n+1} become precisely their zero-lifts. The trivial zero-lift has $q'_I = 0$ for every $I \ni n+1$, so these are all roots of $q'S' \geq 0$ as well. Since $qS \geq 0$ is balanced we have $D^n q = 0$ by Lemma 2.3. So the corresponding rows of A^{n+1} are roots of $q'S' \geq 0$ too. The final row is also and so A^{n+1} contains $2^{n+1} - 2$ roots of $q'S' \geq 0$. Performing the same column operations as in Proposition 2.1, the resulting block matrix (cf. $\tilde{A}^{n+1}$) shows that A^{n+1} has maximal rank $2^{n+1} - 2$. Since $D^{n+1}_{\{j\}} q' = D^n_{\{j\}} q = 0$, the lifted facet is also balanced. □

The following proposition clarifies the situation for subadditive inequalities, which include the non-balanced Araki-Lieb inequalities in their orbits:

Proposition 2.8. *For all $n \geq 2$, a zero-lift of a subadditive inequality (2.16) gives a facet if and only if, using the symmetry $S_{[n;N]\setminus I} = S_I$, it can be put in the singleton SA*

form

$$S_i + S_j \geq S_{ij}, \qquad i \neq j \in [n; N]. \tag{2.48}$$

Proof. Since singleton SA is a balanced facet of H_2, so is $S_i + S_j \geq S_{ij}$ for H_n by Proposition 2.7, as it can be obtained by iterating trivial zero-lifts and making a Sym_n permutation at the end.

For the converse, if a subadditive inequality is not in the form (2.48), we may write it as $S_I + S_{JK} \geq S_{IJK}$, for non-empty subsets I, J and K. This inequality is the sum of three valid inequalities for H_n: the general SA inequality (2.16), the general MMI inequality (2.29) and $S_I + S_K \geq S_{IK}$. Therefore, it is not a facet of H_n. □

Apart from nonnegativity and the Araki-Lieb inequality associated to (2.48), all known facets of H_n are balanced. While balance is sufficient for trivial zero-lifts to preserve facets, a stronger condition is needed for general zero-lifts. A balanced inequality $qS \geq 0$ is *superbalanced* if every inequality in its symmetry orbit under Sym_{n+1} permutations of $[n; N]$ is balanced [68, 66]. Since balance is invariant under permutations of $[n]$, it is in fact only necessary to check if exchanges of every $i \in [n]$ with N yield balanced inequalities. Orbits of superbalanced inequalities are referred to as superbalanced. For example, SA in (2.16) for $I, J \subseteq [n]$ is balanced but not superbalanced and MMI in (2.28) is superbalanced. According to results stated in [66], besides the singleton SA orbit, every orbit of facets of H_n for $n \geq 2$ is superbalanced.

We can now generalize Proposition 2.7 to arbitrary zero-lifts. For an inequality $qS \geq 0$ in $\mathbb{R}^{2^n-1}$, let

$$\tilde{q}_I = \sum_{I \subseteq J \subseteq [n]} q_J. \tag{2.49}$$

Lemma 2.4. *An inequality $qS \geq 0$ in $\mathbb{R}^{2^n-1}$ is superbalanced if and only if $\tilde{q}_I = 0$ for*

every $I \subseteq [n]$ *with* $|I| \leq 2$.

Proof. That $\tilde{q}_{\{i\}} = 0$ for all $i \in [n]$ is just the definition of balance, which is an invariant property under permutations of $[n]$. Permutations of $[n; N]$ also allow for reflections $j \leftrightarrow N$ for each $j \in [n]$. Using $S_{[n;N]\setminus K} = S_K$, the S-vector entries S'_I after reflection are related to the S_I before reflection by $S'_J = S_J$ and $S'_{J\cup\{j\}} = S_{[n]\setminus J}$ for $J \not\ni j$. For example, $1 \leftrightarrow N$ for $n = 3$ gives $S' = (S_{123}, S_2, S_3, S_{13}, S_{12}, S_{23}, S_1)$. The coefficients q_I in $qS \geq 0$ behave accordingly. Under a $j \leftrightarrow N$ reflection, (2.49) gives $\tilde{q}'_{\{j\}} = \tilde{q}_{\{j\}}$, while for $i \neq j$ one gets

$$\tilde{q}'_{\{i\}} = \sum_{i\in J\subseteq[n]\setminus\{j\}} (q'_{J\cup\{j\}} + q'_J) = \sum_{i\in J\subseteq[n]\setminus\{j\}} (q_{[n]\setminus J} + q_J) = \sum_{j\in J\subseteq[n]\setminus\{i\}} q_J + \sum_{i\in J\subseteq[n]\setminus\{j\}} q_J. \quad (2.50)$$

The first sum is over all q_J such that $J \ni j$ but $J \not\ni i$, so it differs from $\tilde{q}_j$ precisely by $\tilde{q}_{\{i,j\}}$. Similarly for the second sum exchanging $i \leftrightarrow j$, so

$$\tilde{q}'_{\{i\}} = \tilde{q}_{\{i\}} + \tilde{q}_{\{j\}} - 2\tilde{q}_{\{i,j\}}. \quad (2.51)$$

After the exchange $j \leftrightarrow N$, $q'S \geq 0$ is balanced if and only if $\tilde{q}'_{\{i\}} = 0$ for all $i \in [n]$. Therefore $qS \geq 0$ is superbalanced if and only if $\tilde{q}_{\{i\}} = \tilde{q}'_{\{i\}} = 0$ for all $i \in [n]$. Applied to (2.51), this means $qS \geq 0$ is superbalanced if and only if $\tilde{q}_{\{i\}} = \tilde{q}_{\{i,j\}} = 0$ for all $i, j \in [n]$. □

We now show that any zero-lift of a superbalanced facet can actually be built solely out of trivial zero-lifts combined with permutations of the extended terminals, both of which preserve facets. Superbalance is needed for such permutations to preserve balance and thus keep Proposition 2.7 applicable. It is also important in what follows that, as is clear from Lemma 2.4 and the form of (2.49), balance and superbalance are properties

which are shared by inequalities related by trivial zero-lifts. We first observe that the trivial zero-lift from H_n to H_{n+1} can be thought of as treating the new terminal $n+1$ as a duplication of the sink (since the sink does not appear anywhere in $qS \geq 0$, neither does $n+1$ in $q'S' \geq 0$). But by the symmetry of H_n under Sym_{n+1} permutations of $[n; N]$, we could analogously consider letting $n+1$ duplicate any other terminal. Let us call such a generalization of a trivial zero-lift where any one extended terminal becomes a doubleton and the rest remain singletons a *simple zero-lift*.

Theorem 2.5. *If $qS \geq 0$ is a superbalanced facet of H_n, then any zero-lift $q'S' \geq 0$ is a superbalanced facet of H_{n+1}.*

Proof. If $qS \geq 0$ is a facet inequality over H_n involving a subset of terminals K with $|K| < n$, then put it in canonical form as an inequality over $H_{|K|}$. Iterating Proposition 2.5, note that $H_{|K|}$ is a projection of H_n. Since q is in canonical form, its coefficients are not changed in projecting it to $H_{|K|}$. A standard result of polyhedral theory is that facets project to facets, so $qS \geq 0$ is a superbalanced facet of $H_{|K|}$.

Starting from $qS \geq 0$, one can get to $q'S' \geq 0$ as follows. If the original zero-lift had $\{i\} \mapsto I_i$, perform $|I_i| - 1$ simple zero-lifts appending terminals to i, and repeat for every $i \in [n]$. If less than $n + 1 - |J|$ steps were required, reach all the way to H_{n+1} via trivial zero-lifts. At that point, a suitable permutation of $[n+1]$ yields $q'S' \geq 0$.

We now show that every simple zero-lift used above can in fact be built solely out of permutations and trivial zero-lifts. In particular, the simple zero-lift involving $I_i = \{i, j\}$ is equivalently accomplished by exchanging $i \leftrightarrow N$, performing a trivial zero-lift, and then exchanging the new sink back with i. If the original inequality is superbalanced, the trivial zero-lift in this process is applied to a balanced inequality. Using Proposition 2.7, one ends up with a facet if one started with a facet. Furthermore, the latter is superbalanced

if the former is. Hence one can go from $qS \geq 0$ to $q'S' \geq 0$ via superbalance- and facet-preserving steps. □

2.5 Integer programs for testing realizability and validity

A direct test of the realizability of an S-vector in K_N can be performed by a feasibility test of a mixed integer linear program (ILP). Similarly, an inequality $qS \geq 0$ can be tested to see if it is valid for all S-vectors that can be realized in K_N. As noted earlier, the polyhedral approach described so far does not force the minimum in (2.3) to be realized by one of the inequalities (2.18a). However, using binary variables this can be achieved and the feasibility of the resulting system tested using ILP solvers such as CPLEX, glpsol or Gurobi.

For any $N > n \geq 3$, we build a set of constraints, $\text{ILP}_{N,n}$, whose feasible solution is the set of all suitably-normalized, valid (S, w) pairs on n terminals realizable in K_N. Firstly, note that for each $\varnothing \neq I \subseteq [n]$, the number of cuts W in K_N that contain I is $2^{N-|I|-1}$. For each such W and I, we introduce a binary variable $y_{W,I}$. Specifically, we consider the following system:

$\text{ILP}_{\mathbf{N,n}}$

For all $\varnothing \neq I \subseteq [n]$ and cuts $W \subseteq [N-1]$ in K_N such that $I = W \cap [n]$,

$$S_I \leq \|C(W)\|, \tag{2.52}$$

$$\|C(W)\| \leq S_I + |W|\,(N-|W|)\,y_{W,I}, \tag{2.53}$$

$$\sum_{W\cap[n]=I} y_{W,I} = 2^{N-|I|-1} - 1, \tag{2.54}$$

$$y_{W,I} \in \{0,1\}, \tag{2.55}$$

$$0 \leq w(e) \leq 1, \qquad \forall\, e \in E_N. \tag{2.56}$$

Proposition 2.9. *A pair (S,w) is valid in K_N with all edge weights at most one if and only if there exists assignments to variables y so that $\{S,w,y\}$ is a feasible solution to* $\mathrm{ILP}_{N,n}$.

Proof. Suppose that (S,w) is a valid pair in K_N with all edge weights at most one. We will show that y variables can be chosen so that $\{S,w,y\}$ constitutes a feasible solution to $\mathrm{ILP}_{N,n}$. Firstly, by assumption w satisfies (2.56). Next, since (S,w) is a realization in K_N, the upper bounds in (2.52) are valid. For each $\varnothing \neq I \subseteq [n]$, choose one $W_I \subseteq [N-1]$ so that W_I realizes a minimum in (2.3), and set $y_{W_I,I} = 0$. All other y variables for this I are set to 1, thus satisfying (2.54) and (2.55). Since $y_{W_I,I} = 0$, the corresponding equation (2.53) gets zero as the second term in its right-hand side, and thus combines with (2.52) into the required equation. The remaining inequalities to verify are those in (2.53) when $y_{W,I} = 1$. Their validity follows from the fact that the cut W in K_N contains $|W|(N-|W|)$ edges, each of weight at most one.

Conversely, let $\{S,w,y\}$ be a feasible solution of $\mathrm{ILP}_{N,n}$. For each $\varnothing \neq I \subseteq [n]$, (2.54) implies that there is a single variable, which we label $y_{W_I,I}$, having value zero. The other y values for this I are one. Together with (2.52), this implies that $S_I = \|C(W_I)\|$ and

that S_I satisfies (2.3). So (S, w) is a valid pair realized in K_N with all edge weights at most one. □

We make use of this ILP formulation in two ways. Firstly, it can be used to test whether or not an S-vector is realizable in K_N for a given N. To do this, we pre-assign the values from the given S-vector to the corresponding S variables in $\text{ILP}_{N,n}$, rescaled to values smaller than 1. We may then run an ILP solver to test whether there is a feasible solution. If so, the values or the variables w will give a realization in K_N. Otherwise, one concludes that the given S-vector cannot be represented in any $K_{N'}$ with $N' \leq N$. Secondly, we may use the ILP to test whether an inequality $qS \geq 0$ is invalid for some S-vector realized in K_N for a given N. This can be done by minimizing $z = qS$ over $\text{ILP}_{N,n}$ and seeing if the optimum solution is negative. The computation can be terminated when the first feasible solution with $z < 0$ is found, at which point $qS \geq 0$ is proven invalid. We could prove that an inequality $qS \geq 0$ is valid over H_n by testing it with $\text{ILP}_{m(n),n}$, but this ILP would be very large with current bounds on $m(n)$.

In its first formulation, the ILP allows one to find the minimum value N_{min} of N for which an S-vector is realizable in K_N. Given an S-vector, we call any such $K_{N_{min}}$ a *minimum realization*. At fixed n, we define $m_{ext}(n)$ as the smallest integer such that all extreme rays of H_n, and hence of $H_{m_{ext}(n),n}$, are realizable in $K_{m_{ext}(n)}$ (cf. the definition of $m(n)$). For $1 \leq n \leq 5$, the ILP shows that $m_{ext}(n)$ is

$$2, \quad 3, \quad 5, \quad 6, \quad 11. \tag{2.57}$$

Combining all extreme-ray graphs into a larger one by identifying them all at $[n; N]$ (cf. conically combining S-vectors), one can also see that for $1 \leq n \leq 3$, $m(n)$ takes values 2, 3 and 5. Namely, no bulk vertices are needed for $n = 1, 2$, and just a single one comes

into play for $n = 3$.

The case $n = 4$ is less trivial. There are two star-graph orbits of 5 extreme rays each, see Figure A.1 in Appendix A.1.2. These are 10 extreme rays realizable in K_6, which contains a single bulk vertex. The other extreme rays of H_4 involve no bulk vertices. Hence, a convex combination of 15 extreme rays may require a total of 10 bulk vertices at most, which with the terminals and sink gives $m(4) \leq 15$. This can be further improved as follows. The Bell-pair extreme rays span a subspace of dimension 10, and the star-graph extreme rays are confined to its 5-dimensional orthogonal complement. Thus at most 5 star-graph extreme rays are needed to conically span any interior ray of H_4, improving the bound down to $m(4) \leq 10$. It turns out that the simplicity of the specific extreme-ray graphs for $n = 4$ in fact allows us to obtain the definite value $m(4) = 6$. The reason for this is that the S-vector of any conical combination of these particular extreme-ray star graphs of H_4 can itself also be realized on a star graph. This follows from the observation that all $n = 4$ extreme-ray star graphs have identical minimal min-cuts: for every $\varnothing \neq I \subseteq [n]$, they all have $W_I = I$ for $|I| = 1, 2$ and $W_I = I \cup \{n\}$ for $|I| = 3, 4$. Pictorially, this allows one to stack them all on top of each other, adding up their edge weights, so as to realize any combination of these star graphs by a star graph.

This discussion illustrates some strategies for obtaining tighter upper bounds on $m(n)$ based on knowledge of extreme rays or the value of $m_{ext}(n)$. Recall that the number of bulk vertices in K_N is $N - n - 1$. Regardless of how many extreme rays H_n has, any interior ray may be a conical combination of at most $2^n - 1$ of them. Since we can realize all extreme rays in $K_{m_{ext}(n)}$, we have

$$m(n) \leq (m_{ext}(n) - n - 1) \times (2^n - 1) + n + 1. \tag{2.58}$$

For instance, since we know $m_{ext}(5) = 11$, this gives $m(5) \leq 161$, which is considerably better than the bound $m(5) < M(5) = 2546$ given earlier.

We can do even better for $n = 5$ by using explicit results about the dimensionality of the span of specific extreme-ray orbits. The Bell pairs take care of 15 dimensions which are not reached by any other extreme ray without introducing any bulk vertices. There is a single orbit that requires $N = 11$, and it consists of 75 extreme rays spanning a subspace of dimension 10 of the remaining 16 of $H_5 \subset \mathbb{R}^{31}$. The largest-$N$ orbit spanning the other 6 dimensions has $N = 8$ and 360 extreme rays. Hence the worst-case scenario would require 10 graphs with $N = 11$ and other 6 with $N = 8$. The total number of vertices carried by a combination of such graphs thus gives the bound $m(5) \leq 74$. This is better than the more general one attained by (2.58), but requires complete knowledge of all extreme-ray graphs, not just of the number $m_{ext}(n)$.

Problem 2.4. *Find tighter bounds on $m_{ext}(n)$. In particular, does $\log_2 m_{ext}(n)$ admit an upper bound that is polynomial in n?*

2.6 Computing H- and V-representations of H_n

This section provides two methods for computing complete descriptions of H_n. We illustrate them for $n = 5$ and describe H_5 in detail in Appendix A.1. A partial description of H_5 was first obtained by [42] and later completed by [13]. Whereas their searches were mostly random/heuristic and did not give minimum realizations of all extreme rays, here we obtain the same results via systematic approaches. The first method is a general formalization of the strategy used earlier for $n = 3, 4$, whereas the second one constructs H_n starting from knowledge of H_{n-1}.

To initialize the first one, set $k = 2$.

Method 1

(a) Generate the H-representation `P{n+k}-n.ine` of $P_{n+k,n}$ using (2.18). Convert this to a V-representation `P{n+k}-n.ext`.

(b) Delete the $2^n - 1$ trivial extreme rays (see Theorem 2.2) and extract the $2^n - 1$ coordinates corresponding to the variables of the S-vectors. Remove redundant rays to obtain the V-representation `H{n+k}-n.ext` of $H_{n+k,n}$. This is an inner approximation of H_n.

(c) Compute the H-representation `H{n+k}-n.ine` of $H_{n+k,n}$ from `H{n+k}-n.ext`. Using the ILP method of Section 2.5 with $N \geq n+1$, reject facet orbits that are invalid for K_N, continuing until either a facet is rejected or N is too large for the ILP to solve.

(d) Test any remaining facet orbits for which the validity is unknown using the proof-by-contraction method of Section 2.4.1. Generate the full orbits of the facets proved valid, getting a cone `HV{n+k}-n.ine` which is an outer approximation of H_n.

(e) Compute the extreme rays `HV{n+k}-n.ext` of `HV{n+k}-n.ine`. The orbits of S-vectors that appeared in `P{n+k}-n.ext` give extreme rays of H_n by Theorem 2.2(b). The remaining orbits can be checked by the ILP method of Section 2.5 with $N \geq n+1$ until finding a realization or N being too large for the ILP to solve. If all extreme-ray orbits can be realized, then `HV{n+k}-n.ine` is an H-representation of H_n and `HV{n+k}-n.ext` is its V-representation by Proposition 2.2(c). Otherwise, increment k and return to step (a).

Applying this to $n = 5$, one finds that $P_{7,5}$ has 83 facets and 194 extreme rays in 52 dimensions. The resulting $H_{7,5}$ has 142 extreme rays in 31 dimensions, and its H-representation consists of 8952 facets in 30 orbits. All but 8 of them are easily eliminated in step (c) and then proved valid in step (d). These orbits give 372 facets which define `HV7-5.ine`. Correspondingly, `HV7-5.ext` has 2267 extreme rays falling into 19 orbits. All of the extreme rays are realizable for $N \leq 11$, so the procedure terminates after a single iteration. Note that we obtain a minimum realization of each extreme ray either in step (a) or (e), wherever it appears first.

The vertex/facet enumeration problems in steps (a), (c) and (e) utilized the code Normaliz[8] v.3.4.1 on mai20[9]. Steps (a) and (c) took only a few seconds, and step (e) took 23 minutes. Step (d) was run on a laptop[10] using a Mathematica v.12.1 implementation[11] of the proof-by-contraction method. Most runs were very fast, taking less than 4 seconds, and all finished in no more than 16 minutes. The ILP runs in step (e) were performed with CPLEX[12] v.12.6.3, also on mai20, and normally completed in under 1 minute, the longest run taking 18 minutes. The filtration by symmetry generally takes just a few seconds.

The second method is more sophisticated and involves working with both outer and inner approximations of H_n, refining them until they are equal. The outer approximation is initialized by choosing any set of valid inequalities for H_n, not necessarily facets, whose intersection is full dimensional. The inner approximation is initialized by choosing any feasible set of rays, not necessarily extreme, whose convex hull is also full dimensional. A strong way to initialize the outer approximation is to zero-lift the superbalanced facets of

[8] https://www.normaliz.uni-osnabrueck.de
[9] mai20: 2× Xeon E5-2690 (10-core 3.0GHz), 20 cores, 128GB memory.
[10] Dell XPS 15 7590, i7-9750H CPU @ 2.60GHz, 6 cores, 12 threads, 32GB memory.
[11] Available upon request.
[12] https://www.ibm.com/analytics/cplex-optimizer

H_{n-1} in all possible ways, add to them singleton SA, and generate their full orbits under Sym_{n+1}. By Theorem 2.5 and Proposition 2.8, these are all facets of H_n and define a cone `OH1-n.ine`. For a strong inner approximation, we zero-lift the extreme rays of H_{n-1} and generate their full orbits under Sym_{n+1}, which are all extremal in H_n by Proposition 2.5. Since this is not always full-dimensional, we add the full orbits of the S-vectors from Proposition 2.1 not already included, and remove redundancies. In general, this may only add the single orbit of size $n+1$ generated by $S^{[n]}$. The resulting cone `IH1-n.ext` is an inner approximation of H_n. Set the iteration counter $k = 1$.

Method 2

Outer	Inner
Compute the V-representation `OHk-n.ext` of `OHk-n.ine`. Check one extreme ray from each orbit to see if it is realizable by the ILP method of Section 2.5. If all rays are realizable, then **exit**. The full orbits of the realizable extreme rays define `OHVk-n.ext`.	Compute the H-representation `IHk-n.ine` of `IHk-n.ext`. Apply to it steps (c) and (d) of Method 1, retaining inequalities proved by the contraction method of Section 2.4.1. If all inequalities are valid, then **exit**. The full orbits of the valid facets define `IHVk-n.ine`.
↓	↓
Merge `IHVk-n.ine` (and any other known valid inequalities) with `OHk-n.ine` and remove redundancies to get `OH{k+1}-n.ine`.	Merge `OHVk-n.ext` (and any other known realizable rays) with `IHk-n.ext` and remove redundancies to get `IH{k+1}-n.ext`.
↓	↓
Increment k and return to the first step of each respective subroutine.	

Note that the inner and outer procedures can be run in parallel. After they both finish the first step, the newly computed data are exchanged, improving both the outer an inner approximations. If **exit** occurs, the corresponding `ine` and `ext` descriptions give H- and V-representations of H_n, respectively. In each subroutine, the second step allows for the incorporation of valid inequalities and/or rays obtained by other means, such as Method 1.

Upon initialization, the starting cones `OH1-5.ine` and `IH1-5.ext` respectively consist of 80 facets in 3 orbits (that of singleton SA and 2 of MMI, cf. Appendix A.1.1), and 66 extreme rays in 4 orbits (cf. Figure A.1 in Appendix A.1.2, and the $J = [n]$ star orbit).

We start with $k = 1$ and describe steps in parallel. In the outer run, `OH1-5.ext` has 3205 extreme rays in 29 orbits, out of which 16 can be shown to be realizable with $N \leq 11$. Their orbits yield 1457 feasible rays defining `OHV1-5.ext`. In the inner run, `IH1-5.ine` has 157153 facets in 346 orbits, out of which one can show 8 are valid and easily reject the rest. Their orbits yield 372 valid inequalities defining `IHV1-5.ine`. In the second step it turns out that the outputs of the first step dominate in both cases. So after the merges, `OH2-5.ine` equals `IHV1-5.ine` and `IH2-5.ext` equals `OHV1-5.ext`.

Setting $k = 2$, in the outer run `OH2-5.ext` has 2267 extreme rays in 19 orbits, all of which are realizable with $N \leq 11$. Hence **exit** is triggered, and the algorithm terminates returning `OH2-5` as the result for H_5. If we continue the inner run we find that `IH2-5.ine` has 1182 facets in 11 orbits, out of which one can show 8 are valid and easily reject the rest. These are the same 8 orbits as before and so the algorithm exits in the first outer step with $k = 3$.

Conversions between cone representations again require vertex/facet enumeration. Those in the first iteration are immediate. Using Normaliz on mai20, the computations of `IH2-5.ine` and `OH3-5.ext` took about 75 seconds and 25 minutes, respectively. The

cost of other computations was similar to that of their counterparts in Method 1.

Although the inner steps of Method 2 may appear similar to Method 1, they are in fact quite distinct. In the latter, the starting cone `H{n+2}-n.ext` only contains S-vectors realizable in K_{n+2}. Many of these will be non-extremal in H_n and therefore absent from `IH1-n.ext`. Among those which are extremal, some may not be obtainable by zero-lift and thus not included in `IH1-n.ext` either. On the other hand, `IH1-n.ext` contains all extreme rays of H_n coming from zero-lifts. These will generally include plenty which are not realizable in K_{n+2} and hence not be contained in `H{n+2}-n.ext`. For example, for $n = 6$, `IH1-6.ext` includes zero-lifts of extreme rays in the 5 orbits of H_5 which are realizable in K_N with $N \geq 9$ (see Table A.2 in Appendix A.1.2), none of which can possibly be in `H8-6.ext`.

Both methods may run into fundamental and/or practical issues. For $n = 5$, one is fortunate that the contraction method successfully proves valid the 8 facet orbits of H_5. However, it remains a logical possibility that for larger n this proof method is not a necessary condition for validity of some facets of H_n (cf. Problem 2.3 at the end of Section 2.4.1). Specifically in Method 1, it so happens that all rays in `HV7-5.ext` are realizable using the ILP of Section 2.5. For larger n, in practice it could be that even if all rays at step (e) were realizable, the value of N required could be too high for the ILP to be solved. Without good bounds on $m_{ext}(n)$, this possibility cannot be easily eliminated. Alternatively, it could be that some rays are indeed not realizable, meaning that the facet description in `HV{n+k}-n.ine` is incomplete. This would require incrementing k and at least one further iteration. As for Method 2, we unfortunately have no proof of convergence using the strong starting inputs suggested without the option to generate and add additional valid inequalities and/or feasible rays in the second step. There are various heuristic methods available to generate such additional inputs. Another compli-

cation that affects these methods is the need to solve large convex hull/facet enumeration problems. All of these issues arise in one form or another in both methods already in the study of H_6.

The successful termination of either method relies on the finding of an H/V-representation of an inner/outer approximation of H_n containing all of its facets/extreme rays. For instance, observe that in Method 1 all facets of H_5 were already discovered in step (a) and computed explicitly in step (c) (along with other non-valid inequalities) from $H_{N,5}$ for just $N = 7$. Similarly, Method 2 converged more easily through an inner approximation `IH1-5` whose H-representation also contained all facets of H_5. That H_n is easier to obtain from an H-representation of an inner approximation is no accident. This is because smaller N for K_N is needed to span all facets than to realize all extreme rays of H_n. This motivates the definition of $m_{ine}(n)$ as the smallest integer such that the H-representation of $H_{m_{ine}(n),n}$ contains all facets of H_n.

It is easily seen that $m_{ine}(n) = m_{ext}(n)$ for $1 \leq n \leq 4$ and that $m_{ine}(n) \leq m_{ext}(n)$ for larger n. For $n = 5$, the cone $H_{6,5}$ turns out to miss some facets of H_5, but $H_{7,5}$ does contain them all as we have seen in Method 1. This shows that $m_{ine}(5) = 7$, contrasting with the extreme rays, which have $m_{ext}(5) = 11$. More generally, when Method 1 terminates, we have $m_{ine}(n) = n + k$ and a minimum realization of each extreme ray, from which one also obtains $m_{ext}(n)$. This makes the importance of $m_{ine}(n)$ manifest and motivates the following problem:

Problem 2.5. *Find tighter bounds on* $m_{ine}(n)$. *In particular, does* $\log_2 m_{ine}(n)$ *admit an upper bound that is polynomial in* n*?*

2.7 Conclusion

Many of the important questions about the HEC remain open. As stated formally throughout the chapter, these include obtaining an explicit description of either the H- or V-representation of H_n, and finding the complexity of testing feasibility of rays and validity of inequalities. The current bounds on the size of the complete graph that can realize all extreme rays of H_n seem far from being tight, at least according to the limited experimental results that we have. Similarly, our findings suggest that much smaller graphs may be sufficient to span all facets of H_n, which strongly motivates understanding better the relative complexity of the H- and V-representations of the HEC. In this work, we have laid the foundations for further exploration of these key questions. Additionally, we have provided sharp computational tools which allowed us to completely describe H_5 after just a few hours of computation.

Chapter 3

The holographic entropy cone from marginal independence

3.1 Introduction

In recent years there has been significant progress in understanding how in the gauge/gravity duality [16, 75, 17] the semiclassical bulk physics is encoded in the boundary theory. Drawing on [76, 77], it was shown [78, 79] that if one has access to a subsystem of the boundary, one can then reconstruct local operators inside a bulk region known as the entanglement wedge [80]. However, this reconstruction of bulk operators assumes the knowledge of the classical geometric background. To obtain a more complete understanding of the holographic encoding of the bulk physics, one would like to characterise which states of the boundary field theory are dual to classical geometries, and for all such cases, decode the bulk metric directly from the boundary data.

Intuition from the scale/radius duality suggests that the deeper in the bulk we wish to see, the more non-local (in a suitable sense) the CFT probe we need. One particularly

convenient class of boundary observables are built from lightlike objects. Perhaps the simplest and most accessible ones are the 'bulk-cone singularities' [81] whereby endpoints of null geodesics through the bulk (which can be used to reconstruct the conformal metric in the bulk region probed by such geodesics, cf. e.g. [82, 6]) are visible directly in the singularities of boundary correlators. This can be generalized to the so-called 'bulk-point singularities' of Landau diagrams corresponding to n-particle scattering [83], used for the construction of the 'light-cone cuts' [45, 46, 10], which offer a much more elegant extraction from much more complicated and harder-to-access CFT data. However to reach even deeper, into causally inaccessible regions, one needs a CFT probe implemented by a spacelike construct in the bulk. One particularly natural such geometrical object is a codimension-2 extremal surface.[1] Knowing the proper areas for a family of such surfaces can be inverted to extract the bulk metric (including the conformal factor), again within the bulk regions reached by such surfaces.[2]

Having given a bulk motivation for codimension-2 extremal surfaces as providing a particularly natural and deep probe of the bulk geometry, it is intriguing to note that even from the CFT viewpoint, such surfaces arise very naturally, in the context of holographic entanglement entropy. Here a key role was played by the celebrated RT/HRT formula [32, 36] (collectively shortened to HRRT), which computes the von Neumann entropy of boundary subsystems in terms of the area of certain bulk surfaces. This "geometrisation" of correlations [39], which more recently has also been observed for other information quantities (such as for example [87, 88]), seems to indicate that from the boundary point

[1] From the purely geometrical level, the motivation is that higher-dimensional surfaces probe deeper (when comparing amongst various-dimensional bulk extremal surfaces anchored within a fixed-radius region of the boundary of a given asymptotically-AdS spacetime) [84].

[2] Early proof of principle was carried out in e.g. [85] and an argument for uniqueness of the recovered metric in 4-dimensions was given in [86], which also reviews further approaches to bulk metric reconstruction.

of view, a possible characterisation of bulk *geometric states* could be formulated in terms of certain features of their entanglement structure. Motivated by this, we would like to use the HRRT formula to extract as much information as possible about the entanglement structure of geometric states in holography.

In general, one way to investigate the entanglement structure of a given class of states utilizes an analysis of constraints, which typically take the form of inequalities satisfied by certain information quantities. Since we are interested in the implications of the HRRT formula, here we will focus on inequalities satisfied by linear combinations of von Neumann entropies of various boundary subsystems.

Two important examples of these classes of inequalities are *subadditivity* (SA) and *strong subadditivity* (SSA), the saturation of either of which has a clear implication for the entanglement structure of a given density matrix. The saturation of SA, or equivalently the vanishing of the mutual information, is associated with the absence of any form of correlation between a specified pair of subsystems, and the corresponding factorisation of the density matrix. As we will see, this fact will play a central role in our construction. The saturation of SSA on the other hand is associated with a particular entanglement structure, commonly known as quantum Markov chain, which is central for the theory of quantum error correction and recovery maps [89].[3,4]

It is important to notice that while the saturation of SA and SSA is associated with specific entanglement structures, the inequalities themselves do not characterise any particular structure because they are satisfied by all quantum states. For restricted classes of states, however, the von Neumann entropies of various subsystems might satisfy addi-

[3]See [90] for a discussion of this property in QFT.

[4]The fact that saturation of SSA will not play a fundamental role in our characterization of geometric states is related to the fact that by MMI (see below) such saturation can only be achieved when SA is also saturated (see [40] for more details).

tional inequalities. This can be easily seen for example for the case of classical probability distributions, where the von Neumann entropy reduces to the Shannon entropy, which in addition to the inequalities mentioned above also satisfies monotonicity.

The fact that in holography the entanglement structure of geometric states is constrained, beyond that of arbitrary quantum states, became evident with the work of [40], which proved an inequality known as *monogamy of mutual information* (MMI).[5] It is then interesting to ask what are all other such inequalities, and a systematic search was initiated in [42]. This work introduced the notion of the *holographic entropy cone* (HEC) and proved that this cone is polyhedral for an arbitrary number of parties N, implying that for any N there exists only a finite number of non-redundant[6] inequalities which define the *facets*. It also derived a set of new inequalities for five parties, which was later proved to be the complete set in [13]. For more than five parties, it was shown in [91] that the holographic entropy cone is indeed contained in the quantum one [56] for any N, but the detailed structure of the HEC remains mostly unknown.[7]

One limitation of [42] is that, while it offered a tool that can be used to prove if a given inequality is valid (via so-called contraction maps), it did not provide a constructive way of deriving such candidate inequalities in the first place, or even determine if they correspond to facets at all.[8] A step forward in this direction was accomplished by [8], in the form of two systematic algorithms for the construction of the HEC for any

[5] It is straightforward to see that outside the holographic setting, this inequality can be violated, for example, by a 4-party GHZ state.

[6] An inequality is non-redundant if it is not implied by other inequalities.

[7] A new family of inequalities for every odd N was also found in [42], and argued to be non-redundant with respect to SA and SSA and among themselves. However it remains unclear whether they genuinely are all facets of the HEC for every N. Computational efforts to construct the complete HEC for $\mathsf{N} = 6$ are also ongoing. At the time of writing, more than 4122 orbits of extreme rays and 182 orbits of facets have been found.

[8] Indeed, the proof-by-contraction method is unable to ascertain that a given inequality is *not* valid. In other words, while the method provides a sufficient condition for an inequality to be valid, it remains an open question whether it is also necessary [8].

number of parties. These algorithms were devised so as to recursively converge towards tighter inequalities, terminating with the finding of all facets of the HEC as one of its outputs. However, even if one were able to somehow find explicit formulae for all the facet inequalities of the HEC, it would still remain totally obscure where they come from and what they mean.

A first attempt at circumventing this limitation was presented in [67, 68], which introduced the notion of *proto-entropies* and defined the *holographic entropy polyhedron*. Motivated by the physical requirement of cut-off independence, this work suggested focusing more on topological features of the relevant RT surfaces, specifically their connectivity, rather than on their actual areas. This ultimately translated the search for the inequalities into the search for certain generating configurations, called building blocks, making the problem more combinatorial in nature. One of the features of this approach is that any inequality found by this procedure is guaranteed by construction to be a facet of the polyhedron. Intuitively, one can think of this procedure as being akin to deriving the facets of a polyhedral cone by first finding its extreme rays. In the present chapter we continue in this direction. By combining some of the techniques of [42] based on graph models of holographic entanglement with the ideas of [67, 68] based on connectivity of entanglement wedges, we take a significant step towards the derivation of the HEC for an arbitrary number of parties. However in doing so, we will also slightly change our perspective on the problem.

As evident from these earlier works, one of the main problems one has to face while trying to derive the facets of the HEC is the complexity of the combinatorics, which is typically characterised by a doubly exponential scaling in the number N of parties involved. In addition, the explicit form of the inequalities might be highly dependent on N, which can make it extremely hard to find a convenient parametrisation of all

the inequalities. And even if one could circumvent these complications, it is far from clear how to interpret the inequalities, for example by searching for explicit structures of density matrices which are ruled out by them. Ultimately, if there exists a general lesson to be learned about the entanglement structure of geometric states, it might be very hard to understand what it is merely by looking at a very large number of complicated and seemingly unrelated expressions. For all these reasons we will not look for an explicit derivation of the inequalities. Instead, we will argue that the holographic entropy cone can be derived, at least in principle, from the solution to a much simpler problem.

For any density matrix on a given number of parties N, and an arbitrary purification of it, one can consider a pair of subsystems (possibly composite and possibly including the purifier) and compute their mutual information to determine whether these subsystems are correlated or not. One can then repeat this analysis for any pair of subsystems to determine what was called in [12] the *pattern of marginal independence* (PMI) of the density matrix. Any PMI, specifying which pairs of subsystems are marginally independent (i.e. have vanishing mutual information) and which are not, can be viewed as a linear subspace in a certain vector space called the *entropy space*. In particular, the PMI is the supporting subspace of a face of a cone, the *subadditivity cone* (SAC), built only from the instances of SA for that N.[9] With this structure in hand, one can conversely ask for which PMI does there exist a density matrix corresponding to it. This was dubbed in [12] the marginal independence problem, and for the restricted class of states corresponding to geometric states in holography, the *holographic marginal independence problem* (HMIP).

The main goal of this work is to argue that the HEC can be fully reconstructed from the solution to the HMIP, and that the solution to this problem amounts to establishing

[9]Said more explicitly, the supporting subspaces of the *facets* of the SAC are the hyperplanes of vanishing mutual information (i.e. the saturation of some instance of SA), and their intersections then form the supporting subspaces of the *faces* of the SAC.

which extreme rays of the SAC can be realized by geometric states. Importantly, as we will explain, if one wants to construct the HEC for a given number of parties N, it will not in general be sufficient to know the solution to the HMIP for the same number of parties. We will argue however that there always exists a finite $\mathsf{N}_{\max}(\mathsf{N})$ such that the N-party HEC can be constructed from the solution of the $\mathsf{N}_{\max}$-party HMIP.

We will not be able to provide a definite proof that this reconstruction is possible, but we will formulate and discuss certain conjectures on graph models realizing the extreme rays of the HEC which imply that this is the case, and provide evidence in their support. We stress that the focus of this work however is not on any specific algorithm for an explicit reconstruction of the HEC, but rather on the possibility of the reconstruction from *only* the seemingly limited information contained in the solution to the HMIP. A conclusive proof of such a possibility would in fact amount to the proof of a deep equivalence between the information contained in the set of *all* holographic entropy cones (for all values of N) and in the set of *all* PMIs that can be realized in holography. Any deeper question about constraints on the entanglement structure of geometric states should then be formulated in terms of these more fundamental objects. Furthermore, since for the reason mentioned above this equivalence would in general not be attained for any specific value of N, the significance of certain specific objects which are manifestly N-dependent (like the holographic entropy inequalities) should be questioned, as these objects might be significantly affected by structural artifacts of the formulation.

A priori it might seem surprising that all the holographic entropy inequalities can be derived from a simpler structure which only involves SA – after all, SA is a universal relation which does not 'know about' holography. Intuitively, one can imagine that this is ultimately related to the fact that in quantum field theory the values of the entropies are typically immaterial because of the cut-off dependence, and that the saturation of SA (at

leading order in N) is sensitive solely to the connectivity of the entanglement wedge. And even though certain 'balanced' combinations of entropies such as the mutual information can be often ascribed a finite value independent of the cutoff [92], such numerical data can typically be 'dialed' by for example deforming the state or the subsystem specification. Since the HEC is specified by the limit to which such dialing can be pushed, it should not involve rescalable numerical values. In other words, for delimiting the HEC, it should only matter whether such finite quantities are zero or non-zero, which is indeed borne out in our results. Even if the starting point in the development of our framework will be the graph models from [42], where the edge weights can be dialed at will, any dependence on a specific choice will be effectively modded out by the fact that we will formulate all our results purely in terms of equivalence classes of graph models inspired by the proto-entropies of [67, 68].

The structure of the chapter is as follows. In Section 3.2 we review some of the basic definitions which were already used in previous works. In Section 3.3 we introduce the main tools which will be used in later sections, in particular the notion of an equivalence class of graph models, specified by a *min-cut structure* on a *topological graph model* of holographic entanglement, and of its corresponding *min-cut subspace*. Section 3.4 reviews the concept of marginal independence from [12], and establishes a connection between patterns of marginal independence and min-cut subspaces for a certain class of tree graphs. Section 3.5 analyses how min-cut structures and subspaces transform when one varies the number of parties, and generalizes the results for tree graphs from Section 3.4. All these tools will then be used in Section 3.6, where we introduce our conjectures and discuss what evidence we have to support them, how they are related to each other, and their implications for the derivation of the holographic entropy cone. We conclude in Section 3.7 with a discussion of the main questions that still need to be answered in

order to obtain a full characterization of the holographic entropy cone for an arbitrary number of parties, and comments on other future directions. Throughout the text we will occasionally decorate various symbols to stress particular choices of the corresponding objects which satisfy additional requirements or have important additional features.

3.2 Basic definitions and notation

In this section we briefly review some of the main definitions, most of which were already used in previous works [56, 42, 68, 67, 9, 8]. In Section 3.2.1 we introduce the concept of entropy cones in quantum mechanics. In Section 3.2.2 we review the holographic set-up, the definition of the holographic entropy cone and the concept of proto-entropies. In Section 3.2.3 we review the basic definitions of the graph models of holographic entropies. Finally, in Section 3.2.4 we list some useful immediate consequences of the main definitions. For more details the reader is referred to the original works.

3.2.1 Entropy cones

Consider a Hilbert space which is a tensor product of N factors,

$$\mathcal{H} \coloneqq \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_{\mathsf{N}}, \tag{3.1}$$

and a density matrix ρ acting on it. For any non-empty subset $\mathcal{I} \subseteq [\mathsf{N}] \coloneqq \{1, 2, \ldots, \mathsf{N}\}$ of these factors, the von Neumann entropy is defined as

$$S_{\mathcal{I}} \coloneqq S(\rho_{\mathcal{I}}) = -\operatorname{Tr}\left(\rho_{\mathcal{I}} \log \rho_{\mathcal{I}}\right), \tag{3.2}$$

where

$$\rho_{\mathcal{I}} \coloneqq \mathrm{Tr}_{\mathcal{H}_{[\mathsf{N}]\setminus\mathcal{I}}}\,\rho, \qquad \text{with} \qquad \mathcal{H}_{\mathcal{I}} \coloneqq \bigotimes_{\ell\in\mathcal{I}} \mathcal{H}_\ell, \tag{3.3}$$

is the reduced density matrix, or *marginal*, for the subsystem $\mathcal{I}$. The *entropy vector* corresponding to the density matrix ρ is the ordered[10] collection of entropies of all its marginals, namely,

$$\mathbf{S}(\rho) = \{S_{\mathcal{I}} \text{ for all } \mathcal{I}\}. \tag{3.4}$$

The vector space where entropy vectors live is $\mathbb{R}^{\mathsf{D}}$, where $\mathsf{D} = 2^{\mathsf{N}} - 1$, and will be referred to as *entropy space*. The labeling of the Hilbert space factors in (3.1) will be called a *coloring*, the subscripts $\ell \in [\mathsf{N}]$ are referred to as *colors* and any non-empty set of colors $\mathcal{I}$ is a *polychromatic* index.

For a fixed number of parties N, the collection of all entropy vectors for all possible Hilbert spaces and density matrices was shown by [56] to be a convex cone[11] known as the N-party *quantum entropy cone* ($\mathrm{QEC}_{\mathsf{N}}$).[12] By construction, the $\mathrm{QEC}_{\mathsf{N}}$ is clearly symmetric under an arbitrary permutation of the N parties, as can be seen by just permuting Hilbert space factors. In fact, it exhibits a larger symmetry group of permutations of $[\mathsf{N}+1]$, as we now explain.

For a density matrix ρ, a *purification* is any pure state $|\psi\rangle$ in an enlarged Hilbert space

$$\underline{\mathcal{H}} \coloneqq \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_{\mathsf{N}} \otimes \mathcal{H}_{\mathsf{N}+1} \tag{3.5}$$

[10] We will follow our previous convention of ordering the entropies first by cardinality of $\mathcal{I}$ and then lexicographically, though the actual order will not play a significant role in what follows. In all specific examples, we will use alternate (and more conventional) notation of letters $A, B, C, \ldots$ instead of numbers $1, 2, 3, \ldots$ to denote colors, with the letter O reserved for the purifier.

[11] A closed convex cone is a set of vectors such that for any two vectors $\mathbf{v}_1, \mathbf{v}_2$ in the set, the *conical combination* $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2$ (where $\alpha, \beta \geq 0$) also belongs to the set.

[12] More precisely, it is the topological closure of this set which is a convex cone, while the set itself has a more complicated structure.

such that

$$\rho = \mathrm{Tr}_{\mathcal{H}_{\mathsf{N}+1}} |\psi\rangle \langle\psi| \ . \tag{3.6}$$

We will refer to the additional auxiliary subsystem $\mathcal{H}_{\mathsf{N}+1}$ as the *purifier*[13] and denote a non-empty subset of $[\mathsf{N}+1]$ by an underlined index $\underline{\mathcal{I}}$. Occasionally we will take complements of (not necessarily underlined) polychromatic indices, and we will always define these complements with respect to the set $[\mathsf{N}+1]$.[14] Since for any pure state the entropy of a subsystem is equal to the entropy of its complement, the entropies of all the subsystems of $|\psi\rangle$ are already encoded in the entries of $\mathbf{S}(\rho)$.[15] Any permutation of the $\mathsf{N}+1$ factors in (3.5) will map an N-party entropy vector to another N-party entropy vector, resulting in the extended symmetry mentioned above. In what follows, when we consider permutations of entropy vectors or inequalities, we will always mean permutations of $[\mathsf{N}+1]$.

For $\mathsf{N} = 2, 3$, the $\mathrm{QEC}_{\mathsf{N}}$ is known to be a polyhedral cone and can therefore be specified by a finite set of inequalities.[16] In the $\mathsf{N} = 2$ case, the facets are given by the permutations of *subadditivity* (SA),[17]

$$S_1 + S_2 \geq S_{12}, \tag{3.7}$$

[13]The purifier will often be referred to as color $\mathsf{N}+1$.

[14]Note that according to this definition the set of non-underlined indices is not closed under this operation, however to simplify the notation we will write the complement of $\mathcal{I}$ as $\mathcal{I}^{\mathsf{C}}$ instead of $\underline{\mathcal{I}^{\mathsf{C}}}$.

[15]By convention, given a pair of complementary subsystems $(\underline{\mathcal{I}}, \underline{\mathcal{I}^{\mathsf{C}}})$ we have $S_{\underline{\mathcal{I}}} = S_{\underline{\mathcal{I}^{\mathsf{C}}}}$ and we denote the entropy of each of them by the index which does *not* include the purifier.

[16]A polyhedral cone is said to be *pointed* when it does not contain any non-trivial linear subspace, and in what follows all polyhedral cones that we will consider will be pointed. Any pointed polyhedral cone can equivalently be described as the *conical hull* (the set of all possible conical combinations) of the set of its extreme rays.

[17]These include the *Araki-Lieb inequality* $S_1 + S_{12} \geq S_2$.

while for $\mathsf{N} = 3$ there exists a new inequality known as *strong subadditivity* (SSA),[18]

$$S_{12} + S_{23} \geq S_2 + S_{123}. \tag{3.8}$$

It is important to notice that for $\mathsf{N} = 3$, the permutations of (3.8) constitute only a proper subset of the full set of facet-defining inequalities of the cone. The additional facets correspond to certain particular instances of SA. For example, one can easily verify that (3.7) specifies a facet while the "lift"

$$S_1 + S_{23} \geq S_{123} \tag{3.9}$$

does not, being just the sum of (3.7) and (3.8).[19] For $\mathsf{N} \geq 4$ the QEC_N is essentially unknown; however, for any N, one can easily construct an "outer bound" (a larger cone that contains it) by considering all instances of SA for all possible pairs of disjoint subsets of $[\mathsf{N}+1]$. For any given N, these define the N-party subadditivity cone (SAC_N), an object which will play a central role in our derivation of the holographic entropy cone (see also [12]).[20]

3.2.2 Holographic constructions

Having introduced an N-party entropy space and the QEC_N therein, we now consider the construct of an entropy cone in the context of holography. A natural split-

[18] Its permutations include *weak monotonicity* $S_{12} + S_{23} \geq S_1 + S_3$.

[19] Geometrically, the saturation of this weaker inequality (3.9) corresponds to a hyperplane which intersects the boundary of the HEC only along a lower dimensional subspace.

[20] The reader might already wonder if one could not derive a more stringent bound by also including SSA. This is of course correct, but as we will see it is SA, rather than SSA, which plays a more fundamental role.

ting of the Hilbert space is achieved[21] by partitioning the space on which the CFT lives, which we now specify in more detail to indicate the generality of the setup. Consider a (not necessarily connected) asymptotically AdS manifold $\mathcal{M}$ with M boundaries, $\partial\mathcal{M} = \bigcup_{m=1}^{\mathsf{M}} \partial\mathcal{M}_m$.[22] For each boundary component $\partial\mathcal{M}_m$, consider a Cauchy slice Σ_m and a partition of it into an arbitrary number of connected regions $\mathcal{A}_m^i$. Given a number of parties N, we introduce a surjective *coloring* of these regions

$$\beta : \{\mathcal{A}_m^i \mid \forall m, i\} \to [\mathsf{N}+1]. \tag{3.10}$$

In other words, each Hilbert space factor $\mathcal{H}_\ell$, which we label by a color, is associated to some collection of these regions. Given a polychromatic index $\mathcal{I}$, the set of regions which under β receive a color $\ell \in \mathcal{I}$ is the preimage $\beta^{-1}(\mathcal{I})$, and will be referred to as the *subsystem* associated to $\mathcal{I}$ (or even more simply the subsystem $\mathcal{I}$).[23] A choice of such a manifold, Cauchy slice, partition and a coloring defines what we will call an N-party *holographic configuration*, denoted by $\mathcal{C}_{\mathsf{N}}$.[24]

Using the HRT prescription [36], we can associate an entropy vector $\mathbf{S}(\mathcal{C}_{\mathsf{N}}, g)$ to any given holographic configuration.[25] However, notice that typically the entropies will not be finite, which occurs whenever an extremal surface is anchored on the AdS boundary, in

[21] Even if strictly speaking the Hilbert space of a QFT does not factorize, we assume that this splitting is a valid approximation, see [93] and references therein for more details on this issue.

[22] The dimension of $\mathcal{M}$ will not play any role.

[23] We use similar notation and terminology for indices $\underline{\mathcal{I}}$ which include the purifier.

[24] Throughout our construction, we will be working in the regime where the holographic dual describes a classical bulk spacetime, i.e., we will not consider stringy or quantum corrections; see the discussion in Section 3.7 for additional comments.

[25] The second argument in $\mathbf{S}(\mathcal{C}_{\mathsf{N}}, g)$ is a shorthand meant to indicate the dependence on the bulk spacetime metric g_{ab}. Although it may seem more natural to use the CFT quantity ρ, we choose to leave this implicit (as describing any density matrix which gives the bulk g_{ab}) to emphasize that in the present context the entanglement entropy is given by a geometrical construct (namely the relevant set of extremal surfaces).

other words when the corresponding region admits an entangling surface.[26] To circumvent this problem, [42] introduced a cut-off surface, making all the entropies finite at the expense of associating to a configuration an entropy vector $\mathbf{S}_\epsilon(\mathcal{C}_\mathsf{N}, g)$ which is now cut-off dependent.

Given a number of parties N, one can consider the set of all possible (finite) entropy vectors, for all possible N-party configurations and choices of cut-off. It is then easy to show [42] that this set has the structure of a convex cone, and is known as the N-party *holographic entropy cone* (HEC_N).[27] Note that despite being called "holographic", the construction of this cone is purely geometric, and in using the HRT formula we implicitly assumed that for any $\mathcal{M}$ the bulk dynamics is dual to the evolution of a tensor product of M copies of a holographic CFT living on $\partial\mathcal{M}$. In what follows we will always make this assumption, leaving the analysis of this subtlety to future work.[28]

While the use of a cut-off is a convenient computational tool in defining the cone, one might worry that it could render the resulting construct intrinsically ill-defined. In particular, would the cone be meaningful if its building blocks required a specification of a cut-off? Fortunately, we can circumvent this subtlety by realizing the cone's extreme rays by configurations where none of the minimal surfaces is anchored to the boundary, specifically by the bulk geometry corresponding to multi-boundary wormholes and subsystems covering the entire connected pieces of the boundary [42]. In such configurations the corresponding HRT surface areas are therefore finite, with no need for cut-offs.

While this proof makes it clear that the "static" HEC is a physically well motivated

[26] The only case where this does not happen is when each Cauchy surface Σ_m is not partitioned at all, and each region $\mathcal{A}_m$ is then the entire Σ_m.

[27] In [42] the definition was given for the static case, but the same formulation pertains also in the dynamical case. The reader who is already familiar with these constructions is reminded that this is however not the case for the proof of polyhedrality.

[28] See [94] for more details regarding this issue.

object to study, the construction has some limitations. First of all, one would like to extend the result to the "covariant" HEC (see [95, 96, 97, 65] for discussions in this direction), which is not even known to be polyhedral.[29] Second, it would be interesting to understand to what extent the properties of the HEC, even in the static case, depend on multiboundary wormhole solutions and on the choice of configurations where all entropies are finite. Finally, it is interesting to investigate how specific substructures of the cone (for example certain internal regions or portions of its boundary) are related to different holographic configurations, especially the more typical ones which have divergent entropies.

Motivated by this, the work of [67] introduced the concept of *proto-entropy* and *proto-entropy vector*, as manifestly cut-off independent objects associated to a configuration and a metric. Given an arbitrary boundary region $\mathcal{A}$ (not necessarily connected), consider a minimal extremal surface $\xi_{\mathcal{A}}$ whose area computes the entropy of $\mathcal{A}$. Such extremal surface can be composed of multiple disjoint pieces, which may or may not be anchored on the boundary. The proto-entropy of $\mathcal{A}$ is then simply defined as the formal sum of the connected components of $\xi_{\mathcal{A}}$. In practice, one should just imagine the HRT prescription without the area functional, with the sum of areas of connected surfaces replaced by a formal sum of the surfaces themselves. The convenience of this construction is that all surfaces are now treated on the same footing, irrespective of whether they are anchored to the boundary (and therefore have infinite area) or not, and no cut-off is ever introduced. Given a pair $(\mathcal{C}_{\mathsf{N}}, g)$, its proto-entropy vector is then defined as the ordered collection of

[29] While there is no well-established definition of such a covariant HEC, the intent is to characterize the entropy vectors for all configurations in all physical time-dependent holographic spacetimes. Since the same subtlety mentioned above and explored in [94] is likely more severe in the time-dependent context, we again retreat to geometrical definition, as a collection of entropy vectors given by the HRT prescription, in spacetimes where the latter applies. Most simplistically, we would then restrict to classical bulk spacetimes obeying the NEC.

the proto-entropies of the boundary subsystems.

Using this formulation, [67, 68] suggested an approach to the derivation of new holographic entropy inequalities where it is the connectivity of the entanglement wedges, rather than the area of the surfaces, that plays a central role. While in the present work we will focus on the graph models of [42] rather than on proto-entropies, the approach to the reconstruction of the HEC presented here will follow the same intuition. We leave a more detailed analysis of the relationship with proto-entropies, and in particular between the HEC and the *holographic entropy polyhedron* to future work.

3.2.3 Graph models

In the previous subsection we gave a definition of a holographic configuration in the most general scenario, where the bulk spacetime is dynamical. We will now restrict to static spacetimes, such that the entropies are computed by minimal (rather than extremal) surfaces. In this case, given a pair $(\mathcal{C}_{\mathsf{N}}, g)$, the minimal surfaces define a partition of the bulk time slice, and we can conveniently describe the configuration by a graph constructed as follows [42].

To each region in the partition of the bulk time slice we associate a vertex. If a region is adjoining to the boundary, we label the vertex by the color of the adjoining boundary region. If two vertices correspond to two adjoining regions of the bulk time slice, they are connected by an edge. Any such edge corresponds to a piece of a minimal surface and will carry a weight equal to its area.[30] Note that such graphs can get quite complicated; for example, even in a simple configuration with regions specified by symmetric distribution of disks on a spatial slice of $\mathbb{R}^{2,1}$, two correlated regions would correspond to a planar

[30] Pieces of RT surfaces which reach the boundary will correspond to edges with infinite weights. If necessary, one can again imagine introducing a cut-off, and work with a graph having all edge weights finite.

graph with 4 vertices and 3 edges, while a graph encoding four pairwise-correlated regions would be a non-planar one with 43 vertices and 90 edges. However, we will not need to consider the specific details of such graphs.

The convenience of this representation stems from the fact that it suggests how to formulate an alternative, but *equivalent*, definition of the (static) HEC which is entirely based on graphs, without the need to consider more explicit holographic configurations as introduced above. In essence, one first defines a "graph model" of holographic entanglement, with a prescription for how to compute an entropy vector starting from it, and then proves that the set of entropy vectors realized by such graph models is equal to the HEC.[31] Since these graph models are the main tool that we will use in the rest of the chapter, we now review their original definition in detail.

Consider an undirected graph $G = (V, E)$ with vertex set V and edge set E of unordered pairs of vertices.[32] With a slight (but by now well-established) abuse of notation, we denote a set of *boundary vertices* by $\partial V \subseteq V$. We will refer to the other vertices (if any[33]) in V as the *bulk vertices*. Analogously to what we discussed for configurations, to each boundary vertex we assign a color via the coloring $\beta : \partial V \to [\mathsf{N}+1]$ and $\beta^{-1}(\mathcal{J})$ is the preimage of the subsystem $\mathcal{J}$. A graph G together with a specification of boundary vertices and a coloring β defines a *topological graph model* of holographic entanglement, which will be denoted by G_{N}.

If this structure is further endowed with a *weight map* $w : E \to \mathbb{R}_{>0}$, we will call it a

[31] Given an arbitrary graph model one can convert it into a "canonical form" using certain entropy-preserving transformations (explicitly reviewed in Section 3.3.3). One then shows that the entropy vector of any graph model in canonical form can be explicitly realized by a multiboundary wormhole configuration. See [42] for more details.

[32] For simplicity the graph is also assumed to be simple (it is not a multigraph) and loopless ($i \neq j$ for all $(i, j) \in E$), although without these restrictions the entropy cone would remain unchanged.

[33] As pointed out by [8], the restriction to only $\partial V = V$ corresponds to the study of the cut function, which when written as a vector is well known to span a polyhedral cone whose facets are the subadditive inequalities [71, 70].

graph model of holographic entanglement, and denote it by $\widetilde{G_{\mathsf{N}}}$. We will often think of the map w as a *weight vector* $\mathbf{w} = w(E)$ in the space of weights $\mathbb{R}^{\mathsf{E}}$ (where $\mathsf{E} \coloneqq |E|$), consisting of all weights $w_e \coloneqq w(e)$ for all $e \in E$ with some ordering.

Any subset $U \subseteq V$ characterizes a bipartition or *cut* of G, which defines a set of *cut edges* $\mathcal{C}(U) \subseteq E$ as

$$\mathcal{C}(U) \coloneqq \{(v, v') \in E \mid v \in U,\ v' \in U^{\complement}\} \tag{3.11}$$

(where $U^{\complement} \coloneqq V \setminus U$ is the complementary set of vertices). For a graph model $\widetilde{G_{\mathsf{N}}}$, the *cut weight* of a cut U is defined as the total weight of its edges

$$\|\mathcal{C}(U)\| \coloneqq \sum_{e \in \mathcal{C}(U)} w(e). \tag{3.12}$$

For any non-empty $\mathcal{I} \subseteq [\mathsf{N}]$, a set $U \subseteq V$ is a *cut homologous to* $\mathcal{I}$, or an $\mathcal{I}$*-cut*, if it contains precisely the boundary vertices colored by $\mathcal{I}$, i.e. if $U \cap \partial V = \beta^{-1}(\mathcal{I})$. We will denote an arbitrary $\mathcal{I}$-cut by $U_{\mathcal{I}}$ and to simplify the notation, we will occasionally denote the corresponding set of cut edges by $\mathcal{C}_{\mathcal{I}}$ (as a shorthand for $\mathcal{C}(U_{\mathcal{I}})$). The minimum cut weight among all $\mathcal{I}$-cuts gives the *entropy* $S_{\mathcal{I}}$ of the associated subsystem, i.e.,

$$S_{\mathcal{I}} \coloneqq \min_{U_{\mathcal{I}}} \|\mathcal{C}(U_{\mathcal{I}})\| \ . \tag{3.13}$$

Any $\mathcal{I}$-cut $U_{\mathcal{I}}$ with minimum cut weight $\|\mathcal{C}(U_{\mathcal{I}})\| = S_{\mathcal{I}}$ is a *min-cut* for $\mathcal{I}$ and will be denoted by $U^*_{\mathcal{I}}$ (similarly, we will occasionally denote the set of cut edges for a min-cut by $\mathcal{C}^*_{\mathcal{I}}$). Notice that from the definition it immediately follows that the complement of a min-cut $U^*_{\mathcal{I}}$ is a min-cut for the complementary subsystem $\mathcal{I}^{\complement}$. Min-cuts however are not necessarily unique, as for holographic configurations. We will come back to this "degeneracy" in later sections, since it will play an important role.

3.2.4 Basic properties of min-cuts

Here we collect some basic results about min-cuts that will be useful later.

For a given topological graph model G_{N}, consider an arbitrary subsystem $\mathcal{I}$,[34] an arbitrary $\mathcal{I}$-cut $U_{\mathcal{I}}$, and the induced subgraph[35] $G_{\mathsf{N}}[U_{\mathcal{I}}]$, with vertex set $U_{\mathcal{I}}$ and boundary vertices $\partial V \cap U_{\mathcal{I}}$ with the coloring inherited from G_{N}. In the special case of a min-cut $U^*_{\mathcal{I}}$, this subgraph is the natural analog of the entanglement wedge (or more precisely the homology region) for the subsystem $\mathcal{I}$. However, as we now show, each connected component of a min-cut $U^*_{\mathcal{I}}$ has to be connected to the boundary, i.e. it has to contain at least one boundary vertex:[36]

Lemma 3.1 (Topological minimality)**.** *If the induced subgraph $G_{\mathsf{N}}[U_{\mathcal{I}}]$ of a topological graph model G_{N} on the vertices of a $\mathcal{I}$-cut $U_{\mathcal{I}}$ is disconnected, and there is a connected component which does not include any of the boundary vertices of G_{N}, then the cut $U_{\mathcal{I}}$ cannot be a min-cut for any graph model $\widetilde{G_{\mathsf{N}}}$ on G_{N}.*

Proof. Denote by $U^{\varnothing}_{\mathcal{I}}$ a connected component of $G_{\mathsf{N}}[U_{\mathcal{I}}]$ such that $U^{\varnothing}_{\mathcal{I}} \cap \partial V = \varnothing$. Then $U'_{\mathcal{I}} = U_{\mathcal{I}} \setminus U^{\varnothing}_{\mathcal{I}}$ is a new $\mathcal{I}$-cut on G_{N} such that

$$\|\mathcal{C}(U'_{\mathcal{I}})\| < \|\mathcal{C}(U_{\mathcal{I}})\|. \tag{3.14}$$

Hence $U_{\mathcal{I}}$ cannot be a min-cut. □

The second basic property of min-cuts is the graph version of the property of HRRT

[34] Throughout this subsection we use non-underlined indices $\mathcal{I}$ rather than $\underline{\mathcal{I}}$ to simplify the notation, but all lemmas obviously also hold for subsystems that include the purifier.

[35] The induced subgraph $G[U]$ of a graph $G = (V, E)$ is the graph $G[U] = (U, F)$ with vertex set $U \subset V$ and edges $F \subseteq E$ connecting only vertices in U.

[36] We are ignoring the trivial case where G_{N} is disconnected and some of its connected components do not include any boundary vertices. However it should be clear that even in these cases, such disconnected components are never cut by a min-cut.

known as "entanglement wedge nesting", which states that entanglement wedges of nested regions must themselves be nested. We will state this property without proof (see [8] for more details):

Lemma 3.2 (Nesting). *For any graph model $\widetilde{G_{\mathsf{N}}}$, subsystems $\mathcal{I}, \mathcal{K}$ with $\mathcal{K} \subset \mathcal{I}$, and min-cut $U^*_{\mathcal{I}}$ for $\mathcal{I}$, there exists a min-cut $U^*_{\mathcal{K}}$ for $\mathcal{K}$ such that $U^*_{\mathcal{K}} \subset U^*_{\mathcal{I}}$.*

Notice that in case of degeneracy, the inclusion $U^*_{\mathcal{K}} \subset U^*_{\mathcal{I}}$ need not necessarily hold for all choices of min-cut pairs.

Using this lemma, one can immediately derive the following result:

Lemma 3.3 (No-crossing). *For any graph model $\widetilde{G_{\mathsf{N}}}$, and subsystems $\mathcal{I}, \mathcal{K}$ with $\mathcal{I} \cap \mathcal{K} = \varnothing$, there exist min-cuts $U^*_{\mathcal{I}}$ and $U^*_{\mathcal{K}}$ such that $U^*_{\mathcal{I}} \cap U^*_{\mathcal{K}} = \varnothing$.*

Proof. Consider the subsystem $\mathcal{I}^{\mathfrak{C}}$. Since $\mathcal{I} \cap \mathcal{K} = \varnothing$, it follows that $\mathcal{K} \subset \mathcal{I}^{\mathfrak{C}}$ and by Lemma 3.2 there exists min-cuts $U^*_{\mathcal{I}^{\mathfrak{C}}}$ and $U^*_{\mathcal{K}}$ such that $U^*_{\mathcal{K}} \subset U^*_{\mathcal{I}^{\mathfrak{C}}}$. The complement of $U^*_{\mathcal{I}^{\mathfrak{C}}}$ is then a min-cut $U^*_{\mathcal{I}}$ such that $U^*_{\mathcal{I}} \cap U^*_{\mathcal{K}} = \varnothing$. □

Finally, we mention another property which is closely related to nesting and will be particularly useful in later proofs:

Lemma 3.4 (Min-cut decomposition). *For any graph model $\widetilde{G_{\mathsf{N}}}$, if the induced subgraph $\widetilde{G_{\mathsf{N}}}[U^*_{\mathcal{I}}]$ for a min-cut $U^*_{\mathcal{I}}$ is composed of two disjoint components,*

$$\widetilde{G_{\mathsf{N}}}[U^*_{\mathcal{I}}] = \widetilde{G_{\mathsf{N}}}[U_{\mathcal{J}}] \oplus \widetilde{G_{\mathsf{N}}}[U_{\mathcal{K}}]$$

with $\mathcal{I} = \mathcal{J} \cup \mathcal{K}$, then $U_{\mathcal{J}}$ and $U_{\mathcal{K}}$ are min-cuts for the corresponding subsystems.

Proof. Proceed by contradiction. Suppose $U_{\mathcal{J}}$ is not a min-cut, so an actual min-cut $U^*_{\mathcal{J}}$ for $\mathcal{J}$ has $\left\|\mathcal{C}(U^*_{\mathcal{J}})\right\| < \|\mathcal{C}(U_{\mathcal{J}})\|$. By hypothesis the union $U^*_{\mathcal{J}} \cup U_{\mathcal{K}}$ defines a cut for $\mathcal{I}$. By

subadditivity,[37] $\left\|\mathcal{C}(U_{\mathcal{J}}^* \cup U_{\mathcal{K}})\right\| \leq \left\|\mathcal{C}(U_{\mathcal{J}}^*)\right\| + \|\mathcal{C}(U_{\mathcal{K}})\| < \|\mathcal{C}(U_{\mathcal{J}})\| + \|\mathcal{C}(U_{\mathcal{K}})\|$, where we used minimality in the last step. But by the disjointness assumption, the right-hand side equals $\|\mathcal{C}(U_{\mathcal{J}}^*)\|$. Hence $U_{\mathcal{J}}^* \cup U_{\mathcal{K}}$ is a cut for $\mathcal{I}$ of smaller weight than the claimed min-cut $U_{\mathcal{J}}^*$, a contradiction (and the same argument of course also applies to $\mathcal{K}$). □

An alternate way of stating Lemma 3.4 is that if the set of cut edges for a min-cut $U_{\mathcal{J}}^*$ can be split into two sets of cut edges for cuts of subsystems bipartitioning $\mathcal{J}$, then each set corresponds to a min-cut for that subsystem. Note that this statement can be trivially iterated when $\widetilde{G_{\mathsf{N}}}[U_{\mathcal{J}}^*]$ is composed of multiple disjoint components.

The min-cut decomposition also makes an immediate connection to SA saturation: if $\widetilde{G_{\mathsf{N}}}[U_{\mathcal{J}}^*] = \widetilde{G_{\mathsf{N}}}\left[U_{\mathcal{J}}^*\right] \oplus \widetilde{G_{\mathsf{N}}}[U_{\mathcal{K}}^*]$ with $\mathcal{I} = \mathcal{J} \cup \mathcal{K}$, then $S_{\mathcal{I}} = S_{\mathcal{J}} + S_{\mathcal{K}}$, so $\mathcal{J}$ and $\mathcal{K}$ have no mutual information and are fully decorrelated.[38]

3.3 The min-cut subspace of a graph model

In the previous section we reviewed the definition of a graph model of holographic entanglement and how an entropy vector is associated to it. We now begin to consider coarser objects associated to graph models. This will be a recurring theme throughout this work, and we will consider an even coarser object in Section 3.4.

In Section 3.3.1 we introduce the notions of a "min-cut structure" and its corresponding "W-cell" in the space of edge weights. These allow us to organize graph models into equivalence classes, and we will introduce the main objects associated to an equivalence class, namely the "S-cell" and its "min-cut subspace". These subspaces will then be used in Section 3.6 to resolve the structure of the holographic entropy cone. The definition

[37] Phrased more generally, this follows from the fact that the min-cut function on a graph is submodular.

[38] The existence of such decomposition of the induced subgraph $\widetilde{G_{\mathsf{N}}}[U_{\mathcal{J}}^*]$ is also a necessary condition for the mutual information to vanish, as we will review later in Section 3.4.2 (see Lemma 3.9).

is not entirely new, as variations of it were already used previously. Here we sharpen this notion and explore its properties in much greater detail, highlighting in particular the role played by "degeneracy", namely the possible coexistence of alternative min-cuts which compute the entropy of a subsystem.

In Section 3.3.2 and Section 3.3.3 we then explain how the S-cell and min-cut subspace of an equivalence class of graph models can be determined from the W-cell. While in Section 3.2 we have defined weights as strictly positive, in some cases it will be useful to consider extremal situations where some of them vanish, and we clarify how to deal with this type of situations in Section 3.3.4. In Section 3.3.5, we present certain important properties of W-cells, S-cells and min-cut subspaces which will be used in later proofs. Finally, in Section 3.3.6, we explain how for disconnected graphs all these constructs can be obtained from those of the connected "building blocks".

Additional important properties of the objects introduced in this section, related to how these transform under recoloring of boundary vertices in a graph, will be analysed in Section 3.5.[39]

3.3.1 Equivalence classes of graph models

Consider a graph model of holographic entanglement $\widetilde{G_{\mathsf{N}}}$, as defined in Section 3.2.3. As we anticipated above, the min-cut $U^*_{\mathcal{I}}$ for an arbitrary subsystem $\mathcal{I}$ is not necessarily unique. We denote the *set of min-cuts for* $\mathcal{I}$ by $\mathcal{U}_{\mathcal{I}}$. A subsystem indexed by $\mathcal{I}$ will be said to be *generic* (*degenerate*) if $\mathcal{U}_{\mathcal{I}}$ has cardinality equal to (greater than) one. If $\mathcal{U}_{\mathcal{I}}$ has more than one element, every $U^*_{\mathcal{I}} \in \mathcal{U}_{\mathcal{I}}$ is referred to as a *degenerate min-cut*, in the

[39] In the discussion, Section 3.7, we will also briefly comment on the relation between the min-cut subspaces defined here, and similar construct introduced in [68, 67] for proto-entropies and the definition of the holographic entropy polyhedron.

sense that any such cut achieves the minimum weight among all possible cuts for $\mathcal{I}$.[40] Similarly, we introduce the following terminology for graph models

Definition 3.1 (Generic and degenerate graph models)**.** *A graph model is* generic *if every subsystem is* generic*, while it is* degenerate *if at least one subsystem is* degenerate*.*

Notice that a "general" graph model may or may not be "generic" according to this definition. The motivation for this choice of terminology is that random graph models typically do not have any degeneracy.

For an arbitrary graph model $\widetilde{G_{\mathsf{N}}}$ we define its *min-cut structure* as follows

Definition 3.2 (Min-cut structure of a graph model)**.** *The* min-cut structure $\mathfrak{m}(\widetilde{G_{\mathsf{N}}})$ *of a graph model* $\widetilde{G_{\mathsf{N}}}$ *is the collection of its min-cut sets for all polychromatic indices, i.e.*

$$\mathfrak{m}(\widetilde{G_{\mathsf{N}}}) \coloneqq \{\mathcal{U}_{\mathcal{I}} \text{ for all } \mathcal{I}\} \tag{3.15}$$

Any two graph models on the same *topological* graph model G_{N} will be considered equivalent if their min-cut structures are equal. This equivalence relation allows us to organize all such graph models into a finite set of equivalence classes, each one corresponding to a distinct min-cut structure $\mathfrak{m}$. We write such an equivalence class as a pair $(G_{\mathsf{N}}, \mathfrak{m})$, stressing the dependence on both the min-cut structure and the underlying topological graph model. Occasionally, we will also write $(G_{\mathsf{N}}, \mathfrak{m})[\widetilde{G_{\mathsf{N}}}]$ to stress that $(G_{\mathsf{N}}, \mathfrak{m})$ is not an arbitrary equivalence class, but the class specified by the representative graph model $\widetilde{G_{\mathsf{N}}}$. Any element of a class is specified by a choice of weights consistent with the min-cut structure $\mathfrak{m}$, and it will be convenient to associate to each equivalence class a region in the space of edge weights $\mathbb{R}^{\mathsf{E}}_{>0}$.

[40] Notice that the cardinality of $\mathcal{U}_{\mathcal{I}}$ is bounded by $1 \leq |\mathcal{U}_{\mathcal{I}}| \leq 2^{|V|-(\mathsf{N}+1)}$, since only the $|V| - (\mathsf{N}+1)$ bulk vertices can be optionally included in the $\mathcal{I}$-cut.

Definition 3.3 (W-cell)**.** *The* W-cell $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m}) \subset \mathbb{R}^{\mathsf{E}}$ *of the min-cut structure* $\mathfrak{m}$ *on the topological graph model* G_{N} *is the set of weight vectors of all graph models* $\widetilde{G_{\mathsf{N}}}$ *in the class* $(G_{\mathsf{N}}, \mathfrak{m})$.

In what follows, when the specification a topological graph model G_{N} is clear from context, and we only need to keep track of the dependence of a W-cell on a min-cut structure $\mathfrak{m}$, we will simply write $\mathcal{W}_{\mathfrak{m}}$ instead of $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})$.

Given a topological graph model G_{N}, one may want to specify a min-cut structure on it more abstractly, by directly listing the min-cut sets $\mathcal{U}_{\mathcal{I}}$ without recurring to graph models and explicit choices of edge weights. Notice however that an arbitrary choice of cuts for each index $\mathcal{I}$ is not necessarily a meaningful min-cut structure, since there may not exist a choice of weights that make all these cuts (and no others) minimal (for example by violating some of the lemmas in Section 3.2.4). We will comment more explicitly on this point in the next section.

Each equivalence class will also be associated to a specific region of entropy space, called the *S-cell* of $(G_{\mathsf{N}}, \mathfrak{m})$, defined as follows

Definition 3.4 (S-cell)**.** *The* S-cell $\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m}) \subset \mathbb{R}^{\mathsf{D}}$ *of the min-cut structure* $\mathfrak{m}$ *on the topological graph model* G_{N} *is the set of entropy vectors of all graph models* $\widetilde{G_{\mathsf{N}}}$ *in the class* $(G_{\mathsf{N}}, \mathfrak{m})$.

By definition, each S-cell is contained in the HEC, since it is a set of entropy vectors that can be realized by graph models. As we will see, different S-cells can have different dimensions, and clearly the union of all S-cells that can be obtained from all possible topological graph models and min-cut structures is the whole $\mathrm{HEC}_{\mathsf{N}}$. However, the set of all S-cells does not form a partition of the $\mathrm{HEC}_{\mathsf{N}}$ since a given entropy vector can

in general belong to multiple S-cells.[41] Furthermore, the relation between equivalence classes and S-cells is not a bijection, since the same S-cell can be associated to distinct equivalence classes.[42]

Having introduced the notion of an S-cell, we now consider a linear subspace of entropy space which is naturally associated to it. Specifically, we define the *min-cut subspace* of a class $(G_{\mathsf{N}}, \mathfrak{m})$ as follows

Definition 3.5 (Min-cut subspace)**.** *The* min-cut subspace $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$ *of the min-cut structure* $\mathfrak{m}$ *on the topological graph model* G_{N} *is the minimal linear subspace which contains the* S-cell $\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m})$*, i.e.*

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \coloneqq \mathrm{Span}(\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m})) \tag{3.16}$$

Notice that even the relation between S-cells and min-cut subspaces is not a bijection, since different S-cells can give the same min-cut subspace.[43]

Having introduced the main definitions, we will now proceed to explain how, given a topological graph model and a min-cut structure, one can compute the S-cell and the min-cut subspace explicitly. We will start from the slightly simpler case of generic graphs and then extend the analysis to situations where some min-cuts are degenerate.

[41] This can easily happen for example with disconnected graph models. We will comment on this type of situation in more detail in Section 3.3.5.

[42] For example one can start from a class $(G_{\mathsf{N}}, \mathfrak{m})$ and add new vertices to G_{N} to obtain a new topological graph model G'_{N} such that the min-cut structure $\mathfrak{m}$ and the sets of cut edges $\mathcal{C}^*_J$ for all subsystems remain unchanged (a trivial way to do so is by adding a disconnected component of only bulk vertices).

[43] This situation can also easily be realized with disconnected graphs, see Section 3.3.5.

3.3.2 Min-cut structures without degeneracy

Given a topological graph model G_{N}, and a *generic* min-cut structure $\mathfrak{m}$, we can determine the W-cell of $(G_{\mathsf{N}}, \mathfrak{m})$ as follows. To each edge of G_{N} we associate a weight *variable* w_e (rather than a weight value as in a graph model $\widetilde{G_{\mathsf{N}}}$). For any polychromatic index $\mathcal{I}$, we can consider all cuts $U_{\mathcal{I}}$, and the min-cut structure $\mathfrak{m}$ specifies which one is the min-cut $U^*_{\mathcal{I}}$. The minimality of $U^*_{\mathcal{I}}$ translates into a set of linear inequalities in the weight variables, which take the form

$$\|\mathcal{C}(U^*_{\mathcal{I}})\| < \|\mathcal{C}(U_{\mathcal{I}})\| \qquad \forall\, U_{\mathcal{I}} \neq U^*_{\mathcal{I}}. \tag{3.17}$$

The collection of all these inequalities in (3.17), for all polychromatic indices $\mathcal{I}$, combined with the condition that each edge weight must be non negative

$$w_e > 0 \qquad \forall\, e \in E \tag{3.18}$$

specifies a polyhedral cone in $\mathbb{R}^{\mathsf{E}}_{>0}$. The cone structure follows from the inequalities being homogeneous, and since there are at least E linearly independent ones (those in (3.18)), the cone is pointed. Furthermore, since all the inequalities are strict, the (non-empty) solution corresponds to the interior of this cone, and the cone is full-dimensional.[44] This is the W-cell $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})$.

As we mentioned earlier, an arbitrary choice of min-cut for each subsystem $\mathcal{I}$ does not necessarily correspond to a meaningful min-cut structure. This is because the system of inequalities described above could have no solutions, in which case the W-cell is just the empty set. In the following we will always ignore these situations, and whenever

[44] The linear span of the vectors inside the cone is the full space.

we consider a min-cut structure on a certain topological graph model, we will implicitly assume that its W-cell is non-trivial.

Having showed how to determine the W-cell $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})$ explicitly, we will now explain how to derive the corresponding S-cell $\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m})$ in entropy space. Notice that for any graph model $\widetilde{G_{\mathsf{N}}}$, each entropy $S_{\mathcal{I}}$ is computed by a sum of weights, each one with unit coefficient, cf. (3.12). For each subsystem $\mathcal{I}$ we introduce an incidence vector $\mathbf{\Gamma}_{\mathcal{I}} \in \{0,1\}^{\mathsf{E}}$,

$$\Gamma^e_{\mathcal{I}} := \begin{cases} 1 & \text{if } e \in \mathcal{C}(U^*_{\mathcal{I}}) \\ 0 & \text{otherwise} \end{cases} \tag{3.19}$$

which specifies which edges participate in the cut. We can then write the entropy as

$$S_{\mathcal{I}} = \Gamma^e_{\mathcal{I}}\, w_e \tag{3.20}$$

where the $\mathsf{D}\times\mathsf{E}$ matrix $\Gamma^e_{\mathcal{I}}$ represents a linear map $\Gamma : \mathbb{R}^{\mathsf{E}} \to \mathbb{R}^{\mathsf{D}}$. The map which associates entropy vectors to weight vectors of fixed topological graph and min-cut structure is the *restriction* $\Gamma|_{\mathcal{W}}$ of Γ to the W-cell $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})$. The S-cell is the image of this restricted map[45]

$$\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m}) = \text{Im } \Gamma|_{\mathcal{W}} \tag{3.21}$$

and the min-cut subspace is

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \text{Span}(\text{Im } \Gamma|_{\mathcal{W}}) = \text{Im } \Gamma \tag{3.22}$$

where the last equality follows from the fact that, as explained above, $\text{Span}(\mathcal{W}) = \mathbb{R}^{\mathsf{E}}$.

[45] Equivalently, one could also define the S-cell as the image of $\mathcal{W}$ under the *unrestricted* map Γ, i.e., $\mathcal{S} = \Gamma(\mathcal{W})$, but we preferred the option in the main text because the matrix Γ is specified by the min-cut structure, which is only defined within $\mathcal{W}$.

In conclusion, in the case of a generic min-cut structure $\mathfrak{m}$ on a topological graph model G_{N}, the min-cut subspace is simply the column space of the matrix Γ.

3.3.3 Min-cut structures with possible degeneracy

We will now generalize the previous construction to the case where some min-cuts could be degenerate. We start again by determining the W-cell of a given min-cut structure $\mathfrak{m}$ and topological graph model G_{N}.

For each degenerate subsystem $\mathcal{I}$, we label the min-cuts $U_{\mathcal{I}}^*$ in the set $\mathcal{U}_{\mathcal{I}}$ with an upper index $\alpha \in \{1, 2, \ldots, |\mathcal{U}_{\mathcal{I}}|\}$. The W-cell is now specified by the following set of inequalities

$$\|\mathcal{C}(U_{\mathcal{I}}^{*\alpha})\| < \|\mathcal{C}(U_{\mathcal{I}})\| \qquad \forall\, U_{\mathcal{I}} \neq U_{\mathcal{I}}^{*\alpha},\ \ \forall\, \alpha \tag{3.23}$$

for each polychromatic index,[46] together with the strict positivity of the weights (3.18), and crucially a set of *degeneracy equations*

$$\|\mathcal{C}(U_{\mathcal{I}}^{*\alpha})\| = \|\mathcal{C}(U_{\mathcal{I}}^{*\beta})\| \qquad \forall\, \mathcal{I},\ \ \forall\, \alpha, \beta \tag{3.24}$$

Like in the generic case, the W-cell is the interior of a polyhedral cone, but now with a crucial difference. Since all inequalities are strict, denoting by $\mathbb{W}$ the proper linear subspace of $\mathbb{R}^{\mathsf{E}}$ which corresponds to the solution of the degeneracy equations (3.24), we have

$$\mathrm{Span}(\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})) = \mathbb{W} \tag{3.25}$$

rather than the full space.

To find the S-cell we can proceed similarly to generic case. For each degenerate

[46] For each generic subsystem, there is just a single index $\alpha \in \{1\}$.

subsystem $\mathcal{I}$, we first choose a "representative" min-cut $U^{*\alpha}_{\mathcal{I}}$, by a specific choice of α. We denote a choice of min-cuts for all $\mathcal{I}$ schematically by $\{\alpha\}$. With this choice we can then construct an incidence vector $\mathbf{\Gamma}^{\alpha}_{\mathcal{I}}$ for each polychromatic index $\mathcal{I}$, and as before a linear map $\Gamma^{\{\alpha\}} : \mathbb{R}^{\mathsf{E}} \to \mathbb{R}^{\mathsf{D}}$. Different choices of representatives will in general give different linear maps, however once restricted to the W-cell, all these maps have the same image, the S-cell of $\mathfrak{m}$, i.e.,

$$\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m}) = \text{Im } \Gamma^{\{\alpha\}}|_{\mathcal{W}} \qquad \forall \{\alpha\} \tag{3.26}$$

To see this, consider an arbitrary weight vector $\mathbf{w} \in \mathcal{W}$ and two maps $\Gamma^{\{\alpha\}}$ and $\Gamma^{\{\alpha'\}}$ for two different choices of representative min-cuts. We want to show that for any such choices

$$\Gamma^{\{\alpha\}}(\mathbf{w}) = \Gamma^{\{\alpha'\}}(\mathbf{w}) \tag{3.27}$$

which we can rewrite as

$$(\Gamma^{\{\alpha\}} - \Gamma^{\{\alpha'\}})(\mathbf{w}) = 0 \tag{3.28}$$

For a row $\mathcal{I}$ of the matrix that appears in the equation above we should then have

$$(\mathbf{\Gamma}^{\alpha}_{\mathcal{I}} - \mathbf{\Gamma}^{\alpha'}_{\mathcal{I}})(\mathbf{w}) = 0 \tag{3.29}$$

If the choice of representatives for the subsystem $\mathcal{I}$ in $\{\alpha\}$ and $\{\alpha'\}$ was the same, (3.29) is trivial. Otherwise it is precisely one of the degeneracy equations that specify $\mathcal{W}$, and it is therefore also satisfied by $\mathbf{w}$.

As for the generic case, the min-cut subspace of a min-cut structure with degeneracies is again defined as the span of the S-cell. Unlike the generic case however, the min-cut

subspace is now generally not equal to the image of any of the unrestricted maps $\Gamma^{\{\alpha\}}$.[47] Hence

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \text{Span}(\text{Im } \Gamma^{\{\alpha\}}|_{\mathcal{W}}) \subseteq \text{Span}(\text{Im } \Gamma^{\{\alpha\}}) \tag{3.30}$$

and typically $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \subset \text{Span}(\text{Im } \Gamma^{\{\alpha\}})$.

Using this construction it is straightforward to prove the following result, which provides a more direct way to determine the min-cut subspace of an arbitrary class $(G_{\mathsf{N}}, \mathfrak{m})$ when one is not interested in the W-cell or the S-cell.

Lemma 3.5. *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and choice of* $\Gamma^{\{\alpha\}}$*, the min-cut subspace of* $(G_{\mathsf{N}}, \mathfrak{m})$ *is the image of* $\mathbb{W}$ *under* $\Gamma^{\{\alpha\}}$

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \Gamma^{\{\alpha\}}(\mathbb{W}) \tag{3.31}$$

Proof. Since the min-cut subspace is defined as the linear span of the S-cell, $\mathcal{S}$ contains some basis $\mathcal{B}_{\mathbb{S}}$ of $\mathbb{S}$. And by (3.30), there exists a collection of vectors in $\mathcal{W}$ which are mapped to $\mathcal{B}_{\mathbb{S}}$, therefore $\Gamma^{\{\alpha\}}(\mathbb{W}) \supseteq \mathbb{S}$ for any $\{\alpha\}$. On the other hand, the W-cell is full dimensional in $\mathbb{W}$, i.e., its linear span is $\mathbb{W}$. Therefore, any $\mathbf{w} \in \mathbb{W}$ can be written as a linear combination of a basis $\mathcal{B}_{\mathbb{W}} \subseteq \mathcal{W}$, which is mapped inside $\mathcal{S}$ by $\Gamma^{\{\alpha\}}$, implying $\Gamma^{\{\alpha\}}(\mathbb{W}) \subseteq \mathbb{S}$. Combining the two inclusions, $\Gamma^{\{\alpha\}}(\mathbb{W}) = \mathbb{S}$. □

Notice that in the generic case, where $\mathbb{W} = \mathbb{R}^{\mathsf{E}}$, (3.31) reduces to (3.22). Because of (3.26), the specific choice of min-cut representatives for a min-cut structure is often immaterial, and in what follows we will often drop the explicit dependence of $\Gamma^{\{\alpha\}}$ on such a choice, writing simply Γ.

We conclude this subsection with an explicit example of these constructs. Consider the graph model $\widetilde{G_3}$ in Figure 3.1. The indicated choice of weights on the underlying

[47] It is tempting to guess that the min-cut subspace of a degenerate min-cut structure is the intersection of all the images of the unrestricted maps for all choices of representatives, i.e., that $\mathbb{S} = \bigcap_{\{\alpha\}} \text{Im } \Gamma^{\{\alpha\}}$. This however is not entirely clear and we leave this question for future work.

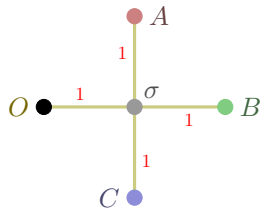


Figure 3.1: The graph model $\widetilde{G_3}$ which generates the extreme ray of the HEC_3 corresponding to the 4-party perfect state. The boundary vertices are $\partial V = \{A, B, C, O\}$, and there is a single bulk vertex, σ. Our drawing conventions for this and all subsequent figures will be to use colors for the boundary vertices (black for the purifier O), and to display all bulk vertices in gray.

topological graph model G_3 specifies the following min-cut structure

$$\begin{aligned} \mathcal{U}_A &= \{\{A\}\} & \mathcal{U}_{AB} &= \{\{A,B\},\{A,B,\sigma\}\} & \mathcal{U}_{ABC} &= \{\{A,B,C\}\} \\ \mathcal{U}_B &= \{\{B\}\} & \mathcal{U}_{AC} &= \{\{A,C\},\{A,C,\sigma\}\} & & \\ \mathcal{U}_C &= \{\{C\}\} & \mathcal{U}_{BC} &= \{\{B,C\},\{B,C,\sigma\}\} & & \end{aligned} \tag{3.32}$$

The degenerate subsystems are AB, AC and BC, and they give rise to the following degeneracy equations

$$\begin{aligned} w_{A\sigma} + w_{B\sigma} &= w_{C\sigma} + w_{O\sigma} \\ w_{A\sigma} + w_{C\sigma} &= w_{B\sigma} + w_{O\sigma} \\ w_{B\sigma} + w_{C\sigma} &= w_{A\sigma} + w_{O\sigma} \end{aligned} \tag{3.33}$$

The solution of the degeneracy equations is the 1-dimensional subspace $\mathbb{W} \subseteq \mathbb{R}^4$ generated by the vector $(1,1,1,1)$. The W-cell is the ray $w(1,1,1,1)$, with $w > 0$, and its image under the map (the rows correspond to entropies ordered as $(A, B, C, AB, AC, BC, ABC)$,

and the columns to the weights ordered as $(A\sigma, B\sigma, C\sigma, O\sigma)$)

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.34}$$

fixed by the choice of representatives

$$\begin{aligned} U^*_{AB} &= \{A, B\} \\ U^*_{AC} &= \{A, C\} \\ U^*_{BC} &= \{B, C\} \end{aligned} \tag{3.35}$$

is the S-cell

$$\mathcal{S} = \lambda\,(1, 1, 1, 2, 2, 2, 1), \qquad \lambda > 0 \tag{3.36}$$

which is the 4-party "perfect state"[48] extreme ray of the HEC_3. In this example the min-cut subspace is therefore 1-dimensional. We will see momentarily that this is always the case for graph models realizing the extreme rays of the HEC.

3.3.4 Vanishing weights

Up to this point we have only considered situations where all the weights in a graph model are strictly positive. However, one may wonder if there could be any issue when allowing some of the weights to vanish, and in this subsection we explore this situation

[48] This is the pure state on four parties which is absolutely maximally entangled, i.e., the state such that each subsystem has maximal entropy.

carefully. The upshot is that if some of the weights vanish in a graph model, all our previous definitions and constructions should be applied to a new graph where all edges with vanishing weight have been deleted.

Consider a topological graph model G_{N}, and suppose that instead of specifying a min-cut structure by a choice of edge weights, we instead try to specify it by a list of min-cut sets for all polychromatic subsystems, which we denote by $\breve{\mathfrak{m}}$. Any such choice will specify a region $\breve{\mathcal{W}}$ of the space of edge weights $\mathbb{R}^{\mathsf{E}}$ via a set of min-cut inequalities (3.23) and degeneracy equations (3.24), together with the condition of strict positivity of the edge weights (3.18). However, as already mentioned before, this region can be empty, in which case $\breve{\mathfrak{m}}$ would not correspond to a valid min-cut structure on G_{N}. For a random choice of min-cut sets, this would typically be the case, since the min-cut sets would easily violate some of the basic properties of min-cuts reviewed in Section 3.2. However, it can also happen that the only reason why $\breve{\mathcal{W}}$ is empty is that some of the degeneracy equations are forcing some of the edge weights to vanish.

Consider now a bipartition $(E_0, E_>)$ of the edge set E and the following constraints on the weights

$$\begin{aligned} w_e &= 0 \qquad \forall\, e \in E_0 \\ w_e &> 0 \qquad \forall\, e \in E_> \end{aligned} \tag{3.37}$$

Suppose that the new region $\mathcal{W}$ specified by (3.37), and precisely the same min-cut inequalities and degeneracy equations that participated in the specification of $\breve{\mathcal{W}}$, is non-empty. According to our definition, even if $\mathcal{W}$ is non-empty, it is not a W-cell for any min-cut structure on G_{N}, since some weights vanish. However, $\mathcal{W}$ can also be specified by an equivalent set of inequalities and equations that can be obtained by simply canceling

the terms w_e for all $e \in E_0$ from all the original min-cut inequalities and degeneracy equations that defined $\breve{\mathcal{W}}$, and by also removing the corresponding equations from (3.37).

Consider now a new topological graph model $\mathring{G}_{\mathsf{N}}$ obtained by simply deleting all edges in E_0 from G_{N}. The space of weights for $\mathring{G}_{\mathsf{N}}$ is now $\mathbb{R}_{>0}^{\mathsf{E}-|E_0|}$ and (3.37) imposes that all remaining weights are strictly positive. Furthermore for each min-cut $U_{\mathcal{J}}^{*\alpha}$ in $\breve{\mathfrak{m}}$ we have

$$\|\mathcal{C}(U_{\mathcal{J}}^{*\alpha})\| - \sum_{e \in E_0(\mathcal{J},\alpha)} w_e = \|\mathring{\mathcal{C}}(U_{\mathcal{J}}^{*\alpha})\| \tag{3.38}$$

where

$$E_0(\mathcal{J},\alpha) = \mathcal{C}(U_{\mathcal{J}}^{*\alpha}) \cap E_0 \tag{3.39}$$

and $\mathring{\mathcal{C}}(U_{\mathcal{J}}^{*\alpha})$ is the set of cut-edges for the min-cut $U_{\mathcal{J}}^{*\alpha}$ on the new graph $\mathring{G}_{\mathsf{N}}$. The equations and inequalities obtained by cancelling the vanishing weights can then be reinterpreted as a set of min-cut inequalities and degeneracy equations for a min-cut structure $\mathring{\mathfrak{m}} = \breve{\mathfrak{m}}$ on the new graph, and the region $\mathcal{W}$ is therefore the W-cell $\mathcal{W}(\mathring{G}_{\mathsf{N}}, \mathring{\mathfrak{m}})$. We will see an example of this reduction procedure in the next section.

Finally, let us briefly comment on the extreme situation where $\breve{\mathcal{W}}$ is the 0-dimensional region that only contains the origin of $\mathbb{R}^{\mathsf{E}}$. In this case the reduction we just described simply produces a new topological graph model $\mathring{G}_{\mathsf{N}}$ with no edges, whose space of weights is now $\mathbb{R}_{>0}^{0} = \{\mathbf{0}\}$. In this space there is only a single W-cell $\mathcal{W}$, which is simply the entire space. The corresponding min-cut structure is generic only in the case where all the vertices in $\mathring{G}_{\mathsf{N}}$ are boundary vertices. Otherwise, the min-cut structure is "maximally degenerate" in the sense that any subset of the bulk vertices can be included in the min-cut of each polychromatic subsystem.

3.3.5 Main properties of W-cells, S-cells and min-cut subspaces

In this subsection we comment on a few important properties of W-cells, S-cells and min-cut subspaces which will be used in later proofs. We begin with a simple observation about the set of W-cells associated to a given topological graph model:

Lemma 3.6. *For any topological graph model G_{N}, the set of* W*-cells associated to all the min-cut structures that can be specified on G_{N} forms a partition of the space of weights $\mathbb{R}^{\mathsf{E}}_{>0}$.*

Proof. Given a topological graph model G_{N}, any choice of weight vector $\mathbf{w} \in \mathbb{R}^{\mathsf{E}}_{>0}$ specifies a min-cut structure uniquely. □

Recall that any W-cell is the interior of a pointed polyhedral cone whose linear span is the subspace $\mathbb{W}$ determined by the degeneracy equations. We now want to consider the boundary of such a cone, and in particular its extreme rays. For a given W-cell $\mathcal{W}$ we denote its closure by $\overline{\mathcal{W}}$. A d-dimensional *face* $\overline{\mathcal{F}}_d$ of $\overline{\mathcal{W}}$ is defined as the intersection of $\overline{\mathcal{W}}$ with a hyperplane $\mathbb{H}$ such that $\overline{\mathcal{W}}$ is contained entirely in one of the half-spaces specified by $\mathbb{H}$ (including $\mathbb{H}$ itself). According to this definition, each face is again a closed polyhedral cone, and we denote its interior[49] by $\mathcal{F}_d$. Notice in particular that the 1-dimensional faces are also closed; they are *closed* extreme rays, since they contain the origin.

In general it is not clear if the interior of a face is by itself a W-cell for some min-cut structure,[50] but we will prove that this is the case for certain extreme rays. We denote a closed extreme ray of $\overline{\mathcal{W}}$ by $\overline{\mathbf{x}}_{\mathcal{W}}$ and its interior (the corresponding *open* ray) by $\mathbf{x}_{\mathcal{W}}$. Moreover, we will say that a ray (either closed or open) is *nowhere-zero* if all the

[49] We define the interior with respect to the subspace topology.

[50] Indeed, the structure of the set of W-cells for a given topological graph model is an interesting object to study. We leave this problem for future work.

components of any vector in its interior are strictly positive. We then have the following lemma:

Lemma 3.7. *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and a nowhere-zero extreme ray* $\overline{\mathbf{x}}_{\mathcal{W}}$ *of* $\overline{\mathcal{W}}(G_{\mathsf{N}},\mathfrak{m})$*, there exists a min-cut structure* $\mathfrak{m}'$ *on* G_{N} *whose* W*-cell is* $\mathbf{x}_{\mathcal{W}}$*.*

Proof. Consider a class $(G_{\mathsf{N}}, \mathfrak{m})$ and a nowhere-zero extreme ray $\overline{\mathbf{x}}_{\mathcal{W}} \in \overline{\mathcal{W}}_{\mathfrak{m}}$, where we have introduced the shorthand notation $\overline{\mathcal{W}}_{\mathfrak{m}} = \overline{\mathcal{W}}(G_{\mathsf{N}}, \mathfrak{m})$. We denote by $\mathbb{F}_1$ the 1-dimensional linear subspace generated by $\overline{\mathbf{x}}_{\mathcal{W}}$. Since $\overline{\mathbf{x}}_{\mathcal{W}}$ is nowhere-zero, by Lemma 3.6 there exists a unique min-cut structure $\mathfrak{m}'$ whose W-cell $\mathcal{W}_{\mathfrak{m}'}$ contains $\mathbf{x}_{\mathcal{W}}$. Suppose now that $\mathbf{x}_{\mathcal{W}}$ is not a W-cell by itself, i.e., that $\mathcal{W}_{\mathfrak{m}'} \neq \mathbf{x}_{\mathcal{W}}$. Then there exists at least another open ray $\mathbf{w} \in \mathcal{W}_{\mathfrak{m}'}$ which corresponds to the same min-cut structure $\mathfrak{m}'$ of $\mathbf{x}_{\mathcal{W}}$. In general the facets of $\overline{\mathcal{W}}$ are supported by two different types of hyperplanes. Some correspond to degeneracy equations for min-cuts in $\mathfrak{m}$, and others are facets of $\mathbb{R}^{\mathsf{E}}_{\geq 0}$. However by the assumption that $\overline{\mathbf{x}}_{\mathcal{W}}$ is nowhere-zero, it follows that $\mathbb{F}_1$ is completely determined by the degeneracy equations only. Since any ray that is not contained in $\mathbb{F}_1$ would violate at least one of these equations, it must be that $\mathbf{w} = \lambda\, \mathbf{x}_{\mathcal{W}}$ with $\lambda > 0$, and $\mathbf{x}_{\mathcal{W}}$ is the whole W-cell of $\mathfrak{m}'$. □

Using this lemma, we can then show that if a class $(G_{\mathsf{N}}, \mathfrak{m})$ has a 1-dimensional min-cut subspace, we can always find another class such that the min-cut subspace is preserved and the new W-cell is just a single ray.

Lemma 3.8. *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *such that* $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$ *is* 1*-dimensional, there exists a class* $(\hat{G}_{\mathsf{N}}, \hat{\mathfrak{m}})$ *such that*

$$\mathbb{S}(\hat{G}_{\mathsf{N}}, \hat{\mathfrak{m}}) = \mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \tag{3.40}$$

and $\mathcal{W}(\hat{G}_{\mathsf{N}}, \hat{\mathfrak{m}})$ *is a single ray.*

Proof. Consider an arbitrary class $(G_{\mathsf{N}}, \mathfrak{m})$ with a 1-dimensional min-cut subspace, and suppose its W-cell $\mathcal{W}_{\mathfrak{m}}$ is *not* a single ray. Since $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$ is 1-dimensional, the S-cell on the other hand is just a single ray, which we denote by $\mathbf{S}$. For any choice of Γ (fixed by a choice of representative min-cuts in $\mathfrak{m}$) we then have

$$\Gamma \mathbf{w} = \lambda_{\mathbf{w}} \mathbf{S} \qquad \forall\, \mathbf{w} \in \mathcal{W}_{\mathfrak{m}} \tag{3.41}$$

where $\lambda_{\mathbf{w}} > 0$ is a scaling factor that depends on $\mathbf{w}$. By linearity this implies that

$$\Gamma \mathbf{w}' = \lambda_{\mathbf{w}'} \mathbf{S} \qquad \forall\, \mathbf{w}' \in \partial\mathcal{W}_{\mathfrak{m}} \tag{3.42}$$

where $\partial\mathcal{W}_{\mathfrak{m}} = \overline{\mathcal{W}}_{\mathfrak{m}} \setminus \mathcal{W}_{\mathfrak{m}}$ denotes the boundary of $\mathcal{W}_{\mathfrak{m}}$, and $\lambda_{\mathbf{w}'} \geq 0$ is a new rescaling factor which depends on $\mathbf{w}'$. Notice that in general $\lambda_{\mathbf{w}'}$ can now vanish. However, any $\mathbf{w} \in \overline{\mathcal{W}}_{\mathfrak{m}}$ can be written as a conical combination of the extreme rays of $\overline{\mathcal{W}}_{\mathfrak{m}}$. Therefore, since for any $\mathbf{w}$ we have $\lambda_{\mathbf{w}} > 0$, there must exist at least one extreme ray $\mathbf{x}_{\mathcal{W}_{\mathfrak{m}}} \in \partial\mathcal{W}_{\mathfrak{m}}$ such that if we apply (3.42), with $\mathbf{w}' = \mathbf{x}_{\mathcal{W}_{\mathfrak{m}}}$, we have $\lambda_{\mathbf{w}'} > 0$.

We now consider one the extreme rays which satisfy this condition, and we have to distinguish two cases, depending on whether $\mathbf{x}_{\mathcal{W}_{\mathfrak{m}}}$ has some vanishing components or not.

If $\mathbf{x}_{\mathcal{W}_{\mathfrak{m}}}$ has no vanishing components, it follows from Lemma 3.7 that it is by itself a W-cell $\mathcal{W}_{\hat{\mathfrak{m}}}$ for some min-cut structure $\hat{\mathfrak{m}}$ on G_{N}. Furthermore, by (3.42), $\mathbf{x}_{\mathcal{W}_{\mathfrak{m}}}$ is mapped by Γ to the same S-cell as all other weight vectors in $\mathcal{W}_{\mathfrak{m}}$. Therefore $(\hat{G}_{\mathsf{N}}, \hat{\mathfrak{m}})$, with $\hat{G}_{\mathsf{N}} = G_{\mathsf{N}}$, has the same min-cut subspace as $(G_{\mathsf{N}}, \mathfrak{m})$.

On the other hand, if $\mathbf{x}_{\mathcal{W}_{\mathfrak{m}}}$ has one or more vanishing components, we can first use the reduction described in Section 3.3.4 (we simply delete the edges with vanishing weights) to obtain a new graph $\hat{G}_{\mathsf{N}}$ and a new ray $\hat{\mathbf{x}}_{\mathcal{W}_{\mathfrak{m}}}$ which has no vanishing weights. Since $\hat{\mathbf{x}}_{\mathcal{W}_{\mathfrak{m}}}$ is determined by all the original degeneracy equations and inequalities that determined

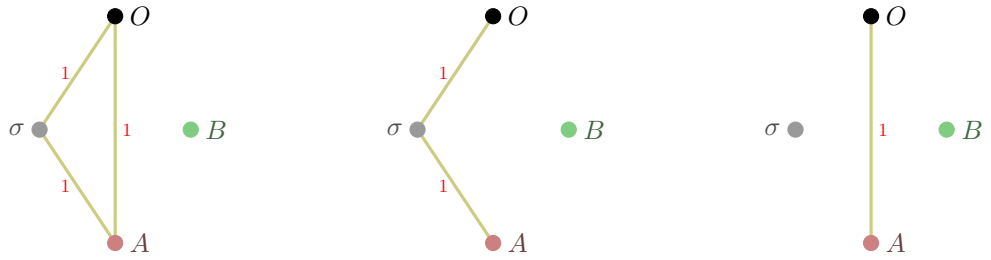


Figure 3.2: An example illustrating Lemma 3.8 and the reduction described in Section 3.3.4 for vanishing weights. The first graph is our starting point. The other two graphs are obtained by deleting from the first one the edges that correspond to the vanishing entries of the extreme rays of the closure of the W-cell specified by it, cf. (3.47).

$\mathbf{x}_{\mathcal{W}_\mathfrak{m}}$, now adapted to the new graph, it is still an extreme ray of the closure of a W-cell $\mathcal{W}_{\hat{\mathfrak{m}}}$ for a min-cut structure $\hat{\mathfrak{m}}$ on $\hat{G}_\mathsf{N}$, and therefore by Lemma 3.7 a W-cell by itself. Finally, we just need to verify that any new map $\hat{\Gamma}$ defined for $\mathcal{W}_{\hat{\mathfrak{m}}}$ on $\hat{G}_\mathsf{N}$ will map $\hat{\mathbf{x}}_{\mathcal{W}_{\hat{\mathfrak{m}}}}$ to the same $\mathbf{S}$. To see this, notice that starting from our initial choice of Γ for the class $(G_\mathsf{N}, \mathfrak{m})$, we can obtain a valid choice of $\hat{\Gamma}$ for $(\hat{G}_\mathsf{N}, \hat{\mathfrak{m}})$ by deleting the columns corresponding to the edges that we have removed from G_N. The equality

$$\hat{\Gamma}\, \hat{\mathbf{x}}_{\mathcal{W}_{\hat{\mathfrak{m}}}} = \Gamma\, \mathbf{x}_{\mathcal{W}_\mathfrak{m}} \tag{3.43}$$

then simply follows from the fact that the columns which have been removed from Γ are precisely the columns which were multiplied by the vanishing components of $\mathbf{x}_{\mathcal{W}_\mathfrak{m}}$. □

As an example of this lemma and the reduction procedure for vanishing weights presented in Section 3.3.4, consider the first graph model in Figure 3.2. The min-cut structure specified by the chosen weights is

$$\mathcal{U}_A = \{\{A\}, \{A, \sigma\}\}, \qquad \mathcal{U}_B = \{\{B\}\}, \qquad \mathcal{U}_{AB} = \{\{A, B\}, \{A, B, \sigma\}\} \tag{3.44}$$

and there is only one degeneracy equation (it is the same for the indices A and AB)

$$w_{AO} + w_{A\sigma} = w_{AO} + w_{O\sigma} \tag{3.45}$$

The solution to this degeneracy equation is the 2-dimensional subspace

$$\mathbb{W} = (0, 1, -1)^{\perp} \subset \mathbb{R}^3 \tag{3.46}$$

where we ordered the weights according to $(AO, A\sigma, O\sigma)$. The W-cell is the interior of the following polyhedral cone in $\mathbb{W}$ (written as embedded in $\mathbb{R}^3_{>0}$)

$$\text{cone}\,\{(0, 1, 1), (1, 0, 0)\} \tag{3.47}$$

Its image under the map

$$\Gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \tag{3.48}$$

fixed by the choice of representatives $U^*_A = \{A\}$ and $U^*_{AB} = \{A, B, \sigma\}$, is the S-cell, which is just the single ray (it is the extreme ray of the HEC_2 corresponding to a Bell pair for the subsystem AO)

$$\mathcal{S} = \lambda\,(1, 0, 1), \qquad \lambda > 0 \tag{3.49}$$

since both elements of (3.47) are mapped to the same vector by (3.48).

We then have a situation where an equivalence class has a 1-dimensional min-cut subspace and a 2-dimensional W-cell, and we can therefore apply Lemma 3.8. The extreme rays of the closure of this W-cell are given in (3.47), and since neither of them is nowhere-zero, we first need to apply the reduction described in Section 3.3.4. Deleting

from the first graph in Figure 3.2 the edges corresponding to the vanishing entries of the vectors in (3.47), we obtain the other graphs shown in Figure 3.2. The non-zero entries of the same vectors specify the weights of the remaining edges. The W-cell of the min-cut structure for the last graph in Figure 3.2 is obviously 1-dimensional (since $\mathsf{E} = 1$), in agreement with Lemma 3.8. On the other hand, for the middle graph in Figure 3.2, the space of edge weights is 2-dimensional. However, there is still a degeneracy, since there are two options for the min-cut for A, namely $\{A\}$ and $\{A, \sigma\}$. Therefore, also for this graph, the W-cell is 1-dimensional, again in agreement with Lemma 3.8. Notice that the generators of these 1-dimensional W-cells are obtained by deleting the vanishing components of the vectors in (3.47), and thus are simply $(1, 1)$ and (1). Finally, the new maps $\hat{\Gamma}$ for the new min-cut structures on the new graphs in Figure 3.2 are respectively

$$\hat{\Gamma} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \hat{\Gamma} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{3.50}$$

which are obtained by deleting the columns of Γ from (3.48) which correspond to the vanishing components of the vectors in (3.47). It is then immediate to verify that once applied to the aforementioned generators of the new 1-dimensional W-cells, these maps give the same S-cell (3.49) as the original graph from Figure 3.2, and therefore also the same-min-cut subspace.

Similarly to W-cells, S-cells also correspond to the interior of a polyhedral cone, and it is again interesting to explore situations where an (open) extreme ray of the closure of an S-cell is an S-cell by itself.[51] We will not answer this question in general, but we

[51] Notice that since S-cells do not form a partition of the HEC, an (open) extreme ray of the closure of an S-cell can in principle be an S-cell by itself, even if it is also in the interior of a different S-cell. This type of questions, like the one mentioned in footnote 50, is related to investigations of substructures of the HEC which are beyond the scope of this work.

will construct a class of examples where this is the case,[52] starting from the prototypical situations where S-cells are 1-dimensional, i.e., graph models realizing the extreme rays of the HEC.

Let us begin by first proving a basic fact about a given graph model $\widetilde{G_{\mathsf{N}}}$ and the face[53] of the $\mathrm{HEC}_{\mathsf{N}}$ that contains the entropy vector of $\widetilde{G_{\mathsf{N}}}$:

Theorem 3.1. *Given a graph model $\widetilde{G_{\mathsf{N}}}$ and a face $\overline{\mathcal{F}}_d$ of the $\mathrm{HEC}_{\mathsf{N}}$ such that the entropy vector $\mathbf{S}(\widetilde{G_{\mathsf{N}}})$ belongs to $\mathcal{F}_d$, the minimal supporting linear subspace $\mathbb{F}_d$ of $\overline{\mathcal{F}}_d$ contains the min-cut subspace $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})[\widetilde{G_{\mathsf{N}}}]$.*

Proof. Given a graph model $\widetilde{G_{\mathsf{N}}}$, all entropy vectors in the S-cell $\mathcal{S}$ of $(G_{\mathsf{N}}, \mathfrak{m})[\widetilde{G_{\mathsf{N}}}]$ can be realized by simply varying the weights in $\widetilde{G_{\mathsf{N}}}$, so clearly $\mathcal{S} \subseteq \mathrm{HEC}_{\mathsf{N}}$. Furthermore, $\mathcal{S}$ is an open set, since it is the image under a linear map of an open set in the space of weights (the W-cell). By definition, a face $\overline{\mathcal{F}}_d$ of the $\mathrm{HEC}_{\mathsf{N}}$ is $\overline{\mathcal{F}}_d = \mathbb{H} \cap \mathrm{HEC}_{\mathsf{N}}$, where $\mathbb{H}$ is a certain hyperplane in $\mathbb{R}^{\mathsf{D}}$ such that the $\mathrm{HEC}_{\mathsf{N}}$ is entirely contained in one of the two half-spaces specified by $\mathbb{H}$ and in $\mathbb{H}$ itself. Since $\mathcal{S}$ is open, and by assumption $\mathcal{F}_d \cap \mathcal{S}$ is non-empty, it must be that $\mathcal{S} \subset \overline{\mathcal{F}}_d$, otherwise there would exist elements of $\mathcal{S}$ on both sides of $\mathbb{H}$, contradicting the fact that $\mathcal{S} \subseteq \mathrm{HEC}_{\mathsf{N}}$ and $\overline{\mathcal{F}}_d$ is a face of the $\mathrm{HEC}_{\mathsf{N}}$. Therefore $\mathbb{S} = \mathrm{Span}(\mathcal{S}) \subset \mathrm{Span}(\overline{\mathcal{F}}_d) = \mathbb{H}$. □

This immediately leads to the following corollary:

Corollary 3.1. *A graph model $\widetilde{G_{\mathsf{N}}}$ realizes an extreme ray of the $\mathrm{HEC}_{\mathsf{N}}$ only if the min-cut subspace $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})[\widetilde{G_{\mathsf{N}}}]$ is 1-dimensional.*

Proof. This is just Theorem 3.1 applied to the specific case where the face $\overline{\mathcal{F}}_1$ is an extreme ray of the $\mathrm{HEC}_{\mathsf{N}}$. □

[52] One can construct these examples by building disconnected graphs starting from graph realizations of the extreme rays of the HEC (see below).

[53] Recall that a face can have any dimension $0 \leq d \leq \mathsf{D}$.

This result will play a central role in the reconstruction of the HEC discussed in Section 3.6. It can be seen to be equivalent to Theorem 2(b) of [8], where it was proven using a different setting wherein 1-dimensional min-cut subspaces constitute the extreme rays of a polyhedral cone whose facet description is known. Because of this, were the converse of Corollary 3.1 to hold, it would provide an explicit derivation of the HEC via direct computation of its extreme rays. Unfortunately though, the converse of Corollary 3.1 is in fact false. Namely, a graph model with a 1-dimensional min-cut subspace is not guaranteed to yield an extreme ray of the HEC, as also shown by [8] through a counterexample in appendix B therein.[54] An example of Corollary 3.1 is instead the graph model $\widetilde{G_3}$ from Section 3.3.3 which realizes the 4-party perfect state extreme ray of the HEC_3 (cf. Figure 3.1).

3.3.6 Disconnected graphs

In this subsection we comment on the construction of a topological graph model and min-cut structure $(G_{\mathsf{N}}, \mathfrak{m})$ via the disjoint union of simpler *building blocks*, and show how the W-cell, S-cell and min-cut subspace of the new graph can be obtained from those of its components. Conversely, the same analysis also clarifies how given an arbitrary class $(G_{\mathsf{N}}, \mathfrak{m})$ where G_{N} is disconnected, one can decompose these structures according to the connected components of G_{N}.

Let us begin with the simple situation where we consider an arbitrary collection of k N-party topological graph models and min-cut structures

$$\left\{\left(G^i_{\mathsf{N}}, \mathfrak{m}^i\right), \forall i \in [k]\right\} \tag{3.51}$$

[54] See [8] for more details on how Corollary 3.1 can nonetheless be used to construct the HEC.

where the boundary vertices of all topological graph models in the collection are colored with the same set of colors $[\mathsf{N}+1]$. We then construct a new topological graph model

$$G_{\mathsf{N}} = \bigoplus_{i\in[k]} G^i_{\mathsf{N}} \tag{3.52}$$

and min-cut structure

$$\mathfrak{m} = \{\bigcup_{i\in[k]} \mathcal{U}^i_{\mathcal{I}} \text{ for all } \mathcal{I}\} \tag{3.53}$$

The space of weights of the new graph is the direct sum of the weight spaces of the individual graphs

$$\mathbb{R}^{\mathsf{E}} = \bigoplus_{i\in[k]} \mathbb{R}^{\mathsf{E}^i} \tag{3.54}$$

and since the weights on any given graph can be varied independently from those of the others, the W-cell is also a *direct sum*[55] of W-cells

$$\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m}) = \bigoplus_{i\in[k]} \mathcal{W}(G^i_{\mathsf{N}}, \mathfrak{m}^i) \tag{3.55}$$

The new S-cell on the other hand is not the direct sum of the individual S-cells, because the W-cells are not mapped to orthogonal subspaces of entropy space. However, since the weights can still be varied independently, the S-cell is the *Minkowski sum*[56] of the individual S-cells

$$\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m}) = \bigdotplus_{i\in[k]} \mathcal{S}(G^i_{\mathsf{N}}, \mathfrak{m}^i) \tag{3.56}$$

[55]The direct sum of cones is the special case of the Minkowski sum (see below) where the individual cones are contained in orthogonal subspaces.

[56]The Minkowski sum of two sets of vectors X, Y is the set of sums $\mathbf{x}+\mathbf{y}$ for all $\mathbf{x}\in X$, $\mathbf{y}\in Y$.

and the min-cut subspace is simply the sum of the individual subspaces

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \bigdotplus_{i\in[k]} \mathbb{S}(G^i_{\mathsf{N}}, \mathfrak{m}^i) \tag{3.57}$$

In the particular case where the individual graphs realize 1-dimensional min-cut subspaces, or equivalently where their S-cells are single rays, the S-cell of the composite graph is simply the conical hull of such rays.[57] A special instance of this construction is when the individual S-cells correspond to extreme rays of the $\mathrm{HEC}_{\mathsf{N}}$. For example, by considering the full list of extreme rays one can construct a new graph whose S-cell is the interior of the $\mathrm{HEC}_{\mathsf{N}}$. But by taking different collections of extreme rays, one can also use this procedure to construct graphs whose S-cells overlap, clarifying as we anticipated in the previous section that S-cells do not form a partition of the $\mathrm{HEC}_{\mathsf{N}}$.[58] Moreover, the same type of construction can also be used to generate graphs with different S-cells but same min-cut subspace, for example by considering two collections of extreme rays which span the supporting subspace of a given face of the $\mathrm{HEC}_{\mathsf{N}}$.

We now want to extend this construction to more general situations, where the boundary vertices of each building block may be colored by a different set of colors. Notice that in the most general scenario the sets of colors of any two building blocks may or may not intersect, even if the boundary vertices of the two topological graph models are labeled by the same number of colors.[59] To deal with this type of situations, it will

[57] As exemplified for Corollary 3.1, for each component graph there is only a single free variable, i.e., a global rescaling of the weights. The fact that the rays are open and the rescaling factors have to be strictly positive implies that the resulting S-cell is an open set (as it should be).

[58] For this last construction the cone has to be non-simplicial, i.e., the number of extreme rays should be strictly larger than its dimension, which happens to be the case for any $\mathsf{N} \geq 4$.

[59] For example, we could have a topological graph model $G_2 = G_2^1 \oplus G_2^2$ where the boundary vertices of G_2^1 are labeled by $\{A, B, O\}$ and those of G_2^2 by $\{A, C, O\}$. Furthermore, if we imagine that the collection of building blocks is the set of connected components of a larger topological graph model that we want to decompose, there is only a single purifier, which does not even need to appear in each component.

be convenient to transform all the building blocks in such a way that their boundary vertices are labeled by precisely the same set of colors. This operation however changes the number of parties, and therefore the dimension of entropy space of a given building block. Because of this, before we can apply the procedure described above, we first need to make sure that we are working in the same space, and we will achieve this by embedding the entropy space of a building block into the entropy space of the entire collection in a precise way.[60]

At same time, the transformation of a given building block should be designed in such a way that the relevant data is not altered in any way, i.e., the same embedding should map the S-cell and min-cut subspace of the original building block to the new the S-cell and min-cut subspace obtained after the transformation. Since this simple construction will be particularly useful in later sections, we explain it in detail for a single graph.

Consider a topological graph model G_{N} (which by itself does not have to be connected), and suppose we want to use it as a building block in a collection that comprises N' colors, with $\mathsf{N}' > \mathsf{N}$. We define:

Definition 3.6 (Standard lift of a topological graph model)**.** *The standard lift of a topological graph model G_{N} to N' parties, with $\mathsf{N}' > \mathsf{N}$, is the topological graph model $G'_{\mathsf{N}'}$ obtained from G_{N} by first labeling by $\mathsf{N}'+1$ all the boundary vertices in G_{N} which were initially labeled by $\mathsf{N}+1$, and then adjoining $\mathsf{N}'-\mathsf{N}$ disconnected boundary vertices labeled by the new colors in $\{\mathsf{N}+1,\ldots,\mathsf{N}'\}$.*

A standard lift of a topological graph model G_{N} to N' parties is naturally associated to the following embedding of $\mathbb{R}^{\mathsf{D}}$ into $\mathbb{R}^{\mathsf{D}'}$

[60] Of course this step is only necessary in the composition of building blocks, and not when we want to analyse the components of a given, disconnected, topological graph model.

Definition 3.7 (Standard embedding)**.** *Given* N *and* $\mathsf{N}' > \mathsf{N}$*, the standard embedding of* $\mathbb{R}^{\mathsf{D}}$ *into* $\mathbb{R}^{\mathsf{D}'}$ *is the embedding specified by the following equations*

$$S_{\mathcal{J}'} = \begin{cases} S_{\mathcal{J}' \cap [\mathsf{N}]} & \forall\, \mathcal{J}' \text{ s.t. } \mathcal{J}' \cap [\mathsf{N}] \neq \varnothing \\ 0 & \text{otherwise} \end{cases} \tag{3.58}$$

We leave it as an exercise for the reader to verify that given a topological graph model G_{N} and its standard lift to N' parties, the S-cell of any min-cut structure, is the standard embedding of the S-cell for the equivalent min-cut structure on G_{N}. Notice that this also implies the analogous result for min-cut subspaces.

Given an arbitrary collection of building blocks, we can then apply the standard lift procedure to each one of them, to obtain a new (equivalent) collection where all building blocks have precisely the same set of colors. The W-cells, S-cells and min-cut subspaces of the various min-cut structures can then be obtained as described above. Finally, given the results of this subsection, in what follows we will typically focus our attention on connected graphs (most notably in Section 3.4.3). It should however always be clear that we can use the construction presented here to generalize any result obtained for connected graph to this more general scenario.

3.4 Min-cut subspaces from marginal independence

In the preceding section we have defined a particular subspace of the entropy space: the min-cut subspace of a graph model. This subspace provides a key construct to characterize the structure of the HEC since it naturally encapsulates the facets, as well as the extreme rays, of the cone. In the present section, we consider another class of subspaces of the entropy space: the "patterns of marginal independence" (PMI). As we

will see, these new subspaces typically contain the min-cut subspaces properly, and one may expect them to be too coarse. Instead, it will turn out that these subspaces are the ones that distill the essential information.

We begin by briefly reviewing in Section 3.4.1 the formal definition of "patterns of marginal independence" from [12]. We then define "holographic PMIs" for min-cut structures on topological graph models in Section 3.4.2. Finally in Section 3.4.3 we show that when a topological graph model has the topology of a tree and all boundary vertices have a different color, the min-cut subspace for any min-cut structure coincides with the PMI. A more general analysis of the relation between min-cut subspaces and PMIs, which involves recolorings of boundary vertices, will be carried out in Section 3.5.

3.4.1 Review of marginal independence and PMIs

Consider an arbitrary N-partite quantum system, described for convenience by a pure state $|\psi\rangle$ on a Hilbert space with $\mathsf{N}+1$ tensor factors, cf. (3.5). Given two non-overlapping subsystems $\underline{\mathcal{I}}, \underline{\mathcal{K}}$ we can measure the total amount of correlation between the corresponding marginals $\rho_{\underline{\mathcal{I}}}, \rho_{\underline{\mathcal{K}}}$ by evaluating their *mutual information*[61]

$$\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) \coloneqq S_{\underline{\mathcal{I}}} + S_{\underline{\mathcal{K}}} - S_{\underline{\mathcal{I}}\cup\underline{\mathcal{K}}} \ . \tag{3.59}$$

If and only if the mutual information vanishes, the two subsystems are *independent* and we have the factorization

$$\rho_{\underline{\mathcal{I}}\cup\underline{\mathcal{K}}} = \rho_{\underline{\mathcal{I}}} \otimes \rho_{\underline{\mathcal{K}}} \ . \tag{3.60}$$

[61] By non-overlapping we mean $\underline{\mathcal{I}} \cap \underline{\mathcal{K}} = \varnothing$. Notice that if $\underline{\mathcal{I}} \cup \underline{\mathcal{K}} = [\mathsf{N}+1]$ the mutual information reduces to twice the entropy of either subsystem.

For any given N-partite state, it is straightforward to determine which pairs of subsystems $(\underline{\mathcal{I}}, \underline{\mathcal{K}})$ are independent, since one simply has to compute all possible instances of the mutual information.

Conversely, one can imagine specifying a certain pattern of independences by listing which pairs of subsystems are independent and which pairs are not (notice that a pattern is "complete", in the sense that each instance of the mutual information is demanded to either vanish or not). For a given pattern, one can then ask if there exists a density matrix which realizes it. This problem was first introduced in [12] and dubbed the *marginal independence problem*.

One obvious restriction to the set of realizable patterns simply comes from the linear dependences among various instances of the mutual information. Since the total number of instances is greater than the dimension of entropy space, the instances are linearly dependent. Writing a particular instance as a linear combination of other instances

$$\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) = \sum_{\alpha} \mathsf{c}_{\alpha}\, \mathsf{I}(\underline{\mathcal{I}}^{\alpha} : \underline{\mathcal{K}}^{\alpha}) \tag{3.61}$$

one immediately sees that if

$$\mathsf{I}(\underline{\mathcal{I}}^{\alpha} : \underline{\mathcal{K}}^{\alpha}) = 0 \qquad \forall \alpha \tag{3.62}$$

then it must be that $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) = 0$. This type of implications restricts the set of meaningful patterns, since a pattern that requires all instances in (3.62) to vanish and $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}})$ to be non-vanishing clearly can never be realized.

A separate kind of restriction comes from physical constraints such as subadditivity. To see how instances of SA constrain the set of possibly realizable patterns, we can

rewrite (3.61) as

$$\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) = \sum_{\alpha_+} \mathsf{c}_{\alpha_+} \mathsf{I}(\underline{\mathcal{I}}^{\alpha_+} : \underline{\mathcal{K}}^{\alpha_+}) + \sum_{\alpha_-} \mathsf{c}_{\alpha_-} \mathsf{I}(\underline{\mathcal{I}}^{\alpha_-} : \underline{\mathcal{K}}^{\alpha_-}), \tag{3.63}$$

where c_{α_+} (c_{α_-}) refer to the positive (negative) coefficients c_α. Suppose now that we partially specify a pattern by requiring that all c_{α_+} terms in (3.63) vanish. Since SA implies that for any state each instance of the mutual information is non-negative, the only way to satisfy (3.63) without violating SA is by also requiring that all other instances appearing in (3.63) vanish.

Given the above fundamental restrictions,[62] [12] defined the marginal independence problem by only considering patterns of independences which are consistent with linear dependences among the instances of the mutual information, and all the instances of SA. Geometrically, one can think of these consistent patterns as corresponding to the faces of the polyhedral cone in entropy space defined by subadditivity. More precisely, one first defines the *subadditivity cone* (SAC) as follows (we already mentioned this object in Section 3.2.1 but we repeat the precise definition here for convenience)

Definition 3.8 (SAC)**.** *The* N*-party* subadditivity cone ($\mathrm{SAC}_{\mathsf{N}}$) *is the polyhedral cone in* $\mathbb{R}^{\mathsf{D}}$ *obtained from the intersection of all the half-spaces specified by the inequalities* $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) \geq \mathsf{0}$ *for all pairs* $(\underline{\mathcal{I}}, \underline{\mathcal{K}})$ *of non-intersecting subsystems.*

One then defines the *patterns of marginal independence* (PMI) as specific linear subspaces of the entropy space which naturally characterize the instances of vanishing mutual information:

[62] There are of course further physical restrictions such as SSA, but it will turn out that considering just SA is particularly useful.

Definition 3.9 (Pattern of marginal independence)**.** *A* pattern of marginal independence (PMI) *is the linear supporting subspace* $\mathbb{P}$ *of a face of the subadditivity cone.*

The reason behind this definition is the following: any PMI is now a geometric object, corresponding to the intersection of a certain set of hyperplanes of the form $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) = 0$, and by construction this set respects the linear dependences among the hyperplanes in the sense described above. Furthermore, since we are only considering subspaces that correspond to faces of the SA cone, any PMI is guaranteed to contain a region of entropy space (the face) such that all entropy vectors in this region respect all instances of SA. All naive patterns which do not respect the linear dependences among the instances of the mutual information, or that do not include any (non-trivial) region of entropy space whose elements respect all instances of SA, are automatically excluded. Notice in particular that all 1-dimensional PMIs are generated by the extreme rays of the $\mathrm{SAC}_{\mathsf{N}}$.

Even if Definition 3.9 excludes a large set of meaningless patterns, it does not guarantee that each PMI contains at least on entropy vector that can be realized by a density matrix. We will say that a PMI $\mathbb{P}$ is *realizable* if there exists a density matrix ρ such that the entropy vector $\mathbf{S}(\rho)$ belongs to $\mathbb{P}$ but not to any lower dimensional subspace $\mathbb{P}' \subset \mathbb{P}$. Notice that the fact that a PMI is realizable does not imply that each vector $\mathbf{S} \in \mathbb{P}$ (even within the intersection of $\mathbb{P}$ and the SAC) is the entropy vector of some density matrix. For any PMI $\mathbb{P}$ *realized* by a density matrix ρ, we will denote by π the map which associates $\mathbb{P}$ to ρ, $\mathbb{P} = \pi(\rho)$. Furthermore, for any PMI $\mathbb{P}$, we denote by $\Pi(\mathbb{P})$ the matrix such that

$$\mathrm{Ker}\ \Pi^{\intercal} = \mathbb{P} \tag{3.64}$$

In other words, the columns of Π are the coefficients of the instances of the vanishing mutual information in $\mathbb{P}$. For example, for $\mathsf{N} = 3$, the 1-dimensional PMI generated by

the perfect state extreme ray of (3.36) corresponds to the matrix

$$\Pi = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \tag{3.65}$$

Occasionally we will informally say that an instance of the mutual information $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}})$ is "in a PMI $\mathbb{P}$", meaning that $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}}) = 0$ is one of the hyperplanes that determine $\mathbb{P}$, or equivalently, that the vector normal to this hyperplane is one of the columns of $\Pi(\mathbb{P})$.

In general, the marginal independence problem asks which PMIs are realizable by a given class of states. In the context of the present work, the states of interest are the geometric states in holographic theories. We will argue in Section 3.6 that the solution to this *holographic* marginal independence problem (HMIP) provides sufficient information to reconstruct the holographic entropy cone. But in order to do this, we first need to clarify what we mean by "holographic PMI" and then to establish a connection between PMIs and min-cut subspaces of graph models.

3.4.2 Patterns of marginal independence for graph models

Having reviewed the definition of a pattern of marginal independence for arbitrary quantum states, we will now introduce a similar definition for graph models, and explain some of its basic properties.

As for quantum states, any graph model $\widetilde{G_{\mathsf{N}}}$ gives an entropy vector, and it is therefore straightforward to determine the corresponding PMI, defined as follows

Definition 3.10 (PMI of a graph model)**.** *Given a graph model* $\widetilde{G_{\mathsf{N}}}$*, its* PMI $\mathbb{P} = \pi(\widetilde{G_{\mathsf{N}}})$ *is the* N*-party* PMI *of smallest dimension that contains the entropy vector* $\mathbf{S}(\widetilde{G_{\mathsf{N}}})$*.*

However, we are as usual interested in min-cut structures on topological graph models, rather than in specific graph models, but in order to be able to work with PMIs of min-cut structures, we first need to check that this is a well defined concept. Let us first recall the necessary and sufficient conditions for an instance of the mutual information to vanish in a graph model:[63]

Lemma 3.9. *Given a graph model* $\widetilde{G_{\mathsf{N}}}$ *and two subsystems* $\underline{\mathcal{I}}$ *and* $\underline{\mathcal{K}}$ *with* $\underline{\mathcal{I}} \cap \underline{\mathcal{K}} = \varnothing$*, the mutual information* $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}})$ *vanishes if and only if there exist min-cuts* $U^*_{\underline{\mathcal{I}}}$*,* $U^*_{\underline{\mathcal{K}}}$ *and* $U^*_{\underline{\mathcal{I}}\cup\underline{\mathcal{K}}}$ *such that*

$$G[U^*_{\underline{\mathcal{I}}\cup\underline{\mathcal{K}}}] = G[U^*_{\underline{\mathcal{I}}}] \oplus G[U^*_{\underline{\mathcal{K}}}] \tag{3.66}$$

This implies that the vanishing of any instance of the mutual information is "detected" by the min-cut structure and that one can therefore think of a PMI as being determined directly by the min-cut structure, specifically:

Corollary 3.2. *For any topological graph model* G_{N} *and min-cut structure* $\mathfrak{m}$ *we have*

$$\pi(\widetilde{G_{\mathsf{N}}}) = \pi(\widetilde{G_{\mathsf{N}}}') \qquad \forall\ \widetilde{G_{\mathsf{N}}}, \widetilde{G_{\mathsf{N}}}' \in (G_{\mathsf{N}}, \mathfrak{m}) \tag{3.67}$$

Proof. Given $(G_{\mathsf{N}}, \mathfrak{m})$ consider any two graph models $\widetilde{G_{\mathsf{N}}}$ and $\widetilde{G_{\mathsf{N}}}'$ in $(G_{\mathsf{N}}, \mathfrak{m})$ and any instance $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}})$ which vanishes for $\pi(\widetilde{G_{\mathsf{N}}})$. By Lemma 3.9 there exist min-cuts $W_{\underline{\mathcal{I}}}$, $W_{\underline{\mathcal{K}}}$ and $W_{\underline{\mathcal{I}}\cup\underline{\mathcal{K}}}$ in $\mathfrak{m}$ such that (3.66) holds, and by the same lemma $\mathsf{I}(\underline{\mathcal{I}} : \underline{\mathcal{K}})$ is also one of the instances that vanish for $\pi(\widetilde{G_{\mathsf{N}}}')$. □

[63] For holographic configurations this is a well know consequence of the HRRT formula.

Having showed that the PMI of a min-cut structure is a well defined concept, from now on we will always work with these objects and denote them by $\pi(G_{\mathsf{N}}, \mathfrak{m})$

$$\pi(G_{\mathsf{N}}, \mathfrak{m}) \coloneqq \pi(\widetilde{G_{\mathsf{N}}}) \qquad \text{for any} \quad \widetilde{G_{\mathsf{N}}} \in (G_{\mathsf{N}}, \mathfrak{m}) \tag{3.68}$$

Nevertheless, recall that not all PMIs are realizable holographically (or even by arbitrary quantum states). This prompts us to introduce the definition of a holographic PMI as one which can be realized by a graph model $\widetilde{G_{\mathsf{N}}}$. In terms of equivalence classes we define it as follows:

Definition 3.11 (Holographic PMI)**.** *A* PMI $\mathbb{P}$ *is* holographic *if there exists a topological graph model* G_{N} *and a min-cut structure* $\mathfrak{m}$ *such that*

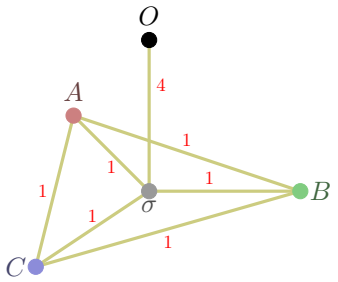
$$\mathbb{P} = \pi(G_{\mathsf{N}}, \mathfrak{m}) \tag{3.69}$$

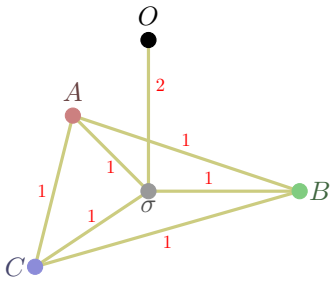
Having defined the PMI of a class $(G_{\mathsf{N}}, \mathfrak{m})$, a natural question to ask is how is such a PMI related to the min-cut subspace. In general we have the following inclusion relation:

Lemma 3.10. *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$*, the* PMI $\pi(G_{\mathsf{N}}, \mathfrak{m})$ *is the lowest dimensional* N*-party* PMI *that contains the min-cut subspace*

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \subseteq \pi(G_{\mathsf{N}}, \mathfrak{m}) \tag{3.70}$$

Proof. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$, Corollary 3.2 implies that any graph model $\widetilde{G_{\mathsf{N}}}$ in this equivalence class has the same PMI, $\pi(\widetilde{G_{\mathsf{N}}}) = \pi(G_{\mathsf{N}}, \mathfrak{m})$. One then has the inclusion $\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m}) \subset \pi(G_{\mathsf{N}}, \mathfrak{m})$, and since $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$ is the linear span of $\mathcal{S}(G_{\mathsf{N}}, \mathfrak{m})$, that $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \subseteq \pi(G_{\mathsf{N}}, \mathfrak{m})$. Furthermore, by Definition 3.10, it also follows that $\pi(G_{\mathsf{N}}, \mathfrak{m})$ is the lowest dimensional N-party PMI such that this inclusion holds. □





$$\Gamma = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 3.3: An example of two graph models, with the same underlying topological model but different (generic) min-cut structures, corresponding to the same PMI but different min-cut subspaces. The min-cut structures are specified by the Γ matrices, where the rows are labeled by polychromatic indices $(A, B, C, AB, AC, BC, ABC)$ and the columns by the edges $(AB, AC, BC, A\sigma, B\sigma, C\sigma, O\sigma)$. In both cases the PMI is the full entropy space $\mathbb{R}^7$ (since no mutual information vanishes), but the graph on the left has the 6-dimensional min-cut subspace $\mathbb{S} = (1, 1, 1, -1, -1, -1, 1)^{\perp}$, while for the graph on the right $\mathbb{S} = \mathbb{R}^7$. Note that the 6-dimensional min-cut subspace of the graph on the left is the hyperplane defined by the vanishing of the tripartite information, corresponding to the saturation of MMI.

Let us now denote by $\mathbb{V}$ and $\mathbb{P}$ the min-cut subspace and PMI of $(G_{\mathsf{N}}, \mathfrak{m})$ respectively. One may wonder if there could exist another topological graph model G'_{N} and min-cut structure $\mathfrak{m}'$ such that $\pi(G'_{\mathsf{N}}, \mathfrak{m}') = \mathbb{P}$, while $\mathbb{S}(G'_{\mathsf{N}}, \mathfrak{m}') \neq \mathbb{V}$. This can easily happen, as exemplified in Figure 3.3. On the other hand, the opposite is not possible, as clarified by the following corollary:

Corollary 3.3. *Given any two classes* $(G_{\mathsf{N}}, \mathfrak{m})$ *and* $(G'_{\mathsf{N}}, \mathfrak{m}')$

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \mathbb{S}(G'_{\mathsf{N}}, \mathfrak{m}') \quad \Longrightarrow \quad \pi(G_{\mathsf{N}}, \mathfrak{m}) = \pi(G'_{\mathsf{N}}, \mathfrak{m}') \tag{3.71}$$

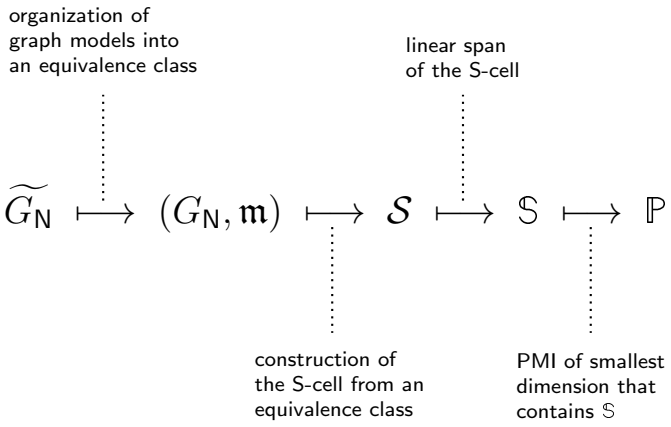


Figure 3.4: A summary of the various constructs that we have associated to graph models. Starting from a graph model $\widetilde{G_{\mathsf{N}}}$, each map along this chain associates to an object a coarser one. As we have exemplified throughout the text, none of these maps is injective.

Proof. By Lemma 3.10, the PMI of a class $(G_{\mathsf{N}}, \mathfrak{m})$ is the lowest dimensional N-party PMI that contains the min-cut subspace $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$, and is therefore uniquely fixed by such subspace. □

As a consequence of Corollary 3.3, a PMI $\mathbb{P}$ is completely determined by a min-cut subspace, and it is interesting to ask in what cases the min-cut subspace and PMI coincide. We will see in the next subsection that this is the case at least for a particular class of topological graph models called "simple trees".

But before we proceed to the next section, let us briefly pause to summarize the list of coarser and coarser objects that along the way we have introduced and associated to graph models. The landscape of these constructs and the maps between them is shown in Figure 3.4. We stress that, as we have clarified with various examples, none of these maps is injective.

3.4.3 The min-cut subspace of a simple tree graph

Having introduced the notion of the pattern of marginal independence for a min-cut structure on a topological graph model, we will now show that for a particular class of such models that we call *simple trees*, the pattern of marginal independence is equal to the min-cut subspace.

The key attribute of such graphs, as we argue below, is that each edge defines a cut for some subsystem, and can thus be naturally associated to the corresponding polychromatic index.[64] This will allow us to view any relation between the edge weights, which determines the min-cut subspace, purely in terms of subsystem entropies; this in turn can be recast in terms of mutual informations, and hence related to the PMI. Let us start from the basic definition

Definition 3.12 (Simple tree graph)**.** *A topological graph model G_{N} with the topology of a tree is* simple *if each boundary vertex is labeled by a different color.*

On a given simple tree, consider an arbitrary edge $e \in E$, and the partition of the vertex set into the two complementary subsets U and $U^{\complement}$ separated by e. If both U and $U^{\complement}$ contain at least one boundary vertex, then e corresponds to a bipartition $(\underline{\mathcal{I}}, \underline{\mathcal{I}}^{\complement})$ of $[\mathsf{N}+1]$ given by

$$\beta(\partial V \cap U) = \underline{\mathcal{I}} \tag{3.72}$$

By convention we define U to be the subset that does not include the purifier, and simply write $\beta(\partial V \cap U) = \mathcal{I}$ (not underlined). The subsystem $\mathcal{I}$ associated to an edge via this prescription will be denoted by $\mathcal{I}(e)$. For an arbitrary tree graph and choice of e, one may also have $\beta(\partial V \cap U) = \varnothing$. This can happen if one or more leaves are not boundary vertices. In this case we write $\mathcal{I}(e) = \varnothing$, with a little abuse of notation since

[64] There is a simple exception that, as we will shortly explain, is irrelevant.

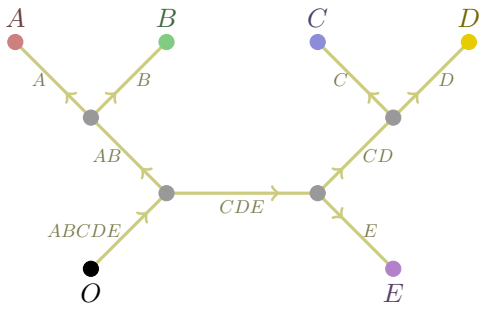


Figure 3.5: An example of a simple tree graph G_5, with explicit edge labeling and orientation indicated.

in this case $\mathcal{I}$ is not a proper polychromatic index according to our definition. While we take into account this possibility for the sake of completeness, notice that by topological minimality (cf., Lemma 3.1) no edge e with $\mathcal{I}(e) = \varnothing$ can belong to the set of cut edges for any min-cut.

With this convention at hand, it is also convenient to introduce a canonical orientation of the edges of the tree which will induce an inclusion relation among the indices $\mathcal{I}(e)$. Specifically, denoting by v_{N+1} the boundary vertex of the graph labeled by the purifier, and by v_L an arbitrary leaf, we consider the path from v_{N+1} to v_L and orient the edges to turn it into a directed path which starts at v_{N+1} and ends at v_L. We will denote such a directed path by $\mathcal{P}(v_{N+1}, v_L)$. By repeating this procedure for all leaves, we fix an orientation for the whole graph.

Using this orientation, we can then introduce a partial order on the set of edges. Given $e, f \in E$ we say that $e < f$ if there exists a leaf L such that $e, f \in \mathcal{P}(v_{N+1}, v_L)$ and e precedes f. We then have

$$\begin{aligned} e < f \quad &\Rightarrow \quad \mathcal{I}(e) \supseteq \mathcal{I}(f) \\ e, f \text{ incomparable} \quad &\Rightarrow \quad \mathcal{I}(e) \cap \mathcal{I}(f) = \varnothing \end{aligned} \tag{3.73}$$

where the equality $\mathcal{I}(e) = \mathcal{I}(f)$ can be attained if e and f are the edges adjoining to a degree-2 bulk vertex.[65] An example of a simple tree graph with the explicit edge labeling and orientation is shown in Figure 3.5.

This relation between edges and "homologous" subsystems allows us to conveniently translate the description of any $\mathcal{I}$-cut from the usual one based on vertices to a new one based on edges. Specifically, from the assumption of simplicity it follows that:

Lemma 3.11. *On a simple tree G_{N}, a non-empty collection of edges $X \subseteq E$ is the set of cut-edges $\mathcal{C}(U_{\mathcal{I}})$ for an $\mathcal{I}$-cut $U_{\mathcal{I}}$, where*

$$\mathcal{I} = \mathcal{I}(X) = \bigtriangleup_{e \in X} \mathcal{K}(e) \tag{3.74}$$

and $\bigtriangleup$ denotes the symmetric difference.[66]

Proof. Given a simple tree G_{N} and a non-empty collection of edges X, we start by constructing a cut U such that $\mathcal{C}(U) = X$. Consider an edge $e \in X$ and a leaf $v_{\scriptscriptstyle L}$ such that $e \in \mathcal{P}(v_{\mathsf{N}+1}, v_{\scriptscriptstyle L})$. Starting from $v_{\mathsf{N}+1}$, we follow the path $\mathcal{P}(v_{\mathsf{N}+1}, v_{\scriptscriptstyle L})$ and label each vertex by U or $U^{\complement}$ as follows. We label $v_{\mathsf{N}+1}$ by $U^{\complement}$ then we follow the path and continue labeling the vertices by $U^{\complement}$ until we reach an edge in X. After we cross the edge we

[65] Note that the implication in (3.73) applies in reverse as well, however most polychromatic subsystems do not have an associated edge (since a simple N-party tree graph has at most $2\mathsf{N}-1$ edges, as compared to $\mathsf{D} = 2^{\mathsf{N}} - 1$ polychromatic subsystems). In particular, two subsystems $\mathcal{I}(e)$ and $\mathcal{I}(f)$ can never be crossing (i.e. have a non-empty intersection which is a proper subset of both).

[66] For any pair of sets (X, Y), the symmetric difference is the disjunctive union

$$X \triangle Y = (X \setminus Y) \cup (Y \setminus X) = (X \cup Y) \setminus (X \cap Y) = Y \triangle X$$

Since the symmetric difference is associative, we can iterate this straightforwardly, so that for n-ary symmetric difference of a collection of sets we have just the elements which are in an odd number of the sets in that collection. For example, for $n = 3$

$$\bigtriangleup_{V \in \{X,Y,Z\}} V = X \triangle Y \triangle Z = (X \setminus (Y \cup Z)) \cup (Y \setminus (X \cup Z)) \cup (Z \setminus (X \cup Y)) \cup (X \cap Y \cap Z)\,.$$

label all the vertices by U until we reach another edge in X. We proceed in this fashion, alternating between $U^{\complement}$ and U each time we cross an edge in X until we reach v_L. Then we repeat the same procedure following other paths, until we have labeled all the vertices in the graph.

Having constructed the desired cut, we now need to determine $\mathcal{I}(X) = \beta(U \cap \partial V)$. For a color $\ell \in [\mathsf{N}]$, denote by v_ℓ the (unique by simplicity) vertex in G_{N} labeled by ℓ. From the construction of U described above, it follows that $\ell \in \mathcal{I}(X)$ if and only if the path $\mathcal{P}(v_{\mathsf{N}+1}, v_\ell)$ includes an odd number of edges in X. Furthermore, the index $\mathcal{K}(e)$ associated to an edge e can be seen as the set of colors labeling the boundary vertices that follow e in any path $\mathcal{P}(v_{\mathsf{N}+1}, v_\ell)$ that includes e. Therefore, a color ℓ appears in the expression at the right hand side of (3.74) if and only if the path $\mathcal{P}(v_{\mathsf{N}+1}, v_\ell)$ includes an odd number of edges in X, concluding the proof.

□

Using this translation from a description in terms of vertices to one in terms of edges, we can then prove a useful property of min-cuts on simple tree graphs which is reminiscent of Lemma 3.4.

Lemma 3.12 (Min-cut decomposition for simple trees)**.** *Let G_{N} be a simple tree with a min-cut structure $\mathfrak{m}$, $\mathcal{I}$ a subsystem, and $\mathcal{C}^*_{\mathcal{I}}$ the set of edges for some min-cut. Then any $X \subseteq \mathcal{C}^*_{\mathcal{I}}$ is the set of edges $\mathcal{C}^*_{\mathcal{K}} = X$ for a min-cut for the subsystem $\mathcal{K} = \mathcal{K}(X)$ given by (3.74).*

Proof. By Lemma 3.11, $X \subseteq \mathcal{C}^*_{\mathcal{I}}$ specifies a $\mathcal{K}$-cut for $\mathcal{K} = \mathcal{K}(X)$, so we just need to show it is minimal. Notice that, similarly, $Y = \mathcal{C}^*_{\mathcal{I}} \setminus X$ specifies a $\mathcal{J}$-cut for $\mathcal{J} = \mathcal{J}(Y)$, and that by assumption $\mathcal{C}^*_{\mathcal{I}} = X \cup Y$ are the min-cut edges for $\mathcal{I} = \mathcal{I}(X \cup Y) = \mathcal{K}(X) \Delta \mathcal{J}(Y)$. To show that X are min-cut edges for $\mathcal{K}$, assume for contradiction that its actual min-cut

edges are X', with weight $\|X'\| < \|X\|$. Then since $\mathcal{I} = \mathcal{K}(X')\Delta\mathcal{J}(Y)$, $X' \cup Y$ specifies an $\mathcal{I}$-cut, and its weight gives the desired contradiction $\|X' \cup Y\| < \|\mathcal{C}^*_{\mathcal{I}}\|$. □

In the particular case where a cut involves two edges, this in turn implies a particularly useful connection to the vanishing of an instance of the mutual information:

Lemma 3.13. *Given a simple tree G_{N}, a min-cut structure $\mathfrak{m}$ and any pair of edges $e, f \in E$, if $\{e, f\} = \mathcal{C}^*_{\mathcal{I}(e)\triangle\mathcal{I}(f)}$ (for some choice of representative min-cut in case of degeneracy), then*

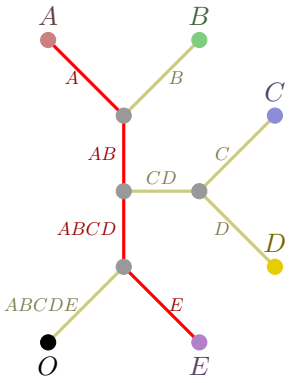
$$S_{\mathcal{I}(e)} + S_{\mathcal{I}(f)} - S_{\mathcal{I}(e)\triangle\mathcal{I}(f)} \tag{3.75}$$

is an instance of the mutual information in the PMI *of $(G_{\mathsf{N}}, \mathfrak{m})$.*

Proof. By Lemma 3.12 we have $\{e\} = \mathcal{C}^*_{\mathcal{I}(e)}$ and $\{f\} = \mathcal{C}^*_{\mathcal{I}(f)}$ (again for some choice of representatives in case of degeneracy) and the combination in (3.75) vanishes. Therefore all we need to show is that there exists a choice of underlined indices $\underline{\mathcal{J}}, \underline{\mathcal{K}}$ such that the expression in (3.59) is equal to (3.75). If e and f are incomparable, cf. (3.73), then $\mathcal{I}(e) \cap \mathcal{I}(f) = \varnothing$ which implies $\mathcal{I}(e)\triangle\mathcal{I}(f) = \mathcal{I}(e) \cup \mathcal{I}(f)$ and we can choose $\underline{\mathcal{J}} = \mathcal{I}(e)$ and $\underline{\mathcal{K}} = \mathcal{I}(f)$. If $e < f$ (if $f < e$ simply swap e and f in what follows), then $\mathcal{I}(e) \supset \mathcal{I}(f)$ which implies $\mathcal{I}(e)\triangle\mathcal{I}(f) = \mathcal{I}(e) \setminus \mathcal{I}(f)$ and we can choose $\underline{\mathcal{J}} = \mathcal{I}(e)^{\complement}$ and $\underline{\mathcal{K}} = \mathcal{I}(f)$. □

Notice that as exemplified in Figure 3.6 the implications of Lemmas 3.12 and 3.13 are stronger than what would follow from a straightforward iteration of Lemma 3.4.

We are now ready to establish the anticipated connection between min-cut subspaces and PMIs for simple tree graphs. As usual we will discuss the case of generic min-cut structures first, and then extend the proof to min-cut structures that might include degeneracies.



$\mathcal{L}_1$: $ACDE$

$\mathcal{L}_2$: ACD, CDE, AE, A, CD, E

$\mathcal{L}_3$: B, $ABCD$, AB, $ABCDE$

$\mathcal{L}_4$: BCD, ABE, $BCDE$, BE

Figure 3.6: An example of a simple tree with a choice of min-cut for $ACDE$, showing the different implications of Lemma 3.4 and Lemma 3.12. The min-cut for $ACDE$ is specified by its set of cut edges $\mathcal{C}^*_{ACDE}$, shown in red in the figure. We have labeled each edge e in the graph by the corresponding polychromatic index $\mathcal{I}(e)$. The right panel shows the full list of subsystems whose min-cuts are fixed by $\mathcal{C}^*_{ACDE}$ according to Lemma 3.12, each one corresponding to a subset of $\mathcal{C}^*_{ACDE}$ (including our starting choice $\mathcal{L}_1$). The min-cut for each subsystem in $\mathcal{L}_2$ can equivalently be obtained by a straightforward application of Lemma 3.4 (after some iteration). For the subsystems in $\mathcal{L}_3$ it is still sufficient to use Lemma 3.4, but one also needs to consider min-cuts for complementary subsystems, which include the purifier (for example, the min-cut for B is fixed by the fact that the min-cut for BEO, which is the complement of the min-cut for ACD, is disconnected). Even complementarity however is not sufficient to determine the min-cuts for the subsystems in $\mathcal{L}_4$ using Lemma 3.4. Notice that we also have for example $\mathsf{I}(AB : E) = 0$, which follows from Lemma 3.13 again via Lemma 3.12, but which is not implied by Lemma 3.4 (since ABE is in the set $\mathcal{L}_4$).

In the generic case, the proof will proceed by showing that due to the tree topology and simplicity, the set of vanishing instances of the mutual information completely determine the min-cut structure.

Lemma 3.14. *For any simple tree G_{N}, and any generic min-cut structure $\mathfrak{m}$ on G_{N}, the min-cut subspace $\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m})$ and the pattern of marginal independence $\pi(G_{\mathsf{N}}, \mathfrak{m})$ coincide.*

Proof. Consider an arbitrary simple tree G_{N} and an arbitrary, but generic, min-cut structure $\mathfrak{m}$. We denote by $\mathbb{S}$ and $\mathbb{P}$ the min-cut subspace and PMI of $(G_{\mathsf{N}}, \mathfrak{m})$ respectively, and by Γ the linear map defined in (3.20). To show that $\mathbb{S}$ and $\mathbb{P}$ coincide we only need to show that $\mathbb{S} \supseteq \mathbb{P}$, since by Lemma 3.10 we have $\mathbb{S} \subseteq \mathbb{P}$.

To show that $\mathbb{S} \supseteq \mathbb{P}$, we will prove that $\mathbb{S}^{\perp} \subseteq \mathbb{P}^{\perp}$, where $\perp$ denotes the orthogonal complement of a linear subspace. Since G_{N} is a simple tree, and $\mathfrak{m}$ is generic, the matrix which represents the map Γ can be put schematically in the following form

$$\Gamma = \begin{pmatrix} \widetilde{\Gamma} & \mathbb{0} & \mathbb{0} \\ \mathbb{1} & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} & \mathbb{0} \end{pmatrix} \tag{3.76}$$

simply by permuting its rows and columns. Here the first, second and third columns of this block matrix correspond respectively to the subsets of edges which participate in at least two cuts in $\mathfrak{m}$, precisely one cut, or no cut at all.

To see that Γ can be put into this form, consider a subsystem $\mathcal{I}$ and the set of cut edges $\mathcal{C}^*_{\mathcal{I}}$. If this set contains at least two edges, we put the row corresponding to the entropy $S_{\mathcal{I}}$ into the top row block in (3.76). In the particular case where no such subsystem $\mathcal{I}$ exists, Γ takes the particularly simple form

$$\begin{pmatrix} \mathbb{1} & \mathbb{0} \end{pmatrix} \tag{3.77}$$

corresponding to the two bottom rows in (3.76). And if all edges participate in at least one cut, we have $\Gamma = \mathbb{1}$. When $\widetilde{\Gamma}$ is non-trivial, the second row in (3.76) is guaranteed to exist by Lemma 3.12, since each single edge in $\mathcal{C}^*_{\mathcal{I}}$ is the min-cut for the corresponding subsystem given by Lemma 3.11. The third row and second column, as well as the third column in (3.76) may be present or not, depending on the specific case. Their presence or absence however does not affect the essence of the rest of the proof.

Since $\mathbb{S} = \text{Im}\ \Gamma$, we have $\mathbb{S}^{\perp} = \text{Ker}\ \Gamma^{\intercal}$, and the generators of $\mathbb{S}^{\perp}$ are the columns of

the matrix

$$\begin{pmatrix} \mathbb{1} \\ -\widetilde{\Gamma}^{\intercal} \\ \mathbb{0} \end{pmatrix} \tag{3.78}$$

where again the last row may or may not be present depending on the specific form of Γ. In order to show the inclusion $\mathbb{S} \supseteq \mathbb{P}$, it is therefore sufficient to show that each column of this matrix is a linear combination of the columns of the matrix $\Pi(\mathbb{P})$ which specifies the PMI $(G_{\mathsf{N}}, \mathfrak{m})$.

In the particular case where $\widetilde{\Gamma}$ is trivial we have $\mathrm{Ker}\, \Gamma^{\intercal} = \{\mathbf{0}\}$, implying that $\mathbb{S} = \mathbb{R}^{\mathsf{D}}$. Since in this case each entropy is computed by a cut of a single edge, by Lemma 3.9 none of the instances of the mutual information vanish, therefore $\mathbb{P} = \mathbb{R}^{\mathsf{D}} = \mathbb{S}$ and the theorem holds.

Going back to the more general case, the labeling by polychromatic indices of the rows in each block of (3.78) is fixed by the construction of Γ. A column q in (3.78) is the vector normal to the hyperplane in $\mathbb{R}^{\mathsf{D}}$ corresponding to the equation

$$S_{\mathcal{I}} = \sum_{e \in \mathcal{C}(U^*_{\mathcal{I}})} S_{\mathcal{K}(e)} \tag{3.79}$$

where $\mathcal{I}$ is the polychromatic index labeling the row q in the block $\mathbb{1}$ of (3.78). In words, these equations say that each entropy $S_{\mathcal{I}}$ such the set $\mathcal{C}^*_{\mathcal{I}}$ contains more than one edge, is equal to the sum of the entropies of the subsystems "homologous" to the edges in the set. This follows already from Lemma 3.12, but (3.78) says that these equations correspond precisely to the generators of $\mathbb{S}^{\perp}$.

We can then rewrite (3.79) as follows

$$\begin{aligned} 0 = &(S_{\mathcal{K}(e_1)} + S_{\mathcal{K}(e_2)} - S_{\mathcal{K}(e_1)\triangle\mathcal{K}(e_2)}) \\ &+ (S_{\mathcal{K}(e_1)\triangle\mathcal{K}(e_2)} + S_{\mathcal{K}(e_3)} - S_{\mathcal{K}(e_1)\triangle\mathcal{K}(e_2)\triangle\mathcal{K}(e_3)}) \\ &+ \dots \end{aligned} \tag{3.80}$$

where by Lemma 3.13 each term in brackets can be recognized as an instance of the mutual information, cf. (3.59). Furthermore, any such instance is guaranteed to belong to the PMI of $(G_{\mathsf{N}}, \mathfrak{m})$ by Lemma 3.12, since the entropies that we added and subtracted in (3.80) all correspond to subsystems homologous to subsets of the edges in $\mathcal{C}^*_{\mathcal{J}}$. This shows that $\mathbb{S}^\perp \subseteq \mathbb{P}^\perp$, completing the proof. □

Finally, we extend this result to arbitrary (not necessarily generic) simple tree graphs. The central idea behind this generalization is again that because of the tree topology, all degeneracy equations correspond to new instances of the mutual information that vanish.

Theorem 3.2. *For any simple tree and min-cut structure, the* PMI *and the min-cut subspace coincide.*

Proof. As in the proof of Lemma 3.14, we only need to show that $\mathbb{S} \supseteq \mathbb{P}$, since by Lemma 3.10 $\mathbb{S} \subseteq \mathbb{P}$, and we will again show that $\mathbb{S}^\perp \subseteq \mathbb{P}^\perp$.

If the min-cut structure is degenerate, the matrix Γ is not uniquely specified. However, as we discussed in Section 3.3.3, we can chose a representative min-cut for each degenerate subsystem, and while the specific Γ will depend on the representatives, the min-cut subspace will not depend on this choice.

Suppose now that we make a choice of representatives for all degenerate subsystems, and therefore of Γ. We can imagine to determine a "partial" PMI by looking only at a

single min-cut for each subsystem as specified by this choice, i.e., we determine the set of instances of the mutual information that vanish according to this subset of the min-cuts. We will denote this partial PMI by $\mathbb{P}_\Gamma$ to stress the dependence on this choice. We can then ignore the degeneracies and follow step by step the proof of Lemma 3.14 to show that $(\text{Span }\Gamma)^\perp \subseteq \mathbb{P}_\Gamma^\perp$.

Because of the degeneracies however, $(\text{Span }\Gamma)^\perp \subseteq \mathbb{S}^\perp$ and $\mathbb{P}_\Gamma^\perp \subseteq \mathbb{P}^\perp$. In order to show that $\mathbb{S}^\perp \subseteq \mathbb{P}^\perp$, we will show that each degeneracy equation adds a new (not necessarily linearly independent) generator to $(\text{Span }\Gamma)^\perp$, and that this generator can be written as a linear combination of the generators in $\mathbb{P}^\perp$, i.e., that each generator of $\mathbb{S}^\perp$ which is not in $(\text{Span }\Gamma)^\perp$ is also in $\mathbb{P}^\perp$.

To see this, we can again follow a similar argument to the one we used in the proof of Lemma 3.14. A degeneracy equation for a subsystem $\mathcal{I}$ is an equation of the form

$$\sum_{e \in \mathcal{C}_\mathcal{I}^{*\alpha}} w_e = \sum_{f \in \mathcal{C}_\mathcal{I}^{*\beta}} w_f \tag{3.81}$$

where we used the more compact notation $\mathcal{C}_\mathcal{I}^{*\alpha} = \mathcal{C}(U_\mathcal{I}^{*\alpha})$. Using Lemma 3.12 we can translate this equation into an equation for the entropies

$$\sum_{e \in \mathcal{C}_\mathcal{I}^{*\alpha}} S_{\mathcal{J}(e)} = \sum_{f \in \mathcal{C}_\mathcal{I}^{*\beta}} S_{\mathcal{K}(f)} \tag{3.82}$$

which can be seen as a new generator in $\mathbb{S}^\perp$. Since both sides of this equation compute

$S_{\mathcal{I}}$, we can also think of (3.82) as a combination of the following two equations

$$\begin{aligned} S_{\mathcal{I}} &= \sum_{e \in \mathcal{C}^{*\alpha}_{\mathcal{I}}} S_{\mathcal{J}(e)} \\ S_{\mathcal{I}} &= \sum_{f \in \mathcal{C}^{*\beta}_{\mathcal{I}}} S_{\mathcal{K}(f)} \end{aligned} \tag{3.83}$$

But each of these equations is of the form (3.79) and can therefore be rewritten as in (3.80). All instances of the mutual information which appear in this decomposition belong to the PMI, therefore the new generator of $\mathbb{S}^{\perp}$ corresponding to (3.82) is a linear combination of the generators of $\mathbb{P}^{\perp}$. Repeating this construction for each degeneracy equation we obtain that $\mathbb{S}^{\perp} \subseteq \mathbb{P}^{\perp}$, concluding the proof. □

We stress that while the simple tree structure is a nice sufficient condition for the equivalence between PMIs and min-cut subspaces, it is by no means necessary. This equivalence is central in our arguments about the reconstruction of the HEC, and it would be interesting to extend it to a larger class of topological graph models and min-cut structures. While we leave this question for future work, in Section 3.6 we will discuss in more detail what kind of generalization of Theorem 3.2 is necessary to achieve the reconstruction, and will see examples of graphs with highly non-trivial topology whose min-cut subspaces and PMIs coincide.

3.5 Varying the number of parties

Up to this point, we have been working with topological graph models and min-cut structures for an arbitrary, but *fixed*, number of parties N. However, in order to uncover some of the deepest relations between min-cut subspaces and PMIs, it is crucial to vary

the number of parties and analyse how the various constructs that we have introduced thus far transform under this operation.

As we will see, the subtle behavior of min-cut subspaces and PMIs under recolorings is in stark contrast with the simple behavior of entropy vectors. This should perhaps make even more evident the fundamental differences between an analysis of holographic constraints based on entropy vectors and graph models, from one which purely relies on equivalence classes, like the one advocated here for the HEC, or the one based on proto-entropies for the holographic entropy polyhedron [68].

In Section 3.5.1 we will analyse *coarse-grainings*, i.e., transformations that reduce the number of parties, and their effect on W-cells, S-cells and min-cut subspaces. The consequences of the opposite transformations, namely *fine-grainings*, will be discussed in Section 3.5.2. Our convention will be to denote by N the original number of parties and by N' the new number of parties, both for coarse-grainings and fine-grainings. *Recolorings* of boundary vertices that reduce, respectively increase, the number of parties will be denoted by $\beta^{\downarrow}$ and $\beta^{\uparrow}$.

3.5.1 Coarse-grainings of equivalence classes

Given an N-party density matrix ρ_{N} and a purification $|\psi\rangle_{\mathsf{N}+1}$ in a Hilbert space of the form (3.5), consider a partition of the set $[\mathsf{N}+1]$ into $\mathsf{N}'+1$ (non-empty) parts, with $\mathsf{N}' < \mathsf{N}$. Each element of the partition is an N-party polychromatic index $\underline{\mathsf{J}}$ which we recolor by a monochromatic index $\ell' \in [\mathsf{N}'+1]$. We capture this coarse-graining by a map $\hat{\phi} : \ell' \mapsto \underline{\mathsf{J}}$, which also tells us which polychromatic indices $\underline{\mathsf{J}} \subseteq [\mathsf{N}+1]$ correspond

to the coarse-grained ones $\underline{\mathcal{I}'} \subseteq [\mathsf{N}'+1]$ through

$$\hat{\phi} : \underline{\mathcal{I}'} \mapsto \bigcup_{\ell' \in \underline{\mathcal{I}'}} \hat{\phi}(\ell') \tag{3.84}$$

This partition of $[\mathsf{N}+1]$ and redefinition of the polychromatic indices corresponds to a redefinition of the Hilbert space $\underline{\mathcal{H}}$ in (3.5) into a new Hilbert space $\underline{\mathcal{H}'}$ with $\mathsf{N}'+1$ factors, each of which is a collection of the original factors in $\underline{\mathcal{H}}$. For an N-party density matrix ρ acting on $\mathcal{H}$ as in (3.1), this transformation gives a new N'-party density matrix ρ' acting on a Hilbert space $\mathcal{H}'$, obtained from $\underline{\mathcal{H}'}$ by ignoring the new factor that contains the original $\mathcal{H}_{\mathsf{N}+1}$ factor (the purifier) in $\underline{\mathcal{H}'}$.

We now want to apply such a transformation to an entropy vector, and to do so we need to use non-underlined indices. Notice that in general a coarse-graining as defined in (3.84) can map the N-party purifier $\mathsf{N}+1$ to an arbitrary N'-party color, not necessarily to the N'-party purifier $\mathsf{N}'+1$. Since our convention is that non-underlined polychromatic indices should not include the purifier, we introduce a new map ϕ that not only implements a coarse-graining as in (3.84), but also such that when it acts on an index $\mathcal{I}'$ it replaces $\hat{\phi}(\mathcal{I}')$ with $[\hat{\phi}(\mathcal{I}')]^{\complement}$ if $\hat{\phi}(\mathcal{I}')$ includes the original $\mathsf{N}+1$ subsystem.

The entropy vector of the new density matrix $\mathbf{S}'(\rho')$ can then be obtained from the entropy vector $\mathbf{S}(\rho)$ of the original density matrix simply as

$$S'_{\mathcal{I}'} = S_{\phi(\mathcal{I}')} \tag{3.85}$$

At the level of entropy vectors, we can therefore think of this transformation as a map

$$\Phi_{\mathsf{N}\to\mathsf{N}'} : \mathbb{R}^{\mathsf{D}} \to \mathbb{R}^{\mathsf{D}'} \qquad \mathbf{S} \mapsto \mathbf{S}' = \Phi_{\mathsf{N}\to\mathsf{N}'}\,\mathbf{S} \tag{3.86}$$

where $\mathsf{D}' = 2^{\mathsf{N}'} - 1$ and $\Phi_{\mathsf{N}\to\mathsf{N}'}$ is the $\mathsf{D}'\times\mathsf{D}$ matrix

$$(\Phi_{\mathsf{N}\to\mathsf{N}'})_{\mathcal{I}'\mathcal{K}} := \begin{cases} 1 & \text{if } \mathcal{K} = \phi(\mathcal{I}') \\ 0 & \text{otherwise} \end{cases} \tag{3.87}$$

This linear map will be referred to as a *color-projection* in what follows.[67]

A coarse-graining of a graph model is defined similarly. Given $\widetilde{G_{\mathsf{N}}}$ one introduces a *recoloring* by a new coloring map $\beta^{\downarrow} : \partial V \to [\mathsf{N}' + 1]$. The recoloring $\beta^{\downarrow}$ induces a coarse-graining of polychromatic indices ϕ as described above, and the new entropy vector $\mathbf{S}'(\widetilde{G_{\mathsf{N}'}})$ is then obtained from $\mathbf{S}(\widetilde{G_{\mathsf{N}}})$ via the color-projection defined in (3.87)

$$\mathbf{S}'(\widetilde{G_{\mathsf{N}'}}) = \Phi_{\mathsf{N}\to\mathsf{N}'}\, \mathbf{S}(\widetilde{G_{\mathsf{N}}}) \tag{3.88}$$

To see this, simply notice that the recoloring does not change the topology of the graph or the weights, and any min-cut of a coarse-grained subsystem $\mathcal{I}'$ is a min-cut for the corresponding original subsystem $\phi(\mathcal{I}')$.

Given the simplicity of the transformation rule for an entropy vector, one may be tempted to conclude that the same transformation also applies to min-cut subspaces. Instead, as we will see, the dimension of a min-cut subspace can even increase under coarse-graining, and this unexpected behavior is another example of a situation where a crucial role is played by the structure of degeneracies.

To understand how this can happen, and in what situations min-cut subspaces do transform analogously to entropy vectors, we need to consider the effect of a coarse-graining on a min-cut structure. As already mentioned above, a recoloring $\beta^{\downarrow}$ only affects

[67] Technically, the map defined in (3.87) is not a projection, since it is not an endomorphism of entropy space. However we will still use this terminology since one can simply consider an embedding of $\mathbb{R}^{\mathsf{D}'}$ into $\mathbb{R}^{\mathsf{D}}$, in which case (3.87) would be a projection.

the labeling of the boundary vertices of a topological graph model, not its topology, and the space of edge weights therefore is the same, before and after the coarse-graining. For a class $(G_{\mathsf{N}}, \mathfrak{m})$ and a coarse-graining ϕ induced by a recoloring $\beta^{\downarrow}$, the new min-cut structure $\mathfrak{m}'$ on the new topological graph model $G_{\mathsf{N}'}$ can be expressed in terms of the original coloring as

$$\mathfrak{m}' = \{\mathcal{U}_{\mathcal{J}'} = \mathcal{U}_{\phi(\mathcal{J}')} \text{ for all } \mathcal{J}'\} \subset \mathfrak{m} \tag{3.89}$$

In words, one can think of deriving $\mathfrak{m}'$ from $\mathfrak{m}$ by simply removing all min-cut sets $\mathcal{U}_{\mathcal{J}}$ for all subsystems $\mathcal{J}$ that are "projected out" by the coarse-graining, i.e., such that there is no $\mathcal{J}'$ with $\phi(\mathcal{J}') = \mathcal{J}$, and then relabeling the elements of $\mathfrak{m}'$ by the new polychromatic indices.

Notice that an immediate consequence of (3.89) is that the relation between the subspaces corresponding to the solutions to the degeneracy equations in $\mathfrak{m}$ and $\mathfrak{m}'$ (cf. (3.24)) is given by

$$\mathbb{W}' \supseteq \mathbb{W} \tag{3.90}$$

This simply follows from the fact that any equation that appears in $\mathfrak{m}'$ also appears in $\mathfrak{m}$, but in general, not vice versa.

From the transformation rule of the min-cut structures we can also derive the relation between the maps Γ and Γ' from $\mathbb{R}^3$ to $\mathbb{R}^{\mathsf{D}}$ and $\mathbb{R}^{\mathsf{D}'}$. In particular, since we are especially interested in the behaviour of degeneracies, we need to clarify if and how the relation between these maps can be affected by different choices of representative min-cuts. This is the content of the next lemma

Lemma 3.15. *For any class $(G_{\mathsf{N}}, \mathfrak{m})$, coarse-graining ϕ to a new class $(G_{\mathsf{N}'}, \mathfrak{m}')$ induced by a recoloring $\beta^{\downarrow}$, and choice of map Γ for $\mathfrak{m}$, there exists a choice of map Γ' for $\mathfrak{m}'$*

such that

$$\Gamma' = \Phi_{\mathsf{N}\to\mathsf{N}'}\,\Gamma \tag{3.91}$$

Proof. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$ consider the map $\Gamma_{\{\alpha\}}$ specified by a choice of representative min-cuts $U^{*\alpha}_{\mathcal{J}}$ for all N-party polychromatic indices $\mathcal{J}$. By (3.89) it follows that

$$U^{*\alpha}_{\mathcal{J}'} := U^{*\alpha}_{\phi(\mathcal{J}')}$$

is a choice of representative min-cuts for all the coarse-grained polychromatic indices $\mathcal{J}'$. Notice that the index α did not change, since for the subsystems which are not removed by the coarse-graining, any possible choice of min-cut in $\mathfrak{m}$ is a valid choice of min-cut in $\mathfrak{m}'$.

Recall now the definition of the matrix Γ given in (3.19). Using for $\mathfrak{m}'$ the choice of representatives induced by the choice for $\mathfrak{m}$ just described, the matrix Γ' can be obtained from Γ as follows. We first delete the rows corresponding to the subsystems $\mathcal{J}$ which are removed by the coarse-graining, and then permute the remaining rows according to the relabeling fixed by $\beta^{\downarrow}$. But this is precisely the transformation performed on Γ by the matrix $\Phi_{\mathsf{N}\to\mathsf{N}'}$ defined in (3.87). □

We are now ready to discuss a general situation where the transformation of min-cut subspaces under coarse-grainings is well behaved, and completely determined by the map $\Phi_{\mathsf{N}\to\mathsf{N}'}$ between entropy spaces. The next result shows that this is the case whenever (3.90) is saturated.

Theorem 3.3 (Color-projections of min-cut subspaces)**.** *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and*

coarse-graining ϕ induced by a recoloring $\beta^{\downarrow}$ to a new class $(G_{\mathsf{N}'}, \mathfrak{m}')$

$$\mathbb{W}' = \mathbb{W} \qquad \Longrightarrow \qquad \mathbb{S}' = \Phi_{\mathsf{N}\to\mathsf{N}'}\, \mathbb{S} \tag{3.92}$$

Proof. For a given class $(G_{\mathsf{N}}, \mathfrak{m})$ consider one of the maps Γ defined in (3.20) (as usual for some choice of representative min-cuts). For any coarse-graining ϕ to a new class $(G_{\mathsf{N}'}, \mathfrak{m}')$, Lemma 3.15 guarantees that there exists a choice of representative min-cuts for $\mathfrak{m}'$ such that the corresponding map Γ' is related to Γ via (3.91). In general the relation between $\mathbb{W}$ and the min-cut subspace is given by Lemma 3.5. Therefore by the assumption that $\mathbb{W}' = \mathbb{W}$ we have

$$\mathbb{S}' = \Gamma'(\mathbb{W}') = (\Phi_{\mathsf{N}\to\mathsf{N}'}\, \Gamma)(\mathbb{W}) = \Phi_{\mathsf{N}\to\mathsf{N}'}\, (\Gamma(\mathbb{W})) = \Phi_{\mathsf{N}\to\mathsf{N}'}\, \mathbb{S}$$

completing the proof. □

A straightforward consequence of this theorem is that min-cut subspaces of generic min-cut structures are always well behaved under coarse-grainings

Corollary 3.4. *For any class $(G_{\mathsf{N}}, \mathfrak{m})$ and coarse-graining ϕ induced by a recoloring $\beta^{\downarrow}$, if $\mathfrak{m}$ is generic then (3.92) holds.*

Proof. Given a $(G_{\mathsf{N}}, \mathfrak{m})$, if $\mathfrak{m}$ is generic we have $\mathbb{W} = \mathbb{R}^{\mathsf{E}}$, since the W-cell $\mathcal{W}(G_{\mathsf{N}}, \mathfrak{m})$ is full dimensional. By (3.89) a coarse-graining cannot introduce new degeneracies. Therefore the W-cell of the coarse-grained min-cut structure also spans $\mathbb{R}^{\mathsf{E}}$, and Theorem 3.3 applies.

□

We have seen that the saturation of (3.90) is a sufficient condition for min-cut subspaces to transform according to (3.92). On the other hand, if the inclusion (3.90) is strict,

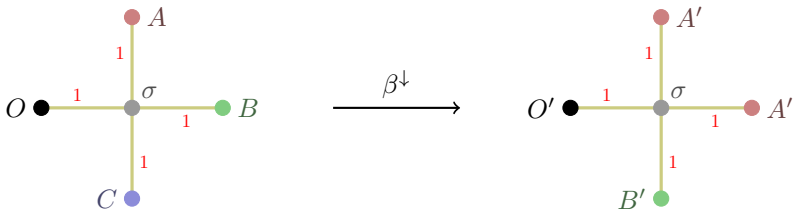


Figure 3.7: The coarse-graining (3.93) of the perfect state graph model $\widetilde{G_3}$ from Figure 3.1.

which can happen if some degeneracy equations are "lost" under the coarse-graining, the relation between the min-cut subspaces is more complicated, and depends on the details of the graph, the min-cut structure and the recoloring. In fact, as already mentioned earlier, the dimension of the min-cut subspace can even grow, which can easily happen for coarse-grainings of highly degenerate min-cut structures.

As a simple example of this non-trivial behavior, consider the graph models depicted in Figure 3.7. The one on the left is the graph that we have already seen before, which realizes the $\mathsf{N} = 3$ perfect state, while the one on the right is obtained via the coarse-graining to $\mathsf{N}' = 2$ specified by

$$\hat{\phi}(A') = AB \qquad \hat{\phi}(B') = C \qquad \hat{\phi}(O') = O \tag{3.93}$$

The entropy ray obtained from this coarse-graining of $\widetilde{G_3}$ can be directly computed from the ray in (3.36), obtaining

$$\mathbf{S}(\widetilde{G_2}) = \lambda\,(2,1,1), \qquad \lambda > 0 \tag{3.94}$$

However we do not want to simply derive the new entropy ray, but also the min-cut subspace of the new topological graph model and min-cut structure. After the recoloring,

the only degenerate min-cut that remains from (3.32) is

$$\mathcal{U}_{A'} = \{\{A'_1, A'_2\}, \{A'_1, A'_2, \sigma\}\} \tag{3.95}$$

(where A'_1, A'_2 denote the two original A, B boundary vertices that have now been recolored) and the only degeneracy equation left from (3.33) is

$$w_{A'_1\sigma} + w_{A'_2\sigma} = w_{B'\sigma} + w_{O'\sigma} \tag{3.96}$$

The solution to this degeneracy equation is the 3-dimensional subspace

$$\mathbb{W}' = (1, 1, -1, -1)^{\perp} \subset \mathbb{R}^4 \tag{3.97}$$

where we ordered the weights according to $(A'_1\sigma, A'_2\sigma, B'\sigma, O'\sigma)$. The W-cell is the interior of the following polyhedral cone in $\mathbb{W}'$ (written as embedded in $\mathbb{R}^4_{>0}$)

$$\text{cone}\,\{(1,0,0,1),\ (0,1,0,1),\ (0,1,1,0),\ (1,0,1,0)\} \tag{3.98}$$

Its image under the map

$$\Gamma = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.99}$$

fixed by the choice of representative $U^*_{A'} = \{A'_1, A'_2\}$ is the S-cell, which is the interior of the following polyhedral cone in $\mathbb{R}^3$

$$\text{cone}\,\{(1,0,1),\ (1,1,0)\} \tag{3.100}$$

This can easily be recognized as the facet of the $\mathrm{HEC}_2 = \mathrm{SAC}_2$ supported by the min-cut subspace (or equivalently, PMI), $\mathsf{I}(B' : O') = 0$. Notice that the straightforward projection of the perfect state entropy ray from (3.36) given in (3.94) is just one of the rays on this facet and has no particular meaning.

The example we just discussed also shows that, like for the min-cut subspace, the dimension of the PMI of a topological graph model and min-cut structure can similarly grow under coarse-graining. Indeed, the reader can easily verify that the PMI of the graph model $\widetilde{G_3}$ of Figure 3.1 is 1-dimensional, while the coarse-grained graph has a 2-dimensional PMI.

3.5.2 Fine-grainings of equivalence classes

For fine-grainings of density matrices and graph models we can proceed similarly as for coarse-grainings. We consider the case of a state $|\psi\rangle_{\mathsf{N}+1}$ in a Hilbert space $\underline{\mathcal{H}}$ as in (3.5), such that some of the factors admit a tensor product structure into "finer components", giving a new Hilbert space $\underline{\mathcal{H}'}$. An N-party density matrix ρ acting on $\mathcal{H}$ in (3.1) can then be seen as a new N'-party density matrix ρ', with $\mathsf{N}' > \mathsf{N}$, acting on $\mathcal{H}'$. From ρ' we can then compute the entropy vector $\mathbf{S}'(\rho')$, which in general will depend on the details of the initial density matrix ρ. For any fine-graining however, we can always chose an appropriate coarse-graining ϕ that "undoes" it. No matter what the details of the initial density matrix ρ are, the entropy vectors will then be related by this coarse-graining as usual

$$\mathbf{S} = \Phi_{\mathsf{N}' \to \mathsf{N}}\, \mathbf{S}' \tag{3.101}$$

As in the case of coarse-grainings, fine-grainings of graph models can be defined similarly to fine-grainings of density matrices. For a given graph model $\widetilde{G_{\mathsf{N}}}$ realizing

an entropy vector $\mathbf{S}$ we consider a boundary recoloring $\beta^{\uparrow} : \partial V \to [\mathsf{N}' + 1]$, which will specify a new *fine-grained graph model* $\widetilde{G_{\mathsf{N}'}}$ from which we obtain a new entropy vector $\mathbf{S}'$. We can then "undo" the fine-graining with a coarse-graining ϕ specified by a boundary recoloring $\beta^{\downarrow}$

$$\beta^{\downarrow} = \beta \tag{3.102}$$

where β is simply the initial coloring of $\widetilde{G_{\mathsf{N}}}$. The entropy vectors before and after the fine-graining induced by $\beta^{\uparrow}$ will then be related as in (3.101).

For fine-grainings of min-cut structures on topological graph models on the other hand, we again need to be more careful. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$, a boundary recoloring $\beta^{\uparrow}$ gives a new topological graph model $G_{\mathsf{N}'}$, but in general this does not automatically induce a fine-graining, since the new min-cut structure might be incomplete. The reason is that in general $\mathfrak{m}$ does not specify the min-cuts for all subsystems $\mathcal{I}'$, and the fine-grainings induced by $\beta^{\uparrow}$ on different graph models $\widetilde{G_{\mathsf{N}}} \in (G_{\mathsf{N}}, \mathfrak{m})$ can correspond to different min-cut structures.

To take this indeterminacy into account, we define the set $\mathfrak{m}\big\uparrow_{\mathsf{N}'}$ of min-cut structures $\mathfrak{m}'$ for $G_{\mathsf{N}'}$ that reduce to the original min-cut structure $\mathfrak{m}$ on G_{N} when we undo the fine-graining, specifically

$$\mathfrak{m}\big\uparrow_{\mathsf{N}'} \coloneqq \{\mathfrak{m}' : \ \mathcal{U}_{\mathcal{I}} = \mathcal{U}_{\phi(\mathcal{I})} \text{ for all } \mathcal{I}\} \tag{3.103}$$

where ϕ is the coarse-graining induced by the recoloring $\beta^{\downarrow}$ given in (3.102). Notice that for any $(G_{\mathsf{N}}, \mathfrak{m})$ and recoloring $\beta^{\uparrow}$, this set is non-empty. The existence of at least one min-cut structure in $\mathfrak{m}\big\uparrow_{\mathsf{N}'}$ is obvious: simply consider any graph model $\widetilde{G_{\mathsf{N}}} \in (G_{\mathsf{N}}, \mathfrak{m})$, apply the recoloring $\beta^{\uparrow}$ and determine the min-cut structure $\mathfrak{m}'$ of the new graph model $\widetilde{G_{\mathsf{N}'}}$. We then define a fine-graining of $(G_{\mathsf{N}}, \mathfrak{m})$ as follows:

Definition 3.13 (Fine-graining of a class $(G_{\mathsf{N}}, \mathfrak{m})$)**.** *A fine-graining of a class* $(G_{\mathsf{N}}, \mathfrak{m})$ *is a pair* $(\beta^{\uparrow}, \mathfrak{m}')$ *of a recoloring* $\beta^{\uparrow}$ *and a min-cut structure from* $\mathfrak{m}\big\uparrow_{\mathsf{N}'}$.

We stress the fundamental difference between a fine-graining of a graph model as defined above, and a fine-graining of an equivalence class as defined here. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$ and a choice of representative $\widetilde{G_{\mathsf{N}}}$, any recoloring $\beta^{\uparrow}$ will automatically specify a fine-graining of $\widetilde{G_{\mathsf{N}}}$, but in general this is only *one of the possible* fine-grainings of the class $(G_{\mathsf{N}}, \mathfrak{m})$ according to Definition 3.13. In fact, for a fixed recoloring $\beta^{\uparrow}$, the induced fine-graining of different choices of representatives will in general correspond to different fine-grainings of $(G_{\mathsf{N}}, \mathfrak{m})$. Importantly, this can even happen when the min-cut subspace of $(G_{\mathsf{N}}, \mathfrak{m})$ is 1-dimensional, and one may be inclined to think that a graph model and its equivalence class are essentially the same object. As we will see (and exemplify) in what follows, the reason for this unexpected behavior is that even when the min-cut subspace is 1-dimensional, the W-cell can be higher dimensional.

As usual, a choice of representative of an equivalence class is a convenient way to specify a min-cut structure, and in Section 3.6 we will often resort to fine-grainings of graph models. However, the reader should always remain aware of the fundamental difference highlighted here, and of the possibility of alternative fine-grainings of an equivalence class specified by a choice of graph model.

Let us now continue our analysis of the set of possible fine-grainings of a topological graph model and a min-cut structure. We have seen in the previous subsection that in certain situations min-cut subspaces are well behaved under coarse-grainings, in the sense that they transform under the same projection which determines a coarse-grained entropy vector. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$, and a recoloring $\beta^{\uparrow}$, it is then natural to ask under what conditions there exists a fine-graining, i.e., a choice of a new min-cut structure $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$, such that Theorem 3.3 applies.

To answer this question, and understand the origin of some of the properties of fine-grainings mentioned above, we first need to analyse in more detail the effect of a recoloring $\beta^{\uparrow}$ on the set of W-cells for a topological graph model G_{N}. As we have seen, a recoloring does not change the topology of G_{N}, and in particular its set of edges, therefore the space of edge weights $\mathbb{R}^{\mathsf{E}}_{>0}$ is unaffected by the recoloring and we can compare W-cells before and after it. Their relation is captured by the following lemma

Lemma 3.16 (Refinement of W-cells)**.** *Given a class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and any recoloring* $\beta^{\uparrow}$*, the set of* W*-cells* $\mathcal{W}_{\mathfrak{m}'}$ *for all* $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ *is a partition of* $\mathcal{W}_{\mathfrak{m}}$*.*

Proof. Given a class $(G_{\mathsf{N}}, \mathfrak{m})$ and any recoloring $\beta^{\uparrow}$, consider a min-cut structure $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ and the corresponding W-cell $\mathcal{W}_{\mathfrak{m}'}$. Denoting by ϕ the coarse-graining induced by the recoloring $\beta^{\downarrow}$ determined by (3.102), it follows from (3.103) that all inequalities (3.23) and degeneracy equations (3.24) that specify the W-cell $\mathcal{W}_{\mathfrak{m}}$ of the original min-cut structure $\mathfrak{m}$ also belong to the set of inequalities and degeneracy equations that specify $\mathcal{W}_{\mathfrak{m}'}$. This implies that $\mathcal{W}_{\mathfrak{m}'} \subseteq \mathcal{W}_{\mathfrak{m}}$, for all $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$. Furthermore, by Lemma 3.6, the different W-cells $\mathcal{W}_{\mathfrak{m}'}$ do not intersect. Therefore it only remains to prove that the union of all W-cells $\mathcal{W}_{\mathfrak{m}'}$ for all $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ is $\mathcal{W}_{\mathfrak{m}}$. Consider a graph model $\widetilde{G_{\mathsf{N}}} \in (G_{\mathsf{N}}, \mathfrak{m})$, specified by a weight vector $\mathbf{w} \in \mathcal{W}_{\mathfrak{m}}$. The recoloring $\beta^{\uparrow}$ induces a fine-graining of $\widetilde{G_{\mathsf{N}}}$ to a new graph model $\widetilde{G_{\mathsf{N}'}} \in (G_{\mathsf{N}'}, \mathfrak{m}_{\mathbf{w}})$ where $\mathfrak{m}_{\mathbf{w}}$ is one of the min-cut structures in $\mathfrak{m}\big\uparrow_{\mathsf{N}'}$, completing the proof. □

We are now ready to establish for what classes $(G_{\mathsf{N}}, \mathfrak{m})$ and recolorings $\beta^{\uparrow}$, there exists a fine-graining $(\beta^{\uparrow}, \check{\mathfrak{m}})$ such that the min-cut subspace of $(G_{\mathsf{N}}, \mathfrak{m})$ can be obtained from that of $(G_{\mathsf{N}'}, \check{\mathfrak{m}})$ via the color projection associated to the coarse-graining ϕ specified by the recoloring $\beta^{\downarrow}$ given in (3.102). The next lemma clarifies that this is always the case.

Lemma 3.17. *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and recoloring* $\beta^{\uparrow}$*, there exists a choice of min-cut structure* $\check{\mathfrak{m}} \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ *such that*

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \Phi_{\mathsf{N}' \to \mathsf{N}}\, \mathbb{S}(G_{\mathsf{N}'}, \check{\mathfrak{m}}) \tag{3.104}$$

Proof. We only need to prove that Theorem 3.3 applies, and for this we only need to show that there exists $\check{\mathfrak{m}} \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ such that $\mathbb{W}_{\check{\mathfrak{m}}} = \mathbb{W}_{\mathfrak{m}}$, where $\mathbb{W}_{\check{\mathfrak{m}}} = \mathrm{Span}\,(\mathcal{W}_{\check{\mathfrak{m}}})$ and $\mathbb{W}_{\mathfrak{m}} = \mathrm{Span}\,(\mathcal{W}_{\mathfrak{m}})$. This follows from Lemma 3.16, since the W-cells $\mathcal{W}_{\mathfrak{m}'}$ for all $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ form a *finite* partition of $\mathcal{W}_{\mathfrak{m}}$, and therefore there must exist at least one min-cut structure $\check{\mathfrak{m}} \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$ such that $\mathrm{Span}\,(\mathcal{W}_{\check{\mathfrak{m}}}) = \mathrm{Span}\,(\mathcal{W}_{\mathfrak{m}})$. □

While this lemma is just a proof of existence of at least one fine-graining such that (3.104) applies, it should be clear from the proof that this fine-graining is in general non-unique. Since these fine-grainings will play a crucial role in the next section, we introduce the following definition

Definition 3.14 (Minimally-degenerate fine-graining)**.** *For any class* $(G_{\mathsf{N}}, \mathfrak{m})$ *and recoloring* $\beta^{\uparrow}$*, a minimally-degenerate fine-graining* $(\beta^{\uparrow}, \check{\mathfrak{m}})$ *is any fine-graining such that*

$$\mathrm{Span}\,(\mathcal{W}_{\check{\mathfrak{m}}}) = \mathrm{Span}\,(\mathcal{W}_{\mathfrak{m}}) \tag{3.105}$$

As explained in the proof of Lemma 3.17, the condition in (3.105) is sufficient to guarantee that (3.104) holds, but a priori it is not necessary. Therefore it is in principle possible that there exists fine-grainings which are not minimally-degenerate according to Definition 3.14, but which are still well behaved in the sense of the coarse-graining in (3.104). Notice that in the particular case of a class $(G_{\mathsf{N}}, \mathfrak{m})$ with a 1-dimensional W-cell, any recoloring $\beta^{\uparrow}$ automatically specifies a minimally-degenerate fine-graining for

the class. This simply follows from the fact that the W-cell $\mathcal{W}_{\mathfrak{m}}$ cannot be partitioned into smaller components, and there is therefore only one possible fine-grained min-cut structure $\check{\mathfrak{m}}$. Furthermore, since $\mathcal{W}_{\mathfrak{m}} = \mathcal{W}_{\check{\mathfrak{m}}}$, (3.105) is trivially satisfied. Equivalently, in this case the class $(G_{\mathsf{N}}, \mathfrak{m})$ has only a single representative $\widetilde{G_{\mathsf{N}}}$ (up to an irrelevant global rescaling of the weights), and the (unique) fine-graining of the class induced by $\beta^{\uparrow}$ necessarily coincides with the fine-graining of $\widetilde{G_{\mathsf{N}}}$.

Everything we have discussed thus far for fine-grainings of a class $(G_{\mathsf{N}}, \mathfrak{m})$, can be iterated until one reaches the largest number of different colors that can be assigned to the boundary vertices of G_{N}. This number is of course given by the cardinality of ∂V, and we denote it by V_{∂}. We define a *maximal recoloring* of G_{N} as any recoloring that attains this bound, taking the form

$$\overline{\beta^{\uparrow}} : \partial V \to [\mathsf{V}_{\partial}] \tag{3.106}$$

Given a class $(G_{\mathsf{N}}, \mathfrak{m})$, any fine-graining of the form $(\overline{\beta^{\uparrow}}, \mathfrak{m}')$, with $\mathfrak{m}' \in \mathfrak{m}\big\uparrow_{\mathsf{N}'}$, will be referred to as a *maximal fine-graining* of $(G_{\mathsf{N}}, \mathfrak{m})$.

An interesting and useful application of maximal fine-grainings and Lemma 3.17 is the generalization of our analysis from Section 3.4.3 concerning the relation between min-cut subspaces and PMIs for tree graphs. The main result of that section was Theorem 3.2, which says that for any simple tree and min-cut structure, the min-cut-subspace and the PMI coincide. Given an arbitrary topological graph model G_{N} with tree topology, we can now turn it into a simple tree via a maximal recoloring $\overline{\beta^{\uparrow}}$. And for any maximal recoloring, Lemma 3.17 guarantees the existence of a minimally-degenerate fine-graining such that (3.104) applies. We have thus proved the following general result for arbitrary topological graph models with tree topology and min-cut structures on them

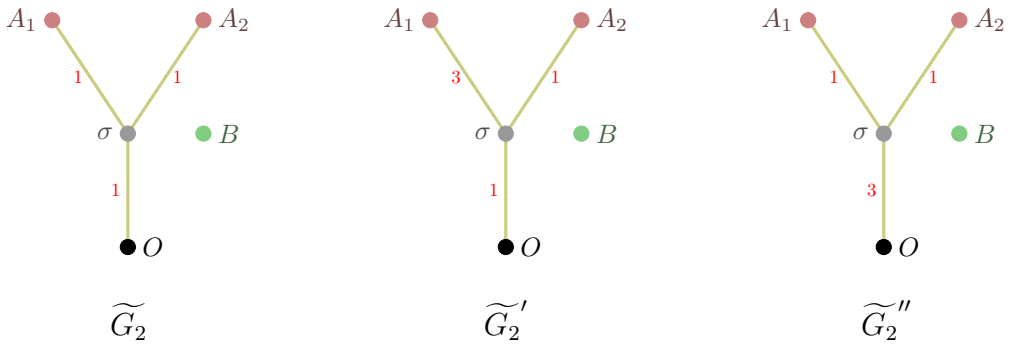


Figure 3.8: A choice of three graph models with the same underlying topology. $\widetilde{G_2}$ and $\widetilde{G_2}'$ have the same min-cut structure and are therefore different representatives of the same class, while $\widetilde{G_2}''$ has a different min-cut structure. Both min-cut structures are generic.

Theorem 3.4. *For any topological graph model G_{N} with tree topology, min-cut structure $\mathfrak{m}$ on G_{N}, and minimally-degenerate maximal fine-graining $(\overline{\beta^{\uparrow}}, \check{\mathfrak{m}})$, the min-cut subspace $\mathbb{S}(G_{\mathsf{N}'}, \check{\mathfrak{m}})$ is given by*

$$\mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) = \Phi_{\mathsf{N}' \to \mathsf{N}}\, \pi(G_{\mathsf{N}'}, \check{\mathfrak{m}}) \tag{3.107}$$

This general result shows that for fixed N many[68] min-cut subspaces can be seen as color-projections of PMIs for a larger number of parties, and it is interesting to ask if all min-cut subspaces can be obtained in this way. As we will discuss in Section 3.6, this question is of particular relevance for the reconstruction of the HEC from the solution to the HMIP.

We conclude this section with an example of fine-grainings in the particular case of an equivalence class whose min-cut subspace is 1-dimensional, illustrating how even in this case there might exists alternative fine-grainings, as well as the interplay between the fine-graining procedure and the application of Lemma 3.8 and the reduction of Section 3.3.4.

[68] Notice that the trees in Theorem 3.4 are allowed to have an arbitrarily large number of vertices for each color, and therefore could in principle encode the same information content as graphs with more complicated topology.

Consider the first two graph models in Figure 3.8. Both of them have the same generic min-cut structure, which takes the form

$$\mathcal{U}_A = \{\{A_1, A_2, \sigma\}\}, \qquad \mathcal{U}_B = \{\{B\}\}, \qquad \mathcal{U}_{AB} = \{\{A_1, A_2, \sigma, B\}\} \tag{3.108}$$

The W-cell for this min-cut structure is the interior of the following polyhedral cone in $\mathbb{R}^3_{>0}$

$$\text{cone}\,\{(1,0,0),(0,1,0),(1,0,1),(0,1,1)\} \tag{3.109}$$

where we ordered the weights according to $(A_1\sigma, A_2\sigma, O\sigma)$. Its image under the map

$$\Gamma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.110}$$

is the 1-dimensional S-cell

$$\mathcal{S} = \lambda\,(1,0,1), \qquad \lambda > 0 \tag{3.111}$$

which is indeed the entropy ray of both graph models $\widetilde{G_2}$ and $\widetilde{G_2}'$.

The min-cut structure of the last graph model in Figure 3.8 is again generic and takes the form

$$\mathcal{U}_A = \{\{A_1, A_2\}\}, \qquad \mathcal{U}_B = \{\{B\}\}, \qquad \mathcal{U}_{AB} = \{\{A_1, A_2, B\}\} \tag{3.112}$$

The W-cell is the interior of another polyhedral cone in $\mathbb{R}^3_{>0}$

$$\text{cone}\,\{(1,0,1),(0,1,1),(0,0,1)\} \tag{3.113}$$

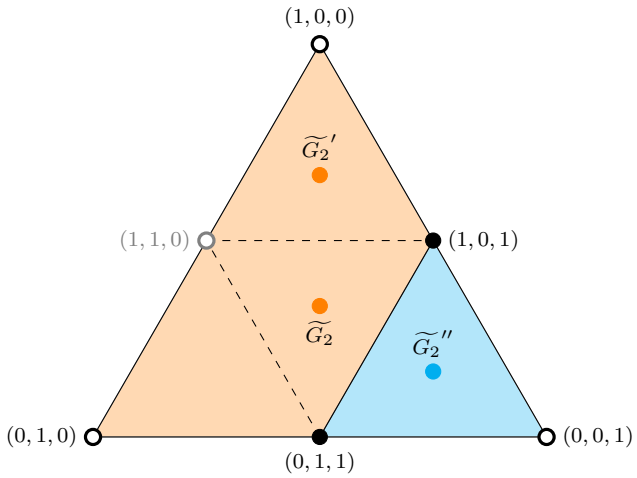


Figure 3.9: A cross section of $\mathbb{R}^3_{>0}$ by an affine $\mathbb{R}^2$ plane orthogonal to the vector $(1,1,1)$. This cross section shows the partition of $\mathbb{R}^3_{>0}$ and its boundary into W-cells for the topological graph model of the graphs in Figure 3.8, before and after the recoloring from (3.115). See the main text for more details.

and its image under the map

$$\Gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \tag{3.114}$$

is the same 1-dimensional S-cell given in (3.111) for the other two graph models.

Figure 3.9 shows the partition into W-cells of the space of edge weights $\mathbb{R}^3_{>0}$ for the topological graph model underlying the three graph models of Figure 3.8. The regions shaded in orange and cyan correspond respectively to the two full-dimensional W-cells given in (3.109) and (3.113). We leave it as an exercise for the reader to verify that the solid edge separating the these two W-cells does indeed correspond, as suggested by the figure, to a degenerate min-cut structure with a 1-dimensional W-cell (see the first graph model in Figure 3.10 for a choice of representative of this class).

The regions on the boundary of $\mathbb{R}^3_{>0}$ instead, correspond to W-cells for min-cut structures on different topological graph models obtained by appropriate deletion of edges, following the reduction described in Section 3.3.4. The vertices shown in solid black in

Figure 3.9 correspond to extreme rays shared by the two full-dimensional W-cells (see the middle graph model in Figure 3.10 for an example). They are 1-dimensional W-cells whose associated S-cells are again given by (3.111), in agreement with Lemma 3.8. On the other hand, the unfilled vertices corresponding to the canonical bases vectors of $\mathbb{R}^3$, are examples of the situation where Lemma 3.8 cannot be used. As explained in the proof of the same lemma, it can happen that an extreme ray of a W-cell is mapped to the null vector by the corresponding Γ map, which is precisely what happens here (cf. the last graph model in Figure 3.10). All other solid edges in Figure 3.9 can also be checked to be 2-dimensional W-cells, again after an appropriate deletion of edges (for now, the dashed edges and the $(1,1,0)$ vertex should be ignored, and the segment connecting $(0,1,0)$ to $(1,0,0)$ should be seen as a single face). Finally, notice that again in agreement with Lemma 3.8, each W-cell (for all dimensions higher than 1) has at least an extreme ray which is not mapped to the null vector by the corresponding Γ matrix. The only exception seems to be the face generated by $\{(0,1,0),(1,0,0)\}$, however in this case even if the W-cell is 2-dimensional, Lemma 3.8 does not apply, since the entire W-cell is mapped to the null vector and the min-cut subspace is not 1-dimensional.

Let us now consider the maximal recoloring of the graphs models in Figure 3.8 spec-

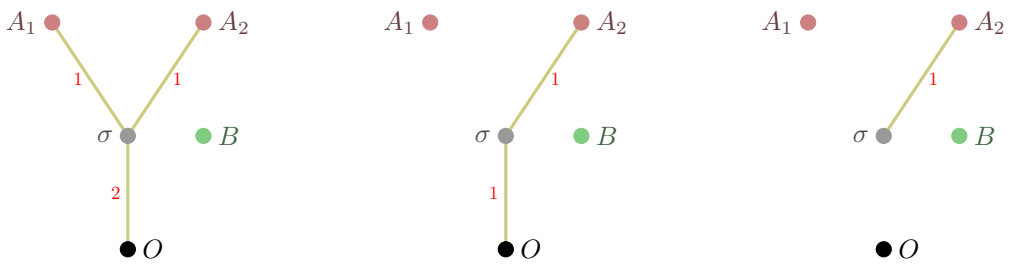


Figure 3.10: Graph models specified by particular choices of weights discussed in the main text.

ified by

$$A_1 \to A, \quad A_2 \to C, \quad B \to B, \quad O \to O \tag{3.115}$$

The new min-cut structures corresponding to the recolored graph models are still generic, and can conveniently be described by their corresponding Γ matrices. They are respectively

$$\Gamma_{\widetilde{G_2}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \Gamma_{\widetilde{G_2}'} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \Gamma_{\widetilde{G_2}''} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \tag{3.116}$$

As one can immediately see, after recoloring, the first graph model in Figure 3.8 has a 3-dimensional min-cut subspace, while the min-cut subspaces of the other two graph models in Figure 3.8 are 2-dimensional, even if the two original graph models belonged to the same W-cell and had a 1-dimensional min-cut subspace.

The W-cells of the first two fine-grained graph models are

$$\begin{aligned} \mathcal{W}_{\widetilde{G_2}} &= \text{cone}\,\{(1,1,0),(0,1,1),(1,0,1)\} \\ \mathcal{W}_{\widetilde{G_2}'} &= \text{cone}\,\{(1,1,0),(1,0,0),(1,0,1)\} \end{aligned} \tag{3.117}$$

These new W-cells are again shown in Figure 3.9, where the dashed edges, and the new vertex $(1,1,0)$, are new W-cells that correspond to new min-cut structures that emerge from the recoloring of the original topological graph model. As one can easily guess from

the figure, there is an additional full-dimensional W-cell

$$\text{cone}\,\{(0,1,0),(1,1,0),(0,1,1)\} \tag{3.118}$$

which can immediately be obtained by swapping A and C in the recolored middle graph model in Figure 3.8. Combined, all these W-cells form a partition of the original W-cell for the first two graph models in Figure 3.8 given in (3.109), in agreement with Lemma 3.16. On the other hand, as also shown in Figure 3.9, the recoloring of the last graph model in Figure 3.8 is not associated to any partition of the original W-cell. The full set of W-cells after the recoloring of the topological graph model can also be seen in Figure 3.11, where we explicitly show the transition between different W-cells as a function of two independent weights (the third can be fixed by rescaling). Notice that there is an additional W-cell that is not visible in Figure B.2, namely the 0-dimensional W-cell corresponding to the origin of $\mathbb{R}^3$.

Finally, let us briefly comment again on the application of Lemma 3.8 to see how it relates to fine-grainings. Starting from $\widetilde{G_2}$ we can apply Lemma 3.8 and select an extreme ray of the closure of the W-cell given in (3.109) such that the min-cut subspace of the new graph model is again the one generated by the S-cell given in (3.111). Suppose that we choose $(0,1,1)$ (cf. the last graph model in Figure 3.10) and then apply the recoloring from (3.115). The W-cell of the resulting graph model is the same as the original one, and the new min-cut structure is degenerate. With the choice of representative $\mathcal{U}_C = \{\{C\}\}$,

the Γ matrix is

$$\Gamma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{3.119}$$

and the (necessarily 1-dimensional) S-cell is

$$\mathcal{S} = \lambda\,(0,0,1,0,1,1,1), \qquad \lambda > 0 \tag{3.120}$$

The coarse-graining associated to the initial coloring of the graph (before the recoloring that induced the fine-graining) specifies the color projection (cf. (3.87))

$$\Phi_{3\to 2} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.121}$$

and one can immediately verify that once applied to (3.120) this transformation gives the original S-cell from (3.111).

3.6 The HEC from marginal independence

In this section we will use the machinery that we have developed thus far to relate the construction of the $\mathrm{HEC}_{\mathsf{N}}$ to the knowledge of set of holographic PMIs.

In Section 3.6.1 we will first look at various graph models which realize the known extreme rays of the $\mathrm{HEC}_{\mathsf{N}}$ up to $\mathsf{N} = 6$, focusing in particular on their PMIs. The reader should not be surprised by the fact that we will look directly at graph models, rather

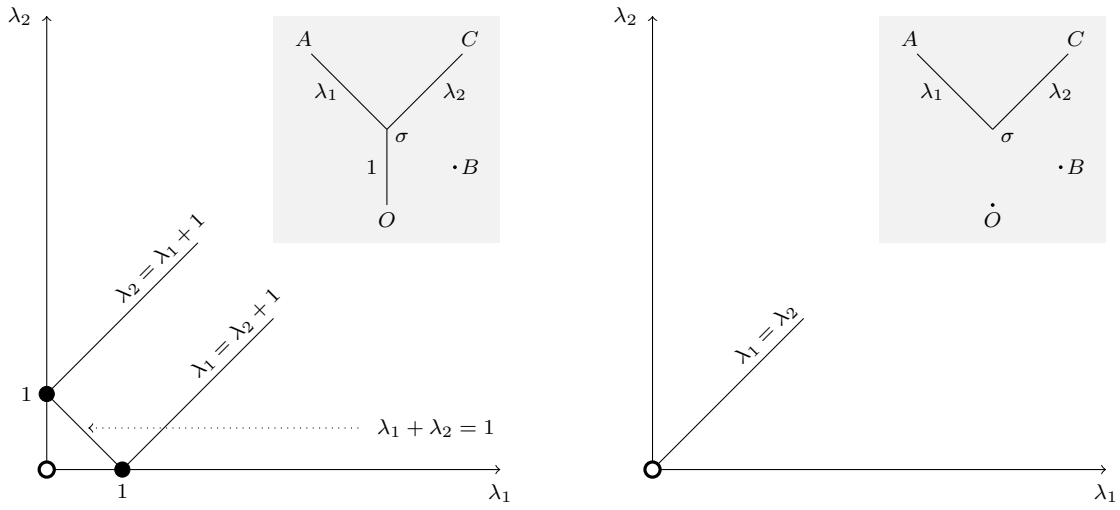


Figure 3.11: The W-cells for the topological graph model underlying the graphs in Figure 3.8, as well as variations of it obtained via the deletion of the appropriate edges. The figure shows the explicit dependence on the values of the weights as well as the degeneracy equations. The shaded insets show the topological graph models associated to the two diagrams, which are distinguished by the presence (left) or absence (right) of the (σ, O) edge.

than equivalence classes, since by Corollary 3.1 each of the graph models discussed here is by itself an equivalence class (up to a global rescaling of the weights). In these cases, an explicit choice of weights should merely be seen as a more compact way of specifying a min-cut structure on the underlying topological graph model.

Based on the evidence presented in Section 3.6.1, in Section 3.6.2 we will then formulate various conjectures about certain properties of the graph models that realize the extreme rays of the $\mathrm{HEC}_{\mathsf{N}}$ for arbitrary N. We will discuss in detail how these conjectures are related to each other, and their implications for the explicit reconstruction of the $\mathrm{HEC}_{\mathsf{N}}$ from the set of holographic PMIs.

In most of the following discussions about extreme rays of the $\mathrm{HEC}_{\mathsf{N}}$, including the formulation of our conjectures (in particular 3.1 and 3.4), we will always implicitly assume that we are focusing only on the restricted set of "genuine N-party" extreme rays,

i.e., on extreme rays which are not embeddings into N-party entropy space of extreme rays for fewer parties. The convenience of this assumption lies in the fact that the components of such extreme rays are strictly positive, which allows us to ignore unnecessary complications related to the connectivity of the graphs. Indeed, if some components of an N-party entropy vector $\mathbf{S}$ vanish, it is always possible to use purity[69] to effectively reduce the number of parties to some $\mathsf{N}' < \mathsf{N}$, and distill the essential information contained in $\mathbf{S}$ into a new N'-party entropy vector $\mathbf{S}'$.

3.6.1 Graph models and PMIs for known extreme rays of the HEC

In Section 3.3.5 we proved that any graph model that realizes an extreme ray of the $\mathrm{HEC}_{\mathsf{N}}$ has a 1-dimensional min-cut subspace, and we have shown an explicit example for $\mathsf{N} = 3$. In that example, the topological graph model G_3 associated to $\widetilde{G_3}$ was a simple tree graph (cf. Definition 3.12). By Theorem 3.2 it follows then that its PMI is also 1-dimensional (it is the min-cut subspace), and that the ray in (3.36) is also an extreme ray of the SAC_3. As it turns out, the same type of situation also occurs for all graph models which realize the extreme rays of the $\mathrm{HEC}_{\mathsf{N}}$, for $\mathsf{N} \in \{2, 3, 4\}$ [42], as the reader can easily verify.

On the other hand, for $\mathsf{N} = 5$, not all realizations of the extreme rays given in [13, 8] were trees. The key observation of this subsection is that one can nonetheless find alternative graph models, which are trees, and realize the same extreme rays. These trees however are no longer simple and this non-simplicity is not just a consequence of particular choices of graph realizations, but rather of the fact that the corresponding

[69] Recall that the von Neumann entropy of a density matrix vanishes if and only if it corresponds to a pure state. Similarly, for graphs, the entropy of a subsystem $\mathcal{I}$ vanishes if and only if the min-cut for $\mathcal{I}$ is disconnected from its complement.

extreme rays of the HEC_5 are not extreme rays of the SAC_5 (their PMIs are higher dimensional, see the next subsection for more details).

In Figure 3.12, we provide a tree graph realization for every HEC_5 extreme ray which was not realized by a tree graph in [8]. Most of these new trees can immediately be obtained from the original graph models from [8] by simply splitting every degree-k boundary vertex into k vertices of the same color.[70] For each of these tree graph models we can then specify a maximal recoloring and consider the associated possible fine-grainings of the corresponding class. As exemplified in the previous section, one possible such fine-gaining is simply obtained via the fine graining of the graph model, i.e., by only applying the recoloring to the boundary vertices, without varying the weights. Despite the fact that in general the dimension of the min-cut subspace can grow after a fine-graining, it turns out that in each case the resulting min-cut subspace is 1-dimensional.[71] Since each of these new graph models is now specified on a simple tree (because the recoloring was maximal), it then follows by Theorem 3.2 that the PMI is also 1-dimensional, and these graph models thus realize extreme rays of the $\mathrm{SAC}_{\mathsf{N}'}$ for the corresponding $\mathsf{N}' = \mathsf{V}_\partial - 1 \geq 5$. For instance, via this procedure, the last graph model in Figure 3.12 lifts to a simple tree graph model for $\mathsf{N}' = 10$, which one can check realizes an extreme ray of the SAC_{10}.[72] Notice that even if the fine grainings specified here may[73] not be minimally-degenerate according to Definition 3.14, the min-cut subspaces before and after the fine-graining are still related by a color-projection, since both S-cells are just single rays.

[70] This is not always sufficient: for example the last graph in Figure 3.12 requires a judicious iterative application of the operations described in Appendix B.1, particularly the Δ-Y exchange operation shown in Figure B.2.

[71] At this point the reader may already wonder whether we could have used Lemma 3.8 in case the resulting min-cut subspace had a higher dimensionality. Indeed this is the case, and we will use this strategy in what follows.

[72] To check extremality of an entropy vector for N parties, one can simply check that it satisfies all instances of SA for that N, and that at least $\mathsf{D} - 1$ linearly independent instances are saturated.

[73] We have not explicitly checked the dimensionality of the W-cells, since it is not crucial here.

For $\mathsf{N} \geq 6$, the HEC_{N} is still unknown, but a large number of its extreme rays is known for $\mathsf{N} = 6$, and we can again explore whether we can realize them with graph models with tree topology. If we succeed, we can then repeat the same procedure we just described for the $\mathsf{N} = 5$ case, to make these tree graph models simple via a maximal recoloring, and check if the resulting min-cut subspace is 1-dimensional. If this is the case, these graph models then realize extreme rays of the $\text{SAC}_{\mathsf{N}'}$ for some $\mathsf{N}' \geq 6$.

At the time of writing, a total of 4122 distinct (orbits of) extreme rays have been discovered. Although only 24 of these can be realized by simple trees, it turns out that as many as 3905 of the others can be realized by graph models that can be immediately turned into non-simple trees by just splitting their boundary vertices (as explained before for the $\mathsf{N} = 5$ case). In other words, only 193 extreme rays are realized by graph models that contain cycles involving only bulk vertices, which can thus not be broken by splitting boundary vertices. However, most of these just contain a single bulk 3-cycle, which is straightforward to break as we show in Appendix B.1. Ultimately, there only remain 14 graphs models which cannot be obviously turned into trees, as they involve larger cycles or more than a single 3-cycle. Figures 3.13 and 3.14 depict graph models realizations containing a bulk 4-cycle for two of the HEC_6 extreme rays. After splitting the boundary vertices and maximally recoloring them, the topological graph models are the same for these two cases. Nevertheless, it is easy to convert the graph of Figure 3.13 into a tree graph (by splitting the bulk cycle at the $\sigma 2$ vertex; cf. the bottom panel), whereas an analogous transformation does not work for the graph of Figure 3.14, even though the original graph looks a bit simpler. For all simple tree graph models obtained via this procedure, the resulting min-cut subspace (and therefore PMI) is indeed 1-dimensional.

In total there are only a few (some of the 14 mentioned above) extreme rays of the HEC_6 which are realized by graph models that we were not immediately able to convert

into tree graphs using these simple operations. While it is possible that some of these extreme rays cannot be realized by any tree graph, we have not attempted a systematic and exhaustive search, and it seems likely that the difficulty in obtaining tree realizations for these extreme rays is just technical rather than fundamental.

3.6.2 Reconstruction of the HEC from marginal independence

We are now ready to discuss the main claim of this work, namely how the reconstruction of the holographic entropy cone is related to the solution of the holographic marginal independence problem. We will not be able to provide a definite proof that such a reconstruction is possible. However, motivated by the observations presented in the previous subsection, we will formulate a series of conjectures each of which will imply that this is indeed the case. We will organize these conjectures following a hierarchy, from the weakest to the strongest, analysing their relations and implications. As we will see, all of these conjectures, except only for the weakest 3.1 (see below), will imply that this reconstruction has a particularly nice form and is intimately related to 1-dimensional holographic PMIs, i.e., holographic extreme rays of the SAC.

Let us begin by explaining more explicitly how the *reconstruction* would work. Since the HEC is a polyhedral cone, it can equivalently be described by providing either the full set of its extreme rays, or the full set of its facets, specified by non-redundant entropy inequalities. Given one description, one can in principle[74] obtain the other using well-known conversion algorithms, and we will focus on the description in terms of extreme rays. Suppose now that we are interested in the $\mathrm{HEC}_{\mathsf{N}^*}$ for a specific number of parties N^*, and we are given the full solution to the HMIP, i.e., we are given the full set of all

[74]There is no known efficient algorithm to perform this conversion, and in practice this is undoable already for $\mathsf{N} = 6$.

holographic PMIs (recall Definition 3.11) for all possible values of N, can we derive all the extreme rays of the $\text{HEC}_{\mathsf{N}^*}$ from this data?

As a warm up, let us first discuss the simpler situation where $\mathsf{N}^* \leq 4$. Suppose that we do not know the extreme rays of the $\text{HEC}_{\mathsf{N}^*}$, but instead we want to try to extract them from the solution to the HMIP for $\mathsf{N} = \mathsf{N}^*$. Such solution is the collection of all PMIs that can be realized holographically for $\mathsf{N} = \mathsf{N}^*$, and it will contain PMIs of different dimensions $1 \leq d \leq \mathsf{D}^*$. For each 1-dimensional PMI it is straightforward to determine an entropy ray that generates it, as one simply needs to pick a ray oriented towards the positive orthant of entropy space. Furthermore, any such ray is automatically an extreme ray of the $\text{HEC}_{\mathsf{N}^*}$ because it is a holographic extreme ray of the $\text{SAC}_{\mathsf{N}^*}$.[75] On the other hand, we will ignore all higher dimensional PMIs, since it is not possible to uniquely determine an entropy ray out of them. We can then build a candidate for the $\text{HEC}_{\mathsf{N}^*}$ by taking the conical hull of the rays thus obtained. The equivalence of the $\text{HEC}_{\mathsf{N}^*}$ with this cone however is just a coincidence related to the fact that we are only considering $\mathsf{N}^* \leq 4$, and as we reviewed in the previous subsection, in this case all extreme rays of the $\text{HEC}_{\mathsf{N}^*}$ correspond to 1-dimensional PMIs.

Let us now focus on $\mathsf{N}^* \geq 5$, and consider an extreme ray $\mathbf{S}$ of the $\text{HEC}_{\mathsf{N}^*}$ which is *not* an extreme ray of the $\text{SAC}_{\mathsf{N}^*}$, and a class $(G_{\mathsf{N}^*}, \mathfrak{m})$ realizing it, i.e., such that

$$\mathbb{S}(G_{\mathsf{N}^*}, \mathfrak{m}) = \mathbb{S} \tag{3.122}$$

where $\mathbb{S}$ is the subspace generated by $\mathbf{S}$. Suppose that there exists a minimally-degenerate fine-graining (cf. Definition 3.14) to a class $(G_{\mathsf{N}}, \check{\mathfrak{m}})$, with $\mathsf{N} > \mathsf{N}^*$ and $\check{\mathfrak{m}} \in \mathfrak{m}\big\uparrow_{\mathsf{N}}$, such that the min-cut subspace and the PMI coincide. Denoting by $\Phi_{\mathsf{N}\to\mathsf{N}^*}$ the corresponding

[75] Notice that any class $(G_{\mathsf{N}}, \mathfrak{m})$ realizing this ray (which exists by assumption) is guaranteed by Lemma 3.10 to have a 1-dimensional min-cut subspace, in agreement with Corollary 3.1.

coarse-graining we then have

$$\mathbb{S} = \Phi_{\mathsf{N}\to\mathsf{N}^*}\, \mathbb{S}(G_{\mathsf{N}}, \check{\mathfrak{m}}) = \Phi_{\mathsf{N}\to\mathsf{N}^*}\, \pi(G_{\mathsf{N}}, \check{\mathfrak{m}}) \tag{3.123}$$

This implies that $\mathbf{S}$ can be obtained from the solution to the HMIP. The reason is that this solution includes $\pi(G_{\mathsf{N}}, \check{\mathfrak{m}})$, from $\pi(G_{\mathsf{N}}, \check{\mathfrak{m}})$ we can obtain $\mathbb{S}$ via a color-projection, and $\mathbf{S}$ can be obtained from $\mathbb{S}$ as described before, since $\mathbb{S}$ is 1-dimensional. This motivates us to formulate our first conjecture

Conjecture 3.1. *For any* N^* *and any extreme ray* $\mathbf{S}$ *of the* $\mathrm{HEC}_{\mathsf{N}^*}$*, there exists a class* $(G_{\mathsf{N}^*}, \mathfrak{m})$ *realizing* $\mathbf{S}$ *and a minimally-degenerate fine-graining to a class* $(G_{\mathsf{N}}, \check{\mathfrak{m}})$ *for some finite* $\mathsf{N} \geq \mathsf{N}^*$ *such that*

$$\mathbb{S}(G_{\mathsf{N}}, \check{\mathfrak{m}}) = \pi(G_{\mathsf{N}}, \check{\mathfrak{m}}) \tag{3.124}$$

and $\pi(G_{\mathsf{N}}, \check{\mathfrak{m}})$ *is nowhere-zero.*[76]

If this conjecture holds, in principle all one needs to do to construct the $\mathrm{HEC}_{\mathsf{N}^*}$ is to consider all possible color-projections from N to N^* of all nowhere-zero PMIs that are realizable holographically, for each $\mathsf{N} \geq \mathsf{N}^*$, pick a non-negative ray from each 1-dimensional subspace obtained by this procedure, and take the conical hull of the resulting set. Notice that not all entropy rays obtained in this way will be extremal, however the ones that are non-extremal are guaranteed to be contained in the $\mathrm{HEC}_{\mathsf{N}^*}$ by the nowhere-zero condition. This follows from the fact that if a holographic PMI is nowhere-zero, it contains at least a (not necessarily extremal) nowhere-zero entropy vector $\mathbf{S}'$ realized by a graph model. Since we are only considering PMIs that are color-projected to 1-dimensional subspaces, and by being nowhere-zero $\mathbf{S}'$ cannot be mapped to the null

[76] Here by a nowhere-zero PMI $\mathbb{P}$ we mean that for any choice of generators of $\mathbb{P}$, none of the components vanishes.

vector by such a projection, the result of the projection must be the subspace generated by the projection of $\mathbf{S}'$, which is inside the $\mathrm{HEC}_{\mathsf{N}^*}$. On the other hand, 3.1 will guarantee that all extreme rays of the $\mathrm{HEC}_{\mathsf{N}^*}$ will be included in the resulting set.

At first sight, this procedure seems daunting, since we are demanding to consider all possible values of $\mathsf{N} \geq \mathsf{N}^*$. In practice however this is never necessary, since for any N^* the $\mathrm{HEC}_{\mathsf{N}^*}$ only has a finite number of extreme rays (it is a polyhedral cone) and can therefore be reconstructed from the solution to the HMIP for $\mathsf{N} = \mathsf{N}_{\max}(\mathsf{N}^*)$ given by

$$\mathsf{N}_{\max}(\mathsf{N}^*) = \max\left\{\mathsf{N}_{\min}(\mathbf{S}),\ \forall\, \mathbf{S}\right\} \tag{3.125}$$

where for each extreme ray $\mathbf{S}$ of the $\mathrm{HEC}_{\mathsf{N}^*}$, $\mathsf{N}_{\min}(\mathbf{S})$ denotes the smallest $\mathsf{N} \geq \mathsf{N}^*$ for which Conjecture 3.1 holds. The reason is that for each extreme ray $\mathbf{S}$ of the $\mathrm{HEC}_{\mathsf{N}^*}$ and PMI $\mathbb{P}$ in the solution to the HMIP for $\mathsf{N} = \mathsf{N}_{\min}(\mathbf{S})$ such that

$$\mathbb{S} = \Phi_{\mathsf{N}_{\min}(\mathbf{S})\to\mathsf{N}^*}\mathbb{P}, \tag{3.126}$$

there is a lift[77] to a new PMI $\mathbb{P}'$ in the solution to the HMIP for $\mathsf{N} = \mathsf{N}_{\max}$ such that

$$\mathbb{P} = \Phi_{\mathsf{N}_{\max}\to\mathsf{N}_{\min}}\mathbb{P}' \tag{3.127}$$

Therefore we can simply obtain $\mathbb{S}$ from $\mathbb{P}'$ as follows

$$\mathbb{S} = \Phi_{\mathsf{N}_{\max}\to\mathsf{N}^*}\mathbb{P}' \tag{3.128}$$

[77] This is a straightforward consequence of the standard lift construction described in Section 3.3.6, and of the fact that the min-cut subspace of a class determines its PMI (cf. Corollary 3.3).

where

$$\Phi_{\mathsf{N}_{\max}\to\mathsf{N}^*} = \Phi_{\mathsf{N}_{\min}\to\mathsf{N}^*}\,\Phi_{\mathsf{N}_{\max}\to\mathsf{N}_{\min}} \tag{3.129}$$

Notice that we are only arguing for the existence of $\mathsf{N}_{\max}(\mathsf{N}^*)$, but we are not providing a way to determine its value for a given N^*. In any concrete reconstruction of the $\mathrm{HEC}_{\mathsf{N}^*}$ one would instead need to know the value of $\mathsf{N}_{\max}(\mathsf{N}^*)$ to make sure that one is considering the solution to the HMIP for a value of N large enough such that it is sufficient to derive the complete set of extreme rays.[78] However, it should by now be clear that the focus of the present work is not on any explicit reconstruction of the HEC, but rather on the equivalence of this problem to the solution of the HMIP.

For small N our evidence strongly supports 3.1 and allows us to derive some tight upper bounds on $\mathsf{N}_{\max}(\mathsf{N}^*)$. For $\mathsf{N} \leq 4$, we have already seen that $\mathsf{N}_{\max}(\mathsf{N}^*) = \mathsf{N}^*$, so the first non-trivial case is $\mathsf{N} = 5$. The graph models shown in the first and third columns of Figure 3.12 correspond to PMIs of dimensions ranging from 2 (e.g., the first one) to 6 (e.g., the last one). To explore Conjecture 3.1, one may consider incremental fine-grainings[79] of these graph models obtained by first splitting boundary vertices of degree higher than one, and then recoloring the boundary vertices. In all cases, one observes that the min-cut subspace remains 1-dimensional upon fine-graining, while the PMI exhibits a non-increasing dimensionality. Remarkably, in most cases one obtains a graph model with a 1-dimensional PMI after only a couple of steps, and in all cases three steps always

[78] In principle one could also proceed as follows. Start from the solution to the HMIP for $\mathsf{N} = \mathsf{N}^*$ and construct a cone as we explained. Then derive its facets and check if they are valid holographic entropy inequalities using the contraction maps of [42]. If all facets are valid inequalities, the reconstruction is complete, otherwise one can increase the value of N and repeat the process. Once $\mathsf{N}_{\max}(\mathsf{N}^*)$ has been reached, the reconstruction is guaranteed to be complete. The problem with this procedure however is that the conversion from extreme rays to facets is inefficient, and that it is currently unknown if all valid holographic inequalities can be proven via contraction maps.

[79] Again, these are particular choices of fine-grainings of the min-cut structure of a graph model specified by the same choice of weights as in the original graph.

suffice.[80] In summary, this allows us to conclude that $\mathsf{N}_{\max}(5) \leq 8$, and that this bound is likely saturated. Although our data for $\mathsf{N}^* = 6$ is incomplete and thus insufficient to derive a similar bound, we highlight that the same qualitative features are observed for all graph models realizing the currently known extreme rays of the HEC_6 (unlike in the previous subsection, here we are not attempting to realize these extreme rays with tree graph models). The splitting of boundary vertices and the straightforward fine-graining of a graph model specified by a maximal recoloring is sufficient to obtain a new graph model such that the corresponding min-cut subspace and PMI coincide.[81]

Notice that while 3.1 in principle allows for fine-grainings to graph models with higher dimensional min-cut subspaces, in all known cases discussed above the equivalence with PMIs was achieved with 1-dimensional min-cut subspaces. This suggests that the reconstruction of the HEC could actually be even simpler, and reliant only on the knowledge of 1-dimensional holographic PMIs. This motivates the following stronger conjecture, consistent with all our data to date:

Conjecture 3.2. *For any* N^* *and any extreme ray* $\mathbf{S}$ *of the* $\mathrm{HEC}_{\mathsf{N}^*}$*, there exists an extreme ray* $\mathbf{R}$ *of the* $\mathrm{SAC}_{\mathsf{N}}$ *for some finite* $\mathsf{N} \geq \mathsf{N}^*$*, a graph model realizing* $\mathbf{R}$ *and a recoloring* $\beta^{\downarrow}$ *such that from the corresponding coarse-graining* ϕ *we have*

$$\mathbf{S} = \Phi_{\mathsf{N}\to\mathsf{N}^*}\,\mathbf{R} \tag{3.130}$$

If this conjecture holds, it is not necessary to know the full solution to the HMIP to

[80] In fact, by taking the currently known extreme rays of the HEC_6 which are also extremal in the SAC_6 (equivalently, extreme rays of the SAC_6 which are holographic), and coarse-graining down to 5 parties, we observe that as many as 16 out of the 19 orbits of extreme rays of the HEC_5 are obtained. In other words, remarkably, 16 of the extreme rays of the HEC_5 descend from the $\mathrm{SAC}_{\mathsf{N}}$ already at $\mathsf{N} = 6$.

[81] Technically this observation is not an exact confirmation that 3.1 holds in these cases, since a priori the fine-grainings of the min-cut structures may not be minimally-degenerate. However, we can also imagine to first use Lemma 3.8 before the recoloring, to obtain a 1-dimensional W-cell. The fine-graining induced by the recoloring is then guaranteed to be minimally-degenerate, and 3.1 holds.

reconstruct the HEC. Instead it is sufficient to know only the 1-dimensional PMIs that are holographic or, equivalently, the extreme rays of the SAC that can be realized by graph models. We could then redefine $\mathsf{N}_{\max}(\mathsf{N}^*)$ as the smallest number of parties from which all $\mathrm{HEC}_{\mathsf{N}^*}$ extreme rays can be recovered as in (3.130). Explicitly, the $\mathrm{HEC}_{\mathsf{N}^*}$ would then be given by

$$\mathrm{HEC}_{\mathsf{N}^*} = \mathrm{cone}\left\{\Phi_{\mathsf{N}_{\max}\to\mathsf{N}^*}\,\mathbf{R},\ \forall\,\mathbf{R}\in\mathcal{R}_{\mathrm{H}}^{\mathsf{N}_{\max}},\ \forall\,\Phi_{\mathsf{N}_{\max}\to\mathsf{N}^*}\right\} \tag{3.131}$$

where $\mathcal{R}_{\mathrm{H}}^{\mathsf{N}}$ is the set of holographic extreme rays of the $\mathrm{SAC}_{\mathsf{N}}$. For instance, from the bound on $\mathsf{N}_{\max}$ mentioned above for $\mathsf{N}^* = 5$ we see that the HEC_5 can be obtained from the conical hull of just those extreme rays of the SAC_8 which are holographic. Notice that from a structural point of view this a highly non-trivial statement, since the projections $\Phi_{\mathsf{N}_{\max}\to\mathsf{N}^*}$ that appear in (3.131) are not arbitrary projections, but color-projections associated to coarse-grainings. Finally, notice that the reconstruction formula (3.131) also gives a full solution to the HMIP for $\mathsf{N} = \mathsf{N}^*$, as one can recover all holographic PMIs from the knowledge of the extreme rays.

Due to its generality however, proving 3.2 might be quite challenging, in particular because the topology of the graph models that realize the holographic extreme rays of the SAC is not restricted in any way. On the other hand, the observations at the end of the previous subsection suggest that there might be an additional simplification. We have seen that for all extreme rays of the HEC_5 we could find realizations by graph models with tree topology. And that for all these tree graph models, the fine-graining specified by a maximal recoloring was automatically giving a new (simple tree) graph model with a 1-dimensional PMI. For instance, the tree graphs exhibited in Figure 3.12 show that all the extreme rays of the HEC_5 can be obtained by coarse-graining extreme

rays of the $\mathrm{SAC}_{\mathsf{N}}$ realizable by simple tree graph models for some $5 < \mathsf{N} \leq 11$.[82] Similar facts seem to hold for the known extreme rays of the HEC_6: by just applying the simple entropy-preserving graph manipulations described in Appendix B.1, we are able to turn almost all graph models realizing these rays into new graph models with tree topology. One of just a handful of exceptions is the graph in Figure 3.14. However, we stress that despite the fact that we have not found a tree form for it to support 3.3, one can easily check that after a splitting of e.g. vertex C alone, and a maximal recoloring, one already obtains a graph model realizing an extreme ray of the SAC_7, in agreement with 3.2.

Based on these observations, it is therefore interesting to explore the implications of the following conjecture.

Conjecture 3.3. *For any* N^**, every (nowhere-zero) extreme ray of the* $\mathrm{HEC}_{\mathsf{N}^*}$ *is realizable by a graph model with tree topology.*

The reader will probably already guess how this new conjecture is related to 3.2 and the reconstruction problem, since we have already seen several examples of this situation. However let us explain more carefully why 3.2 implies that the reconstruction is given by (3.131).

Consider an extreme ray $\mathbf{S}$ of the $\mathrm{HEC}_{\mathsf{N}^*}$, and suppose that 3.3 holds. Then there exists a class $(G_{\mathsf{N}^*}, \mathfrak{m})$ whose S-cell is the ray generated by $\mathbf{S}$, and we denote by $\mathbb{S}$ its (1-dimensional) min-cut subspace. In general the W-cell of $(G_{\mathsf{N}^*}, \mathfrak{m})$ could be higher dimensional, and an arbitrary fine-graining could produce a new class whose min-cut subspace has dimension greater than one (recall the example at the end of Section 3.5.2). However, if this is the case, we can always use Lemma 3.8 to find a new class $(\hat{G}_{\mathsf{N}^*}, \hat{\mathfrak{m}})$ such that the min-cut subspace is still $\mathbb{S}$, while the S-cell is just a single entropy ray.

[82] Specifically, this upper bound comes from the tree counterpart to the only non-planar graph in Figure 3.12, which requires the highest lift of all graphs for obtaining a simple tree: from $\mathsf{N}^* = 5$ to $\mathsf{N} = 11$.

Notice that, as we have already seen in various examples, when we use Lemma 3.8 we may be forced to apply the reduction described in Section 3.3.4 and delete some of the edges. We can now apply Theorem 3.4 to $(\hat{G}_{\mathsf{N}^*}, \hat{\mathfrak{m}})$ to get yet a new class $(\check{G}_{\mathsf{N}}, \check{\mathfrak{m}})$, for some $\mathsf{N} > \mathsf{N}^*$, where $\check{G}_{\mathsf{N}}$ is now a simple tree and the color-projection $\Phi_{\mathsf{N}\to\mathsf{N}^*}$ of its PMI is $\mathbb{S}$. Finally, since the W-cell of $(\hat{G}_{\mathsf{N}^*}, \hat{\mathfrak{m}})$ was 1-dimensional, it is guaranteed by Lemma 3.16 to remain 1-dimensional after the fine-graining, and the PMI of $(\check{G}_{\mathsf{N}}, \check{\mathfrak{m}})$ is therefore 1-dimensional as well.

In summary, assuming 3.3, we have shown how to construct a new class that satisfies 3.2, showing that 3.3 is sufficient for the reconstruction of the HEC via (3.131).

We conclude by mentioning one more conjecture which is not specific to any particular set of faces of the HEC or the SAC, and could be an interesting general property of all min-cut structure on arbitrary topological graph models.

Conjecture 3.4. *For any* N*, topological graph model* G_{N} *and min-cut structure* $\mathfrak{m}$*, there exists a topological graph model* G'_{N} with tree topology *and a min-cut structure* $\mathfrak{m}'$ *such that*

$$\mathbb{S}(G'_{\mathsf{N}}, \mathfrak{m}') = \mathbb{S}(G_{\mathsf{N}}, \mathfrak{m}) \tag{3.132}$$

A priori one may have doubts about this conjecture, considering how special tree graphs are compared to arbitrary graphs. Notice however that the trees in 3.1 are not required to be simple, and can have an arbitrary number of boundary vertices for each color. Finally should be clear that if one could prove 3.4, then 3.3 would follow immediately. By the argument presented above, 3.3 would then imply 3.2, and the reconstruction of the $\mathrm{HEC}_{\mathsf{N}^*}$ from the extreme rays of the $\mathrm{SAC}_{\mathsf{N}_{\max}}$ for all N^* would be given by (3.131).

3.7 Discussion

In this section we briefly summarize our results and what remains to be proven to definitely conclude that the solution to the holographic marginal independence problem contains sufficient information for the reconstruction of the holographic entropy cone. Moreover, we comment on several interesting directions for future investigations which are suggested by the possibility of such a reconstruction, and by the framework that we have developed throughout this work.

Summary: The main goal of this work was to argue that the holographic entropy cone can be reconstructed, for an arbitrary number of parties, from the solution to the holographic marginal independence problem. We have shown how such a reconstruction would work, and argued that it takes a particularly simple form. Indeed, we argued that to be able to reconstruct the HEC one does not even need to know the full solution to the HMIP, but only the set of 1-dimensional holographic PMIs. Ultimately, what this implies is the following

Claim. *The knowledge of all holographic entropy inequalities, for all values of* N*, is equivalent to the knowledge of all holographic extreme rays of the* SAC*, for all values of* N*.*

A summarizing cartoon of this equivalence and of the reconstruction procedure is provided in Figure 3.15. We stress that, as we explained above, this claim does not hold for a fixed value of N when $\mathsf{N} \geq 5$, i.e., knowing the holographic extreme rays of the SAC for N parties does not in general provide enough information to derive all the N-party holographic entropy inequalities.

Completing the proof: To complete the proof of our main claim above it remains to prove 3.2. This problem however might be challenging, since 3.2 is specific to extreme rays, which are unknown (and are precisely what we want to reconstruct), and the graph models realizing them are allowed to have unrestricted topology. We believe that the best way to proceed is to prove 3.3. Although we were not able to check that this conjecture holds for all known extreme rays of the HEC_6, we also did not attempt an exhaustive search, and it seems likely that one simply needs to consider more complicated tree graphs.

While in principle it is possible that 3.3 only holds for extreme rays, it also seems reasonable to expect that it is a more general fact about 1-dimensional min-cut subspaces, in which case proving this conjecture would presumably be easier. In fact, the realizability of min-cut subspaces via tree graphs might be even more general, and one could also attempt to directly prove 3.4. An interesting way to proceed in this direction would be to better understand the structure of the partition of the space of edge weights into W-cells for a given topological graph model. Since this structure also determines the maps to entropy space that produce the min-cut subspaces for the various min-cut structures, if one could directly relate this structure to the topology of the graph, one might then be able to prove that any min-cut subspace can be obtained from a tree.

Finally, we should also contemplate the possibility that our main claim above is false. In such a case it might still be possible to reconstruct the HEC from the solution to the HMIP, but it would not be sufficient to only know the holographic extreme rays of the SAC. To prove that the reconstruction is possible one should then try to prove 3.1, although this conjecture suffers from the same complications we just mentioned for 3.2, since it is specific to extreme rays and the graphs have unrestricted topology. However, one may also wonder if for each equivalence class of graph models (not just those which

realize extreme rays), there always exists a fine-graining to a new one such that the min-cut subspace and the PMI coincide. While stronger then 3.1, this property might be easier to prove by investigating more deeply the relation between min-cut subspaces and PMIs for arbitrary topological graph models and min-cut structures.

Finding the holographic extreme rays of the SAC: If our main claim above holds, the problem of determining the HEC is mapped to the problem of determining which extreme rays of the SAC are holographic. While this is beyond the scope of the present work, we briefly comment on the importance of this question and some intriguing possible answers.

For a fixed number of parties N, consider the following sets of extreme rays of the SAC_{N} (dropping the N-dependence to simplify the notation)

$$\begin{aligned}
\mathcal{R} &\coloneqq \{\text{extreme rays of the SAC}\} \\
\mathcal{R}_{\text{SSA}} &\coloneqq \{\text{extreme rays of the SAC that satisfy SSA}\} \\
\mathcal{R}_{\text{Q}} &\coloneqq \{\text{extreme rays of the SAC that can be realized by quantum states}\} \\
\mathcal{R}_{\text{H}} &\coloneqq \{\text{extreme rays of the SAC that can be realized by graph models}\}
\end{aligned} \tag{3.133}$$

These sets clearly satisfy the following chain of inclusions

$$\mathcal{R}_{\text{H}} \subseteq \mathcal{R}_{\text{Q}} \subseteq \mathcal{R}_{\text{SSA}} \subseteq \mathcal{R} \tag{3.134}$$

and we have $\mathcal{R}_{\text{SSA}} = \mathcal{R}$ only for $\mathsf{N} = 2$, while typically $|\mathcal{R}_{\text{SSA}}| \ll |\mathcal{R}|$.

If our main claim holds, the deep question about the physical origin of the HEC and its structure is distilled to the question about the relation between the sets $\mathcal{R}_{\text{H}}$ and

$\mathcal{R}_{\text{Q}}$. If $\mathcal{R}_{\text{H}}$=$\mathcal{R}_{\text{Q}}$, all extreme rays of the SAC compatible with quantum mechanics would participate in the construction of the HEC, and there would be no other holographic constraint to resolve. While all data currently available points in this direction, it is conceivable that for larger values of N there exist extreme rays of the SAC that can be realized by quantum states but not by geometric states. This would hint at more fundamental holographic constraints which operate at the level of the SAC and would be very interesting to investigate.

Answering this question however could be quite difficult, since very little is known about the structure of the QEC, and to the best of our knowledge, there is no systematic construction of quantum states that realize the extreme rays of the SAC. A more approachable question on the other hand could be whether $\mathcal{R}_{\text{H}}$ and $\mathcal{R}_{\text{SSA}}$ coincide. Interestingly, for $\mathsf{N} \leq 5$, for which the elements of $\mathcal{R}_{\text{SSA}}$ can be determined using standard algorithms, this turns out to be the case. If $\mathcal{R}_{\text{H}} = \mathcal{R}_{\text{SSA}}$, the chain of inclusions (3.134) immediately implies that $\mathcal{R}_{\text{H}} = \mathcal{R}_{\text{Q}}$ and $\mathcal{R}_{\text{Q}} = \mathcal{R}_{\text{SSA}}$, and we would not only obtain a full characterization of the HEC, but also new information about the QEC for arbitrary N. In order to establish if $\mathcal{R}_{\text{H}} = \mathcal{R}_{\text{SSA}}$, the structure of $\mathcal{R}_{\text{SSA}}$ is currently being investigated. The goal is to obtain a sufficiently detailed characterization of the elements of this set, such that they could then be matched with explicit holographic realizations in terms of graph models.

Generalization to HRT: In this work we focused on the static version of the holographic entropy cone, since this is the situation where configurations of HRRT surfaces can be described by graph models. The connection between the HEC and the solution to the HMIP however seems to suggest that the same structure would pertain also to dynamical spacetimes. Intuitively, the reason is that the crucial information about the

configurations of HRRT surfaces is not rooted in the area of the individual surfaces, but rather the connectivity of the entanglement wedges of the various subsystems. It would be interesting to explore this possibility in more detail.[83]

The most convenient contingency would be one wherein one can devise a graph model for the general time-dependent situation, with the subsystem entanglement entropy again given by the sum of the edge weights of the cut edges for a suitably-defined min-cut, in which case the graph side of the HEC construction would be identical. The immediate obstacle with this approach arises from the fact that parts of the HRT surfaces for various regions can be timelike-separated, and more importantly that the surfaces in question have their areas extremized rather than simply minimized.

The holographic entropy polyhedron (HEP): Many of the tools that we used here have a close relationship with those introduced in [68, 67] in the context of the *holographic entropy arrangement*. The reader who is familiar with these works will probably notice for example the similarity between the min-cut structure of a topological graph model and the proto-entropy vector of a holographic configuration. Analogously, the min-cut subspace is closely related to the set of constraints that must be satisfied by the coefficients of an unspecified "information quantity" (defined as an arbitrary linear combination of entropies, or equivalently as an element of the dual of entropy space) such that it vanishes on a full S-cell. When a min-cut subspace $\mathbb{S}$ has dimension $\mathsf{D}-1$ in fact, $\mathbb{S}^{\perp}$ is the vector of coefficients of the only information quantity that vanishes on all vectors in the S-cell whose linear span is $\mathbb{S}$. This is the type of information quantity that was dubbed a "primitive" in [68, 67]. The holographic entropy arrangement was then defined as

[83] For two-dimensional CFT, it was proven in [65] that any holographic entropy inequality which holds for RT surfaces also holds for HRT surfaces. Despite the limitations related to the special properties of a 3-dimensional bulk spacetime, it would be useful to clarify the relation between the present work and the result of [65], where it was SSA rather than SA that played a crucial role.

the set of all hyperplanes in N-party entropy space which correspond to these primitive quantities, and the HEP as the polyhedron carved by the entropy inequalities associated to the primitive quantities that have a definite sign for geometric states.

In light of the results of the present work, it is tempting to identify the holographic entropy arrangement with the set of codimension-1 min-cut subspaces, and the HEP with the HEC. Indeed, as we have shown here, each extreme ray of the HEC corresponds to a 1-dimensional min-cut subspace, and the min-cut subspace of a graph obtained from the disjoint union of multiple graphs is the sum of the min-cut subspaces of the individual graphs (cf. Section 3.3.6). Using this construction one could then combine the graphs realizing the extreme rays of the HEC to obtain new graphs whose min-cut subspaces are the supporting subspaces of the facets of the HEC. All information quantities corresponding to holographic entropy inequalities would then be primitive in the sense of [68, 67], and the holographic entropy cone and polyhedron would coincide.

This simple reasoning however obfuscates some of the subtleties that were at the core of the motivations of [67] for introducing the HEP in the first place. Throughout this work we have constantly seen that a crucial role in determining the structure, or certain transformation properties, of the objects that we introduced has been played by the "pattern of degeneracies" of an equivalence class of graph models. This is especially true for the realization of the extreme rays of the HEC via graph models, which as we have seen, have a 1-dimensional min-cut subspace and are therefore "maximally degenerate". In terms of holographic configurations and extremal surfaces, these degeneracies seem to translate to situations where the configurations are so fine-tuned that the entropies of multiple subsystems are computed by several coexisting surfaces of equal area. In many other contexts however, like bulk reconstruction, one is typically inclined to ignore such fine-tuned cases to avoid worrying about subtleties regarding order of limits [98] by fo-

cusing instead on generic situations. Indeed, this was precisely the approach followed by [67], which defined primitive information quantities by restricting to generic configurations. In the language of this work, resolving this issue seems to boil down to the question of whether the min-cut subspaces which support the facets of the HEC can be realized by generic equivalence classes, rather than via the aforementioned construction based on extreme rays. The precise details about the connection between the two formulations however could be more subtle, and requires more careful investigations.

Finally, a similar fine-tuning is related to the concrete holographic realizations of graph models, which is obtained via 3-dimensional multiboundary wormhole geometries [42] where the subsystems are chosen to cover the entire boundaries. This choice in fact allows for a fine-tuning of the entanglement among the CFTs that live on the different boundaries, which has no counterpart for subsystems of a single boundary, where the entropies are divergent and cannot be meaningfully regulated. For this reason, [67] defined the HEP by further restricting the set of primitive quantities to those that can be realized by single-boundary configurations. Work is in progress to shed light also on this issue, and to resolve the subtle differences between the HEC and the HEP.

Quantum entropies: Our discussion in this chapter has focused on entropies realizable by geometric states in holography, but it would be interesting to explore if some of the tools developed here can also be utilized to analyze properties of quantum entropies for other classes of states. For example, models of entanglement capturing richer patterns of quantum correlations through hypergraphs and topological links have been explored in [11, 9, 69, 4] and much of the structure we studied here admits a natural extrapolation to those settings. It is then interesting to ask if any of our results and conjectures could be suitably generalized. For instance, are min-cuts the only ingredient at the

heart of why the HEC descends from the SAC in the sense of Conjecture 3.2? If so, it would be reasonable to expect an analogous phenomenon with hypergraphs, namely, that the hypergraph entropy cone be obtainable from extreme rays of the SAC that are realizable by hypergraphs. This poses a non-trivial question already at 4 parties, when the hypergraph entropy cone includes an extreme ray which is not an extreme ray of the SAC_4, and it is unclear whether it can be obtained as a color projection of an extreme ray of the $\mathrm{SAC}_{\mathsf{N}}$ for some $\mathsf{N} > 4$ [11].

Quantum corrections: The graph models studied here precisely capture the properties of the RT prescription in holography. What can our results say about holographic entanglement entropy beyond the strictly classical limit where entropies are purely geometric? It would be interesting to explore how quantum contributions from matter fields affect our results. The application of the combinatorial machinery to the study of entropies computed by the quantum extremal surface prescription was already shown to be fruitful in [5]. A similar reasoning could be used here to discern the imprint that the area term leaves on the generalized entropy and the discrete structure potentially emerging from it.

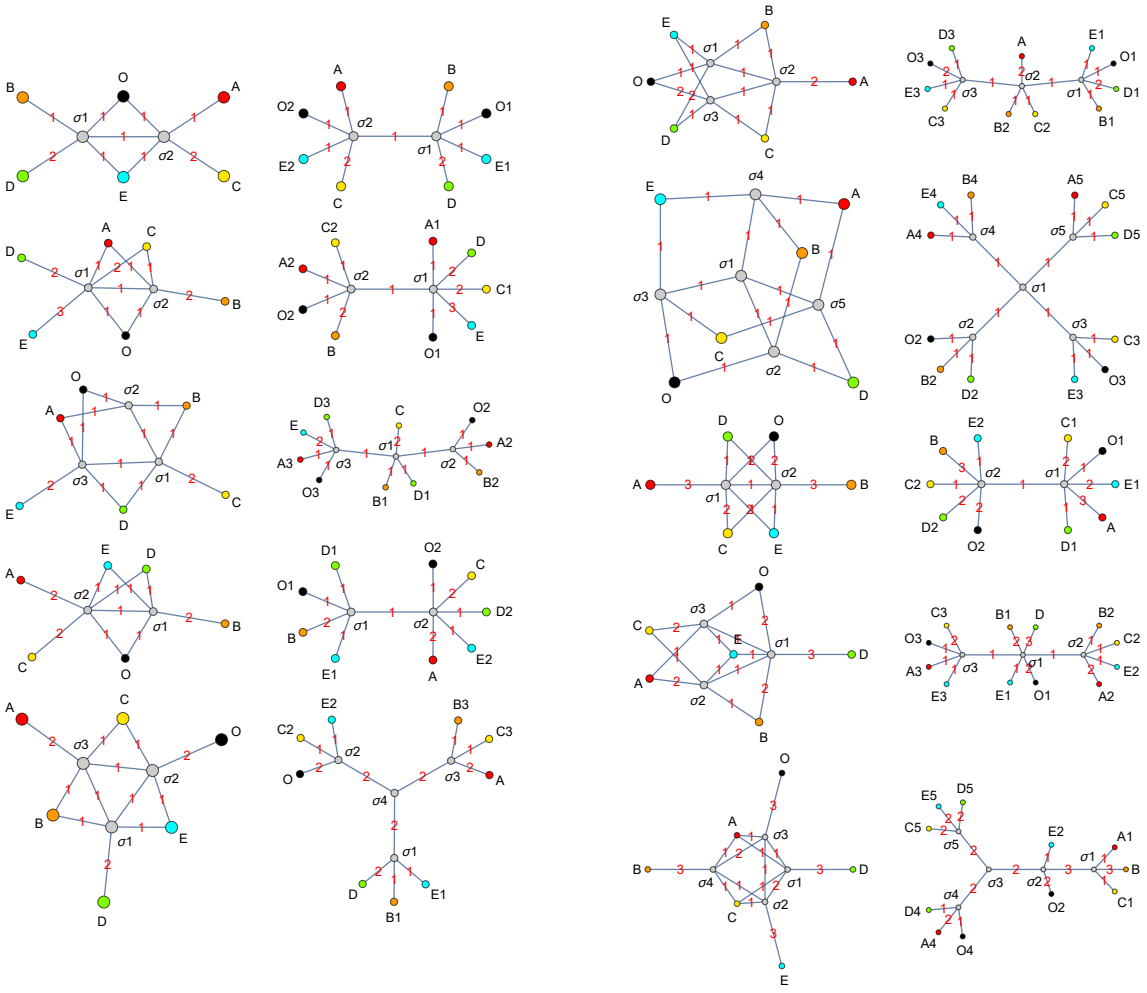


Figure 3.12: The first and third columns show graph realizations of the extreme rays of the HEC_5 which cannot be realized by simple trees. Rather than their original form from [13], here we show the graphs of minimal number of vertices obtained by [8]. To the right of each graph, we provide a non-simple tree realization of the same extreme ray, obtained through repeated application of the graph operations described in Appendix B.1.

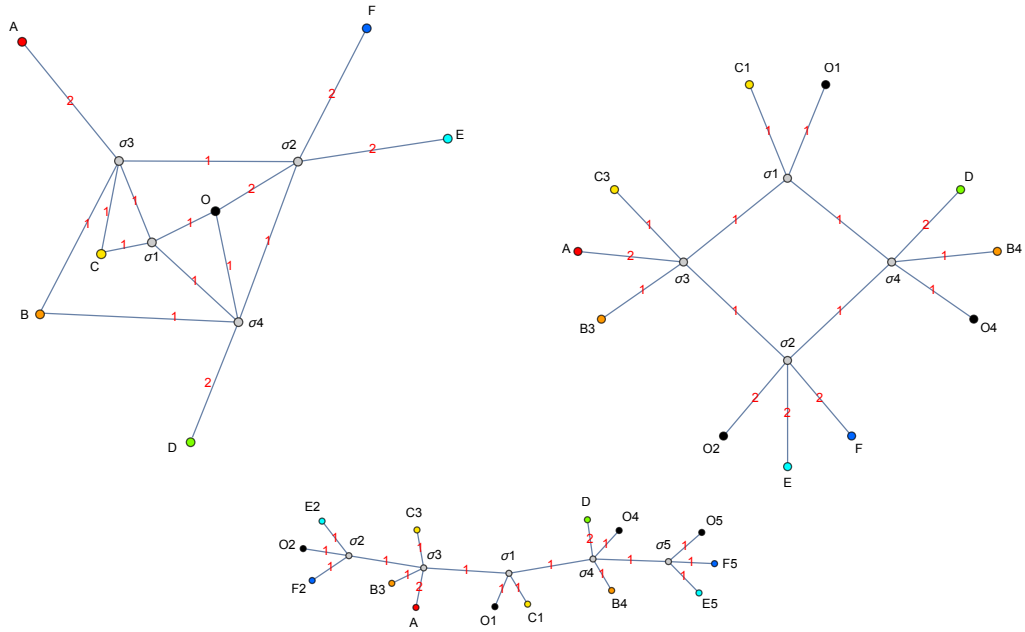


Figure 3.13: Alternative graph realizations of one of the extreme rays of the HEC_6. Notice that one of these realizations is a tree graph.

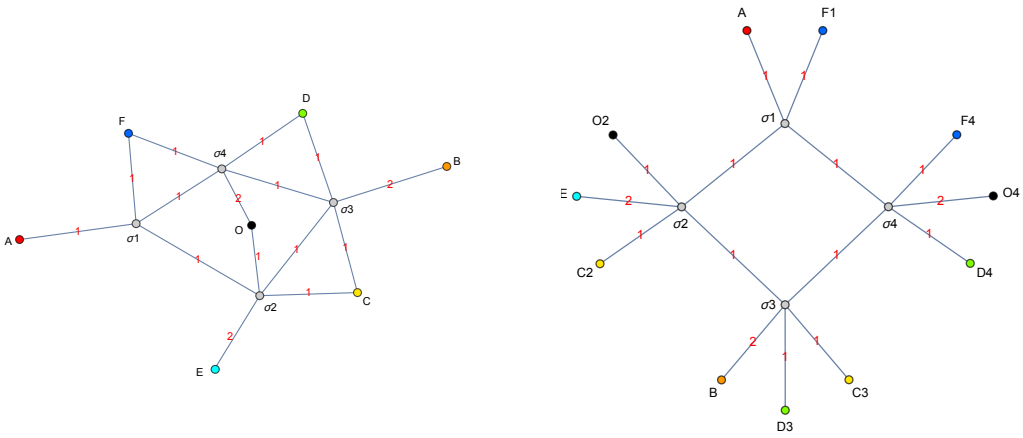


Figure 3.14: Alternative graph realizations of another extreme ray of the HEC_6. Notice that the graph on the right has the same topology as the second graph in Figure 3.13.

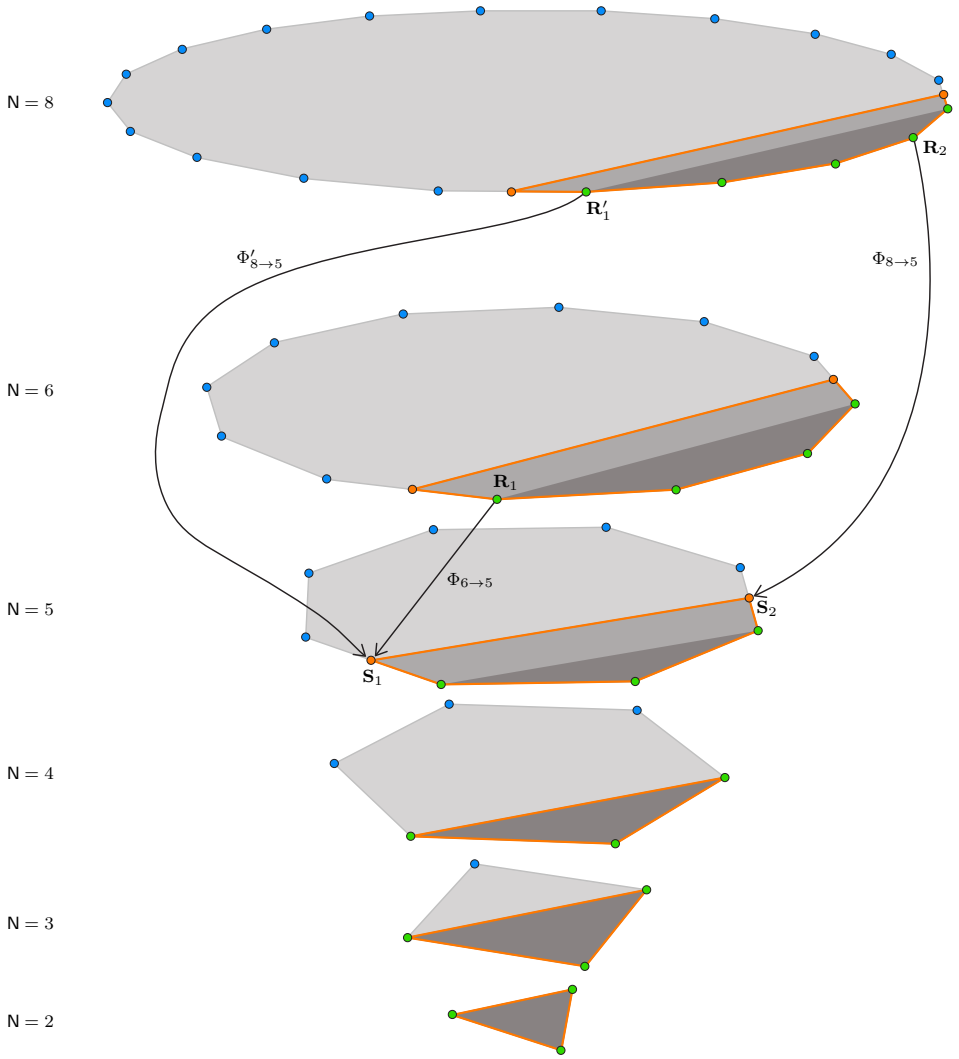


Figure 3.15: A schematic cartoon of a cross-section of the SAC_N and the HEC_N for different values of N. The extreme rays of the SAC_N are represented by the green and blue vertices. The green ones are also extreme rays of the HEC_N, and their conical hull is the darkest shaded region. The blue ones are outside of the HEC_N. The boundary of the HEC_N is highlighted in orange, and the orange vertices represent extreme rays of the HEC_N which are *not* extreme rays of the SAC_N. The arrows indicate color-projections associated with coarse-grainings that implement the reconstruction, mapping the green vertices to the orange. For $\mathsf{N} = 5$ the figure shows two extreme rays $\mathbf{S}_1$ and $\mathbf{S}_2$ that are not extreme rays of the SAC_5 and need to be reconstructed. While $\mathsf{N} = 6$ is sufficient to obtain $\mathbf{S}_1$ as $\mathbf{S}_1 = \Phi_{6\to5}\mathbf{R}_1$, it is not sufficient to reconstruct $\mathbf{S}_2$. On the other hand, $\mathsf{N} = 8$ is sufficient to reconstruct both rays, $\mathbf{S}_1 = \Phi'_{8\to5}\mathbf{R}'_1$ and $\mathbf{S}_2 = \Phi_{8\to5}\mathbf{R}_2$. To avoid clutter, we are suppressing all irrelevant projections/lifts of other extreme rays. The shapes of the cone cross-sections (as well as the numbers of facets and extreme rays of various kinds) are not meant to be taken literally.

Chapter 4

Quantum extremal surfaces and the holographic entropy cone

4.1 Introduction

The holographic principle has served as a remarkable guide for research in quantum gravity [14, 15]. Its most prominent realization, the AdS/CFT correspondence [16, 17], has in recent years elucidated the underlying mechanism for bulk gravitational observables to be encoded holographically on the boundary theory [78, 99, 79, 100]. The derivation of these results has heavily relied on a progressively sophisticated understanding of holographic entanglement entropy.

The program of relating the von Neumann entropy of boundary regions to bulk quantities started with the groundwork of [32], which provided the first version of this entry in the holographic dictionary, known as the Ryu-Takayanagi (RT) formula. According to it, the von Neumann entropy $S(R)$ of a boundary region R is given by the area $\mathcal{A}(\gamma_R)$ of a minimal-area bulk codimension-2 surface γ_R homologous to R relative to ∂R. More

explicitly,

$$S_{\text{RT}}(R) = \min_{\gamma_R} \frac{\mathcal{A}(\gamma_R)}{4G}, \tag{4.1}$$

where the minimization is over surfaces γ_R contained in a time-symmetric slice and anchored to R, i.e., subject to the condition $\partial\gamma_R = \partial R$. This relation underpins the idea of geometry emerging from entanglement, an important theme in our understanding of holography [39]. The RT formula, (4.1), and its time-dependent [36] and quantum generalizations [37, 38] have been exploited in various ways in the context of bulk reconstruction from boundary data [78, 99, 79, 100, 101, 102, 86, 103].

Given the insights that holography provides, it has been of great interest to study necessary and sufficient conditions for field theories to have holographic duals [104]. Within such boundary theories, there is a "code subspace" of states whose bulk duals admit semiclassical geometric descriptions [78, 105]. These so-called holographic states are a special sub-class of all quantum states, satisfying certain additional constraints. The RT formula has allowed us to identify some of these constraints as inequalities that the entropies of boundary subregions need to satisfy. The first such genuinely holographic inequality that was discovered was the monogamy of mutual information (MMI) [40, 95], an inequality on 3 parties which reads

$$S(\text{AB}) + S(\text{AC}) + S(\text{BC}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABC}). \tag{4.2}$$

More generally, for any set of parties $[n] \equiv \{1, 2, \ldots, n\}$, one expects the RT formula to imply inequalities involving various combinations of the entropies of the $2^n - 1$ subsystems associated to the nonempty subsets $\varnothing \neq I \subseteq [n]$. A systematic study of such entropy inequalities obeyed by the RT formula was initiated by [42]. The starting point is to, for each $n \geq 1$, arrange the $2^n - 1$ subsystem entropies of general mixed states on n

parties into entropy vectors $S \in \mathbb{R}^{2^n-1}$. This allows one to think of entropy inequalities as bounding a space of allowed entropy vectors in $\mathbb{R}^{2^n-1}$. [42] was able to prove that this space is a polyhedral cone, i.e., a convex space bounded by a finite set of homogeneous, linear inequalities at each n, which they termed the Holographic Entropy Cone (HEC). Since then, there has been a lot of interest in understanding the HEC [67, 68, 13, 106, 66, 8] and the general structure of entanglement in holographic states [107, 72, 108, 109].

An important caveat, however, is that the RT formula only works to leading order in G when the amount of entropy in bulk matter is small, i.e., when the geometric $\mathcal{O}(1/G)$ term dominates. The perturbative generalization of (4.1) to all orders in G utilizes the notion of a quantum extremal surface (QES) [38]. A bulk surface γ_R anchored to R is called a QES if it extremizes the generalized entropy functional

$$S_{\text{gen}}(R) = \frac{\mathcal{A}(\gamma_R)}{4G} + S_{\text{bulk}}(\sigma_R), \tag{4.3}$$

where the last term is the von Neumann entropy of bulk quantum fields on any achronal homology region σ_R defining the entanglement wedge of R and satisfying $\partial\sigma_R = \gamma_R \cup R$. Accordingly, the QES formula states that $S(R)$ is given by

$$S_{\text{QES}}(R) = \min\left\{\underset{\gamma_R}{\text{ext}}\ S_{\text{gen}}(R)\right\}, \tag{4.4}$$

which instructs one to pick the minimal value over extrema of the generalized entropy if there happens to be more than one QES. This generalization has been of paramount importance in recent computations of the unitary Page curve of black hole evaporation [48, 47], a remarkable development in our understanding of the black hole information problem [50, 110, 49].[1]

[1] For a review of these topics, see Refs. [111, 112].

Importantly, in settings where (4.4) receives contributions from the bulk entropy term S_{bulk} of (4.3) in unrestricted ways, the arguments for the HEC inequalities no longer apply. In fact, one expects that by choosing the bulk matter to have arbitrary entanglement structures, there are no entropy constraints on the boundary quantum state other than the universal ones obeyed by quantum states (such as strong subadditivity). Hence the interesting question to ask is what effect entropic constraints on the bulk matter fields have on the entanglement structure of the boundary state.

To make this problem more precise, let us introduce the notion of a QES entropy cone as the space of all entropy vectors compatible with the QES formula, (4.4), at a moment of time symmetry (analogous to the HEC for RT).[2] As alluded to above, we expect such a QES cone to coincide with the general quantum cone of [56]. In general, the QES cone can be shown to contain the quantum cone as a subset since one can construct situations where the area term can be neglected. Further, it seems plausible that the QES cone is the same as the quantum cone since one can use tensor networks with bulk legs to construct states where the QES formula applies [91].

One could then impose constraints on S_{bulk}, beyond those of quantum mechanics, to obtain a constrained QES cone which will naturally be a subset of the original one. Such constrained QES cones cannot be arbitrarily small: regardless of how strong the constraints one imposes on the bulk entropy are, the QES formula will always be able to probe all possible entropies compatible with HEC inequalities by simply neglecting the bulk entropy term. Hence we find that constrained QES cones will always be supersets of the HEC and, in that sense, be quantum-corrected versions of the HEC. Indeed, we see that by varying the restrictions on the quantum contribution S_{bulk}, constrained QES

[2] A generalization to nontrivial time-dependent settings will be no less subtle than that of understanding the analogue of the HEC for HRT, a problem which remains barely understood except for specific situations such as the low-dimensional case studied in [65].

cones nicely interpolate between the "classical" HEC and the fully quantum entropy cone.

In this chapter, we initiate the program of understanding quantum corrections to the HEC by looking at a couple of specific constrained QES cones. In particular, we investigate the QES cone 1) when S_{bulk} is constrained to obey all HEC inequalities and 2) when S_{bulk} is constrained to obey just MMI. We refer to these as the HEC-constrained QES cone and MMI-constrained QES cone, respectively.

For instance, in order for the boundary to satisfy MMI, which is a 3-party inequality, it was shown in [113] that the bulk state would have to satisfy a specific 7-party inequality, namely,

$$S(\text{ABDG}) + S(\text{ACEG}) + S(\text{BCFG}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABCDEFG}). \tag{4.5}$$

More generally, we will see that any given n-party holographic entropy inequality holds for S_{QES} so long as the bulk state satisfies a certain n'-party inequality for $n' \gg n$, where n' is generally doubly-exponential in n. To make precise the relation between bulk and boundary entropy constraints, here we prove the following two important results:

- **Result 1**: Bulk HEC $\Rightarrow$ Boundary HEC
- **Result 2**: Bulk MMI $\Rightarrow$ Boundary MMI

In more detail, **Result 1** states that the HEC-constrained QES cone is the same as the HEC. We motivate this result in Sec. 4.2 by using the idea of double holography, described in [110, 114]. Intuitively, a doubly holographic setup ensures that although the boundary entropies computed using the QES formula receive a large contribution from the bulk entropy, they can secretly be thought of as being computed by an area in a

higher-dimensional bulk. The standard arguments for proving HEC inequalities obeyed by the RT formula can then be generalized, elucidating a connection between the bulk HEC and the boundary HEC.

Taking inspiration from the lessons learned from double holography, we then prove **Result 1** formally in Sec. 4.3 by employing the formalism of contraction maps pioneered by [42] (see also [8] for more details). After reviewing the formalism in Sec. 4.3.1, we prove **Result 1** in Sec. 4.3.2. The proof involves an elegant reinterpretation of the original contraction maps used to prove boundary HEC inequalities in the context of the RT formula. Despite the way we motivate it, we emphasize that **Result 1** is independent from double holography and applies more generally to any holographic field theory.

In Sec. 4.4, we then initiate a more controlled analysis of the implications of bulk entropy constraints on the boundary state. In particular, we consider the situation in which we constrain the bulk entropy to obey MMI for arbitrary subregions, which defines the MMI-constrained QES cone. Remarkably, we find that this suffices for the boundary state to also satisfy MMI for arbitrary subregions, thereby proving **Result 2**. In other words, the aforementioned 7-party inequality, (4.5), is not only a HEC inequality (as implied by **Result 1**), but moreover it is weak in the sense that it is implied already by the 3-party MMI inequality. Thus, we see that the MMI-constrained QES cone is at least as small as the cone obtained by imposing all instances of MMI on quantum entropies, i.e., the MMI cone.

Finally, we conclude in Sec. 4.5 with a summary of our results and a discussion of future directions of the research program we have initiated here.

4.2 Motivation: Double Holography

In this section we motivate our first result, namely, the fact that the bulk HEC implies the boundary HEC. To do so, we consider a doubly holographic setup where the d-dimensional gravity dual to a holographic quantum system itself has an effective holographic description in a $(d+1)$-dimensional theory of gravity. This setup has been exploited to study questions involving strong quantum effects[3] in d-dimensional quantum gravity, which can be comfortably analyzed by working with classical gravity on a $(d+1)$-dimensional bulk [115, 116, 110, 117, 118, 119, 120, 121, 122, 123, 124].

Consider a d-dimensional holographic boundary conformal field theory (BCFT). The bulk dual to such a theory was described in Refs. [125, 126] and involves a $(d+1)$-dimensional bulk with an end-of-the-world (EOW) brane anchored to the boundaries of the BCFT. In fact, such a holographic theory involves two layers of holography and thus admits essentially three descriptions, as seen in Fig. 4.1:

- **Description 1**: A UV complete, d-dimensional BCFT where the CFT couples to $(d-1)$-dimensional boundary defects which admit a holographic description.
- **Description 2**: An effective d-dimensional theory where the CFT now couples to a d-dimensional gravitating brane replacing the boundary defects of Description 1.
- **Description 3**: A $(d+1)$-dimensional theory of gravity with an EOW brane.

As described in [110, 114], the entropy of a subregion R of the BCFT in Description 1 can be related to quantities in the other descriptions using the QES formula (see Fig. 4.1),

[3] Most notably, in the derivation of the unitary Page curve of black hole evaporation.

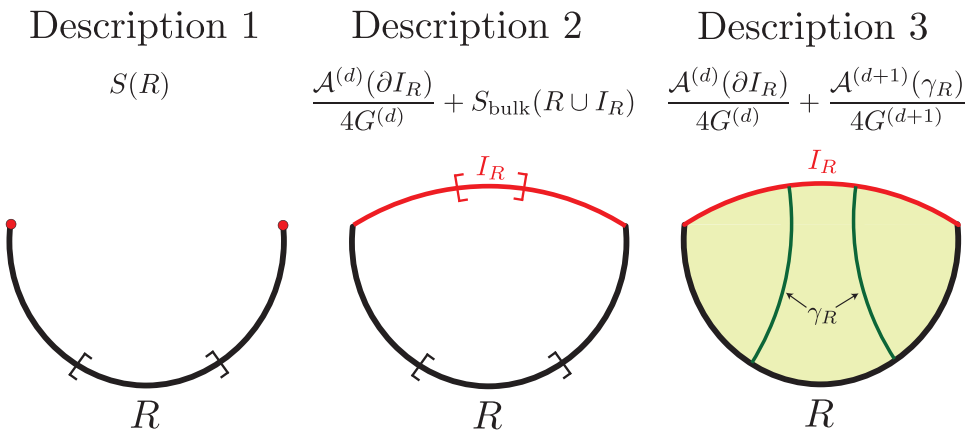


Figure 4.1: The three descriptions in the doubly holographic setup and their respective calculations for the entropy of subregion R. Description 1: A d-dimensional BCFT with a subregion R. Description 2: A d-dimensional CFT coupled to gravity on a d-dimensional brane (red); the entropy is computed using the QES formula including contributions from the island I_R. Description 3: A $(d+1)$-dimensional bulk with an EOW brane; the entropy is computed using the RT surface γ_R (green) which is anchored to subregion R and its island I_R. Note that we have assumed the RT surface to be connected in Description 3, corresponding to R having a nontrivial island in Description 2.

$$S_{\text{QES}}(R) = \min\left\{\operatorname*{ext}_{I_R} \frac{\mathcal{A}^{(d)}(\partial I_R)}{4G^{(d)}} + S_{\text{bulk}}(R \cup I_R)\right\} \tag{4.6a}$$

$$\approx \min\left\{\operatorname*{ext}_{\gamma_R} \frac{\mathcal{A}^{(d)}(\partial I_R)}{4G^{(d)}} + \frac{\mathcal{A}^{(d+1)}(\gamma_R)}{4G^{(d+1)}}\right\}. \tag{4.6b}$$

Here, the first line is a computation in Description 2 using the QES formula and I_R can be thought of as an island of R, living on the d-dimensional brane. Equivalently, the entropy of R could be computed using the quantum maximin procedure which gives the same answer as (4.6a) [113]. The second line is a computation in Description 3 using the simpler RT formula (i.e. neglecting any quantum corrections), where the bulk entropy term $S_{\text{bulk}}(R \cup I_R)$ can be calculated using the $(d+1)$-dimensional area of a surface γ_R homologous to $R \cup I_R$.

By construction, it is clear that all entropies computed in Description 2, including those for regions on the brane, obey all HEC inequalities. This is because this "bulk entropy" can be effectively computed in Description 3 by the RT formula, which satisfies the HEC as defined in [42]. Hence we say that the "bulk HEC" is obeyed. Now, we would like to consider entropies in Description 1, i.e., those given by the full expression in (4.6a), and show that these also lie within the HEC. This should be understood as the "boundary HEC", which explicitly coexists with the bulk HEC in this setup through (4.6).

Since the general idea will go through for all holographic inequalities in a similar fashion, we focus on the simple example of proving strong subadditivity (SSA) for subregions in Description 1. Given subregions A, B and C in Description 1, we would like to show that

$$S(\mathrm{AB}) + S(\mathrm{BC}) \geq S(\mathrm{B}) + S(\mathrm{ABC}). \tag{4.7}$$

A useful guide to the following calculation is provided by Fig. 4.2. We can compute the left-hand side using (4.6) to obtain

$$S(\mathrm{AB}) + S(\mathrm{BC}) = \frac{\mathcal{A}^{(d)}(\partial I_{AB})}{4G^{(d)}} + S_{\mathrm{bulk}}(AB \cup I_{AB}) + \frac{\mathcal{A}^{(d)}(\partial I_{BC})}{4G^{(d)}} + S_{\mathrm{bulk}}(BC \cup I_{BC}) \tag{4.8}$$

$$= \frac{\mathcal{A}^{(d)}(\partial I_{AB})}{4G^{(d)}} + \frac{\mathcal{A}^{(d+1)}(\gamma_{AB})}{4G^{(d+1)}} + \frac{\mathcal{A}^{(d)}(\partial I_{BC})}{4G^{(d)}} + \frac{\mathcal{A}^{(d+1)}(\gamma_{BC})}{4G^{(d+1)}}, \tag{4.9}$$

where γ_{AB} and γ_{BC} are respectively RT surfaces anchored to $AB \cup I_{AB}$ and $BC \cup I_{BC}$. These RT surfaces divide up the $(d+1)$-dimensional bulk and d-dimensional brane into four classes of regions. Each region can be labelled by a bitstring which encodes inclusion

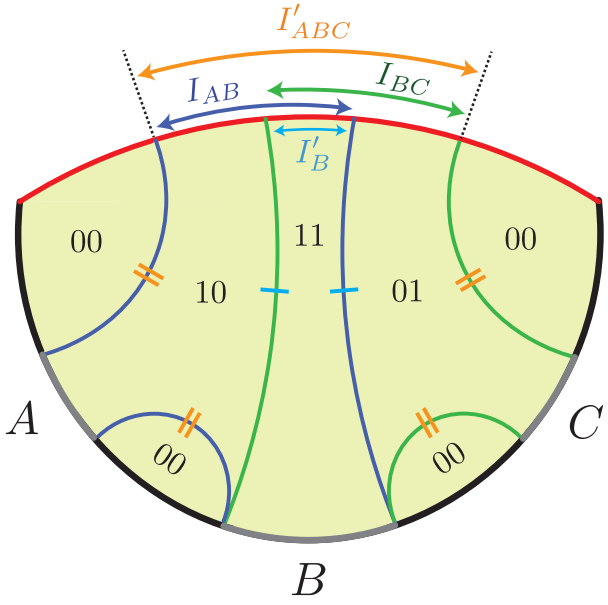


Figure 4.2: The entropies of subsystems AB and BC are computed by RT surfaces γ_{AB} (blue) and γ_{BC} (green), together with brane contributions coming from the boundaries of their associated islands on the brane, I_{AB} and I_{BC}, respectively. The homology regions generating these RT surfaces correspond to unions of bulk regions, namely: γ_{AB} is obtained from the boundary of 10 and 11, while γ_{BC} comes from 11 and 01. Similarly, the islands whose boundaries contribute to the relevant entropies can be formed from analogous unions of brane regions which carry the same labels as their adjacent bulk regions. All these bulk and brane regions can be rearranged into homology regions and islands for the subsystems B and ABC. In particular, 11 alone gives a surface γ'_B (single line) homologous to B with island I'_B, and the union of 10, 11 and 01 gives a surface γ'_{ABC} (double line) homologous to ABC with island I'_{ABC}. Thus the same bulk and brane labels allow one to keep track of homology regions and islands for left-hand and right-hand side subsystems in (4.7).

/ exclusion with respect to the homology regions and islands of AB and BC in the form of binary variables [42]. For instance, the region labelled by 10 in Fig. 4.2 receives this bitstring label because it is included in the homology region of AB (bounded by blue lines) but excluded from that of BC (bounded by green lines).[4] The labelling of $(d+1)$-dimensional bulk regions in Description 3 is associated to a corresponding labelling of d-dimensional brane regions in Description 2 – in Fig. 4.2, brane regions carry the same labels as their adjacent bulk regions.

One can now reproduce the cut-and-paste procedure prescribed in [42], now for both RT and island contributions, in terms of these labelled bulk and brane regions – see Fig. 4.2 for more details. In particular, the RT surfaces and island boundaries can be rearranged into a different set of surfaces, γ'_B and γ'_{ABC}, and island boundaries, $\partial I'_B$ and $\partial I'_{ABC}$, now homologically associated to subsystems B and ABC, respectively. By the minimization in the QES formula, we are guaranteed that these will obey

$$S(\mathrm{AB}) + S(\mathrm{BC}) = \frac{\mathcal{A}^{(d)}(\partial I'_B)}{4G^{(d)}} + \frac{\mathcal{A}^{(d+1)}(\gamma'_B)}{4G^{(d+1)}} + \frac{\mathcal{A}^{(d)}(\partial I'_{ABC})}{4G^{(d)}} + \frac{\mathcal{A}^{(d+1)}(\gamma'_{ABC})}{4G^{(d+1)}} \tag{4.10}$$

$$\geq S(\mathrm{B}) + S(\mathrm{ABC}), \tag{4.11}$$

thereby proving SSA.

Thus, by virtue of double holography, we see that both d-dimensional and $(d+1)$-dimensional contributions can be handled in a coordinated way using standard cut-and-paste arguments. Respectively, these contributions can be seen as being associated to Descriptions 2 and 3. From the perspective of Description 3, the above argument can be understood as a simple generalization of the arguments in [42] with the inclusion of island boundary terms of the form $\frac{\mathcal{A}^{(d)}(\partial I)}{4G^{(d)}}$ in the RT formula. Alternatively, from the

[4] This notation will be described in more detail in Sec. 4.3.1, as it is crucial for a general discussion of the arguments leading to the holographic entropy inequalities.

perspective of Description 2, it can be understood as a generalization where the islands associated to the QES surfaces include bulk entropy contributions of a holographic type, behaving as $(d+1)$-dimensional contributions growing off of the islands. Since Description 3 will no longer be available when abstracting away from double holography, the latter perspective will be the more useful one in what follows. Indeed, the basic strategy of handling bulk entropy contributions in terms of bitstring labels induced by QES surfaces will still be a powerful one and serve as the cornerstone for proving **Result 1** in Sec. 4.3.

Although we considered a time-independent situation, where the cut-and-paste technique of [42] can be directly applied, SSA and MMI can be proved in time-dependent settings. For these inequalities, one can use the maximin techniques of Refs. [95, 113], but now with the inclusion of extra boundary terms.[5]

To summarize, we have exemplified how the usual constructs for holographic entropy inequalities can be applied to both bulk and boundary entropies in a coordinated way in the setting of double holography. Within this framework, it becomes clear that the bulk HEC holds (by double holography) and so does the boundary HEC (because island contributions do not spoil the usual cut-and-paste arguments).

In fact, it is plausible that requiring the bulk HEC to hold for arbitrary subregions (including those on the brane), could imply the existence of a higher-dimensional geometric description where the entropies are computed using the RT formula.[6] This in turn would imply the boundary HEC by using the corresponding doubly holographic picture. Nevertheless, the construction of such a holographic dual given the boundary entropies would be a nontrivial task, if possible at all. Thus, in Sec. 4.3, we bypass the usage of double holography as an intermediate tool. Instead, we elucidate a more direct

[5]Note that there are inequalities that are not provable using the techniques of maximin [96].
[6]In certain cases, uniqueness of such a geometry would be guaranteed by the results of Refs. [86, 103].

mechanism by which the bulk HEC implies the boundary HEC in full generality.

4.3 Holographic Entropy Inequalities

In Sec. 4.3.1, we review the proof technique of contraction maps utilized in [42] to prove holographic entropy inequalities. In particular, we describe this method without resorting to graph theory, in a geometric language that we believe should be more transparent for the AdS/CFT community. In Sec. 4.3.2, we prove **Result 1**, i.e., the fact that the bulk HEC implies the boundary HEC, using a novel application of contraction maps.

4.3.1 Proof-by-Contraction

The *proof-by-contraction* method, originally developed by [42], is a combinatorial method for proving holographic entropy inequalities when using the RT formula. This technique has by now been extended in various directions (see [8] for an upgraded version and [11] for a generalization), but always formalized in the language of graph theory. Since graph models of holographic entanglement[7] have not played any role in our discussion so far, we will refrain from introducing them at this point. Instead, we now proceed to present the proof-by-contraction technique in a general geometric setting, which we believe will provide an explanation of its inner workings that feels more natural to holographers.

This proof method is a generalization of a strategy first employed in [127] for a holographic proof of SSA, and later applied in [40] to the proof of MMI. In these first appearances, the setting was in fact geometric, so it is natural to try to reformulate the proof-by-contraction method in geometric terms. Hence, it will be useful to reproduce

[7] See [42] for more details on these.

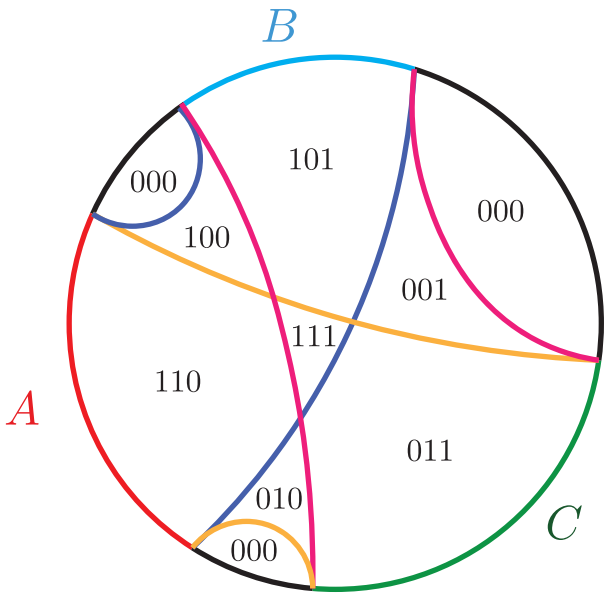


Figure 4.3: Example of a 3-party holographic configuration illustrating the proof-by–contraction method for MMI. The relevant parties are A, B and C, with boundary regions colored in red, light blue and green, respectively – their complement or purifier is shown in black. For the three terms on the left-hand side of the MMI inequality, AB, AC and BC, the RT surfaces are represented by the interior curves in blue, yellow and pink, respectively. The resulting homology regions partition the bulk geometry into subregions labelled by 2^3 different bitstrings in $\{0,1\}^3$. These bitstrings are specified by inclusion/exclusion as explained in the main text.

these early arguments by carrying along the proof of MMI as an example.

Without loss of generality, our MMI example can be described on the geometric configuration shown in Fig. 4.3. We emphasize though that the proof-by-contraction method does not assume anything about the underlying geometry, other than that it be a Riemannian manifold (e.g. arising from a time-symmetric slice of a Lorentzian manifold where the RT formula applies) with standard or asymptotic codimension-1 boundaries.[8]

[8]Geometries may have arbitrarily complicated topology, be non-smooth, involve multiple connected components, have any number of standard manifold or asymptotic boundaries, or even any kind of asymptotics, not necessarily AdS. Similarly, boundary regions may consist of multiple connected components as well, be adjoining or not, be compact or not, or cover entire boundaries. In a holographic context, it may be important to restrict to the boundary being a convex surface [128, 129].

A general candidate inequality will be canonically written as

$$\sum_{l=1}^{L} \alpha_l S_{I_l} \geq \sum_{r=1}^{R} \beta_r S_{J_r}, \tag{4.12}$$

such that the coefficients α_l and β_r are all positive, and L and R are the number of entropy terms appearing on each side. For an inequality written as in (4.12) to hold, it must be the case that its left-hand side (LHS) is guaranteed to be no smaller than its right-hand side (RHS) in any imaginable configuration. The basic idea of the proof-by-contraction is to devise an organizing principle that allows one to compare LHS terms to RHS ones in full generality. This turns out to be amenable to a combinatorial formulation which can be set up step-wise as follows:

1. Collect the RT surfaces and their corresponding homology regions for all L subsystems appearing on the LHS of the given inequality.

2. Partition the bulk into the 2^L subregions allowed by inclusion/exclusion inside each of the L homology regions (some subregions may be empty).

3. Use the inclusion/exclusion subregions to reconstruct the LHS homology regions, and utilize them suitably to construct candidate homology regions for RHS subsystems – the latter will generically not be associated to minimal RT surfaces.

4. Express areas of surfaces bounding homology regions in terms of the subregions contributing to them. For the LHS, these will yield back entropy areas, whereas for the RHS this will yield some possibly nonminimal areas.

5. Devise a diagnostic to compare the LHS entropy areas to the resulting RHS areas associated to the so-constructed homology regions.

We now describe each of these steps in detail, referring to the MMI example when helpful.

Step 1 requires us to compute the RT surfaces associated to all subsystem entropies on the LHS of the inequality. For instance, for MMI these are the interior blue, yellow and pink curves in Fig. 4.3. If γ_I is the RT surface of the region R_I, let σ_I be its corresponding homology region satisfying $\partial\sigma_I = \gamma_I \cup R_I$. We will take these homology regions σ_I to be open sets. Each LHS subsystem I_l has an associated σ_{I_l} subregion.

For step 2, we will use these σ_{I_l} to discretize the geometry as follows. Any bulk point $p \in M$ can be classified by whether it is inside or outside of σ_I for every LHS subsystem I, a binary choice that has to be made L times, one for each such term. Length-L bitstrings $x \in \{0,1\}^L$ can thus be used to encode all regions in M where points can be located in terms of LHS homology regions – see Fig. 4.3. In particular, these subsets obtained by inclusion/exclusion in homology regions can be written as

$$\sigma(x) = \bigcap_{l=1}^{L} \sigma_{I_l}^{x_l}, \qquad \sigma_I^b = \begin{cases} \sigma_I & \text{if} \quad b = 1, \\ \sigma_I^{\complement} & \text{if} \quad b = 0. \end{cases} \tag{4.13}$$

These $\sigma(x)$ sets partition M into 2^L disjoint subregions, some of which may be empty or consist of multiple connected components as shown in Fig. 4.3.

Step 3 employs the subregions in (4.13) to build back the LHS homology regions. Namely, since by construction $\sigma(x)$ is a nonempty subset of σ_{I_l} if and only if $x_l = 1$,

$$\sigma_{I_l} = \bigcup_{x:x_l=1} \sigma(x). \tag{4.14}$$

For example, the homology region for AB in Fig. 4.3 consists of subregions labelled by 110, 100, 111 and 101, all with $x_A = 1$, as required.

Crucially, these $\sigma(x)$ subregions can also be utilized to construct homology regions for the RHS subsystems. For a region $\sigma_{J_r} \subseteq M$ made up from $\sigma(x)$ subregions to be a valid homology region for a boundary region R_{J_r}, there are certain $\sigma(x)$ which one is obliged to include/exclude. In particular, it must be the case that $\sigma(x) \subseteq \sigma_{J_r}$ whenever $\sigma(x)$ is adjacent to R_{J_r}, and $\sigma(x) \not\subseteq \sigma_{J_r}$ whenever $\sigma(x)$ is only adjacent to other boundary regions $R_{J_r'} \neq R_{J_r}$. Once these constraints are satisfied, one is free to include any other $\sigma(x)$ in σ_{J_r}.

To make these conditions more explicit, it will be convenient to think of general pure states on a set of parties $[n+1]$ where $n+1$ labels the complement to all n parties.[9] The desired constraint for valid homology can then be phrased in terms of so-called *occurrence bitstring*. Such a bitstring $x^{(i)}$ for a party $i \in [n+1]$ is defined bit-wise by a binary Boolean function that determines the set of LHS subsystems I_l in which party i occurs, i.e.,

$$x_l^{(i)} = \delta(i \in I_l). \tag{4.15}$$

In words, $x^{(i)}$ is the label specifying the bulk region $\sigma(x^{(i)})$ adjacent to the boundary region labelled by i. In the MMI example, we get $x^A = 110$ (because A occurs in AB and AC, but not in BC), and we see that the bulk region $\sigma(x^A)$ this bitstring labels is indeed the one adjacent to boundary region A in Fig. 4.3. For each $i \in [n+1]$, these are the relevant bulk subregions involved in the homology constraint discussed above. Hence, this condition becomes the requirement that for occurrence bitstrings $\sigma(x^{(i)}) \subseteq \sigma_{J_r}$ if and only if $i \in J_r$ – for any other bitstring one is free to choose whether to include $\sigma(x)$ as part of σ_{J_r} or not. These conditions and choices can all be captured by introducing a general map $f : \{0,1\}^L \to \{0,1\}^R$ specifying RHS homology regions via (cf. (4.14) for

[9] Recall the notation $[n+1] \equiv \{1, 2, \ldots, n+1\}$, and the standard identification $1 \leftrightarrow A$, $2 \leftrightarrow B$, $\ldots$, where for the purifier we reserve $n+1 \leftrightarrow O$.

the LHS)

$$\sigma_{J_r} = \bigcup_{x:f(x)_r=1} \sigma(x), \tag{4.16}$$

and subject to a constraint on all occurrence bitstrings for $i \in [n+1]$ of the form

$$f(x^{(i)})_r = \delta(i \in J_r), \tag{4.17}$$

guaranteeing that (4.16) is a valid homology region for J_r. The freedom in including other $\sigma(x)$ subsets in σ_{J_r} corresponds to varying the map f while preserving (4.17). As noted previously, the RHS homology region σ_{J_r} resulting from a choice of map f will generically not be bounded by a minimal surface for J_r in the RT sense. For instance, applying (4.16) to the RHS subsystem $J_4 = ABC$ in the MMI inequality, the constraint (4.17) would instruct us to include $\sigma(x^{(i)}) \subseteq \sigma_{ABC}$ for all $i \in [3]$, but not for $i = 4$ (since O does not appear in ABC). In Fig. 4.3, this means σ_{ABC} must include regions labelled by 110, 101 and 011, and exclude the ones labelled by 000, which we see makes perfect sense as minimal requirements for any ABC homology region built out of $\sigma(x)$ subregions. In addition, $f(x)_4$ would be free to take any value for any non-occurrence bitstring, e.g. $f(010)_4 = 1$ and $f(111)_4 = 0$ would respectively correspond to having $\sigma(010) \subseteq \sigma_{ABC}$ and $\sigma(111) \not\subseteq \sigma_{ABC}$.

Having understood how bulk subregions $\sigma(x)$ can be used to construct homology regions for both LHS and RHS subsystems, we can now proceed to step 4. The crucial observation is that this partitioning of M into $\sigma(x)$ subregions also induces a discretization of the RT surfaces which will allow for a cut-and-paste argument to compare areas of LHS and RHS surfaces. The subsets in (4.13) can be used to chop RT surfaces into

pieces by adjacency using the following object,

$$\gamma(x,x') = \overline{\sigma(x)} \bigcap_{\text{codim-1}} \overline{\sigma(x')}, \tag{4.18}$$

where overlines denote closures and we have defined an intersection operator which yields the empty set if and only if the intersection of the two sets is of codimension higher than 1. In other words, $\gamma(x,x')$ will be empty unless the subsets $\sigma(x)$ and $\sigma(x')$ are adjacent to each other along a codimension-1 surface, so $\gamma(x,x')$ will always be a subregion of an RT surface lying between bulk regions labelled by x and x'.[10] Furthermore, we can identify which RT surface $\gamma(x,x')$ is part of, namely: $\gamma(x,x')$ is a nontrivial portion of γ_{I_l} if and only if x and x' differ in the l^{th} bit x_l (i.e., they lie on opposite sides of γ_{I_l}) and $\gamma(x,x') \neq \varnothing$ (i.e., $\sigma(x)$ and $\sigma(x')$ are adjacent).

The power of this is that we can now reconstruct the RT surface of any LHS subsystem I_l by piecing together appropriate portions $\gamma(x,x')$ by iterating over bulk subregions through

$$\gamma_{I_l} = \bigcup_{x,x':x_l\neq x'_l} \gamma(x,x'). \tag{4.19}$$

Furthermore, since no two $\gamma(x,x')$ have a codimension-1 intersection, (4.19) can also be used to compute entropies via the RT formula in (4.1),

$$S_{\text{RT}}(I_l) = \frac{1}{4G} \sum_{x,x'} |x_l - x'_l| \mathcal{A}\left[\gamma(x,x')\right], \tag{4.20}$$

where $|x_l - x'_l|$ is a convenient indicator function implementing the condition that $x_l \neq$

[10] At first sight, it could naively seem that two bitstrings x and x' will be labelling adjacent regions if and only if they differ by a single bit. This is in fact not true: bitstrings x and x' differing in multiple bits may be labelling adjacent regions if different subsystems share connected components of RT surfaces. In such cases, crossing the relevant RT surface would be associated to flipping more than one bit.

x_l' algebraically. Notice that all we needed to obtain (4.20) was a specification of the homology region of I_l in terms of bulk subregions $\sigma(x)$. We have the same ingredients for the RHS subsystems from (4.16), except the homology regions we can construct for RHS subsystems this way are not guaranteed to be bounded by surfaces of minimal area. This means that the homology regions we build for the RHS will yield

$$S_{\text{RT}}(J_r) \leq \frac{1}{4G} \sum_{x,x'} |f(x)_r - f(x)'_r| \mathcal{A}\left[\gamma(x,x')\right]. \tag{4.21}$$

We are now in a position to complete the argument with step 5 by comparing LHS and RHS terms in a candidate holographic inequality as written in (4.12). It will be convenient to introduce the notion of a *weighted Hamming distance* d_w which, given a weight vector $w \in \mathbb{R}^m$, for any pair of bitstrings $y, y' \in \{0,1\}^m$, is defined as

$$d_w(y,y') = \sum_{k=1}^{m} w_k |y_k - y'_k|. \tag{4.22}$$

With this notation in hand, and using (4.20), the LHS of (4.12) translates into

$$\sum_{l=1}^{L} \alpha_l S_{\text{RT}}(I_l) = \frac{1}{4G} \sum_{x,x'} d_\alpha(x,x') \mathcal{A}\left[\gamma(x,x')\right], \tag{4.23}$$

whereas using (4.21), the RHS of (4.12) obeys

$$\frac{1}{4G} \sum_{x,x'} d_\beta(f(x), f(x')) \mathcal{A}\left[\gamma(x,x')\right] \geq \sum_{r=1}^{R} \beta_r S_{\text{RT}}(J_r). \tag{4.24}$$

A successful comparison between (4.23) and (4.24) that proves (4.12) would correspond

to establishing that

$$\sum_{x,x'} d_\alpha(x,x')\mathcal{A}\left[\gamma(x,x')\right] \geq \sum_{x,x'} d_\beta(f(x),f(x'))\mathcal{A}\left[\gamma(x,x')\right]. \tag{4.25}$$

An obvious sufficient condition[11] for this to hold is that f be a *contraction map* for the distance functions d_α and d_β, i.e., that for every $x, x' \in \{0,1\}^L$,

$$d_\alpha(x,x') \geq d_\beta(f(x),f(x')). \tag{4.26}$$

This whole discussion gives rise to the desired proof-by-contraction method, which can finally be stated as follows:

Theorem 4.1 (Proof-by-contraction)**.** *An inequality of the form of* (4.12) *holds for the RT formula if there exists a contraction map* $f : \{0,1\}^L \to \{0,1\}^R$ *for the weighted Hamming distances* d_α *and* d_β *obeying the occurrence bitstring conditions in* (4.17) *for all* $i \in [n+1]$.

We exemplify this proof method in Tab. 4.1 by exhibiting a contraction map for the MMI inequality given in (4.2). One can visualize the cut-and-paste strategy that this contraction map encodes as follows (see Fig. 4.4). Every bitstring in the domain $\{0,1\}^3$ of Tab. 4.1 labels a distinct (possibly disconnected) region in Fig. 4.3, as specified by RT surfaces of LHS subsystems through (4.13). The contraction map f is then used to form homology regions for RHS subsystems as in (4.16). For instance, we see in Tab. 4.1 that for $J_1 = A$, $f(x)_1 = 1$ only for $x = 110$. This means $\sigma_A = \sigma(110)$, the minimal

[11] At face value this condition seems too strong to be necessary. Although this intuition turns out to be correct for (4.26) as written, the question becomes more subtle when β is made into an all-1 vector by simply expanding RHS terms in (4.12) into multiple copies of unit coefficient. Upon this innocuous-looking manipulation, it is believed that (4.26) becomes a necessary condition – see [8] for more details on this.

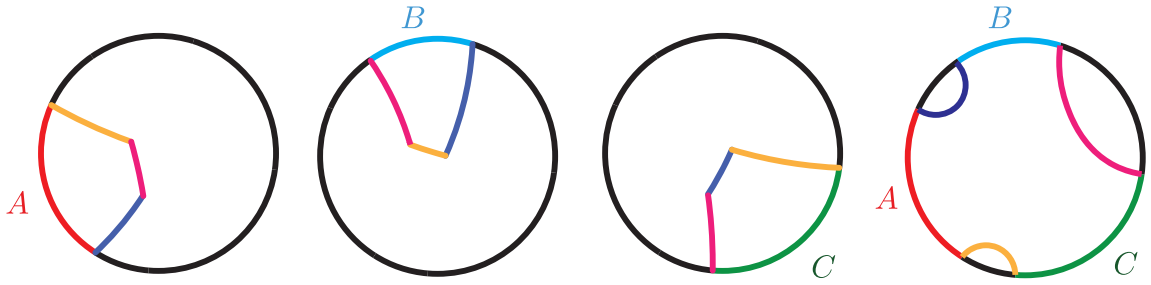


Figure 4.4: Homology regions specified by the unique contraction map for MMI in Tab. 4.1, exemplified by the configuration in Fig. 4.3. From left to right, subfigures show the RHS homology region and bounding surfaces that the cut-and-paste procedure yields for subsystems A, B, C and ABC.

homology region one can form for A as given by the occurrence bitstring, which results in the non-minimal surface shown on the left of Fig. 4.4. In contrast, for $J_4 = ABC$ the contraction map instructs us to include every single bulk region other than the one homologous to O, yielding now the minimal surface shown on the right of Fig. 4.4. One can easily check that no other choice of images for f would obey the contraction property.

	AB	AC	BC	A	B	C	ABC
O	0	0	0	0	0	0	0
	0	0	1	0	0	0	1
	0	1	0	0	0	0	1
C	0	1	1	0	0	1	1
	1	0	0	0	0	0	1
B	1	0	1	0	1	0	1
A	1	1	0	1	0	0	1
	1	1	1	0	0	0	1

Table 4.1: Tabular representation of the (unique) contraction map that proves validity of the MMI inequality (4.2) in holography. Occurrence bitstrings as defined in (4.15) and their images, fixed by (4.17), are indexed by their associated party label on the left-most column, including the one for the purifier O. Columns are indexed by the bitstring entry they label, with a vertical double-line separating domain from codomain of the contraction map. For the domain, each I_l labels entry x_l for $l \in [L]$ of $x \in \{0,1\}^L$ and, for the codomain, each J_r labels entry y_r for $r \in [R]$ of $y \in \{0,1\}^R$. Every row represents one entry of the map $f : x \mapsto y$ by listing input bits in x followed by output bits in y.

4.3.2 Bulk HEC implies Boundary HEC

The previous subsection has taught us how to prove a holographic entropy inequality when the entropies are given by the RT formula, (4.1). We would now like to investigate the case in which the entropies are instead given by the QES formula in (4.4). Since the generalized entropy, (4.3), receives contributions from bulk entropies, we will have to understand how this affects the proof-by-contraction method. Besides using the QES rather than RT surfaces, the partitioning of the bulk into regions via inclusion/exclusion carries through identically.

The first departure we observe involves (4.20), where we now have to include bulk entropy contributions as well. Making use of (4.14), the QES entropy can be written in terms of bitstrings as

$$S_{\text{QES}}(I_l) = \frac{1}{4G}\sum_{x,x'} |x_l - x'_l| \mathcal{A}\left[\gamma(x,x')\right] + S_{\text{bulk}}\left(\bigcup_{x:x_l=1}\sigma(x)\right). \tag{4.27}$$

Let f now be a map specifying homology regions for RHS subsystems as originally introduced above (4.16). Then, similar to (4.21), the minimality condition in the definition of the QES entropy guarantees that

$$S_{\text{QES}}(J_r) \leq \frac{1}{4G}\sum_{x,x'} |f(x)_r - f(x)'_r| \mathcal{A}\left[\gamma(x,x')\right] + S_{\text{bulk}}\left(\bigcup_{x:f(x)_r=1}\sigma(x)\right). \tag{4.28}$$

Following steps analogous to those below (4.21), we find that the desired inequality in

the form of (4.12) that we wish to prove becomes (cf. (4.25))

$$\sum_{l=1}^{L} \alpha_l S_{\text{bulk}} \left(\bigcup_{x:x_l=1} \sigma(x) \right) - \sum_{r=1}^{R} \beta_r S_{\text{bulk}} \left(\bigcup_{x:f(x)_r=1} \sigma(x) \right) \geq - \sum_{x,x'} \left(d_\alpha(x,x') - d_\beta(f(x),f(x')) \right) \mathcal{A}\left[\gamma(x,x')\right]. \tag{4.29}$$

Now, suppose that the inequality in (4.12) is a valid HEC inequality for RT and thus, f can be chosen to be a contraction map consistent with Thm. 4.1. Upon choosing such an f, the right-hand side of (4.29) is guaranteed to be non-positive and the inequality can be collapsed down to a purely bulk entropy inequality of the form

$$\sum_{l=1}^{L} \alpha_l S_{\text{bulk}} \left(\bigcup_{x:x_l=1} \sigma(x) \right) \geq \sum_{r=1}^{R} \beta_r S_{\text{bulk}} \left(\bigcup_{x:f(x)_r=1} \sigma(x) \right). \tag{4.30}$$

If this bulk inequality holds, then we would be guaranteed that its boundary counterpart, (4.12) with entropies computed by the QES, would hold as well. But (4.30) looks like a very complicated entropy inequality: since the unions run over $x \in \{0,1\}^L$, the number of distinct regions $\sigma(x)$ now playing the role of parties will generally be as large as $2^L - 1$[12]. For instance, when (4.12) is the MMI inequality in (4.2), the bulk inequality in (4.30) that the contraction map in Tab. 4.1 yields is precisely the 7-party inequality in (4.5). Hence the relevant question is whether there is any natural condition on the bulk entropy from which inequalities of the form of (4.30) would follow. The answer is yes, and the condition is that S_{bulk} itself obey HEC inequalities as well!

To show this, we will once again employ the proof-by-contraction method. In particular, notice that if S_{bulk} obeys HEC inequalities, then it must obey any inequality which

[12]It is not 2^L because the purifier O always has an all-zero occurrence vector on both sides, so its associated $\sigma(x)$ subregion does not appear anywhere.

can be proved this way. Hence, what we will show is that (4.30) itself admits a contraction map. Indeed, the contraction map that proves (4.30) is the very same f that defines it and which recall was assumed to prove the original boundary inequality (4.12) for the RT formula. By showing that f is a contraction map for (4.30), we will have shown that the boundary entropy S_{QES} obeys HEC inequalities if the bulk entropy S_{bulk} does so too.

The proof is straightforward and just requires thinking about the occurrence bitstrings of (4.30). The bulk regions $\sigma(x)$ for every $x \in \{0,1\}^L$ that is not the all-zero bitstring are the $2^L - 1$ parties that make up this inequality. Now, the occurrence bitstring of a given party $\sigma(x)$ is nothing but x itself – this is simply because $\sigma(x)$ shows up for every l for which $x_l = 1$, and does not show up otherwise. So every single domain bitstring $x \in \{0,1\}^L$ is itself an occurrence bitstring, including the all-zero bitstring for the purifying region. Suppose we now want to build homology regions for the RHS subsystems via some map $\tilde{f}$. The homology constraint prescribed by (4.17) fixes the image of all occurrence bitstrings to be precisely $\tilde{f}(x) = f(x)$, as follows from (4.30). However, this completely fixes the LHS-to-RHS map to be $\tilde{f} = f$, leaving no residual freedom. Since f was born as a contraction map for (4.12), f itself provides the proof-by-contraction that proves (4.30).

In practice, looking at MMI as an example, all we really had to do was to take Tab. 4.1 and assign a label to every unlabelled row. Understanding these as parties, and the LHS and RHS bitstrings as their occurrence bitstrings and respective images then tells us which parties to append to each column label. If we were to label rows top-to-bottom by *O-F-E-C-D-B-A-G*, the resulting column labels are precisely the ones which would correspond to (4.5). With this relabelling, Tab. 4.1 itself proves (4.5) to be a valid HEC inequality.

Altogether, we have proven the desired result:

Theorem 4.2. *If S_{bulk} obeys all HEC inequalities, then S_{QES} obeys all HEC inequalities too.*

This is a very general result which extends the relevance of the HEC from RT to the QES prescription.[13] However, it should be clear from our discussion that Thm. 4.2 gives a sufficient but highly unnecessary condition on S_{bulk} for S_{QES} to obey any particular HEC inequality. In a more controlled analysis, we could ask: what natural condition on S_{bulk} guarantees that S_{QES} obey a certain HEC inequality? This is precisely what we turn to next for the case of MMI.

4.4 Bulk MMI implies Boundary MMI

The results in this section are simple enough to state and prove, and thus we simply present them without elaborating on the techniques used to obtain them.

For subsets $\varnothing \neq I, J, K \subseteq [n+1]$, the following bits of notation will be useful:

$$I_2(I:J) \equiv S(I) + S(J) - S(I \cup J), \tag{4.31a}$$

$$Q(I;J:K) \equiv S(I \cup J) + S(I \cup K) - S(I) - S(I \cup J \cup K), \tag{4.31b}$$

$$I_3(I:J:K) \equiv S(I \cup J) + S(I \cup K) + S(J \cup K) - S(I) - S(J) - S(K) - S(I \cup J \cup K). \tag{4.31c}$$

[13] A subset of this result, i.e. bulk HEC implies boundary MMI, has been explored in [130] using the formalism of bit threads.

These may be recognized as defining arbitrary instances of the following inequalities:[14]

$$\text{SA:} \qquad I_2(I:J) \geq 0, \tag{4.32a}$$

$$\text{SSA:} \qquad Q(I;J:K) \geq 0, \tag{4.32b}$$

$$\text{MMI:} \qquad -I_3(I:J:K) \geq 0. \tag{4.32c}$$

The first two, subadditivity (SA) and strong subadditivity (SSA), are universal quantum entropy inequalities (i.e., valid for all quantum states), whereas the third one, monogamy of mutual information (MMI), is a holographic one (i.e., valid only when the RT formula applies). In what follows it is implicitly assumed that S_{bulk} obeys SA and SSA.

The first result is a negative but expected one (see e.g. [113]): that S_{bulk} obey the universal inequalities of SA and SSA is not enough for S_{QES} to obey MMI. In other words,

Proposition 4.1. *Boundary MMI does not hold in general.*

Proof. Consider three sufficiently small and distant boundary subregions A, B and C such that their geometric contributions to S_{QES} for any subsystem I built out of them factorizes and takes the form $S_{\text{QES}}(I) = \sum_{i\in I} S_{\mathcal{A}}(i) + S_{\text{bulk}}(I)$. Evaluating (4.31c) on such a configuration, all $S_{\mathcal{A}}$ contributions cancel out and only S_{bulk} ones remain. Choosing the bulk state among these subregions and their complement to be of 4-party GHZ-type, one achieves $S_{\text{bulk}}(I) = S_0$ for all I subsystems. The upshot is a configuration where the bulk state clearly obeys SA and SSA, but for which S_{QES} yields $I_3(\text{A:B:C}) = 2S_0 \geq 0$, thus violating boundary MMI. □

We can also understand this argument in terms of a violation of the 7-party condition in (4.5). Namely, the configuration used in the proof above would correspond to only

[14] Restrictions on the choices of subsystems $\varnothing \neq I, J, K \subseteq [n+1]$ are needed to prevent these from trivializing.

having the occurrence bitstrings 000, 110, 101 and 011 labelling nonempty regions in Fig. 4.3. In terms of the 7-party inequality derived from (4.30) using Tab. 4.1, this results in (4.5) with D, E, F and G being empty regions. This reduces (4.5) to the form of MMI for S_{bulk}, which is clearly violated by a 4-party GHZ state.

In searching for sufficient conditions on S_{bulk} for S_{QES} to obey MMI, we are thus led to impose MMI in the bulk as well. A natural way to proceed is thus to combine all instances of SA, SSA and MMI in (4.32) for 7 parties and check if these imply (4.5) as a bulk inequality. This leads us to **Result 2**:

Proposition 4.2. *Bulk MMI implies boundary MMI*

Proof. The 7-party bulk inequality (4.5) which implies boundary MMI is [113]

$$S(\text{ABDG}) + S(\text{ACEG}) + S(\text{BCFG}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABCDEFG}). \tag{4.33}$$

Consider the special case that $D = E = F = \varnothing$. What remains is

$$S(\text{ABG}) + S(\text{ACG}) + S(\text{BCG}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABCG}), \tag{4.34}$$

which follows from combining the two inequalities

$$S(\text{AG}) + S(\text{BCG}) \geq S(\text{A}) + S(\text{BC}), \tag{4.35}$$

$$S(\text{ABG}) + S(\text{ACG}) + S(\text{BC}) \geq S(\text{AG}) + S(\text{B}) + S(\text{C}) + S(\text{ABCG}), \tag{4.36}$$

where the first is weak-monotonicity for the regions A, BC, G and the second is MMI for the regions AG, B, C. (Note that weak-monotonicity can be written as SSA by including the purifier. Let $DEFO$ purify $ABCG$. Then (4.35) is SSA for the regions $A, DEFO, G$.)

Now let D be arbitrary. What remains is

$$S(\text{ABDG}) + S(\text{ACG}) + S(\text{BCG}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABCDG}), \tag{4.37}$$

which follows from (4.34) combined with

$$S(\text{ABDG}) + S(\text{ABCG}) \geq S(\text{ABG}) + S(\text{ABCDG}). \tag{4.38}$$

This holds from SSA on the regions ABG, C, D. Similarly, now let E be arbitrary to obtain

$$S(\text{ABDG}) + S(\text{ACEG}) + S(\text{BCG}) \geq S(\text{A}) + S(\text{B}) + S(\text{C}) + S(\text{ABCDEG}), \tag{4.39}$$

which follows from (4.37) combined with

$$S(\text{ACEG}) + S(\text{ABCDG}) \geq S(\text{ACG}) + S(\text{ABCDEG}). \tag{4.40}$$

This follows from SSA on the regions ACG, BD, E. Adding back F works the same way. It returns us to the 7-party inequality (4.5), which follows from (4.39) combined with

$$S(\text{BCFG}) + S(\text{ABCDEG}) \geq S(\text{BCG}) + S(\text{ABCDEFG}). \tag{4.41}$$

This follows from SSA on the regions BCG, ADE, F. This completes the proof, which can be compactly summarized in the following expression:

$$-I_3(\text{AG:B:C}) + Q(\text{A;DEFO:G}) + Q(\text{ABG;C:D}) + Q(\text{ACG;BD:E}) + Q(\text{BCG;ADE:F}) \geq 0. \tag{4.42}$$

□

The finding of such an expression is nontrivial but can also be easily tackled as a linear programming problem. We emphasize that there is no obvious reason why Prop. 4.2 should hold: that the same contraction map proves both (4.2) and (4.5) (as in the proof of Thm. 4.2) is a priori logically unrelated to the fact that the former implies the latter. Indeed, this is particularly remarkable because only MMI, which effectively is just a 3-party inequality, suffices to prove validity of (4.5), which is a 7-party inequality.

4.5 Discussion

Summary: So far, the exploration of holographic entropy inequalities had focused on using the RT formula and its geometric nature. However, we have learned in the past few years that quantum effects incorporated using the QES formula can play a crucial role, e.g., in the presence of black holes. Thus, the goal of this work is to open up a research program of understanding the connection between bulk and boundary entropy inequalities.

To summarize, we have derived relations between constraints imposed on bulk entropies and the corresponding constraints on boundary entropies by relating them via the QES formula. The first result we showed in this direction is that the bulk HEC imposes the boundary HEC in a nontrivial way. A generic n-party boundary HEC inequality requires a bulk inequality that could contain an exponentially large number of parties. Nevertheless, the entire collection of bulk HEC inequalities guarantees the full set of HEC inequalities on the boundary. Secondly, we showed that imposing MMI on arbitrary subregions in the bulk leads to arbitrary boundary subregions satisfying MMI as well. This result is a first step in carrying out a more controlled study of the interplay

between specific bulk and boundary constraints.

Assumption: Note that, in stating both of our main results, we assumed the QES prescription is valid. This is typically expected to be true when it is applied in semi-classical bulk states, i.e., states of quantum fields on fixed, possibly curved backgrounds. However, it has recently been pointed out that there are certain semiclassical states for which the QES prescription gives the wrong entropy at leading order in $1/G$ [131]. One could worry that in such states, our results will not hold, for example that bulk MMI will no longer imply boundary MMI. We expect that this is actually not a problem. That is, even with the leading order corrections of [131], we expect that our results continue to be valid. The basic idea is that different parts of the bulk wavefunction will each satisfy the QES prescription, obeying our results, and the inequalities in the HEC are such that their mixture will therefore also obey our results. We leave a detailed analysis of this for future work.

Conjectures: Our results here can be interpreted as the two extremities of a general set of such connections between boundary and bulk entropies. We have analyzed the HEC-constrained QES cone, which we showed to be equal to the HEC, and also the MMI-constrained QES cone, which could be at most as large as the MMI cone. There are various ways in which one could imagine interpolating between our two results. We conjecture a list of possibilities here:

- If a given n-party inequality no stronger than HEC ones is satisfied in the bulk, then the same n-party inequality is satisfied on the boundary.

- If the complete set of n-party HEC inequalities is satisfied in the bulk, then every n-party HEC inequality is satisfied on the boundary.

- If the complete set of k-party HEC inequalities is satisfied in the bulk, then there is some n_k such that every n-party HEC inequality is satisfied on the boundary for all $k \leq n \leq n_k$.

In the first possibility above we needed to exclude inequalities stronger than HEC ones because those cannot possibly by satisfied by any constrained QES cones since, as we mentioned before, all such cones automatically contain the HEC due to the area term. On the other hand, notice that we did not restrict this possibility to just HEC inequalities – indeed, it is possible that, in full generality, any weaker bulk constraint implies its boundary counterpart. Such a situation may be hard to prove but easy to falsify.

When focusing on HEC inequalities, the three possibilities above can be easily seen to go from strong to weak, in the sense that each of them would imply the subsequent one. For instance, the second one implies the trivial case $n_k = k$ for the third one, which in principle could be weaker by having $n_k \geq k$. It would be interesting to explore these and other logical possibilities in future work.

Hypergraphs: Holographic states can be understood as states whose entanglement structure admits a suitably discretized representation in terms of graphs where the RT formula computes the entropy of a boundary subregion in terms of the minimum cut across the graph. A simple generalization of this class of states is to hypergraphs, where nodes can be connected using hyperedges instead of regular edges. A k-hyperedge is a connection that groups k nodes simultaneously, where $k = 2$ corresponds to usual edges. These states also satisfy a nontrivial entropy cone if one posits that the entropy of a boundary subregion is computed using a generalized RT formula, i.e., using the minimum cut in a hypergraph [11, 9]. Further, such hypergraph states can be explicitly constructed as stabilizer states using random tensor networks with k-party GHZ links

instead of Bell-pair-like bonds [69].

There is also a simple generalization of the QES formula to hypergraphs, taking a form analogous to (4.4). In the hypergraph, the homology region σ_R is described by the collection of nodes defining the minimum cut for a set of boundary nodes R. The area term in the generalized entropy is implemented by the total weight of the minimum cut, which is given by the sum of the weights of all hyperedges that connect nodes in σ_R to those in its complement. Finally, the bulk entropy contribution can also be realized in a random tensor network in the form of bulk dangling legs on every node, analogous to [91].

Proving entropy inequalities for the RT formula on hypergraph states involves a generalization of the contraction-map technique used in Sec. 4.3. Roughly, apart from requiring that the map f described in Sec. 4.3.1 obey a contraction property for pairs of bitstrings, the proof of inequalities on hypergraphs involves additional multi-bitstring contraction conditions (see [11] for more details). Intuitively, these extra conditions make it strictly harder for such an f to exist, thereby explaining why hypergraph entropies obey weaker inequalities than graph ones and thus attain richer entanglement structures. Nevertheless, we can again run a similar argument to that in Sec. 4.3.2 to show in an analogous fashion that the bulk inequalities required to prove any specific hypergraph cone boundary inequality follow from the same contraction map upon relabelling parties. Thus, one can again see that the bulk hypergraph cone implies the boundary hypergraph cone.

It would also be interesting to probe the relation between specific hypergraph inequalities in the bulk and boundary, in the spirit of Sec. 4.4. For instance, in [11] it was

shown that minimum cuts on hypergraphs obey the Ingleton inequality [132]

$$S(\mathrm{AB})+S(\mathrm{AC})+S(\mathrm{AD})+S(\mathrm{BC})+S(\mathrm{BD}) \geq S(\mathrm{A})+S(\mathrm{B})+S(\mathrm{CD})+S(\mathrm{ABC})+S(\mathrm{ABD}). \tag{4.43}$$

However, since this inequality involves $L = 5$ terms, its hypergraph contraction map f would lead to a $2^L - 1 = 31$-party bulk inequality, so it seems rather nontrivial to prove a result analogous to the one proved for MMI in Sec. 4.4.

Part II

Causality

Chapter 5

Bulk reconstruction of metrics with a compact space asymptotically

5.1 Introduction

A central question in holography [16, 17, 75, 18] is to understand how spacetime geometry emerges from the dual field theory. The standard discussions of entanglement wedge reconstruction do not address this since they depend on a choice of code subspace which represents small fluctuations about a given semiclassical bulk geometry (or perhaps a finite number of such geometries). The idea of geometry emerging from entanglement [39] has led to various attempts to determine the bulk metric from measures of entanglement [133, 134, 135, 136, 137, 138, 101, 139], in particular via the geometrization in the bulk of the von Neumann entropy of boundary regions [32, 36]. Recently, it has been shown that a bulk geometry (if it exists) is uniquely determined by second order variations of the area of two-dimensional extremal surfaces anchored to a certain family of regions on

the boundary [86].[1]

In [45, 46] a very different approach toward reconstructing the bulk geometry was presented. This involves special cross-sections of the conformal boundary of an asymptotically Anti-de Sitter (AdS) spacetime. These cross-sections are called light-cone cuts, and can be thought of as the intersection of the past (or future) light cones of bulk points with the boundary.[2] (A more precise definition will be given in the next section.) It was shown that knowledge of these light-cone cuts is sufficient to determine the conformal metric in the bulk, i.e. the metric up to an overall local rescaling, for most points causally related to the boundary.[3] It was further shown how to determine these light-cone cuts from singularities in certain time-ordered Lorentzian correlators in the dual field theory which originate precisely from bulk locality [147, 83].

The results in [45, 46], as well as those which employ entanglement entropy, apply to spacetimes which are asymptotically AdS. However, the most well studied models of holography require spacetimes to approach $\mathrm{AdS}_n \times \mathbb{S}^k$. The goal of this chapter is to extend the analysis of light-cone cuts to these more general spacetimes. (We will always assume $n > 2$, since for $n = 2$ the light-cone cut consists of isolated points and does not determine the conformal metric.) It is easy to see that a naive, straightforward attempt to apply light-cone cuts to spacetimes with compact extra dimensions will fail to determine the bulk geometry. However, we will show that there is a generalization of light-cone cuts that we call "extended cuts", that indeed determine the conformal metric

[1] These areas correspond to the entanglement entropy of boundary regions when the bulk is four-dimensional. In other dimensions, they are related to expectation values of Wilson loops in some cases, but their general holographic interpretation is not well understood [86, 140, 141].

[2] Analogous cross-sections of null infinity in asymptotically flat spacetimes were first introduced in [142, 143], and shown to encode the conformal metric of such spacetimes in [144]. However, holography or the presence of internal spaces played no role in these discussions.

[3] An explicit implementation of the light-cone cut approach to bulk reconstruction was explored in [145]. The light-cone cut formalism was also used in [146] to covariantize the notion of bulk depth and relate it to energy scales in the dual field theory.

of the full higher-dimensional spacetime. We will then propose a method for obtaining these extended cuts from the dual field theory.

The basic idea behind our extended cuts is the following. Every null geodesic which reaches the boundary of AdS approaches a fixed point on $\mathbb{S}^k$. This is simply because a geodesic with asymptotic motion on the sphere acts like a massive particle in AdS and will not reach the boundary. Thus for every point on our light-cone cut, we can associate a point on $\mathbb{S}^k$. Our extended cut is just the original light-cone cut $C(p)$ together with a map $C(p) \to \mathbb{S}^k$ specifying the asymptotic location on $\mathbb{S}^k$ of the null geodesics from p to $C(p)$. We show in Section 5.3 that this map is precisely the extra information that is needed to reconstruct the full bulk conformal metric of generic spacetimes (if the spacetime has symmetries that asymptotically act only on the internal space, this approach may fail). In Section 5.4, we propose a way to determine the extended cuts from the dual field theory, using correlation functions involving the operators dual to the Kaluza-Klein modes of the higher-dimensional bulk field. Our approach does not use any particular property of $\mathbb{S}^k$ and should work equally well for a general compact internal space.

To completely determine the bulk geometry, one also needs to know the conformal factor. This remains an open problem in general, however it is known how do to this for some special asymptotically AdS spacetimes [46].

5.2 Review of Light-Cone Cuts

In this section we review the construction in [45, 46] for obtaining the conformal metric from the dual field theory. Two metrics g and $\bar{g}$ are conformally related if there is a positive function f such that $\bar{g}_{\mu\nu} = f^2 g_{\mu\nu}$. Points in spacetimes with conformally

related metrics clearly have the same light cone, but one does not need to know the entire light cone (or even an open subset of it) to determine the conformal metric at those points; a sufficient number of null vectors will do. This can be seen as follows. In a D-dimensional spacetime, take D linearly independent null vectors ℓ_i at a point p. Since the ℓ_i all have zero norm, the conformal metric at p is fixed by their inner products. To determine them, take a new collection of null vectors, η_k, and expand them in terms of the null basis ℓ_i:

$$\eta_k = \sum_i M_{ki}\ell_i. \tag{5.1}$$

Using the fact that each η_k has zero norm, we obtain a set of algebraic equations for the inner products $\ell_i \cdot \ell_j$:

$$0 = \eta_k \cdot \eta_k = \sum_{i,j=1}^{D} M_{ki}M_{kj}(\ell_i \cdot \ell_j) \qquad \text{no sum on } k. \tag{5.2}$$

While it is not always true that such equations have a solution, we are guaranteed a solution here precisely because these equations describe a Lorentzian metric which by construction exists. By choosing at least $D(D-1)/2$ vectors η_k, the solution will be unique up to an overall constant rescaling of all inner products. This determines the conformal metric at p. Repeating this local construction at each point in a spacetime region U determines the conformal metric on U.

Our goal is to determine these null vectors at p from boundary data. Due to gravitational lensing, the light cone of a bulk point p can develop caustics. When this happens, some null geodesics reach points that are timelike related to p. Since we want boundary points that are null-related to p we proceed as follows.

Let (M, g) be an asymptotically locally AdS spacetime (without compact extra dimensions) with conformal boundary ∂M, and denote its conformal compactification by

$(\bar{M}, \bar{g})$. Recall that the causal past $J^-(p)$ of a point $p \in M$ is the set of points in M which can be reached by a past-directed causal curve starting at p. $J^+(p)$ is defined similarly with "future" replacing "past". A spacetime is said to be AdS-hyperbolic if there exist no closed causal curves and for any two points $p, q \in M$, the set $J^+(p) \cap J^-(q)$ is compact in the conformally compactified spacetime $\bar{M}$ [95]. We will assume our spacetime is C^2 differentiable, maximally extended, connected, and AdS-hyperbolic. The future/past light-cone cut $C^{\pm}(p)$ of a point $p \in M$ is defined as the intersection of the boundary of the causal future/past of p, $\partial J^{\pm}(p)$, with the conformal boundary ∂M, i.e.

$$C^{\pm}(p) \equiv \partial J^{\pm}(p) \cap \partial M. \tag{5.3}$$

This is illustrated in Fig. 5.1. We will use $C(p)$ to denote either the future or past cut of a bulk point p. Light-cone cuts are not differentiable everywhere since they can have cusps due to caustics. However, it can be shown that the cusps form a set of measure zero within the cut (cf. Proposition 5.1 in Section 5.3.2).

It was shown in [45] that light-cone cuts satisfy the following properties:

1. $C(p)$ is a complete spatial slice of the conformal boundary.

2. There is a one-to-one, onto map from past light-cone cuts to points in the future of the boundary, even inside black holes. (A similar statement holds for future cuts.)

3. Two distinct cuts cannot agree on an open set.

4. If $C(p)$ and $C(q)$ intersect at precisely one point, and both cuts are C^1 at this point, then p and q are null-separated.

Using these properties, it is easy to construct the bulk conformal metric given the light-cone cuts. Property 2 says that the set of past cuts represents all points to the

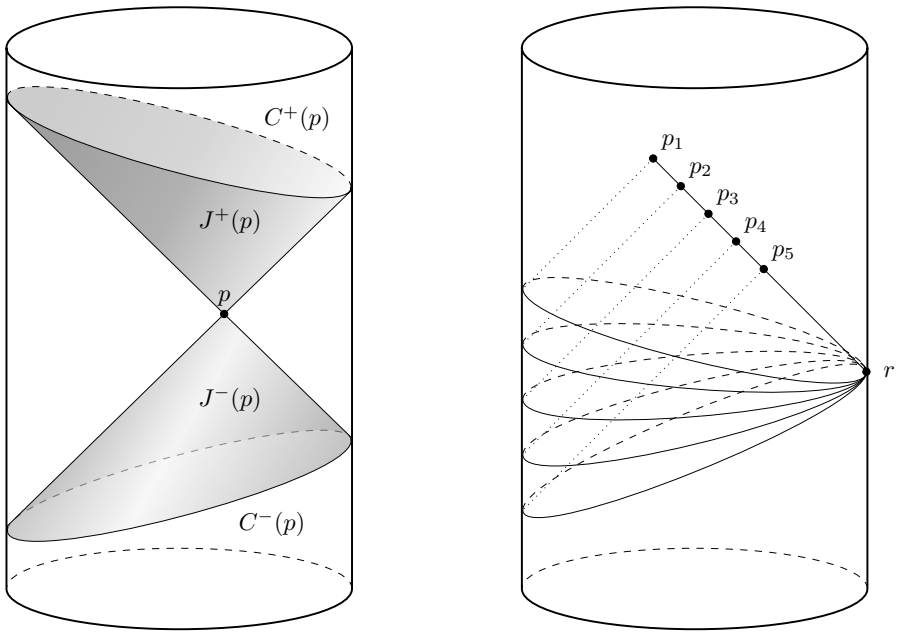


Figure 5.1: In the left figure, the shaded region illustrates the future and past light cones $\partial J^{\pm}(p)$ of a bulk point $p \in M$ in causal contact with the boundary in an asymptotically AdS spacetime M. Their intersections with the conformal boundary ∂M define the future and past light-cone cuts $C^{\pm}(p)$, which are complete spatial slices of ∂M (cf. properties 1 and 2). The right figure shows how a sequence of null-related bulk points $\{p_i \in M\}$ corresponds to a set of light-cone cuts which all intersect at a single point $r \in \partial M$ (cf. property 4).

future of the boundary. Property 4 says that given a light-cone cut $C(p)$, the set of cuts $C(q)$ which are tangent to $C(p)$ at a regular point $r \in C(p)$ represents a null curve passing through p, as illustrated in Fig. 5.1. Repeating this for $D(D+1)/2$ cut points r allows one to reconstruct the conformal metric at p.[4] It is clear that a basis of null vectors ℓ_i at p can be obtained this way, since the light-cone cut $C(p)$ enables one to reconstruct an open subset of the light cone at p.

The second half of the construction is a procedure for determining $C(p)$ from the dual field theory without using the bulk geometry. This is achieved using the notion of bulk-point singularities, first argued for in [147] and later studied in [83].[5] Given D boundary points in a D-dimensional spacetime, the only subset of M which can be null-related to all of them are individual points. It was shown in [147, 83] that a time-ordered Lorentzian $(D+1)$-point correlator on the boundary of AdS is singular when there exists a momentum-preserving scattering point in the bulk that is null-related to all of them (i.e. if one can draw a position-space Landau diagram with null lines in the bulk).[6] This is the case if, for example, one chooses two points in the past cut $x_1, x_2 \in C^-(p)$, and $D-1$ points in the future cut $x_i \in C^+(p)$ of a bulk point p, in a manner similar to Fig. 5.2. Then, physically, high energy quanta from x_1 and x_2 can scatter at p conserving energy-momentum and send high energy quanta to the remaining x_i in the future, which results in a singular correlator. In special cases, only derivatives of the correlator will diverge and the correlator itself may remain finite. However, for most operators, the

[4] One needs D points for the basis vectors ℓ_i, and $D(D-1)/2$ for the null vectors η_k used to determine the inner products.

[5] These are singularities which appear in correlators for holographic states of the dual field theory at large gauge group rank N and large 't Hooft coupling λ. At finite N or finite λ, these singularities are resolved and only expected to manifest themselves as resonances with a holographic origin [83, 148]. Large N and λ are assumed in what follows.

[6] Even though a single bulk point p can be fixed by the condition that it is null-related to D boundary points, one needs at least one extra point in the correlator to ensure that energy-momentum is conserved at p.

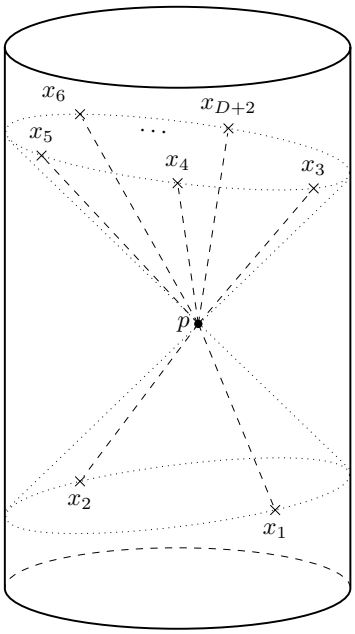


Figure 5.2: Position-space Landau diagram for a boundary correlator with a bulk–point singularity from $p \in M$ used to obtain light-cone cuts from the dual field theory. For a D-dimensional bulk, the D boundary points in the future already specify p as the unique bulk point null-related to all of them. The two points in the past can be rotated around maintaining momentum conservation at the interaction vertex (and hence the divergence in the correlator) to trace out the light-cone cut.

correlator itself will diverge and we will use such operators below.

To use this to find the light-cone cuts we need two modifications. First, we consider correlation functions in certain excited states, not the ground state, so the dual spacetime is only asymptotically locally AdS and not pure global AdS. Second, we consider $(D+2)$-point correlators, with two points x_1 and x_2 in the past and D points $x_3, \ldots, x_{D+2}$ in the future (see Fig. 5.2). In this case, if there is a bulk point p null-related to all the x_i to the future, it will remain fixed if we move the ones in the past. Starting with a configuration of points where the correlator diverges, we can thus move x_1 and x_2 in a coordinated manner keeping the correlator singular to trace out the past cut of p.[7]

[7]One could actually work with $D+1$ points and still move one vertex in a limited way to trace out part of the light-cone cut, but one has more freedom to trace out the entire cut by adding an additional

5.3 Extended Light-Cone Cuts

In this section we extend the discussion of light-cone cuts to spacetimes that have a compact space asymptotically such as $\text{AdS}_n \times \mathbb{S}^k$. The presence of this compact space implies that most of the null geodesics on the light cone of a bulk point p end up crossing other null geodesics and entering the interior of $J(p)$. Only a small subset of these null geodesics stay on the boundary of $J(p)$ and form the light-cone cut. To illustrate this, consider the three-dimensional flat spacetime $ds^2 = -dt^2 + dz^2 + d\chi^2$, with χ periodically identified. Starting at any point p, all null geodesics with $\dot{\chi} \neq 0$ will go around the $\mathbb{S}^1$ and reach points timelike related to p. The only ones that stay on $\partial J(p)$ are those with $\dot{\chi} = 0$. More generally, for spacetimes locally asymptotic to $\text{AdS}_n \times \mathbb{S}^k$, the light-cone cut has bulk codimension $k+2$ rather than 2. This means that one cannot recover an open subset of the light cone of a bulk point p. Fortunately, as reviewed above, one does not need an open subset of the light cone to recover the conformal metric at p. All one needs is a basis of null vectors and some additional null vectors. As we discuss below, this can be obtained in generic spacetimes from a simple generalization of the light-cone cut.

For asymptotically locally $\text{AdS}_n \times \mathbb{S}^k$ spacetimes, one way to understand the reduction in the size of the light-cone cut is by noting that the conformal boundary of $\text{AdS}_n \times \mathbb{S}^k$ is degenerate, in the sense that it is codimension $k+1$ rather than 1 [149]. Indeed, the $\mathbb{S}^k$ factor of the direct product shrinks to zero size and leaves a boundary manifold ∂M which is locally isometric to the conformal boundary of just the AdS_n part.

The presence of a degenerate boundary turns out to invalidate most results proven in [46, 45]. Fortunately, it is possible to recover them with appropriate generalizations of

point. One must also minimize the time difference between the points in the past and future to avoid caustics along the null geodesics from the bulk point to the boundary.

the framework. To motivate the solution, let us first understand the complications that arise when the boundary is degenerate. In particular, consider the following two results from [45] (cf. properties 3 and 4 reviewed in Section 5.2) and counterexamples to them already in the simple case of global $\mathrm{AdS}_n \times \mathbb{S}^k$:

- *$C(p) \cap C(q)$ contains a nonempty open set if and only if $p = q$*: For any two points p and q on $\mathrm{AdS}_n \times \mathbb{S}^k$ with the same global coordinates on the AdS part one has $C(p) = C(q)$, even if they have different coordinates on the sphere. More precisely, thinking of the compactification space $\mathbb{S}^k$ as a fiber of a trivial bundle $\pi : \mathrm{AdS}_n \times \mathbb{S}^k \to \mathrm{AdS}_n$, this means that $C(p) = C(q)$ for any $p, q \in \mathrm{AdS}_n \times \mathbb{S}^k$ with the same base space point $\pi(p) = \pi(q)$, implying that light-cone cuts do not distinguish points on the fibers.

- *If $C(p)$ and $C(q)$ intersect at precisely one point, and both cuts are C^1 at this point, then p and q are distinct and null-separated*: To falsify this claim, consider an arbitrary point $p \in \mathrm{AdS}_n \times \mathbb{S}^k$ and another null-separated point $q \in \partial J(p)$ such that the null-geodesic between p and q reaches ∂M at some point $r \in C(p) \cap C(q)$. It is easy to see that their light-cone cuts will indeed intersect precisely only at r, and that both cuts will be C^1 at this point (since the spacetime is pure AdS_n). Now take another point $\tilde{q}$ which is at the same AdS location as q, but at a different point on the sphere. Since the metric on the sphere is Euclidean, p and $\tilde{q}$ will be spacelike-separated. But from the counterexample to the previous claim, one still has $C(\tilde{q}) = C(q)$. Altogether, this shows that $C(p)$ and $C(\tilde{q})$ intersect at precisely one point, both cuts are C^1 at this point, but p and $\tilde{q}$ are spacelike-separated, thus contradicting the statement above.

As anticipated, the existence of these counterexamples can be traced back to the fact

that the light cone $\partial J(p)$ of a bulk point $p \in M$ degenerates asymptotically in essentially the same way the conformal boundary does. More precisely, suppose a boundary observer wanted to resolve the compact dimensions by introducing a regulated boundary ∂M_ϵ at a finite UV cutoff $0 < \epsilon \ll 1$, with $\lim_{\epsilon\to 0} \partial M_\epsilon = \partial M$. On ∂M_ϵ, the dimensions of $\mathbb{S}^k$ are restored and one has $\operatorname{codim} \partial M_\epsilon = 1$, the dimensionality only dropping by k in the strict limit $\epsilon \to 0$. Similarly, intersecting $\partial J(p)$ with the regulated boundary ∂M_ϵ, one sees that the corresponding regulated light-cone cut $C_\epsilon(p) = \partial J(p) \cap \partial M_\epsilon$ is now bulk-codimension 2, the dimensionality only decreasing by k in the strict limit $\epsilon \to 0$.

Crucially, under the pertinent assumptions, all results proven in [46, 45] apply now to regulated light-cone cuts. However, because the dual field theory does not gain any dimensions, we need to find a way to retain this information in the limit $\epsilon \to 0$. Unsurprisingly, this will require supplementing the standard cuts $C(p)$ with some information from $C_\epsilon(p)$. Precisely how the $\epsilon \to 0$ limit of $C_\epsilon(p)$ can be used to extend $C(p)$ sufficiently for the light-cone cut reconstruction to succeed is the subject of this section.

5.3.1 Asymptotics of spacetimes with degenerate boundaries

The first step is to have an elementary understanding of how null geodesics behave asymptotically in spacetimes with an internal space. Henceforth, the bulk spacetime M is assumed to be asymptotically locally isometric to $\mathrm{AdS}_n \times \mathbb{S}^k$, whose metric in global coordinates reads

$$g = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_{n-2}^2 + \ell^2 d\Omega_k^2 \quad \text{where} \quad f(r) = 1 + \frac{r^2}{\ell^2}. \tag{5.4}$$

Here ℓ is the radius of curvature of AdS_n, and the shorthand Ω_d is used to collectively refer to all coordinates on $\mathbb{S}^d$. Define dimensionless time τ and radial ρ coordinates via

$\tau = t/\ell$ and $r = \ell \tan \rho$, so that (5.4) becomes

$$g = \frac{\ell^2}{\cos^2 \rho} \left(-d\tau^2 + d\rho^2 + \sin^2 \rho \, d\Omega_{n-2}^2 + \cos^2 \rho \, d\Omega_k^2 \right). \tag{5.5}$$

Since null geodesics are only sensitive to the causal structure, which depends just on the conformal class of the metric, consider a Weyl rescaling $g \mapsto \bar{g} = \omega^2 g$, with $\ell\omega = \cos\rho$. This gives

$$\bar{g} = -d\tau^2 + d\rho^2 + \sin^2 \rho \, d\Omega_{n-2}^2 + \cos^2 \rho \, d\Omega_k^2, \tag{5.6}$$

which is simply time cross $\mathbb{S}^{n+k-1}$. Noting that the conformal boundary ∂M corresponds to the limit $\rho \to \pi/2$, it is now evident how the induced metric on $\mathbb{S}^k$ degenerates in the strict asymptotic limit. In fact, this is no different from the way in which the metric degenerates at the origin $\rho = 0$ in these coordinates. More explicitly, letting $\rho = \frac{\pi}{2} - \epsilon$ with $0 < \epsilon \ll 1$ and expanding locally in a neighborhood of ∂M, one finds

$$\bar{g} = -d\tau^2 + (1 - \epsilon^2/2) \, d\Omega_{n-2}^2 + d\epsilon^2 + \epsilon^2 \, d\Omega_k^2 + \mathcal{O}\left(\epsilon^4\right). \tag{5.7}$$

The (ϵ, Ω_k) sector above provides a convenient chart on the space orthogonal to ∂M. The AdS_n sector of (5.6) takes the familiar form of one half of the Einstein static universe, and the metric induced on ∂M,

$$\bar{g}_{\partial M} = -d\tau^2 + d\Omega_{n-2}^2, \tag{5.8}$$

reveals the usual boundary topology $\mathbb{R} \times \mathbb{S}^{n-2}$ of conformally compactified AdS spacetimes.

The leading behavior of null geodesics in M near the conformal boundary can be

extracted from g in (5.5) in the limit $\rho \to \pi/2$. Since null geodesics are conformally invariant, we can actually work with (5.6). Let γ be a null geodesic curve with affine parameter λ and tangent vector field $N = \dot{\gamma}$. The Killing symmetries of (5.6) give rise to several conserved quantities along γ. If we choose coordinates on the spheres so that the geodesic is moving in the φ direction on $\mathbb{S}^{n-2}$ and ψ direction on $\mathbb{S}^k$, then we get the following conserved charges:[8]

$$E = \dot{\tau}, \qquad L_{n-2} = \sin^2\rho\,\dot{\varphi} \qquad \text{and} \qquad L_k = \cos^2\rho\,\dot{\psi}, \tag{5.9}$$

One can fix an arbitrary overall factor in N by setting $E = \pm 1$, where the sign determines the time orientation. The general asymptotic form of N can thus be written

$$N^a = \pm(\partial_\tau)^a + \dot{\rho}\,(\partial_\rho)^a + \frac{L_{n-2}}{\sin^2\rho}(\partial_\varphi)^a + \frac{L_k}{\cos^2\rho}(\partial_\psi)^a, \tag{5.10}$$

where the null condition $N^2 = 0$ constrains ρ to obey

$$\dot{\rho}^2 = 1 - \frac{L_{n-2}^2}{\sin^2\rho} - \frac{L_k^2}{\cos^2\rho}. \tag{5.11}$$

The limit $\rho \to \pi/2$ in (5.11) makes it immediately clear that γ can only reach ∂M if $L_k = 0$. This means that null geodesics only reach the conformal boundary if they approach a fixed point on $\mathbb{S}^k$ at infinity. This is easily understood from the perspective of Kaluza-Klein reduction, where a non-zero L_k would physically correspond to a massive test particle on the dimensionally-reduced spacetime, which of course cannot reach the conformal boundary.

[8] These are only conserved charges in global $\mathrm{AdS}_n \times \mathbb{S}^k$, and will not actually be conserved along γ on M in general. More appropriately, these quantities should be thought of as the asymptotic charges carried by γ as it reaches ∂M.

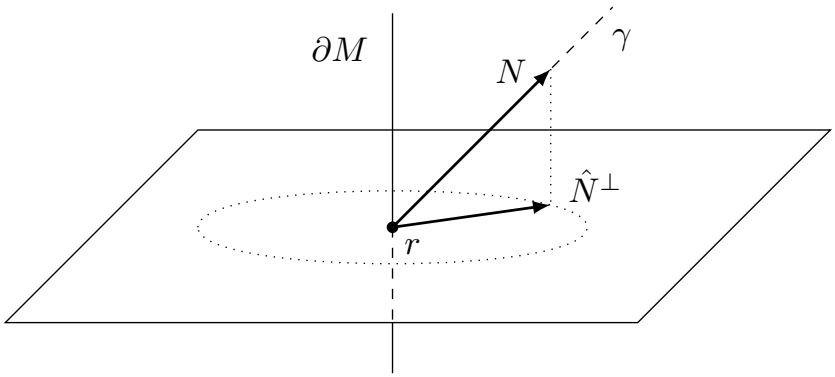


Figure 5.3: Illustration of the tangent space normal to the conformal boundary at some $r \in \partial M$. The vertical line represents the conformal boundary ∂M, and the normal plane corresponds to the radial and $\mathbb{S}^k$ bulk dimensions. A null geodesic γ reaches ∂M with future-directed tangent vector N. The unit-norm orthogonal projection of this vector $\hat{N}^{\perp}$ gives a point on $\mathbb{S}^k$ which corresponds to the asymptotic location of γ on the compact space.

Expanding about ∂M as in (5.7), the asymptotic form of N becomes

$$N^a = \pm(\partial_\tau)^a - \sqrt{1 - L_{n-2}^2}\,(\partial_{\epsilon(\Omega_k)})^a + L_{n-2}\,(\partial_\varphi)^a + \mathcal{O}\left(\epsilon^2\right), \tag{5.12}$$

where the notation $\partial_{\epsilon(\Omega_k)}$ is introduced to make it explicit that the direction of the radial vector ∂_ϵ on the (ϵ, Ω_k) sector is parameterized by the angular coordinates Ω_k on the asymptotic $\mathbb{S}^k$, like in ordinary spherical coordinates. The corresponding parametric form of its asymptotic integral curve is thus, to leading order,

$$\gamma(\epsilon) = \left(\tau^\infty \mp \epsilon,\, \frac{\pi}{2} - \sqrt{1 - L_{n-2}^2}\,\epsilon,\, \varphi^\infty - L_{n-2}\,\epsilon,\, \Omega_k^\infty\right) + \mathcal{O}\left(\epsilon^2\right), \tag{5.13}$$

where coordinates with superscripts ∞ denote asymptotic values and ∂M is reached at $\epsilon = 0$. Note that the limiting Ω_k^∞ will always be well defined despite the fact that the spherical coordinate system (ϵ, Ω_k) degenerates at its origin $\epsilon = 0$. In particular, $L_k = 0$ implies that $\partial_{\epsilon(\Omega_k)} = \partial_{\epsilon(\Omega_k^\infty)}$ in (5.12).

5.3.2 Definition of extended light-cone cuts

In order to recover the higher-dimensional bulk conformal metric, we will need the point on $\mathbb{S}^k$ associated with the null geodesic going from p to $C(p)$. Since $\mathbb{S}^k$ shrinks to zero size on the boundary, we will use the $\epsilon \to 0$ limit of $\partial_{\epsilon(\Omega_k^\infty)}$. The latter can be characterized geometrically as the unit vector $\hat{N}^\perp \in \mathbb{R}^{k+1}$ along the projection of the null tangent vector N orthogonal to ∂M. As a unit vector in a $(k+1)$-dimensional vector space, one can identify $\hat{N}^\perp$ with a point on $\mathbb{S}^k$, the coordinates of which are Ω_k^∞ (see Fig. 5.3).

Unfortunately, more than one null geodesic may connect p to $C(p)$, so the assignment of a point on $\mathbb{S}^k$ may not be unique. This motivates the following definition:

Definition 5.1 (Regular light-cone cut point)**.** *A regular light-cone cut point $r \in C(p)$ for some bulk point $p \in M$, is a cut point such that there exists a unique null geodesic from p to r.*

It is tempting to think of a point r that fails to be regular as belonging to some caustic on the light cone, as is the case in spacetimes without internal spaces. While this will commonly be true here too, one should bear in mind that the null geodesics that connect p and r might actually stay at finite proper distance apart on $\mathbb{S}^k$, only coinciding strictly at the conformal boundary. If this happened to be the case for all null geodesics connecting p to r, these points would not be conjugate points, and thus it would not be correct to think of r as arising from some bulk caustic. To account for this subtlety, it will be useful to dispense with the notion of caustics and use only what happens to be relevant from the boundary perspective in identifying whether a cut point is regular. Two null vectors N_1 and N_2 at r clearly define inequivalent null geodesics if and only if one is not a rescaling of the other. Hence the failure of a light-cone cut point $r \in C(p)$

to be regular can be characterized by the existence of at least two null geodesics γ_1 and γ_2 from p to r with respective tangent vector fields N_1 and N_2 satisfying $(N_1 \cdot N_2)_r \neq 0$. It will thus be intuitive to refer to a non-regular cut point as a *cusp point*.

Let $G(p) \subseteq C(p)$ be the subset of regular points in the light-cone cut of $p \in M$. On this subset, there exists a well-defined map $\Phi : G(p) \to \mathbb{S}^k$ associating a point on the unit k-sphere to every regular point. Explicitly, as remarked above, this map may be written

$$\Phi(r) = \hat{N}_r^{\perp}, \tag{5.14}$$

where an isomorphism between the unit $\mathbb{S}^k$ and the space of $(k+1)$-dimensional unit vectors is implied (see Fig. 5.4). In contrast, there is no guarantee that an analogous map on the set of cusp points $E(p) \equiv C(p) \setminus G(p)$ would be well-defined due to potential multi-valuedness on $\mathbb{S}^k$.

Definition 5.2 (Extended light-cone cut)**.** *The extended future/past light-cone cut* $\mathcal{C}^{\pm}(p)$ *of a point* $p \in M$ *is defined on the set of regular points* $G^{\pm}(p) \subseteq C^{\pm}(p)$ *as*

$$\mathcal{C}^{\pm}(p) = \bigcup_{r \in G^{\pm}(p)} (r, \Phi(r)) \,.$$

These extended cuts $\mathcal{C}(p)$ may be thought of as a generalization of the standard cuts $C(p)$ where every suitable point, namely every $r \in G(p)$, is further endowed with the point on $\mathbb{S}^k$ at which the null geodesic from p to r ends up (see Fig. 5.4).

Since the reconstruction strategy relies on the existence of regular points on which the map (5.14) is defined, it is important to check whether $G(p)$ contains sufficiently many points at one's disposal. An important step in this direction is accomplished by the following proposition, which as proven in Appendix C.1 and applies to light-cone cuts

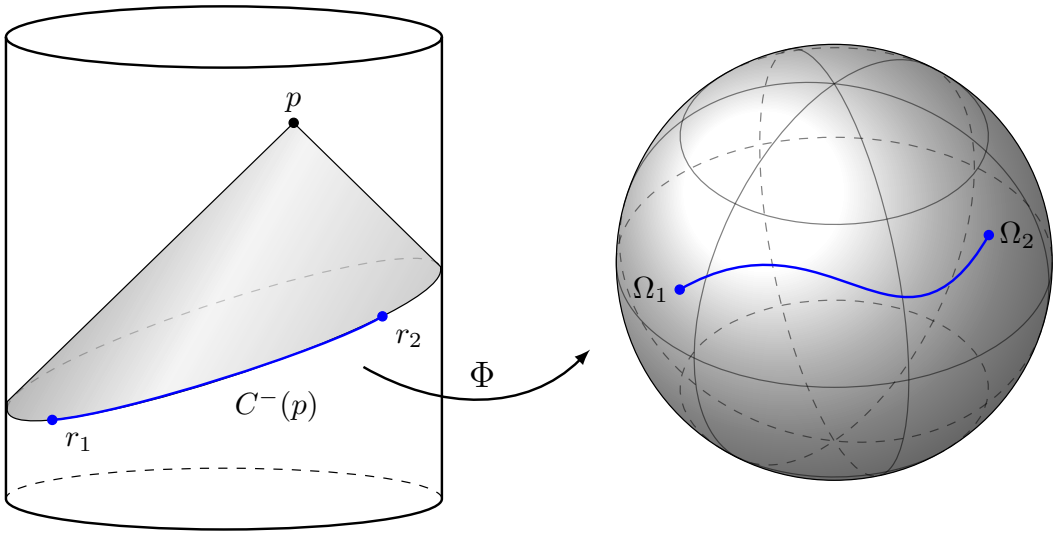


Figure 5.4: The map Φ defined in (5.14) takes (regular) points in the cut $C(p)$ of a point $p \in M$ and maps them to $\mathbb{S}^k$. For instance, the blue segment between points $r_1, r_2 \in C^-(p)$ maps to the blue segment between points $\Omega_1, \Omega_2 \in \mathbb{S}^k$.

in spacetimes with degenerate boundaries:

Proposition 5.1. *Every light-cone cut $C(p)$ is differentiable everywhere except on a set of measure zero.*

A cut $C(p)$ can be non-differentiable at r only if there is more than one null geodesic from p to r. So any point r at which $C(p)$ fails to be differentiable will be a cusp point $r \in E(p)$, and thus the set of all non-differentiable cusp points is of measure zero in $C(p)$.

Since there may be cusp points where $C(p)$ is differentiable, this is not enough to conclude anything about the measure of $E(p) \subseteq C(p)$. However, differentiability at cusp points is only possible if all geodesics from p to r happen to have tangent vectors at r with the same normalized projection onto ∂M. Fortunately, given one vector, such a condition on the second is satisfied only by a set of measure zero and thus the set of all differentiable cusp points is expected to be of measure zero in $E(p)$.

Putting together the conclusions of the last two paragraphs, one expects that the

union of all differentiable and non-differentiable cusp points, which is nothing but the set of all cusp points $E(p)$, is of Lebesgue measure zero as a subset of $C(p)$. This implies that its complement, i.e. the set of all regular points $G(p)$, is of full measure, everywhere dense and that its closure $\bar{G}(p) = C(p)$.

The key property of the extended cut that we will use is the following:

Proposition 5.2. *Each point $(r, \Phi(r))$ on the extended cut $\mathcal{C}(p)$ determines the unique null geodesic from r to p.*

This result is proven in Appendix C.1 and provides the connection to previous results in [46, 45].

5.3.3 Recovering the bulk conformal metric from extended cuts

The following results apply to standard light-cone cuts $C(p)$ and their proofs are identical to those in [45], so they are omitted:[9]

Proposition 5.3. *$C(p)$ is a complete spatial slice of ∂M.*

Proposition 5.4. *For any $p \in J^{\pm}[\partial M]$, there exists precisely one past/future cut $C^{\mp}(p)$.*

The following results, in contrast, are generalizations of results in [45] which now apply to extended light-cone cuts (see Appendix C.1 for proofs):[10]

Proposition 5.5. *$\mathcal{C}(p) \cap \mathcal{C}(q)$ contains more than one point if and only if $p = q$.*

Theorem 5.1. *If $\mathcal{C}(p) \cap \mathcal{C}(q)$ contains exactly one point, then p and q are distinct and null-related.*

[9] These results correspond to parts (1) and (2) of the Proposition in [45].

[10] These results are analogous to (a stronger version of) part (3) of the Proposition and Theorem 1 in [45].

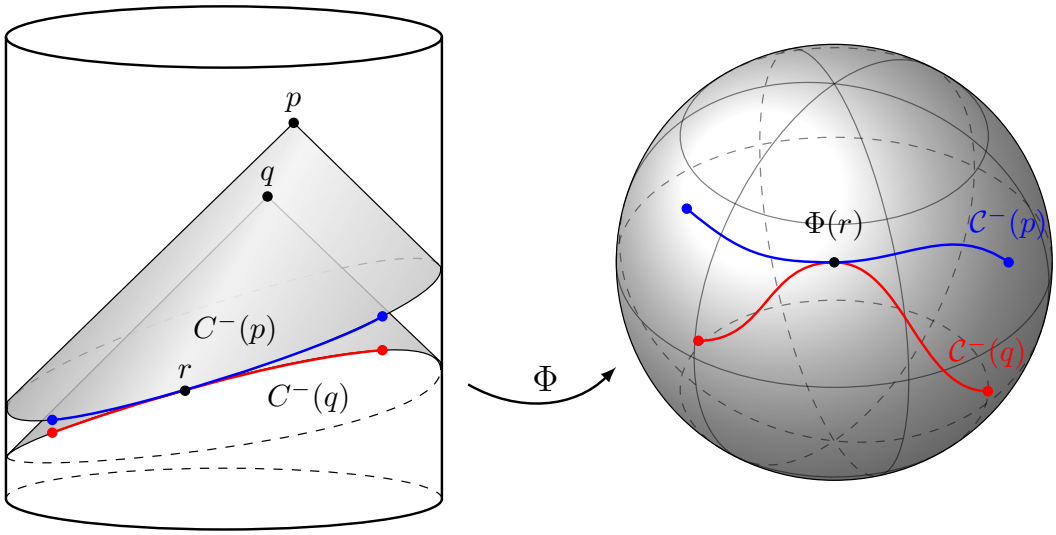


Figure 5.5: By Theorem 5.1, two extended cuts $\mathcal{C}^-(p)$ and $\mathcal{C}^-(q)$ intersect at precisely one point $(r, \Phi(r))$ only if they correspond to distinct, null-related points $p, q \in M$. As illustrated, this requires both that their standard cuts $C(p)$ and $C(q)$ intersect at precisely one point $r \in \partial M$, and also that their images under the map Φ intersect precisely at $\Phi(r) \in \mathbb{S}^k$.

Actually, a slightly stronger version of Theorem 5.1 is proven in Appendix C.1. The idea of the proof is simply that the common point on both extended cuts defines an ingoing null geodesic that must go through both p and q, and hence they must be null related.

From Proposition 5.4, the extended past cuts provide a copy of the space $J^+[\partial M]$. From Theorem 5.1, we can determine a class of null directions at each point $p \in J^+[\partial M]$, by looking for extended cuts $\mathcal{C}^-(q)$ which intersect $\mathcal{C}^-(p)$ at precisely one point. This situation is depicted in Fig. 5.5. One cannot recover all null directions at p but only those corresponding to null geodesics which stay on $\partial J^-(p)$. To obtain the conformal metric, one needs a basis of null directions. So the key question is whether the null directions we can reconstruct form a basis. This is not obvious since the tangent space at p is $n + k$ dimensional, and we only have access to an $n - 2$ dimensional space of

null directions associated to points of $\mathcal{C}(p)$. For instance, the answer would be negative in a spacetime which is globally a product such as $\mathrm{AdS}_n \times \mathbb{S}^k$, since $\mathcal{C}(p)$ would yield null geodesics that are everywhere orthogonal to $\mathbb{S}^k$. However, for a generic spacetime without any symmetries acting only on the internal space asymptotically, one expects the $n-2$ dimensional space of null geodesics corresponding to $\mathcal{C}(p)$ to span the tangent space, and not all be orthogonal to any vector at p. Hence we expect that one can generically construct a basis of null vectors ℓ_i at p. One can then choose the additional null vectors η_k and determine the conformal metric as described in Section 5.2.

5.4 Data from the Dual Field Theory

In the present context, there is no obstruction to obtaining the standard light-cone cuts from the perspective of the boundary theory via the method reviewed in Section 5.2 and originally presented in [46, 45].[11] Nevertheless, as observed in Section 5.3, knowledge of the cuts $C(p)$ is not sufficient for the reconstruction of the higher-dimensional bulk metric when the latter has a degenerate conformal boundary. The additional information needed for such reconstruction to succeed is encoded in the extended light-cone cuts $\mathcal{C}(p)$ and given by the map $\Phi : G(p) \to \mathbb{S}^k$ from regular cut points to the asymptotic k-sphere. The main focus of this section is to address the problem of how to obtain this extra ingredient solely from the dual field theory. We will propose a procedure to recover this map to the sphere using only field theory correlators.

[11] Notice, though, that the required number of correlator insertions to obtain a bulk-point singularity now needs to account for the bulk dimensions, not the boundary dimensions. In other words, one needs at least $n+k+1$ operators, not just $n+1$. See Section 5.4.3 for more details.

5.4.1 Higher-dimensional bulk-to-boundary propagator

From the bulk perspective, the action that describes some matter field φ is naturally defined on all $D = n+k$ dimensions of the bulk spacetime. Accordingly, the bulk-to-bulk propagator $\mathcal{G}$ takes as input the coordinates X of bulk points in some higher-dimensional chart, i.e. $X \in \mathbb{R}^D$.[12] In particular, if φ obeys an equation of motion of the form $\mathbb{P}_X\varphi(X) = J(X)$ for some source term J, then $\mathcal{G}$ is defined as the Green function of $\mathbb{P}_X$,

$$\mathbb{P}_X\mathcal{G}(X,\tilde{X}) = \frac{1}{\sqrt{\det g}}\delta^D(X-\tilde{X}). \tag{5.15}$$

Although it is a natural object, $\mathcal{G}$ rarely appears in the literature (see [151] for an exception in global $\mathrm{AdS}_n \times \mathbb{S}^k$). Instead, propagators are commonly obtained after dimensionally-reducing spacetime and Kaluza-Klein expanding on the compact dimensions. The result is an infinite family of simpler propagators associated to the infinite Kaluza-Klein tower of modes which, holographically, correspond to operators of definite conformal dimension. However, in a completely general spacetime, there is no well-defined way of discriminating the compact dimensions far from the conformal boundary. Hence, one cannot hope to learn much about the higher-dimensional spacetime geometry from the perspective of boundary correlators unless one understands how all such Kaluza-Klein mode propagators combine into the higher-dimensional propagator $\mathcal{G}$ and its bulk-to-boundary analogue $\mathcal{K}$. The goal of this section is to define and understand these higher-dimensional propagators and demonstrate how they may be used to obtain the map Φ for the construction of the extended cuts.

Although $\mathcal{G}$ is a perfectly well-defined object, it turns out to be nontrivial to ob-

[12] In what follows, it suffices to work with retarded and advanced propagators. Under an appropriate notion of global or AdS hyperbolicity, these are well-defined and unique in general time-dependent spacetimes [150].

tain an explicit, compact expression for it for a general minimally-coupled Klein-Gordon scalar field even in global $\mathrm{AdS}_n \times \mathbb{S}^k$. Without simplifying assumptions, the latter can be expressed as an infinite Kaluza-Klein series expansion as in (C.18). However for a conformally flat choice of radii in $\mathrm{AdS}_n \times \mathbb{S}^k$ and a specific mass term for the scalar corresponding to the Weyl invariant coupling to the scalar curvature, this infinite sum can be recast into the very simple form of (C.22) [151]. One example of this is the massless dilaton in $\mathrm{AdS}_5 \times \mathbb{S}^5$ (with equal radii), since the scalar curvature vanishes.

On the other hand, the bulk-to-boundary propagator $\mathcal{K}$ is a more subtle object. For local $\mathrm{AdS}_n \times \mathbb{S}^k$ asymptotics, a canonical choice of coordinates near the conformal boundary is Fefferman-Graham (z, x) on the AdS_n part [152] and standard hyperspherical coordinates Ω on the $\mathbb{S}^k$ part. Accordingly, in some neighborhood of the conformal boundary one may set $X = (z, x, \Omega)$. Despite Ω being a degenerate coordinate for any point on the conformal boundary, corresponding to $z = 0$, the limiting value of Ω exists along some curves as $z \to 0$ (cf. the discussion at the beginning of Section 5.3.2). From this standpoint, one would expect that some generalization of the extrapolate dictionary should allow one to obtain the bulk-to-boundary propagator $\mathcal{K}$ given the bulk-to-bulk propagator $\mathcal{G}$. In particular, one would hope to construct an object like $\mathcal{K}(\tilde{X}; x, \Omega)$ as some limit $z \to 0$ along curves of constant (x, Ω) of $\mathcal{G}(\tilde{X}; z, x, \Omega)$, where $\tilde{X}$ are the coordinates of an arbitrary bulk point.

There is a subtlety, though: because the boundary is a conformal boundary, one generally only considers the $z \to 0$ limit of propagators of definite scaling dimension, for which it is clear which power of z the leading term carries. Asymptotically, this scaling dimension is associated to Kaluza-Klein modes arising from the dimensional reduction of the $\mathbb{S}^k$. But by virtue of being higher-dimensional, the propagator $\mathcal{G}$ incorporates all such modes, and therefore the extrapolation of it to $\mathcal{K}$ via the $z \to 0$ limit should take

care of all of them at once. Due to these complications, we shall take a more axiomatic approach in defining $\mathcal{K}$.

As a bulk-to-boundary propagator, $\mathcal{K}$ will be defined to be the kernel of $\mathbb{P}_X$, i.e. the solution to the homogeneous equation

$$\mathbb{P}_{\tilde{X}}\mathcal{K}(\tilde{X};\, x, \Omega) = 0, \tag{5.16}$$

and subject to some choice of boundary conditions at ∂M. These conditions are imposed on the limit in which the bulk point approaches the conformal boundary too. In this limit, $\tilde{X} = (\tilde{z}, \tilde{x}, \tilde{\Omega})$ is again an appropriate chart and as $\tilde{z} \to 0$ one may work with the intuition that $\mathrm{AdS}_n \times \mathbb{S}^k$ provides. In particular, by dimensionally reducing near the conformal boundary, one can decompose $\mathcal{K}$ into contributions from lower-dimensional propagators for all possible Kaluza-Klein modes $\mathcal{K}_\Delta$ of definite scaling dimension Δ. Thus, at least for $\tilde{X}$ near the boundary, we have

$$\mathcal{K}(\tilde{X};\, x, \Omega) = \sum_{L=0}^{\infty} \mathcal{K}_{\Delta_L}(\tilde{X};\, x, \Omega). \tag{5.17}$$

The dependence on $\mathbb{S}^k$ is not arbitrary, but fixed by the choice of boundary conditions. For the usual Dirichlet conditions one would demand that $\lim_{\tilde{z}\to 0} \tilde{z}^{\Delta-d}\mathcal{K}_\Delta(\tilde{z}, \tilde{x}, \tilde{\Omega};\, x, \Omega) \propto \delta^d(x - \tilde{x})$, where $d = n - 1$. In the higher-dimensional analogue, the $\mathbb{S}^k$ coordinates really correspond to physical, compact dimensions, and the Dirichlet conditions should be imposed on those too. This motivates accounting for all Kaluza-Klein modes L in the definition of boundary conditions via

$$\lim_{\tilde{z}\to 0} \sum_{L=0}^{\infty} \tilde{z}^{\Delta_L - d}\mathcal{K}_{\Delta_L}(\tilde{z}, \tilde{x}, \tilde{\Omega};\, x, \Omega) = \frac{1}{\sqrt{\det g_{\mathbb{S}^k}}}\delta^d(x - \tilde{x})\delta^k(\Omega - \tilde{\Omega}). \tag{5.18}$$

We can now continue the propagator $\mathcal{K}$ deeper inside the bulk as a kernel of $\mathbb{P}_X$ using either retarded or advanced evolution. The result is our desired bulk-to-boundary propagator in the full spacetime. This approach is followed in Appendix C.2.2 to obtain the general form of the bulk-to-boundary propagator for the Klein-Gordon scalar field in global $\mathrm{AdS}_n \times \mathbb{S}^k$, expressed as an infinite series in (C.25) (cf. the bulk-to-bulk series in (C.18)).[13] In the particular case of Weyl-invariant matter, it is again possible to resum this series expansion and obtain a compact expression, namely (C.33).

5.4.2 The compact space from the dual field theory

The asymptotic form of a scalar field φ on an asymptotically locally $\mathrm{AdS}_n \times \mathbb{S}^k$ spacetime admits a Kaluza-Klein expansion over the $\mathbb{S}^k$ in scalar hyperspherical harmonics $Y_L^{I_L}$ of the form[14]

$$\varphi(z, x, \Omega) = \sum_{L=0}^{\infty} \sum_{I_L} Y_L^{I_L}(\Omega) \varphi_L^{I_L}(z, x). \tag{5.19}$$

According to the holographic dictionary, the term that leads asymptotically of the non-normalizable branch of every mode

$$\lim_{z \to 0} z^{\Delta_L - d} \varphi_L^{I_L}(z, x) = \phi_L^{I_L}(x), \tag{5.20}$$

becomes a source of a local boundary operator $\mathcal{O}_L^{I_L}(x)$ of definite conformal dimension Δ_L. Introducing a generic bulk field φ involving arbitrarily many Kaluza-Klein modes thus corresponds to turning on arbitrarily heavy operators on the boundary theory. Explicitly, the bulk partition function is equal to a field theory partition function involving

[13] Since this spacetime is static, Euclidean propagators are used in Appendix C.2.

[14] For more details on how the harmonic functions $Y_L^{I_L}$ are defined see Appendix C.2.1.

a complicated operator sum $\mathcal{O}_\phi$ of the form

$$\mathcal{O}_\phi(x) = \sum_{L=0}^{\infty} \sum_{I_L} \phi_L^{I_L}(x) \mathcal{O}_L^{I_L}(x). \tag{5.21}$$

As a boundary operator in its own right, $\mathcal{O}_\phi$ creates a bulk field with a conformal asymptotic profile $\phi(x, \Omega)$ which is given by contributions from all sources

$$\phi(x, \Omega) = \sum_{L=0}^{\infty} \sum_{I_L} Y_L^{I_L}(\Omega) \phi_L^{I_L}(x). \tag{5.22}$$

Following this intuition and using a quantum mechanical language, at any fixed boundary coordinate x, the insertion of $\mathcal{O}_\phi(x)$ produces a particle which is thrown into the bulk localized at a point in ∂M with coordinates x and whose wavefunction is spread over the asymptotic $\mathbb{S}^k$ according to $\phi(x, \Omega)$ as a function of Ω. More explicitly, the action of the operator $\mathcal{O}_\phi(x)$ on the vacuum state $|0\rangle$ of the boundary theory creates a state $|\phi_x\rangle = \mathcal{O}_\phi(x)\,|0\rangle$. When projected onto the position basis Ω of $\mathbb{S}^k$, this state reads $\langle \Omega | \phi_x \rangle = \phi(x, \Omega)$, whereas when projected onto the basis of eigenfunctions $Y_L^{I_L}$ of $\Box_{\mathbb{S}^k}$, it reads $\langle Y_L^{I_L} \mid \phi_x \rangle = \phi_L^{I_L}(x)$.

Consider the following object, a generalization of which will be relevant in the next subsection:

$$\Pi(\tilde{X};\, x) = \int_{\mathbb{S}^k} d\Omega\; \phi(x, \Omega) \mathcal{K}(\tilde{X};\, x, \Omega). \tag{5.23}$$

For instance, in global $\mathrm{AdS}_n \times \mathbb{S}^k$, using the Kaluza-Klein expanded form of $\mathcal{K}$ in (C.25),

$$\Pi(\tilde{z}, \tilde{x}, \tilde{\Omega};\, x) = \sum_{L=0}^{\infty} \sum_{I_L} Y_L^{I_L}(\tilde{\Omega})\; \phi_L^{I_L}(x) K_{\Delta_L}(\tilde{z}, \tilde{x};\, x), \tag{5.24}$$

where K_{Δ_L} is the usual L-mode bulk-to-boundary propagator, given in (C.23). The

bilocal field Π in (5.23) can be thought of as the response function of a boundary probe ϕ at x smeared over the $\mathbb{S}^k$ to a localized bulk source at $\tilde{X}$ propagated through spacetime via $\mathcal{K}$. This interpretation will naturally follow from a more complicated but closely related construct in Section 5.4.3 that comes out of a correlation function which boundary observers have access to. Although the right-hand side of (5.23) is integrated over Ω, note that Π depends on the profile of ϕ as a function of Ω and is thus sensitive to dependencies on the asymptotic $\mathbb{S}^k$. More precisely, if a boundary observer who can measure Π had complete control over ϕ, by tuning the boundary profile to be $\phi(x,\Omega) = \delta^k(\Omega - \Omega')$ parametrized by Ω', it would be possible for them to scan over Ω' and reproduce $\mathcal{K}$ precisely. However, note that by completeness of the spherical harmonics, such a choice of ϕ would correspond to picking $\phi_L^{I_L}(x) = {Y_L^{I_L}}^*(\Omega')$, which according to (5.21) would build $\mathcal{O}_\phi$ out of operators $\mathcal{O}_L^{I_L}$ of all dimensions L, including arbitrarily heavy ones.

More realistically, one might want to only use light operators and get as good an approximation to $\mathcal{K}$ as possible. With this goal, consider letting $\phi_L^{I_L} = \delta_{L\tilde{L}}\delta^{I_L\tilde{I}_{\tilde{L}}}$ in (5.22) (which corresponds to simply $\mathcal{O}_\phi = \mathcal{O}_{\tilde{L}}^{\tilde{I}_{\tilde{L}}}$), and label the resulting right-hand side in (5.23) by $\Pi_{\tilde{L}}^{\tilde{I}_{\tilde{L}}}$. This allows one to invert (5.23) by writing $\mathcal{K}$ as a harmonic series

$$\mathcal{K}(\tilde{X};\, x, \Omega) = \sum_{L=0}^{\infty} \sum_{I_L} \Pi_L^{I_L}(\tilde{X};\, x)\, {Y_L^{I_L}}^*(\Omega), \tag{5.25}$$

where the correlators in the sum are effectively the Fourier coefficients of the expansion. For an approximation to $\mathcal{K}$, one may want to employ L modes only up to some finite cut-off $L_\infty < \infty$. It should be noted that (5.25) applies to any asymptotically locally $\mathrm{AdS}_n \times \mathbb{S}^k$ spacetime (cf. (5.17) and comments below).

We would like to obtain the position of a local bulk source solely from the boundary perspective using the bulk-to-boundary propagator. It is pertinent at this point to make

clear the semantic distinction between *localizing* and *locating*. We do not want to create a perturbation *localized* on $\mathbb{S}^k$, which would require the whole tower of Kaluza-Klein modes. Instead, what we want is to *locate* a source that already is localized on $\mathbb{S}^k$, which need not require such high-L physics. Indeed, in the tractable case of global $\mathrm{AdS}_n \times \mathbb{S}^k$, we now show that using Π it is possible from the boundary perspective to find the exact location on $\mathbb{S}^k$ of a localized bulk source employing just $L = 1$ operators.

Let $\mathcal{O}_\phi$ only involve light operators in the fundamental representation of $SO(k+1)$ such that only $L = 1$ harmonics contribute to ϕ. With homogeneous sources, a general expression for the latter is obtained by writing the coefficients $\phi_L^{I_L} = \delta_{L,1} Y_1^{I_1 *}(\Omega)$ parameterized by a point Ω on $\mathbb{S}^k$. Suggestively writing $\Pi(\tilde{X};\, x) = \Pi_1(\tilde{X};\, x, \Omega)$ for this choice of ϕ, (5.23) becomes[15]

$$\begin{aligned} \Pi_1(\tilde{z}, \tilde{x}, \tilde{\Omega};\, x, \Omega) &= \sum_{I_1} Y_1^{I_1 *}(\Omega) \int_{\mathbb{S}^k} d\Omega'\; Y_1^{I_1}(\Omega') \mathcal{K}(\tilde{X};\, x, \Omega') \\ &= (k+1) K_{\Delta_1}(\tilde{z}, \tilde{x};\, x) \cos\theta, \end{aligned} \tag{5.26}$$

where θ is the angular separation between coordinates $\tilde{\Omega}$ and Ω on $\mathbb{S}^k$. Therefore, a boundary observer that is able to vary Ω will find $\Omega = \tilde{\Omega}$ precisely at the maximum of Π_1, corresponding to $\theta = 0$. This shows that, from the boundary perspective, the function Π_1 of $L = 1$ modes allows one to *locate* the exact position on $\mathbb{S}^k$ at which a *localized* bulk source resides.

5.4.3 Recovering the extended cut

As observed in previous sections, even with a compact space asymptotically, one can determine the standard light-cone cuts $C(p)$ from bulk-point singularities in certain

[15] Note that Ω here has been introduced as just a parameter for the choice of coefficients $\phi_L^{I_L}$.

boundary correlators. The only change is the number of operators in the correlator. The light cone of an arbitrary boundary point permeates the bulk as a submanifold of bulk-codimension one. In a generic spacetime, the intersection of the light cones of ℓ arbitrary boundary points will generically be a submanifold of bulk-codimension ℓ (or the empty set when $\ell > D$.)[16] So one needs at least D operators to single out a point in the bulk. In this section we further refine the usage of these correlators in order to obtain the map Φ. In particular, the strategy will be to find Φ from the prefactor of the leading divergent term of bulk-point singular correlators, which exhibits a suitable dependence on the asymptotic $\mathbb{S}^k$.

We start with a divergent correlator as used to find the standard light-cone cuts $C^{\pm}(p)$ of some bulk point $p \in M$,[17]

$$\left\langle \mathcal{T} \left\{ \prod_{i=1}^{D+2} \mathcal{O}(x_i) \right\} \right\rangle \tag{5.27}$$

where $\mathcal{T}$ denotes time ordering. As argued above, the choice of any D such points x_i in the correlator above singles out p as the unique bulk point that is null-related to all of them. As shown in Fig. 5.6, we place these D points on the future cut and add two points on the past cut. By moving these two points in a way that keeps the correlator divergent (which requires maintaining momentum conservation at p), we can trace out the past cut. This does not depend on the choice of operator insertions. Let us denote all but one of these boundary points collectively by $\mathbf{x} = \{x_i \in C^{\pm}(p) \mid i = 2, \ldots, D+2\}$ and use $L = 0$ scalar operators $\mathcal{O}$ at all of these $D+1$ points.[18] To obtain the extended

[16] As stated, it is important for this result to be generic that the spacetime compactification does not factorize exactly or have exact symmetries, and that the boundary points be chosen arbitrarily. Global $\mathrm{AdS}_n \times \mathbb{S}^k$ is thus clearly non-generic.

[17] Any set of operators referred to henceforth shall be assumed to correspond to some local interaction term in the action of the bulk theory. For example, in 10-dimensional supergravity, there is a coupling between the dilaton and 3-form, $e^{-2\phi}H_3^2$. Expanding out the exponential yields $\phi^{10}H_3^2$ interaction terms.

[18] One may want to consider more general insertions $\mathcal{O}_{\phi_i}$ at each boundary point, but this is unneces-

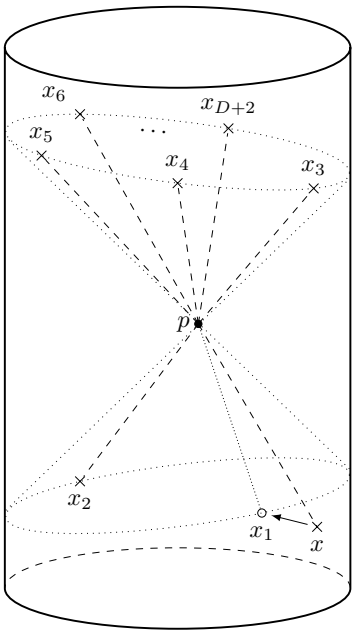


Figure 5.6: Configuration of boundary points in the correlator (5.28) used to obtain the extended light-cone cuts from the dual field theory. By choosing suitable operators at x and looking at the coefficient of the divergence when x approaches the cut point $x_1 \in C^-(p)$, one can obtain the map Φ from regular cut points to the $\mathbb{S}^k$.

cut, it will be convenient to work with a probe point x near the remaining point x_1. We choose an operator $\mathcal{O}_\phi$ at x as in (5.21) which is sensitive to the $\mathbb{S}^k$.

As a result of the existence of the null-related, momentum-preserving point p, the $(D+2)$-point, time-ordered Lorentzian correlator

$$F_\phi(x) \equiv \left\langle \mathcal{T} \left\{ \mathcal{O}_\phi(x) \prod_{i=2}^{D+2} \mathcal{O}(x_i) \right\} \right\rangle \tag{5.28}$$

will develop a bulk-point singular contribution in the limit $x \to x_1$ [83]. This divergent limit of interest is illustrated in Fig. 5.6. Written out in a particularly convenient form,

sary.

for some choice of normalization, the correlation function (5.28) reads

$$F_\phi(x) = \int_{\mathbb{S}^k} d\Omega\ \phi(x,\Omega) \int_M d^D\tilde{X}\ \Psi_{\mathbf{x}}(x,\Omega;\ \tilde{X}), \tag{5.29}$$

where the integrand of the bulk-point integral is

$$\Psi_{\mathbf{x}}(x,\Omega;\ \tilde{X}) = \lambda\ \mathcal{K}(x,\Omega;\ \tilde{X}) \int_{\mathbb{S}^k} \prod_{i=2}^{D+2} d\Omega_i\ \mathcal{K}(x_i,\Omega_i;\ \tilde{X}), \tag{5.30}$$

and λ is the coupling of some local interaction involving the $D+2$ fields of interest.

The dominant bulk-point singular contribution from p to (5.29) manifests itself as the highest-order pole in $\Psi_{\mathbf{x}}$, precisely located at the coordinates X_p of p, close to which the function $\Psi_{\mathbf{x}}$ will be governed by a power-law divergence in the proper distance between $\tilde{X}$ and X_p. To intuitively see why this is the case, observe first that propagators generally behave as inverse powers of proper distances between the points in their arguments, here with coordinates (x_i,Ω_i) and $\tilde{X}$. Importantly, because the asymptotic $\mathbb{S}^k$ trivializes on ∂M, this proper distance does not depend on the value of Ω_i for the boundary point. Now, since all x_i are null-related to X_p, for $\tilde{X}$ in a small neighborhood of X_p, to leading order the proper distance $s(x_i,\tilde{X})$ between any boundary point x_i and $\tilde{X}$ will be proportional to $s(\tilde{X},X_p) \approx \left\|\tilde{X}-X_p\right\|$, where the use of the Minkowski metric in the last approximation is justified by local flatness at X_p. The dependence on the choice of boundary points $\mathbf{x}$ is thus relegated simply to the specification of the unique bulk point p in this equation (cf. the rank argument in [83]) and the form of the residue of the pole of $\Psi_{\mathbf{x}}$ at X_p. The order of the dominant pole Δ_{D+2} depends on the operator insertions and details of the

spacetime metric.[19] Pulling out the leading divergent factor in $\Psi_{\mathbf{x}}$, one may write

$$\Psi_{\mathbf{x}}(x, \Omega;\, \tilde{X}) = \frac{\psi(x, \Omega;\, \tilde{X})}{\left\| \tilde{X} - X_p \right\|^{\Delta_{D+2}}}, \tag{5.31}$$

where now the function ψ is finite and non-zero at $\tilde{X} = X_p$. To leading order in the distance $\|x - x_1\|$ off the light-cone cut, the integral of (5.31) over $\tilde{X}$ will be dominated by the zeroth order term of ψ in a series expansion about $\tilde{X} = X_p$ and evaluated at $x = x_1$. This leads to

$$F_\phi(x) = I(x) \int_{\mathbb{S}^k} d\Omega\; \phi(x, \Omega) \psi(x_1, \Omega;\, X_p), \tag{5.32}$$

where $I(x)$ captures the bulk-point singularity as $x \to x_1$ from the integral over $\tilde{X}$,[20]

$$I(x) \propto \|x - x_1\|^{-(\Delta_{D+2} - D)}. \tag{5.33}$$

The previous section showed that in global $\mathrm{AdS}_n \times \mathbb{S}^k$ it was possible to locate the unique direction specified by Ω_1 in which the null geodesic from X_p arrives at x_1 using the object defined in (5.23). The reason for this could be traced back to the fact that the higher-dimensional propagator $\mathcal{K}$ in (5.26) had a global maximum at $\Omega = \Omega_1$. By causality, this fact is expected to extend to arbitrary spacetimes, where now the general function ψ in (5.32) is the object peaked at $\Omega = \Omega_1$.[21]

[19] In the case of global $\mathrm{AdS}_n \times \mathbb{S}^k$, the symmetries lead to Δ_{D+2} being just a sum over the largest scaling dimension of each of the boundary operator insertions.

[20] The order of the pole agrees with the result in [83] if one identifies $\Delta_{D+2} \to (D+1)\Delta$ (corresponding to $D+1$ external vertices rather than $D+2$), and $D \to d+1$ (corresponding to no internal space).

[21] In $\mathrm{AdS}_n \times \mathbb{S}^k$, it sufficed to use light modes with $L = 1$ to locate this point since in this highly symmetric case, all nontrivial Kaluza-Klein modes are peaked at the same point. In a general spacetime, it is still expected that ψ will have a global maximum at $\Omega = \Omega_1$, but no single L mode need be peaked there. Put differently, ψ will generically exhibit no symmetries in Ω and higher L will be required to locate Ω_1.

In analogy to the previous section, choosing ϕ to to consist of a single $Y_L^{I_L}$, define

$$F_L^{I_L}(x) = I(x) \int_{\mathbb{S}^k} d\Omega \, Y_L^{I_L}(\Omega) \psi(x_1, \Omega; \, X_p), \tag{5.34}$$

which, up to $I(x)$, may be thought of as the Fourier coefficients of an expansion of ψ into hyperspherical harmonic functions. Inverting this relation leads to

$$\psi(x_1, \Omega; \, X_p) = \lim_{x \to x_1} \frac{\psi_0}{F_0(x)} \sum_{L=0}^{\infty} \sum_{I_L} F_L^{I_L}(x) {Y_L^{I_L}}^*(\Omega), \tag{5.35}$$

where F_0 corresponds to (5.34) for $L = 0$ and is introduced to cancel out the common bulk-point singular factor of every term in the series. The constant ψ_0, given by

$$\psi_0 = \frac{1}{\operatorname{vol} \mathbb{S}^k} \int_{\mathbb{S}^k} d\Omega \, \psi(x_1, \Omega; \, X_p), \tag{5.36}$$

is irrelevant and will be left undetermined.[22] The upshot is that, up to an overall constant, ψ can be reconstructed to arbitrary precision by computing the terms in the series in (5.35) for increasingly high L values. Since the right-hand side is built solely out of boundary correlators, this information is in principle accessible to boundary observers. Once obtained, the location of the global maximum of ψ in Ω, namely Ω_1, determines the desired map Φ to the asymptotic $\mathbb{S}^k$. More explicitly, one obtains $\Phi(x_1) = \Omega_1$ from the solution to $\psi(x_1, \Omega; \, X_p) = \max \psi$, where the specific choice of the $D + 1$ additional light-cone cut points $\mathbf{x}$ may be ignored since it is arbitrary so long as they belong to the same choice of past and future light-cone cuts $C(p)$ of point p at X_p (see Fig. 5.6).

[22] If ψ_0 vanishes identically so will F_0, and one may just use a different L mode to cancel out singular factors.

5.5 Discussion

Most discussions of bulk reconstruction in holography consider asymptotically AdS spacetimes and ignore the extra compact directions. This was true for the discussion of light-cone cuts in [45, 46]. We have considered asymptotically locally $\mathrm{AdS}_n \times \mathbb{S}^k$ spacetimes and defined a generalization of light-cone cuts that we call extended cuts. We then showed that in the region of spacetime causally connected to the boundary, one can generically recover the full higher-dimensional conformal metric just from the location of the extended cuts. Finally, we proposed a procedure for determining these extended cuts from the dual field theory. Note that at no time did we need to use any bulk equations of motion, or impose any restrictions on the matter content (such as energy conditions).

Our proposal for determining the extended cuts from the dual field theory is not very practical since it requires considering the entire tower of Kaluza-Klein modes to precisely locate the bulk points. But the lesson is that the information is there in principle.[23] It would be interesting to find a more efficient way to determine the extended cuts.

Although we have focused on the case where the bulk metric asymptotically approaches $\mathrm{AdS}_n \times \mathbb{S}^k$, our reconstruction should work equally well for spacetimes that asymptotically approach $\mathrm{AdS}_n \times K$, where K is any compact Einstein space. To see this, note that null geodesics that remain on $\partial J(p)$ will again approach a fixed point on K, and we can again define our extended cut to be the light-cone cut $C(p)$ together with a map $C(p) \to K$. The arguments in Section 5.3 then apply to show that the conformal metric can be reconstructed from the location of these extended cuts. One difference with $\mathbb{S}^k$ is that when we conformally rescale the asymptotic metric, the result will take the

[23] In practice, from the perspective of the dual field theory, one would probably first want to know how many extra dimensions the bulk spacetime has. This interesting question was recently addressed in [153].

form (5.6) with $d\Omega^2$ replaced by the metric on K, which will be singular at the conformal boundary. This should not be a problem since our arguments only require that each point on the extended cut defines a unique ingoing null geodesic in the bulk. Since we know the geodesic starts at a fixed point on K, and the bulk metric reduces to pure AdS asymptotically when a point on K is held fixed, the geodesic leaves the boundary exactly as it would in pure AdS. The arguments in Section 5.4 also extend to this case since the hyperspherical harmonics on $\mathbb{S}^k$ can be replaced by the eigenfunctions of the Laplacian on K which form a complete basis of functions. Scalar fields can be expanded in terms of these functions, yielding the usual infinite tower of massive Kaluza-Klein modes in the asymptotic AdS region. Holography requires that there is a operator dual to each of these modes which we can use in our correlators.

It is natural to ask how quantum or stringy corrections affect our arguments. It was argued in [83] that bulk-point singularities would still be present when perturbative $1/N$ or $1/\lambda$ corrections to holography are included, but not for finite N or λ. More recently, the stringy resolution of these singularities has been quantified under some general assumptions in [148]. Since bulk-point singularities are a key ingredient in our approach, we note that exact reconstruction of the bulk conformal metric is possible with perturbative but not finite quantum or stringy corrections.

We close with a few open questions. First, to recover the full bulk metric and not just the conformal metric, we clearly need a procedure to obtain the conformal factor. One would like this to be independent of the bulk equations of motion. Second, general arguments on bulk reconstruction [78] show that one should be able to reconstruct the higher-dimensional metric on the entire entanglement wedge of the boundary. The light-cone cut approach to bulk reconstruction only applies to points in the causal wedge, since they have to be in causal contact with the boundary both to the past and future. Actually,

not all points in the causal wedge are accessible since momentum must be conserved near the vertex. This means that points just outside the horizon of an eternal black hole are excluded since they are causally connected to infinity only through a narrow cone.

To expand the reach of the light-cone cuts, we either need more general ways to obtain the extended cuts from the dual field theory, or we need to use other methods (perhaps combined with light-cone cuts). The recent work of [154] relating bulk scattering and holographic entanglement suggests a plausible direction to connect the light-cone cut approach to bulk reconstruction with those based on entropic measures, thereby hinting at a potentially synergistic combination of the two. It is nevertheless important to note that it is unlikely that the standard holographic entanglement entropy as given by the prescriptions in [32, 36] could on its own be used for higher-dimensional bulk reconstruction. In a variety of nontrivial examples, it has been shown that these prescriptions applied to dimensionally-reduced spacetimes give entropies that agree with those obtained by performing the extremization problem on the full higher-dimensional spacetime, which suggest that the latter carries no more information about the extra dimensions than the former does [44]. Intuitively, this is a consequence of the boundary condition that instructs the higher-dimensional extremal surfaces to wrap uniformly around the compact dimensions asymptotically. However one might be able to generalize these ideas, perhaps along the lines of [155, 156, 43], to probe the higher-dimensional geometry. If a suitable boundary interpretation of this generalized entropy is available, one could perhaps use e.g. some upgraded version of the arguments in [86] to prove uniqueness of the higher-dimensional metric and potentially come up with an entropy-based reconstruction strategy.

Chapter 6

Boundary causality violating metrics in holography

6.1 Introduction

A bulk description of nonperturbative quantum gravity is not yet available. A popular approach is to consider a path integral over metrics. We investigate some aspects of this approach. The metric could turn out to be only a low-energy approximation of other fundamental degrees of freedom. Our results would plausibly still apply, since the issues we address arise already in metrics with curvature well below the Planck (and string) scale.

We begin with a Lorentzian formulation of holography, in which one integrates over asymptotically anti-de Sitter (AdS) spacetimes and matter fields (with certain boundary conditions) to compute correlation functions in a dual quantum field theory. One often works in a large N or semiclassical limit and only includes classical supergravity solutions and small perturbations of them. Classically, spacetimes satisfy the null energy condition

so the Gao-Wald theorem [157] ensures that the bulk preserves boundary causality. In other words, the fastest way to send a signal between two observers on the boundary is via a path that stays on the boundary. No trajectory that enters the bulk can arrive sooner. Semiclassically, the achronal averaged null energy condition ensures boundary causality [158]. (A necessary and sufficient geometric condition is given in [159].)

However, in full quantum gravity (finite N), the bulk path integral includes metrics that violate boundary causality. That is, two boundary points that are spacelike-separated on the boundary can nevertheless be timelike-separated with respect to some bulk metrics. We will see that these boundary causality violating metrics are not "rare" – they include open sets in the space of metrics. These metrics describe causally well-behaved bulk geometries – there are no closed timelike curves or any causal pathology. So there is no reason to exclude them from the bulk path integral. Hence one might worry that they could contribute to causality violations in the dual field theory, which would be a problem.

Consider, for example, the commutator of an operator $\mathcal{O}$ at two spacelike-separated boundary points: $[\mathcal{O}(x), \mathcal{O}(y)]$. This must vanish identically, but holography says that it should be given by a bulk path integral over metrics and a matter field ϕ dual to $\mathcal{O}$. If one computes the commutator $[\phi(p), \phi(q)]$ in each metric g, and then follows the standard limiting procedure of taking $p \to x$ and $q \to y$, one generically expects a nonzero answer whenever x and y are causally related with respect to g. This is not yet a contradiction: the integral of this over metrics g weighted by e^{iS_g} could still vanish exactly.

However, it is not just one quantity that must integrate to zero. In principle, $\mathcal{O}$ can be any operator in the dual field theory, and the commutator could be multiplied by any other operator. Their corresponding bulk expressions would all be nonzero in some metrics g, but their integrals over g must still vanish identically. Since the gravitational

weighting e^{iS_g} and measure are independent of the operator insertions, this seems unlikely.

To resolve this puzzle, we give a prescription for the bulk path integral in which the boundary commutator is *not* given by a limit of the naive bulk commutator. In our approach, the bulk calculation always vanishes when the points are spacelike-separated on the boundary, despite the presence of boundary causality violating metrics.

We first give some examples of simple metrics which violate boundary causality. We then review the "differentiate dictionary" and the "extrapolate dictionary" for operators in holography. After these preliminaries we explain our resolution of this causality puzzle.

6.2 Specific examples

For $b \in \mathbb{R}$, consider the family

$$g_{(b)} = -f_b(r)\,\mathrm{d}t^2 + f_b(r)^{-1}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega_2^2. \tag{6.1}$$

One requires $f_b(r) \approx r^2/\ell^2$ at large r for AdS asymptotics, and $f_b(0) = 1$ and $f_b'(0) = 0$ for differentiability at $r = 0$. We connect this family at $b = 0$ to AdS_4 by setting $f_0(r) = r^2/\ell^2 + 1$. In pure AdS, two boundary points are null-separated on the boundary if and only if they are null-separated in the bulk. Hence we explore boundary causality in (6.1) by comparing to AdS. Restricting to radial null geodesics for convenience, the boundary-to-boundary crossing time is

$$t_\infty = 2\int_0^\infty \frac{\mathrm{d}r}{f_b(r)}. \tag{6.2}$$

Hence (6.1) violates boundary causality if $t_\infty < t_\infty^{\mathrm{AdS}} = \pi\ell$; e.g., $f_b(r) > f_0(r)$ everywhere would suffice. In Einstein-Hilbert gravity, (6.1) gives the same action as AdS_4 so long

as $\lim_{r\to\infty}(f_b(r) - f_0(r))/r = 0$. Setting $\ell = 1$, a simple example of this is (6.1) with $f_b(r) = r^2 + 1 + b\, r^2(1 + r^4)^{-1}$, which violates boundary causality for $b > 0$.

These static configurations only contribute to a path integral where the initial and final surfaces have induced metric $\mathrm{d}r^2/f_b(r)+r^2\mathrm{d}\Omega_2^2$. One might ask if there exist classical solutions interpolating between such surfaces. The answer is yes: global AdS gives the desired induced metric with $f_b(r) \geq f_0(r)$ on spacelike surfaces with $t = b\, h(r)$. For AdS asymptotics, h must fall off like $1/r^2$ or faster. As an example, taking $h(r) = f_0(r)^{-1}$ one gets $f_b(r) = f_0(r)^3\,(f_0(r)^2 - 4b^2r^2)^{-1}$, defined for $b^2 < 1$.

Alternatively, to match a standard static surface in AdS, one can simply make b time-dependent near the surface, going to zero on it. More generally, one can match any given induced metric by making $g_{(b)}$ appropriately time-dependent without affecting the causality violating region. Furthermore, since the action is a global property of the spacetime, in this process one can make the action take any desired value, e.g., the same as pure AdS.

Although we have focused on four-dimensional examples, it is clear that similar examples exist in all dimensions greater than two. In particular, if x is timelike-related to y in one metric, it remains timelike-related in any nearby metric. Hence boundary causality is violated in open sets in the space of metrics.

6.3 Operator dictionaries

AdS/CFT is generally formulated as an equivalence between bulk and boundary partition functions [75, 17]:

$$\mathcal{Z}_{\text{bulk}}[\phi_0] = \mathcal{Z}_{\text{bdy}}[\phi_0]. \tag{6.3}$$

Here, $\mathcal{Z}_{\rm bdy}$ is defined on some fixed Lorentzian manifold $(\mathcal{B},\gamma)$, whereas $\mathcal{Z}_{\rm bulk}$ path-integrates over spacetimes $(\mathcal{M},g)$ having $(\mathcal{B},\gamma)$ as their conformal boundary. The symbol ϕ_0 is a placeholder for all boundary conditions for dynamical fields on the bulk side, and for sources of operator deformations of the action on the boundary side. To make progress, most work in the literature studies (6.3) perturbatively in $1/N$, where $\mathcal{Z}_{\rm bulk}$ is amenable to a saddlepoint approximation (plus corrections). Here we test one aspect of (6.3) as a statement of holography in full quantum gravity, i.e., nonperturbatively in $1/N$.

To study real-time correlation functions, we assume (6.3) applies in Lorentzian signature. From the standpoint of the CFT, $\mathcal{Z}_{\rm bdy}[\phi_0]$ is a standard field-theoretic generating functional. Explicitly [17],

$$\mathcal{Z}_{\rm bdy}[\phi_0] = \left\langle \exp\left\{ i \int_{\mathcal{B}} \sum_{\alpha} \phi_0^{\alpha}\, \mathcal{O}_{\alpha} \right\} \right\rangle_{\rm bdy} , \tag{6.4}$$

where $\langle\,\cdot\,\rangle_{\rm bdy}$ denotes the boundary path integral over CFT fields, and each $\mathcal{O}_\alpha$ is a gauge-invariant operator sourced by ϕ_0^α. In light of (6.3), the bulk object $\mathcal{Z}_{\rm bulk}[\phi_0]$ may also be regarded as a generating functional of boundary correlators of the form

$$G \equiv \langle \mathcal{O}_n(x_n) \cdots \mathcal{O}_1(x_1) \rangle_{\rm bdy} . \tag{6.5}$$

If G is time-ordered with respect to boundary time, then

$$G = \frac{1}{i^n} \frac{\delta}{\delta\phi_0^n(x_n)} \cdots \frac{\delta}{\delta\phi_0^1(x_1)} \left. \frac{\mathcal{Z}_{\rm bulk}[\phi_0]}{\mathcal{Z}_{\rm bulk}[0]} \right|_{\phi_0=0} . \tag{6.6}$$

The rationale so far is that of [75, 17], later coined the "differentiate dictionary" in [160].

In practice though, one usually takes a leap and follows the logic of [161, 162] to argue that a boundary correlator should be nothing but an appropriate limit of bulk

ones. More specifically, if ϕ_α is the bulk field dual to $\mathcal{O}_\alpha$ and the latter has conformal dimension Δ_α, one may propose [1]

$$G \overset{?}{\propto} \int \mathcal{D}g\, e^{iS_g} \lim_{z\to 0} z^{\Delta_\Sigma} \langle \phi_n(x_n, z) \cdots \phi_1(x_1, z) \rangle_g \,, \tag{6.7}$$

where $\Delta_\Sigma = \sum_\alpha \Delta_\alpha$, $\langle \cdot \rangle_g$ denotes the bulk path integral over quantum fields in a fixed metric g, and (x, z) are Fefferman-Graham coordinates [152]. (These provide an unambiguous asymptotic location for the bulk operators.) Alternatively, one could consider (6.7) with $z \to 0$ taken outside the integral. While physically both seem like reasonable proxies for boundary correlators, it is unclear whether they are equal or mathematically consistent with (6.6).

When gravity is treated semiclassically, these two options do agree. In this case, they were referred to as the "extrapolate dictionary" in [160], where their consistency with (6.6) was carefully studied. Their findings established a nontrivial perturbative equivalence between these two dictionaries for self-interacting scalar fields in a fixed background with Euclidean AdS asymptotics. We deviate from [160] by studying Lorentzian physics and at a nonperturbative level.

From the standpoint of the CFT, local operators at spacelike boundary separation commute. In the bulk, there appears to be a problem with the extrapolate dictionary stemming from the presence of off-shell contributions to the gravitational path integral that violate boundary causality.

Similar problems arise regardless of when one takes the $z \to 0$ limit, so consider (6.7) just for definiteness. For a commutator of operators at spacelike separation one has

[1] Time-ordering has been intentionally not included in the bulk correlator since the notion of time order in bulk is subtle. This is addressed in the next section.

$\langle[\mathcal{O}(x),\mathcal{O}(y)]\rangle_{\text{bdy}} = 0$, so (6.7) becomes

$$\int \mathcal{D}g\, e^{iS_g} \lim_{z\to 0} z^{2\Delta} \langle[\phi(x,z),\phi(y,z)]\rangle_g \stackrel{?}{=} 0. \tag{6.8}$$

Since the commutator vanishes in all metrics that respect boundary causality, the integral over g can only receive contributions from bulk metrics for which x and y are causally related. As explained above, these contain open sets of metrics and generically yield nonzero commutators. However, if the extrapolate dictionary was valid, the integral would have to vanish. Furthermore, any operator-valued function of the commutator that vanishes on the boundary when the commutator does should also vanish in the bulk. So there are infinitely many integrals like (6.8) which should give zero. Since e^{iS_g} and the measure are the same while the integrands change, we do not expect all such path integrals to be identically zero. It thus appears that the extrapolate dictionary is inconsistent with boundary causality.

6.4 Resolution

There was a subtlety omitted from the above discussion. With a proper interpretation of this subtlety, both dictionaries are consistent with boundary causality.

Consider the commutator

$$\langle[\phi(p),\phi(q)]\rangle_g \equiv \langle\phi(p)\phi(q)\rangle_g - \langle\phi(q)\phi(p)\rangle_g\,, \tag{6.9}$$

with each term computed via a path integral. Although we have used the bulk notation $\langle\cdot\rangle_g$, this also applies to the boundary by replacing $(\mathcal{M},g)$ with $(\mathcal{B},\gamma)$. Now, the path integral on the original manifold always computes time-ordered correlators. To compute

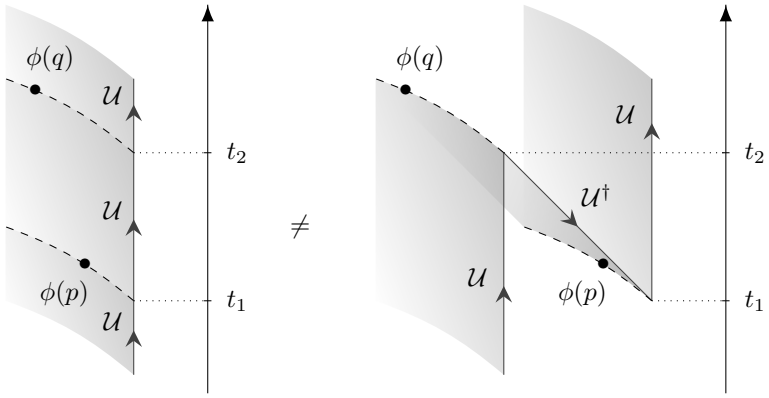


Figure 6.1: Inequivalent timefolds for computing a two-point function for a field ϕ inserted with different time orderings. Point q is to the future of p, each respectively lying at times t_1 and $t_2 > t_1$. The left figure represents the time-ordered correlator $\langle\phi(q)\phi(p)\rangle$, where time-evolution is implemented by $\mathcal{U}$ everywhere. The right figure shows the nontrivial timefold required for the out-of-time-order correlator $\langle\phi(p)\phi(q)\rangle$, where backwards time-evolution via $\mathcal{U}^\dagger$ is needed between insertions. Generically, these give different results.

an out-of-time-order correlator, one needs to manufacture from the original manifold a *timefold* as follows. Consider an n-point function for a field ϕ,

$$\langle\phi_+|\phi(p_n)\cdots\phi(p_2)\phi(p_1)|\phi_-\rangle_g\,, \tag{6.10}$$

where $\phi_\pm$ are initial and final states on Cauchy surfaces $\Sigma_\pm$. Starting at Σ_-, one should use the evolution operator $\mathcal{U}$ up to a Cauchy surface containing p_1 and insert $\phi(p_1)$. One should then evolve to p_2 for the next insertion. If p_2 is to the future of p_1, one could just use $\mathcal{U}$ again. However, if p_2 is to the past of p_1, one should evolve backwards with $\mathcal{U}^\dagger$ for the insertion of $\phi(p_2)$. This procedure must be repeated for all n insertions before finally evolving to Σ_+. The whole process produces a timefold, a new zigzagged spacetime implementing the correct ordering of operators (see Fig. 6.1). In the path integral, backward components receive weighting e^{-iS_g}, instead of the usual e^{iS_g} on forward ones.

If two points are timelike-related, the commutator is a difference between two cor-

relators computed on distinct timefolds, so will be generically nonzero. In the case of spacelike-separated points, however, the two points can be inserted on the same Cauchy surface. Hence the same timefold qualifies for a computation of either ordering, yielding a vanishing commutator. Again, the discussion here applies to both boundary commutators $\langle[\mathcal{O}(x),\mathcal{O}(y)]\rangle_{\rm bdy}$ and bulk commutators $\langle[\phi(x,z),\phi(y,z)]\rangle_g$ on a specific metric, and therefore one has both boundary and bulk timefolds.

In general, the field theory $\mathcal{Z}_{\rm bdy}$ is defined on some manifold $\mathcal{B}$, and $\mathcal{Z}_{\rm bdy}[\mathcal{B},\phi_0]$ is the generating functional of time-ordered correlators on $\mathcal{B}$. To compute an out-of-time-order correlator, one needs to extend $\mathcal{Z}_{\rm bdy}$ to an appropriate timefold spacetime $\mathcal{F}$ constructed out of $\mathcal{B}$ as described above. Then $\mathcal{Z}_{\rm bdy}[\mathcal{F},\phi_0]$ allows one to compute the desired correlator. The statement of holography in terms of partition functions must account for this. Hence one has to more explicitly rewrite (6.3) as

$$\mathcal{Z}_{\rm bulk}[\mathcal{F},\phi_0]=\mathcal{Z}_{\rm bdy}[\mathcal{F},\phi_0], \tag{6.11}$$

where $\mathcal{Z}_{\rm bulk}[\mathcal{F},\phi_0]$ path-integrates over all metrics with the timefold $\mathcal{F}$ as conformal boundary, and boundary conditions ϕ_0 appropriately distributed on $\mathcal{F}$.

We will call the regions of spacetime separated by the creases the "sheets" of the timefold. For a unitary theory, if x and y are spacelike-separated, $\langle\mathcal{O}(y)\mathcal{O}(x)\rangle$ can be computed on either a trivial timefold (i.e. the original spacetime with no foldings), or a nontrivial timefold with $\mathcal{O}(x)$ on some sheet and $\mathcal{O}(y)$ on some other sheet with the same result. This is just the statement that evolution forward and back – without operator insertions – is the identity. It is not obvious that the corresponding bulk partition functions will agree since we do not know (independent of holography) that the gravitational path integral describes unitary evolution. This leads to a potential ambiguity

in the definition of $\mathcal{Z}_{\rm bulk}$. Since we are trying to understand how a basic property of the boundary theory follows from the bulk path integral, we have to resolve this ambiguity. We adopt a *minimal-timefold* approach, introducing a fold only when needed to represent operators at causally-related points where the later operator appears first in the correlator. When folding minimally, backward evolution just needs to sweep causal diamonds between insertions at non-time-ordered points (including small neighborhoods for spacelike creases) [2].

We now show that with the minimal timefold, both the differentiate and extrapolate dictionaries predict no violation of boundary causality. Consider first the differentiate dictionary. Equation (6.6) tells us how the time-ordered correlators on the boundary are computed from the bulk, but we have so far not included out-of-time-order correlators in this equality. Extending the dictionary to general n-point functions, one has

$$G = \frac{1}{i^n}\frac{\delta}{\delta\phi_0^n(x_n)}\cdots\frac{\delta}{\delta\phi_0^1(x_1)}\left.\frac{\mathcal{Z}_{\rm bulk}[\mathcal{F},\phi_0]}{\mathcal{Z}_{\rm bulk}[\mathcal{F},0]}\right|_{\phi_0=0}, \tag{6.12}$$

where $\mathcal{F}$ is a minimal *boundary* timefold constructed so as to order the operator insertions as given in G (cf. (6.6) when the left-hand side is time-ordered, in which case $\mathcal{F}$ is just $\mathcal{B}$). If x and y are spacelike-separated boundary points, the minimal timefold for a two-point correlator is trivial. Since this is the same for both terms in the commutator, the only difference between them is the order in which one takes the derivatives with respect to ϕ_0. Since these derivatives commute, the commutator obviously vanishes. Note that this is completely independent of whether x and y are causally related with respect to some bulk metrics. All we use is that the integral over all metrics and matter fields is some functional of the boundary conditions ϕ_0 on a trivial timefold. If x and y are

[2] We thank Raphael Bousso for pointing out the role of causal diamonds.

causally related on the boundary, the situation is different. In that case, each term in the commutator requires a different minimal timefold: a trivial one for the time-ordered correlator, and a nontrivial one for the out-of-time-order correlator. In the latter, the source for the earlier operator is moved to the second sheet, so the commutator can be nonzero.

We now turn to the extrapolate dictionary as in (6.7). For a bulk metric relating x and y causally, one would expect $\langle\phi(x,z)\phi(y,z)\rangle_g \neq \langle\phi(y,z)\phi(x,z)\rangle_g$ since they are computed on different bulk timefolds. This nonzero commutator on g is the origin of the causality puzzle. However, there is a subtlety: The boundary condition on g must be the same as for the differentiate dictionary. Namely, one must have a trivial boundary timefold when x and y are spacelike-related on the boundary. If one computes the commutator by a path integral over fields on the bulk metric g, the causality puzzle is resolved as follows. Recall that a field path integral on a trivial timefold always gives the time-ordered correlator. So when the two operators are in the asymptotic region, inside the bulk path integral each of the two terms in the commutator (6.9) should come with a time-ordering with respect to g, $\langle\mathcal{T}\phi(x,z)\phi(y,z)\rangle_g = \langle\mathcal{T}\phi(y,z)\phi(x,z)\rangle_g$, yielding a vanishing commutator. Indeed, if the two points are timelike-related in the bulk, in the limit $z \to 0$ a non-time-ordered correlator would require a bulk metric with a nontrivial timefold asymptotically, violating our boundary condition. In this case, the naive bulk commutator $\langle[\phi(x,z),\phi(y,z)]\rangle_g$ plays no role in computing the CFT commutator $\langle[\mathcal{O}(x),\mathcal{O}(y)]\rangle_{\text{bdy}}$.

Even with no timefold on the boundary, the bulk path integral includes timefolds that trivialize asymptotically [3]. Intuitively, the amount of time that one evolves backward goes to zero as $z \to 0$. This has no effect on the differentiate dictionary, since derivatives only

[3]This is related to the fact that the bulk path integral should impose gravitational constraints, necessitating integration over both positive and negative lapse [163]. We thank Don Marolf for discussion about this.

act on the boundary conditions. However, for the extrapolate dictionary, it introduces an ambiguity in the location of the bulk operators ϕ. If there are different sheets of the bulk timefold, one has to specify which sheet the operator is on, in addition to giving its location on the sheet. If we use (6.7) one expects this ambiguity to have no effect since we take the limit $z \to 0$ for each metric where the timefold becomes trivial.

In contrast, there is a potential problem with taking $z \to 0$ after integration. For any nonzero z, there are bulk metrics where for timelike-separated points (x, z) and (y, z) there is a timefold that is trivial asymptotically, but both the time-ordered and out-of-time-order correlators can be obtained by placing the operators on different sheets [4]. In other words, the bulk contains a timefold that allows a nonzero commutator, but in the region outside the location of the operators, the timefold decays and completely disappears at the boundary. Under these conditions, it would appear that the integrand can be nonzero, and one again is left with the puzzle of why the integral over metrics vanishes exactly.

However, the above occurs only if we allow bulk operators to change sheets depending on their ordering in a correlator. Since the choice of sheet is not fixed by the boundary correlator that we are trying to reproduce, one can adopt the rule that one fixes the ambiguity in the extrapolate dictionary by picking one sheet of a timefold for each operator independent of the location of the operator in the correlator. With this understanding, the integrand remains zero for each g (even before taking $z \to 0$), consistent with boundary causality.

Finally, we extend the ordering prescription for the bulk path integrand to general

[4] We are using the fact that since the bulk timefolds do not have to be minimal, a nonzero commutator can be obtained from a single timefold, by moving the earlier operator from one sheet to another. For example, on the right-hand side of Fig. 6.1, one can compute the commutator by moving $\phi(p)$ to the first sheet.

n-point functions, allowing for timefolds inside the bulk. This upgrades (6.7) to

$$G \propto \int \mathcal{D}g\, e^{iS_g} \lim_{z\to 0} z^{\Delta_\Sigma} \left\langle \mathcal{P}\phi_n(x_n, z) \cdots \phi_1(x_1, z) \right\rangle_g \tag{6.13}$$

and analogously if $z \to 0$ is taken after integration. Here g is restricted to metrics that asymptote to the minimal boundary timefold required to order the field theory correlator correctly and $\mathcal{P}$ is the ordering operator enforced by the field path integral on g, which reduces to the time-ordering operator $\mathcal{T}$ when g is a trivial timefold (cf. the order followed by the arrows in Fig. 6.1).

6.5 Discussion

We have explored whether boundary causality violating metrics in the bulk path integral can lead to possible violations of causality in the dual field theory. We showed that the answer is no, if we choose the minimal timefold on the boundary. This follows from both the differentiate dictionary as given in (6.12) and the extrapolate dictionary in the form (6.7). It also applies to the latter if one integrates first, with a suitable rule for how to place operators on bulk timefolds that are trivial at the boundary. The basic reason for this is that a nonzero commutator for two asymptotic bulk operators requires a bulk timefold that is nontrivial at infinity, but the minimal timefold for two spacelike-separated points on the boundary is trivial. So the bulk dual of the commutator of two field theory operators at spacelike-separated points vanishes, and is not the limit of the naive bulk commutator.

The bulk dual of the commutator of two stress energy tensors on the boundary does not involve bulk matter fields. Nevertheless our argument using the differentiate dictio-

nary still applies. Since the boundary metric is the source for the stress tensor, one should compute the path integral over bulk geometries with a general boundary metric. One then functionally varies the boundary metric at the location of the stress tensors. If they are spacelike-related, there is no timefold on the boundary and hence the commutator vanishes. (There is an issue with even defining the extrapolate dictionary nonperturbatively in this case arising from the apparent need to split the metric degrees of freedom into background and fluctuations.) Our argument also applies to fermionic operators which anticommute at spacelike separations. For the differentiate dictionary, the sign difference comes from the sources for these operators being also fermionic, which makes their variations anticommute.

As explained, off-shell boundary causality violating metrics do not pose a problem to microcausality of the field theory – local operators still commute at spacelike separation. Nonetheless, such metrics do contribute to general field theory correlators. It would be interesting to understand what, if any, are the implications of this quantum gravity effect. We do not believe there are examples of holography in which a bulk classical solution violates boundary causality, but if there are, the boundary theory would have to be either acausal or non-unitary.

Besides microcausality, there are other properties the field theory is expected to have, which in general impose further constraints on the bulk state classically and semiclassically. For example, the invariance of von Neumann entropy under unitary transformations requires that the causal wedge be inside the entanglement wedge [95, 80]. (This is stronger than and implies boundary causality [158].) Another requirement states that, for two causally-related bulk points, boundary regions encoding each of them cannot be totally spacelike [164]. (This extends our discussion from local operators to operators supported in subregions.) As in our discussion, in a nonperturbative gravitational path

integral there will be configurations violating this. It would be interesting to understand how the bulk path integral preserves these properties.

Part III

Factorization

Chapter 7

A tale of two saddles

7.1 Introduction

The contributions of Euclidean wormhole saddles to gravitational systems have been as puzzling as they have been illuminating (see e.g. [165, 166, 167, 168, 169, 24, 49, 50, 170, 171, 172, 173, 174, 175, 176]). On the one hand, in the gravitational replica trick for the von Neumann entropy [35, 37, 49, 50] these wormholes provide an explanation for the quantum extremal surface (QES) formula [38], and consequently for the consistency of semiclassical black hole evaporation with unitarity [47, 48]. On the other hand, the presence of wormholes gives rise to an apparent lack of factorization that raises questions about whether a low-energy description of gravity can be contained in a single theory or must be emergent from an ensemble [165, 166, 167, 169, 24, 170, 173, 172, 174, 177, 171, 122, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190].

Given the crucial role these wormholes play in ensuring unitarity, are there other quantities aside from measures of entropy to which wormholes contribute in the semiclassical regime? This seems likely: we might expect that the imprint of such an important

aspect of gravity would be detectable with more general and simpler observables than entropies. The most fundamental such object in an investigation into the basic import of non-factorizing Euclidean geometries is the generating functional.

How, then, does the factorization problem manifest at the level of the generating functional? An answer to this question would go significantly further in resolving the puzzles raised by the inclusion of replica wormholes in the gravitational path integral than the study of any one individual quantity, be it entropy or any other observable.

This query was initially raised in the context of gravity in [51], which found significant modifications to the behavior of the generating functional of Jackiw-Teitelboim (JT) gravity (at nonperturbatively low temperatures) depending on whether connected topologies were permitted to contribute. In fact, in the absence of replica wormholes the generating functional gives rise to a negative entropy at low temperature, while the inclusion of wormholes in a nonperturbative completion of JT does in fact give rise to a positive thermodynamic entropy all the way to zero temperature [191, 192, 193, 194, 195, 196, 197, 198, 199].

However, these advances offer limited insight in the quest towards an understanding of the role of non-factorization in the emergence of semiclassical gravity. First, the results of [191, 192, 193, 194, 195, 196, 197, 198, 199] are highly nonperturbative and thus do not shed light on the nature and behavior of observables in the *semiclassical* regime. Second, in working with pure JT gravity, we restrict to a theory which is known to feature ensemble averaging, rendering the factorization problem less mysterious than in theories where there is no obvious ensemble. In particular, we would like to understand replica wormholes in higher-dimensional theories of gravity.

In the semiclassical regime, the contribution of replica wormholes to the generating functional can be tractably investigated using a replica trick. This replica trick differs

substantially from the more familiar one for the von Neumann entropy, so we now briefly review it. In a theory of gravity with Euclidean action I defined by a boundary geometry or conformal geometry (B, h) (e.g. this may be an asymptotically AdS boundary), we schematically denote the gravitational path integral by

$$\mathcal{P}(B) = \sum_M \int \mathcal{D}g\, \mathcal{D}\psi\, e^{-I[g,\psi]}, \tag{7.1}$$

where the sum is over manifolds M of different topologies with boundary B, the integral is over metrics g on M inducing (B, h) on their boundary, and ψ represents any additional matter fields[1]. The aforementioned factorization puzzle arises from noting that the contribution of different topologies to $\mathcal{P}(B)$ implies that $\mathcal{P}(B^n) \neq \mathcal{P}(B)^n$, and hence $\mathcal{P}(B)$ cannot be interpreted as the partition function $Z(B)$ of a standard quantum system on B, which would factorize on B^n. Instead, this observation suggests an interpretation of $\mathcal{P}(B)$ as some coarse-graining (possibly an averaging) over quantum mechanical partition functions defined by B:[2]

$$\mathcal{P}(B) = \overline{Z(B)}. \tag{7.2}$$

It prima facie appears that we are simply out of luck in any attempt to directly infer the implications of replica wormholes on the generating functional: a naïve semiclassical computation of the generating functional would just correspond to taking $\ln \mathcal{P}(B) = \ln \overline{Z(B)}$; the gravitational path integral here involves only a single boundary and thus no Euclidean wormholes can be included.

[1] If matter fields are present, their boundary values should also be fixed.

[2] We will be agnostic on the precise protocol that produces $\overline{Z(B)}$ from $Z(B)$, though in certain special cases (such as pure JT gravity) it can be understood precisely; see e.g. [122, 183, 185, 184, 200, 180, 201, 202, 203, 198, 204, 205, 190] for a number of possible interpretations in more general contexts.

However, $\ln Z(B)$ admits an alternative calculation via a replica trick:

$$\ln Z(B) = \lim_{m\to 0} \frac{1}{m} \left(Z(B^m) - 1\right), \tag{7.3}$$

where B^m is the union of m copies of the boundary B. The advantage of this rewriting is that it immediately reveals an alternative semiclassical calculation in which replica wormholes could contribute to the generating functional: using (7.2), an "overline" of (7.3) gives

$$\overline{\ln Z(B)} = \lim_{m\to 0} \frac{1}{m} \left(\mathcal{P}(B^m) - 1\right). \tag{7.4}$$

While at the fine-grained quantum mechanical level the replica trick (7.3) for $\ln Z(B)$ is no different from the direct calculation thereof, the calculations of $\ln \overline{Z(B)}$ and $\overline{\ln Z(B)}$ from the gravitational path integral do not necessarily agree. This discrepancy is easiest to see in the context where $\mathcal{P}(B)$ actually computes an ensemble average: in that case, $\ln \overline{Z(B)}$ – the so-called *annealed* generating functional, which we shall denote by Γ_A – allows the parameters of the ensemble to equilibrate with the dynamical fields, whereas $\overline{\ln Z(B)}$ – the *quenched* generating functional Γ_Q – freezes the ensemble in each computation of the generating functional before averaging. Since the parameters in such an ensemble interpretation (to which we do not necessarily subscribe) are distinguished by the fact that they are not dynamical, it is clearly the latter that is of physical relevance, and hence the question of whether Euclidean wormholes contribute to the semiclassical generating functional amounts to whether saddles with connected topologies exist and dominate over their disconnected counterparts in the $m \to 0$ limit above.

This is what motivates our work in this paper: we look for replica wormholes that dominate over the disconnected topology in a (strictly classical) saddle-point approximation. The natural starting point is the gravitational replica trick technique of Lewkowycz

and Maldacena (LM) [35], which can be applied to explore the analytic continuation to small m in the replica trick (7.4) for the generating functional. Recall that in a saddle-point approximation for $\mathcal{P}(B^m)$, this technique considers on-shell geometries (M_m, g_m) with a $\mathbb{Z}_m$ symmetry. Quotienting by this symmetry produces a single-boundary "quotient" geometry $(\widehat{M}_m, \widehat{g}_m)$ containing a (not necessarily connected) codimension-two conical defect with opening angle $2\pi/m$. This is analytically continued in m away from the integers (while still imposing the equations of motion), giving an on-shell action $\widehat{I}_m$ which can be used to approximate the path integral: $\mathcal{P}(B^m) \approx e^{-m\widehat{I}_m}$. To compute, say, the von Neumann entropy, the limit $m \to 1$ is then taken in the corresponding replica trick; working perturbatively about the $m = 1$ geometry then recovers the QES formula.

In the present context, the appropriate limit is $m \to 0$, which introduces a number of subtleties not present in the computation of the von Neumann or Rényi entropies [35, 206] (in which case m is a positive integer). First, as a consequence of the divergent conical surplus in the $m \to 0$ limit, there appears to be no $m = 0$ geometry about which we can work perturbatively. Second, in contrast with the entropy replica trick, standard examples of the quenched generating functional (e.g. spin glasses) often require replica symmetry breaking (RSB). We may therefore expect the need for incorporating RSB when evaluating the replica trick (7.4) in semiclassical gravity. We will anticipate this and include an algorithmic way of breaking replica symmetry into our protocol for computing Γ_Q; indeed, this will prove to be necessary in the examples discussed below.

Our prescription for computing Γ_Q from the gravitational replica trick is presented in Section 7.2. In short, the procedure consists of two steps:

1. First, we break replica symmetry by allowing for partially connected wormholes where the m boundaries cluster into m_1-boundary wormholes, with m_1 some divisor of m, as shown in Figure 7.1. These wormholes have symmetry group $(\mathbb{Z}_{m_1})^{m/m_1} \times$

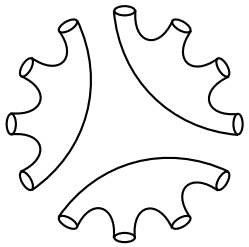


Figure 7.1: An illustration of the kinds of replica symmetry-breaking wormholes we allow in our analysis of the replica trick (7.4). Here $m = 12$ and $m_1 = 4$.

$\mathbb{S}_{m/m_1}$, which naturally interpolates between the $\mathbb{S}_m$ symmetry of the fully-disconnected phase and the $\mathbb{Z}_m$ one of the fully-connected wormhole. Quotienting by this symmetry group gives a single-boundary geometry with conical defects of opening angle $2\pi/m_1$.

2. We then analytically continue m and m_1 away from the integers and use (7.4) to obtain a formula for the quenched generating functional:

$$\Gamma_Q \equiv \overline{\ln Z} \approx - \min_{m_1 \in [0,1]} \widehat{I}_{m_1}. \tag{7.5}$$

In practice, evaluating the action $\widehat{I}_{m_1}$ in order to compute Γ_Q requires solving for the quotient metric $\widehat{g}_m$ explicitly. In general theories of gravity this is a difficult task, but it is simplified substantially in two-dimensional models. Consequently, in exploring the contributions of replica wormholes to Γ_Q we focus our attention on JT gravity and JT gravity coupled to matter, in which $\widehat{g}_m$ describes a two-dimensional geometry of constant negative curvature and hence the geometric degrees of freedom are simply the moduli in the space of such geometries. In the quotient geometry, these moduli reduce to a single degree of freedom that determines the proper distance D between the conical defects. This simplification makes an explicit construction of the quotient geometry tractable, and in Section 7.3 we discuss two methods of constructing it: we obtain $\widehat{g}_m$ either by

patching together appropriate regions of the Poincaré disk, or by solving the Liouville equation in the presence of defects. The former method is very explicit, while the latter is more akin to how we would need to proceed to obtain $\widehat{g}_m$ in higher dimensions.

With the quotient geometry constructed, we can then compute the quotient action $\widehat{I}_m$ in our specific models, which involves solving the equation of motion for the "boundary degree of freedom" (i.e. the Schwarzian mode). This procedure is most straightforward in pure JT, but as is well-known, Euclidean wormholes do not exist as saddles in pure JT due to the tendency of the modulus D to "pinch off" the wormhole throats. Nevertheless, it is possible to investigate the behavior of off-shell "constrained" wormholes in which D is fixed by hand; we perform this investigation in Section 7.4 as an illustrative example that highlights the rather complicated structure of $\widehat{I}_m$ at $m < 1$. In particular, the continuation of $\widehat{I}_m$ to complex m is an infinitely-sheeted Riemann surface, and the equations of motion are crucial for determining which of these sheets gives the correct answer for $\widehat{I}_m$.

In order to stabilize the modulus D to get genuinely classical wormholes, however, we need to support them by coupling JT to some matter. In Section 7.5 we therefore couple JT to a massless scalar field and construct the resulting action $\widehat{I}_m$. In order for the matter field to exhibit a nontrivial stress tensor – as is necessary to stabilize the wormholes – we turn on boundary sources for the matter field; these sources break the $U(1)$ Euclidean time-translation symmetry, so the states that we consider are not thermal states. Nevertheless, the parameter β that sets the length of the Euclidean time circle (i.e. the "inverse temperature") is a tunable boundary condition, and we find that at sufficiently small β^{-1} relative to the strength of our sources the matter is able to support classical wormholes. In particular, these wormholes exist for $m < 1$. Moreover, these new saddles have smaller action than the disconnected saddle that contributes to Γ_A, which

immediately suggests that the quenched generating functional of JT gravity coupled to classical matter has a phase in which it is dominated by replica wormholes. The inclusion of these wormholes results in quantitatively and qualitatively different behavior of the quenched generating functional Γ_Q from its annealed counterpart Γ_A.

Of course, whether or not these new saddles genuinely contribute to the quenched generating functional depends on their stability properties under a given definition of the path integral. We find that while the disconnected saddle is stable against all Euclidean perturbations, the new connected saddles at $m < 1$ (with lower action) are not: they are stable when restricted to perturbations that admit a particular analytic continuation to (real) Lorentzian time, but unstable to arbitrary Euclidean perturbations. Hence in a purely Euclidean treatment that forgets about the Lorentzian origins of the theory, the quenched generating functional appears to reproduce its annealed version; but in a treatment that imposes a real Lorentzian section (or perhaps that rotates the contour of integration in the path integral to an appropriately "Lorentzian" one), the quenched and annealed generating functionals may differ at low temperatures. In fact, these statements appear to be robust under quantum corrections: in Section 7.6, we compute quantum corrections to the matter action perturbatively around $m = 1$ and find that for $m < 1$ these quantum corrections exhibit a stabilizing effect on the wormholes. This is to be contrasted with the situation for $m > 1$, where a Casimir effect actually has a destabilizing influence [207, 208, 209].

These observations naturally prompt questions of how to determine what the "right" saddles to include in the quenched generating functional are; while we do not answer this question, we discuss some interesting facets and avenues of exploration in Section 7.7.

7.2 The Replica Trick for the Generating Functional

Under the interpretation (7.2) of the gravitational path integral $\mathcal{P}(B^m)$, the replica trick (7.4) is a trivial identity; its nontrivial content arises from how $\mathcal{P}(B^m)$, which is only well-defined for integer m, is to be continued in m to a neighborhood of $m = 0$. In general this analytic continuation is not unique and must be treated carefully. For instance, if $\mathcal{P}(B^m)$ exhibits appropriate behavior in the right-half complex m plane, Carlson's theorem can be used to ensure a unique analytic continuation; or if the "ensemble average" genuinely corresponds to an average over an appropriate distribution of partition functions, one could try to use Carleman's condition in a similar way. However, these approaches are often insufficient to provide a unique continuation to $m = 0$; see e.g. [186] for a discussion of some of the difficulties involved.

Instead, a fruitful approach is to work in a saddle-point approximation wherein the path integral $\mathcal{P}(B^m)$ is dominated by a single geometry obeying the saddle-point equations: that is, the classical equations of motion. By continuing the equations of motion to non-integer m in a controlled way, one obtains a unique analytic continuation of $\mathcal{P}(B^m)$ to non-integer m which typically captures the correct physics. In condensed matter contexts, the replica trick has been used in this way for decades; see e.g. [210]. In the gravitational context, this analytic continuation of the equations of motion amounts to the Lewkowycz-Maldacena (LM) construction for computing the von Neumann entropy gravitationally [35] with some important differences that we now describe.

7.2.1 The Saddle-Point Approximation

Let us review the gravitational replica trick of LM, but adapted to (7.4) rather than to the von Neumann entropy. The calculation is done in a saddle-point approximation,

where the path integral $\mathcal{P}(B)$ is approximated by the on-shell action of a classical solution:

$$\mathcal{P}(B) \approx e^{-I[g^{\text{clas}}]}, \tag{7.6}$$

where g_{clas} solves the classical equations of motion (and we have suppressed any matter fields ψ – they are treated classically in the same way as gravity, or semiclassically by including the one-loop effective matter action to the above). Now, when B consists of several disconnected regions, as in the replica trick (7.4), we should consider all possible topologies of bulk manifold M and find the configuration with smallest action. In the case of $\mathcal{P}(B^m)$, the m boundaries exhibit an $\mathbb{S}_m$ permutation symmetry. A common assumption is that the on-shell bulk solution that approximates $\mathcal{P}(B^m)$ is highly symmetric as well. We will review the construction of the gravitational replica with this in mind and leave all discussion of further replica symmetry breaking – which is crucial for our construction – to the subsequent section.

Let us first consider the maximally symmetric saddles, which are the disconnected solutions. In these saddles, the bulk manifold M consists of m disconnected pieces that "fill in" each boundary B. In this case the on-shell action is just $mI[g_{m=1}]$, where $g_{m=1}$ is the on-shell metric defined by a single boundary. The analytic continuation in m is then trivial, and gives a contribution to the quenched generating functional of

$$\Gamma_Q \equiv \overline{\ln Z} \supset -I[g_{m=1}] = \ln \overline{Z} \equiv \Gamma_A, \tag{7.7}$$

where we identify the average of the partition function as $\overline{Z} \approx e^{-I[g_{m=1}]}$. This contribution is just the annealed generating functional, and is the standard way of computing e.g. free energies in gravitational theories.

The second type of saddle that is typically considered in this context is the wormhole

solution that connects various copies of B in an arrangement exhibiting a $\mathbb{Z}_m$ symmetry, as shown in the left diagram of Figure 7.2. Denoting the metric on the fully-connected manifold as g_m, the on-shell action is simply $I[g_m]$. In order to continue this on-shell action away from integer m, the wormhole geometry is quotiented by the $\mathbb{Z}_m$ symmetry, yielding a quotient manifold $\widehat{M}_m$ with metric $\widehat{g}_m$; the codimension-two surfaces of fixed points of the $\mathbb{Z}_m$ isometry in the full geometry become conical defects in $(\widehat{M}_m, \widehat{g}_m)$ with opening angle $2\pi/m$. All m-dependence appears only in the opening angle about these defects, and so m can be sensibly continued away from the integers while still imposing the equations of motion. Note that $\widehat{g}_m$ does not solve the equations of motion obtained from the action $I[g]$ at the defects; one can either impose the equations of motion everywhere away from the defects, or modify the action by the contribution of a cosmic brane with tension proportional to $1-1/m$ sourcing the defects, as in [206]. Either way, the resulting contribution to the quenched generating functional from this replica-symmetric saddle would be

$$\Gamma_Q \supset \lim_{m\to 0} \frac{1}{m}\left(e^{-mI[\widehat{g}_m]} - 1\right) = -\lim_{m\to 0} \widehat{I}_m, \tag{7.8}$$

where we have introduced the shorthand notation $\widehat{I}_m \equiv I[\widehat{g}_m]$. Note that we have left the $m \to 0$ limit explicit because it is not clear whether $\widehat{g}_0$ is a well-defined geometry. In particular, while for $m > 1$ the conical defect is an angular deficit, for $m < 1$ the defect is an excess, and in fact as $m \to 0$ the excess angle becomes arbitrarily large[3].

The contribution (7.8) to Γ_Q captures the effect of replica wormholes in the path integral. If this contribution is subdominant to the conventional one (7.7) from the disconnected topology, then Γ_Q simply reproduces the annealed calculationΓ_A. But if wormholes dominate over the disconnected contribution,Γ_Q and Γ_A may differ substan-

[3]From the cosmological brane point of view, for $m > 1$ the brane tension $T_m \propto (1 - 1/m)$ is positive, making it gravitationally attractive; for $m < 1$ the brane tension is negative, making it gravitationally repulsive.

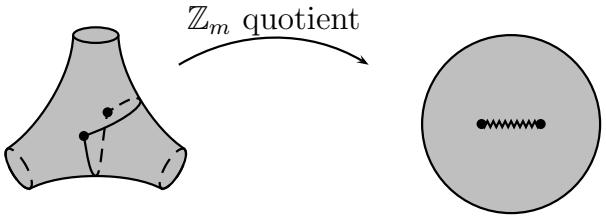


Figure 7.2: The LM construction used to obtain the analytic continuation in m for the replica trick (7.4). Upon quotienting by $\mathbb{Z}_m$, the two black lines in the left diagram are stitched together at the branch cut in the right diagram.

tially.

It is worth remarking on the differences between this procedure and the analogous one for the von Neumann entropy. The first difference is clearly the number of replicas: in the derivation of the RT/HRT formula for von Neumann entropy [32, 36], one ultimately takes the $m \to 1$ limit of Rényi entropies. This limit allows us to work perturbatively around the "original geometry" $m = 1$ to derive a formula for the von Neumann entropy that does not require an explicit construction of the quotient geometry $\widehat{g}_m$. On the other hand, our $m \to 0$ limit for Γ_Q genuinely requires a computation $\widehat{g}_m$ well away from $m = 1$. This is more akin to the gravitational computation of Rényi entropies [206], which requires one to compute $\widehat{I}_m$ at integer $m > 1$. The second difference is that in computations of von Neumann (or Rényi) entropies in gravitational theories, correlations in the boundary conditions of the gravitational path integral render the symmetry group of the boundary to be $\mathbb{Z}_m$ (e.g. in the derivation of the RT formula, B consists of an m-sheeted branched cover whose branch points correspond to the entangling surface). Hence it is quite natural to take the infilling bulk solution to share this $\mathbb{Z}_m$ symmetry as well. On the other hand, in the replica trick (7.4), the boundary B^m consists of m *completely disconnected* pieces with symmetry group $\mathbb{S}_m$. This symmetry is broken to $\mathbb{Z}_m$ by the wormhole. In general we expect that connected geometries that preserve the full $\mathbb{S}_m$ symmetry of the boundary do not exist, so in this sense we may think of the wormhole as

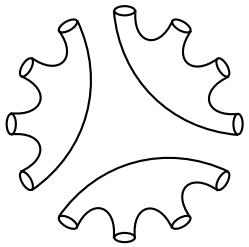


Figure 7.3: An illustration of our replica symmetry-breaking wormholes; here $m = 12$ and $m_1 = 4$, giving a wormhole with $m/m_1 = 3$ connected components.

a kind of mild form of replica symmetry breaking (RSB). But if some amount of breaking the $\mathbb{S}_m$ of the boundaries is inherent to the wormholes, it is natural to wonder whether we can proceed further by generalizing this breaking in a controlled way. Indeed it can, and incorporating such RSB will be crucial to our later analysis. We now briefly discuss a form of "one-step gravitational RSB" motivated by the structure of the Parisi ansatz for spin glasses [210].

7.2.2 Replica Symmetry Breaking

The one-step RSB procedure that we define incorporates wormholes that are not maximally connected[4]. For a given integer m, we take m_1 to be a positive integer that divides m. We then consider wormholes that connect the m boundaries into groups of m_1, as illustrated in Figure 7.3 for the case $m = 12$, $m_1 = 4$. Hence m_1 is a parameter that encodes various possible wormhole topologies, and so we should ultimately minimize the action with respect to it to find the dominant contribution.

The bulk geometry consists of m/m_1 disconnected pieces. Assuming that each of these pieces has the same $\mathbb{Z}_{m_1}$ symmetry discussed above, the symmetry group of the bulk geometry is $(\mathbb{Z}_{m_1})^{m/m_1} \times \mathbb{S}_{m/m_1}$, with the $\mathbb{S}_{m/m_1}$ factor corresponding to the freedom

[4]The contributions of various topologies of off-shell wormholes to gravitational computations of von Neumann entropy were considered in e.g. [50].

to permute the disconnected pieces amongst each other. We may quotient the geometry by this symmetry to conclude that the total on-shell action is $m\widehat{I}_{m_1}$. We may now analytically continue m away from the integers. Because m_1 was constrained to be a divisor of m, continuing in m naturally leads us to continue in m_1 as well. However, since m_1 was constrained to range between one and m, we preserve this constraint even after the analytic continuation: that is, we take m_1 to be an arbitrary real number in the range $m_1 \in [1, m]$. Hence for arbitrary m, the path integral is approximated by

$$\mathcal{P}(B^m) \approx \max_{m_1 \in [1,m]} e^{-m\widehat{I}_{m_1}}, \tag{7.9}$$

and hence the replica trick (7.4) gives

$$\Gamma_Q \approx - \min_{m_1 \in [0,1]} \widehat{I}_{m_1}. \tag{7.10}$$

Note that in taking the limit $m \to 0$, we have maintained the constraint that $m_1 \in [1, m] \to [0, 1]$. If the minimization is achieved by $m_1 = 1$, then Γ_Q coincides with the annealed result Γ_A in (7.7); if the minimization is achieved by $m_1 = 0$ (or perhaps more carefully, in the limit $m_1 \to 0$), then we recover the quenched generating functional (7.8) obtained from the completely connected wormholes. We interpret any other value of m_1 as breaking replica symmetry.

Some comments are in order. First, though this procedure seems ad hoc (in particular, the analytic continuation of m_1 and its restriction to lie in the interval $[1, m]$ even after continuing m to zero), it is precisely analogous to one-step RSB in spin glasses. As with most instances of the replica trick, ultimately we should interpret (7.10) as a *prescription* for computing the quenched generating functional, rather than as a derivation. Its validity can only be determined on physical grounds. Second, values of m_1 different from m

actually correspond to a *larger* symmetry group than that of the fully-connected wormhole $m_1 = m$, so what we are calling "replica symmetry breaking" can really be thought of as a "symmetry enhancement" relative to the fully-connected wormhole. Regardless, any value of m_1 different from 1 still corresponds to a breaking of the $\mathbb{S}_m$ symmetry of the boundaries. Finally, there is a potential subtlety: the action functional could formally change signs for sufficiently small m. In that situation, presumably we would need to *maximize* the action instead of minimizing it. This type of behavior is characteristic of spin glasses, where the number of degrees of freedom formally becomes negative in the $m \to 0$ limit[5]. Whether or not an analogous change of sign happens in the gravitational contexts we are considering will presumably depend on the details of the gravitational theory. At least for the theories we will consider later in this paper, this phenomenon does not happen and we expect to need to minimize over m_1.

As a final note, in many cases of interest one takes B to have topology $\Sigma \times S^1$, where the circle S^1 has length β. If all boundary sources exhibit a $U(1)$ symmetry corresponding to translations around the S^1, $Z(B)$ can be interpreted as a canonical partition function at inverse temperature β, and the quenched generating functional is related to the quenched free energy:

$$F_Q \equiv -\beta^{-1}\Gamma_Q. \tag{7.11}$$

However, in later sections we will consider boundary sources that break the $U(1)$ isometry

[5]More precisely, in a spin glass the degrees of freedom the action is to be extremized over are the off-diagonal components of an $m \times m$ correlation matrix $q_{\alpha\beta}$ encoding correlations between replicas. The action takes the form of a sum over the components of $q_{\alpha\beta}$, and the number of off-diagonal components is $m(m-1)/2$. When analytically continuing in m, one sets the off-diagonal components of $q_{\alpha\beta}$ to all be equal to some q, giving the action an explicit overall factor of $m(m-1)/2$, which becomes negative when $m < 1$. Due to this change in the overall sign of the action, the physical minimization of the action with respect to the individual components of $q_{\alpha\beta}$ corresponds to the *maximization* of the action with respect to the degree of freedom q in the replica ansatz once m is taken less than one.

of the boundary thermal circle. In such a case F_Q no longer has an interpretation as the free energy of a thermal state, and for this reason we restrict our investigation to the generating functional Γ_Q.

7.3 The Quotient Geometry in Two Dimensions

In certain two-dimensional asymptotically (nearly) AdS models of gravity like JT, the geometry has constant negative curvature. Moreover, in a classical (or semiclasscal) limit, contributions from higher genera are suppressed. These simplfications render the construction of the quotient geometry $(\widehat{M}_m, \widehat{g}_m)$ quite tractable, and we now describe it in this context. In later sections we will rely on this construction to investigate the behavior of the $m \to 0$ limit invoked in the computation of the quenched generating functional in pure JT gravity and in JT gravity coupled to matter.

It will be useful to describe two different ways of obtaining the quotient geometry. The first exploits the fact that since $\widehat{g}_m$ has constant negative curvature, it is locally AdS_2; hence $\widehat{g}_m$ can be constructed by stitching together appropriate regions of exact AdS_2, i.e. patches of the Poincaré disk. This approach has the advantage of giving an exact and explicit form of $\widehat{g}_m$, but at the cost of requiring more than one coordinate chart to cover the entire quotient manifold. On the other hand, the second approach is analogous to the procedure in higher dimensions: we solve the equations of motion directly. In this case we are only able to obtain an approximate solution for $\widehat{g}_m$, but we are able to cover the entire quotient manifold with just one coordinate chart. As we will see, the two constructions are useful in different contexts.

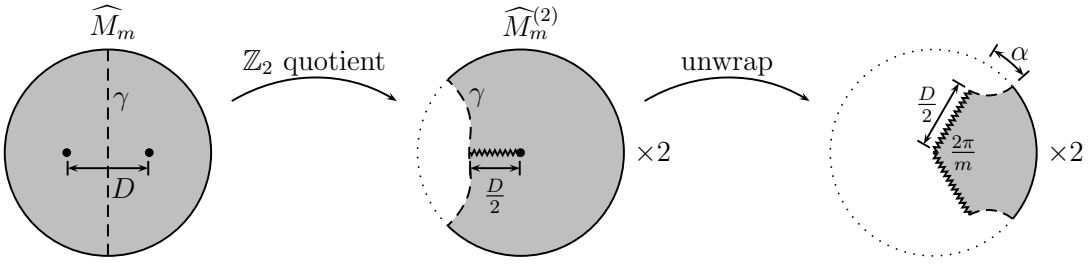


Figure 7.4: In two dimensions, the quotient space $(\widehat{M}_m, \widehat{g}_m)$, shown left, can be mapped to the unit disk in the complex plane with two conical defects on the real axis separated by a proper distance D. The imaginary axis is a geodesic γ about which the geometry exhibits a reflection symmetry. Quotienting by this symmetry gives the geometry $\widehat{M}_m^{(2)}$ which contains only a single defect; it corresponds to a portion of the Poincaré disk with a single defect bounded by a geodesic γ, shown center. Unwrapping the defect by cutting along the jagged line turns $\widehat{M}_m^{(2)}$ into a subregion of the Poincaré disk (with no defect), shown right. The geodesic γ consists of two segments traversing an angle α, related to D by (7.14).

7.3.1 Patchwise Construction

In two dimensions, the quotient manifold has the topology of a disk with two conical defects. Per the replica trick, the locations of these defects in the geometry should be fixed dynamically. However, all but one degree of freedom in the locations of the defects can be gauge-fixed: if we think of $\widehat{M}_m$ as a subset of the complex plane with the disk topology, the automorphisms of $\widehat{M}_m$ can be used to fix three (real) degrees of freedom in the locations of the defects. It is convenient to thus take $\widehat{M}_m$ to be the unit disk in the complex plane and to place the defects on the real axis symmetrically about the origin, leaving the proper distance D between them as the single dynamical degree of freedom, as shown on the left of Figure 7.4. Note that in this construction, the imaginary axis corresponds to a geodesic γ about which the quotient geometry exhibits a $\mathbb{Z}_2$ reflection symmetry. There is an additional $\mathbb{Z}_2$ reflection symmetry about the real axis, so we expect the quotient geometry to exhibit a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry.

Due to the reflection symmetry across γ, we may construct the quotient geometry

by stitching together the left and right halves of the disk along γ; each of these halves corresponds to a geometry $(\widehat{M}_m^{(2)}, \widehat{g}_m^{(2)})$ that contains only a *single* conical defect and is bounded by a geodesic, as shown in the middle of Figure 7.4. This geometry is simply a portion of the Poincaré disk with a defect. We can then "unwrap" the defect by cutting $\widehat{M}_m^{(2)}$ along the indicated jagged line, ending up with the wedge-shaped region of the Poincaré disk (with *no* defect) shown at right in Figure 7.4[6]. Concretely, in complex coordinates in which the metric on the Poincaré disk takes the form

$$ds^2 = \frac{4\, dz\, d\bar{z}}{(1 - z\bar{z})^2}, \tag{7.12}$$

the two jagged lines are the rays $\mathrm{Arg}(z) = \pm\pi/m$, which are to be identified. Likewise, the geodesic γ consists of two segments of the circles defined by

$$\left| z - e^{\pm i\pi/m} \sec\alpha \right| = \tan\alpha, \tag{7.13}$$

with α a parameter that sets the size of these circles, related to the proper distance D between the defects as

$$\sin\alpha = \mathrm{sech}(D/2). \tag{7.14}$$

Note that for $m > 2$, this construction implies that D is bounded from below: in order for the geodesics that define γ to neither intersect one another nor exclude the origin, we

[6] An alternative way of deriving this construction is to switch the order of the quotients: start with the original m-boundary geometry (M_m, g_m) for integer m and first quotient by the $\mathbb{Z}_2$ symmetry that maps the two fixed points of the replica symmetry to one another. The resulting geometry is the Poincaré disk with m geodesic boundaries corresponding to the fixed points of the $\mathbb{Z}_2$, identical to the left diagram of Figure D.3 in the Appendix. Then quotienting by $\mathbb{Z}_m$ immediately gives $\widehat{M}_m^{(2)}$ in the form of the wedge shown at right of Figure 7.4.

must have

$$0 < \alpha < \min\left\{\frac{\pi}{2}, \frac{\pi}{m}\right\}. \tag{7.15}$$

So the minimum proper distance between the defects is $D_{\min} = 2\,\mathrm{arcsech}(\sin(\pi/m))$ when $m > 2$. More intuitively, for $m > 2$ the mass of the defect in the middle diagram of Figure 7.4 becomes sufficiently large that even when the geodesic γ orbits once around it, it is only able to reach a closest approach distance of $D_{\min}/2$; achieving $D < D_{\min}$ would require γ to self-intersect.

It is worth noting that there is a qualitative difference in this construction between $m \geq 1$ and $m < 1$: for $m \geq 1$, the opening angle $2\pi/m$ about the conical defects is less than (or equal to) 2π, so $\widehat{M}_m^{(2)}$ really is a subregion of the Poincaré disk, as shown in the left diagram of Figure 7.5. On the other hand, for $m < 1$ the opening angle $2\pi/m$ is greater than 2π, so we must instead interpret $\widehat{M}_m^{(2)}$ as being a subregion of a *covering* of the Poincaré disk – in terms of the standard angular coordinate θ on the disk defined by $z = re^{i\theta}$, we no longer identify θ with $\theta + 2\pi$. This turns the disk into a Riemann surface with infinitely many sheets, as shown in the right diagram of Figure 7.5. In particular, $\widehat{M}_m^{(2)}$ includes arbitrarily many sheets of this Riemann surface as $m \to 0$.

The upshot is that this patchwise construction of $(\widehat{M}_m, \widehat{g}_m)$ is very convenient in models that can be reduced to local boundary dynamics. In such cases we can use the near-boundary behavior of $\widehat{g}_m^{(2)}$ to construct a local boundary action, and then simply impose appropriate boundary conditions at the (boundaries of the) geodesic γ at which the two copies of $\widehat{M}_m^{(2)}$ are stitched together. We will use such a construction in Section 7.4 to study pure JT, as well as in Appendix D.4 to study a model of JT coupled to end-of-the-world branes.

However, in models that cannot be reduced to local boundary dynamics, as in the

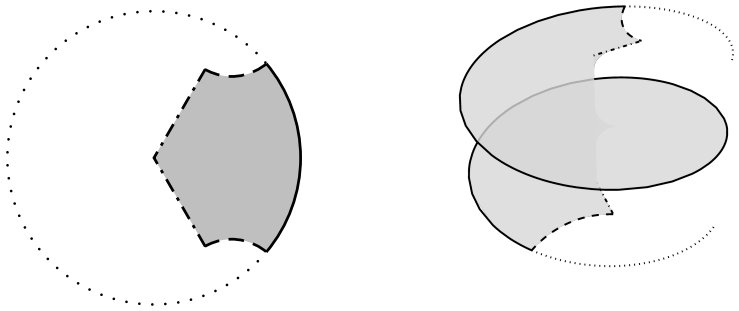


Figure 7.5: For $m \geq 1$, the opening angle about the conical defects is smaller than 2π, so the manifold $\widehat{M}_m^{(2)}$ used in the construction of the quotion geometry consists of the wedge-shaped subregion of the Poincaré disk shown at left (with the dot-dashed lines identified). But for $m < 1$, the opening angle about the defect is greater than 2π, so to construct $\widehat{M}_m^{(2)}$ we must "unwrap" the disk into an infinitely-sheeted Riemann surface. $\widehat{M}_m^{(2)}$ is then a subregion of this Riemann surface, shown at right.

case of JT coupled to a massless scalar that we study in Section 7.5, the need to impose nontrivial boundary conditions *everywhere* along the stitching geodesic γ makes this patchwise construction less useful. For such cases, it is instead desirable to construct the quotient geometry $\widehat{M}_m$ in a *single* coordinate chart. We turn to such a construction next.

7.3.2 Liouville Construction

To construct the quotient geometry in a single coordinate chart, we take $\widehat{M}_m$ to be a subregion Ω of the complex plane with disk topology, with two points z_1, z_2 chosen to be the locations of the conical defects. As remarked above, there is ample gauge freedom in these choices which we must fix. In particular, we are free to conformally map Ω to any other region of the complex plane with disk topology, as well as to change the locations of z_1 and z_2 using automorphisms of Ω. Since such automorphisms contain three real degrees of freedom, the choice of Ω, z_1, and z_2 contains only a single real physical degree of freedom. For instance, as in the left diagram of Figure 7.4 we could take Ω to be the unit disk and set $z_1 = -a$, $z_2 = a$ with $a \in (0, 1)$ the physical modulus that controls

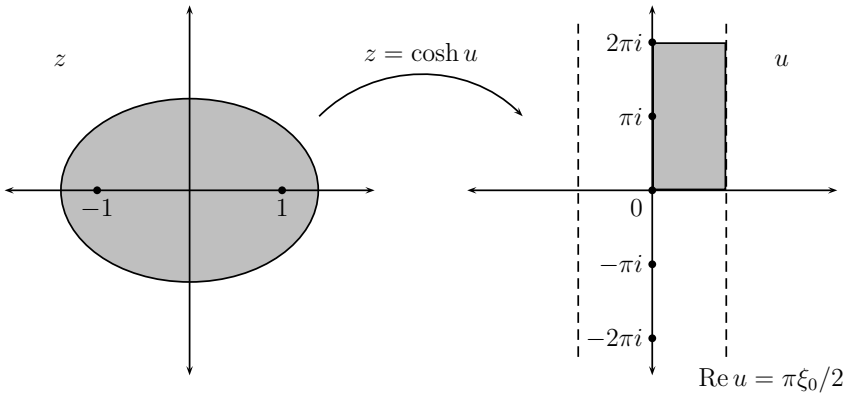


Figure 7.6: The elliptical domain used in constructing the quotient geometry (left) and the rectangle to which we map it (right); the marked points are the locations of the defects, placed at the foci $z = \pm 1$ of the ellipse. The rectangle can be extended to an infinite strip (dashed lines) containing infinitely many images of these defects.

the proper distance between the defects. For our purposes, it is instead convenient to take Ω to be the interior of an ellipse with eccentricity ϵ with foci at $z = \pm 1$ and to place z_1 and z_2 at these foci, as shown in the left diagram of Figure 7.6. In this case the physical degree of freedom is $\epsilon \in (0, 1)$, with ϵ near 0 and 1 corresponding to the defects being close together and far apart, respectively. Also note that this choice of Ω naturally manifests the expected $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry of the quotient geometry through the reflection symmetries across the principal axes of the ellipse.

The advantage of this choice of Ω is that it can easily be mapped to a coordinate rectangle by converting to elliptic coordinates (ξ, ϕ) through

$$z = \cosh u, \qquad u \equiv \xi + i\phi. \tag{7.16}$$

In the original z plane, curves of constant ξ correspond to confocal ellipses of eccentricity $\operatorname{sech} \xi$ with foci at $z = \pm 1$, while curves of constant ϕ correspond to confocal hyperbolae with foci at $z = \pm 1$. Hence the ellipse Ω is mapped to the rectangular region $\phi \in [0, 2\pi)$, $\xi \in [0, \operatorname{arcsech} \epsilon)$ and the defects are mapped to $u = 0$ and $i\pi$, as shown

in the right diagram of Figure 7.6. In what follows we will define $\xi_0 \equiv (2/\pi)\,\mathrm{arcsech}\,\epsilon$, so that the conformal boundary is at $\xi = \pi\xi_0/2$. Note that the map to the rectangle in the u plane doubles the angle around the defects, so if the total angle around them was $2\pi/m$ in the original ellipse, it is $4\pi/m$ in the rectangle.

With the quotient manifold thus fixed, we can solve for the quotient metric $\widehat{g}_m$ on it. We work in conformal gauge in which the metric on the rectangle takes the form

$$ds^2 = e^{2\sigma(u,\bar{u})}\,du\,d\bar{u}. \tag{7.17}$$

Since $\widehat{g}_m$ has constant curvature $R = -2$ everywhere except at the defects, the equation of motion for σ is the Liouville equation [49]

$$-4\partial_u\partial_{\bar{u}}\sigma + e^{2\sigma} = 2\pi\left(1 - \frac{2}{m}\right)\left(\delta_0(u,\bar{u}) + \delta_{i\pi}(u,\bar{u})\right), \tag{7.18}$$

where we define the complex delta function as $\delta_{u_*}(u,\bar{u}) \equiv \delta(\xi - \xi_*)\delta(\phi - \phi_*)$ (with $u \equiv \xi + i\phi$ and $u_* \equiv \xi_* + i\phi_*$ as above). The delta functions on the right can be thought of as contributions to the curvature localized to the defects; note that their strength is proportional to $1 - 2/m$ due to the doubling of the angle in going to the rectangular domain. As shown in Appendix D.1, extending the problem to the entire strip $\xi \in (-\pi\xi_0/2, \pi\xi_0/2)$, $\phi \in (-\infty, \infty)$ and defining $\tilde{\sigma}$ via

$$\sigma \equiv \tilde{\sigma} + \ln\left(\frac{1}{\xi_0}\sec\left(\frac{u+\bar{u}}{2\xi_0}\right)\right) + \frac{H}{\nu}, \tag{7.19a}$$

$$\text{where } H \equiv \sum_{k=-\infty}^{\infty} \ln\left|\tan\left(\frac{u - ik\pi}{2\xi_0}\right)\right| \text{ and } \nu \equiv \frac{m}{2-m}, \tag{7.19b}$$

the Liouville equation becomes

$$-\left(\partial_\xi^2+\partial_\phi^2\right)\tilde{\sigma}+\frac{1}{\xi_0^2}\sec^2\left(\frac{\xi}{\xi_0}\right)\left(e^{2H/\nu}e^{2\tilde{\sigma}}-1\right)=0, \tag{7.20a}$$

$$\tilde{\sigma}\left(\xi=\pm\frac{\pi\xi_0}{2},\phi\right)=0, \qquad \tilde{\sigma}(\xi,\phi+2\pi)=\tilde{\sigma}(\xi,\phi). \tag{7.20b}$$

Although we are unable to solve for $\tilde{\sigma}$ analytically, in practice it is straightforward to obtain $\tilde{\sigma}$ numerically as discussed in Appendix D.2.1, and indeed we will make use of these numerical solutions later. However, it *is* possible to construct an analytic approximation for $\tilde{\sigma}$ which we will use extensively in Section 7.5.

To obtain this approximate solution, we will work at small ξ_0 and small ν – in fact, it will be convenient to take the relative scaling of ν and ξ_0 to be such that $e^{-\pi/2\xi_0}\ll\nu$. If ξ_0 is small, the sum defining H is rapidly convergent since the k^{th} term is sharply peaked around $\phi=k\pi$. Although (7.20a) is not linear in H or $\tilde{\sigma}$, we consequently expect it to approximately linearize. This motivates us to define

$$\tilde{\sigma}(\xi,\phi)\equiv\sum_{k=-\infty}^{\infty}f(\xi,\phi-k\pi)+\delta\tilde{\sigma}(\xi,\phi), \tag{7.21}$$

where f is a solution to

$$-\left(\partial_\xi^2+\partial_\phi^2\right)f+\frac{1}{\xi_0^2}\sec^2\left(\frac{\xi}{\xi_0}\right)\left(\left|\tan\left(\frac{\xi+i\phi}{2\xi_0}\right)\right|^{2/\nu}e^{2f}-1\right)=0, \tag{7.22a}$$

$$f(\pm\pi\xi_0/2,\phi)=0, \quad f(\xi,\phi\to\pm\infty)=0 \tag{7.22b}$$

and $\delta\tilde{\sigma}$ is a correction term which must vanish at $\xi=\pm\pi\xi_0/2$. Just as with each term in H, we expect f to be sharply localized around $\phi=0$ at small ξ_0.

If f localizes around $\phi=0$, then $\delta\tilde{\sigma}$ must be small. To see this, note that (7.20a)

becomes

$$-(\partial_\xi^2+\partial_\phi^2)\delta\tilde{\sigma}+\frac{1}{\xi_0^2}\sec^2\left(\frac{\xi}{\xi_0}\right)\left\{\exp\left[2\sum_{k=-\infty}^{\infty}F(\xi,\phi-k\pi)\right]e^{2\delta\tilde{\sigma}}-1 - \sum_{k=-\infty}^{\infty}\left(e^{2F(\xi,\phi-k\pi)}-1\right)\right\}=0, \quad (7.23)$$

where we have defined

$$F(\xi,\phi)\equiv\frac{1}{\nu}\ln\left|\tan\left(\frac{\xi+i\phi}{2\xi_0}\right)\right|+f(\xi,\phi). \quad (7.24)$$

Per our expectations on f, F should also be sharply peaked around $\phi=0$ when ξ_0 is small. Though (7.23) is an exact rewriting of the Liouville equation, in order to estimate the magnitude of $\delta\tilde{\sigma}$ we may linearize in terms that are small. Without loss of generality we restrict our attention to the region $\phi\in(-\pi/2,\pi/2]$, since we can extend the solution to the entire strip by symmetry about $\phi=k\pi/2$. In this region, $F(\xi,\phi-k\pi)$ is small for all $k\neq 0$, and hence (7.23) gives

$$\xi_0^2\cos^2\left(\frac{\xi}{\xi_0}\right)\left(\partial_\xi^2+\partial_\phi^2\right)\delta\tilde{\sigma}-e^{2F(\xi,\phi)}\left(e^{2\delta\tilde{\sigma}}-1\right) = 2\left(e^{2F(\xi,\phi)+2\delta\tilde{\sigma}}-1\right)\sum_{k\neq 0}F(\xi,\phi-k\pi)+\cdots, \quad (7.25)$$

where the dots denote subleading terms in $F(\xi,\phi-k\pi)$ for $k\neq 0$. In the region $\phi\in(-\pi/2,\pi/2]$, the right-hand side is small; if it were to vanish, then clearly so would $\delta\tilde{\sigma}$.

Hence $\delta\tilde{\sigma}$ must be small, as claimed; linearizing in it, we find

$$\left[\xi_0^2 \cos^2\left(\frac{\xi}{\xi_0}\right)\left(\partial_\xi^2+\partial_\phi^2\right) - 2e^{2F(\xi,\phi)}\right]\delta\tilde{\sigma} \\ = 2\left(e^{2F(\xi,\phi)}-1\right)\sum_{k\neq 0} F(\xi,\phi-k\pi)+\cdots, \quad (7.26)$$

where the dots now also include subleading terms in $\delta\tilde{\sigma}$. Consequently we expect the magnitude of $\delta\tilde{\sigma}$ to scale like the magnitude of the source on the right-hand side.

All that remains is to solve (7.22) for f to construct the solution (7.21) and to quantify the size of $\delta\tilde{\sigma}$ (as well as to verify the expectations that led to (7.26) in the first place). At small ν, f can be constructed perturbatively in ν using a matched asymptotic expansion, as shown in Appendix D.1.1. We ultimately find that

$$f = -\ln\left[\left(\frac{1}{y}+\nu\right)\sinh y\right] + y + \nu\left(1+y\left(1-\coth y\right)\right) \\ + \frac{\nu^2}{6}\left(3y^2\operatorname{csch}^2 y + 2y^3(1-\coth y) - 3\right) \\ + \frac{\nu^3}{3}\left(1+y^3(1-\coth^3 y)+y^4\operatorname{csch}^2 y\right)+\mathcal{O}(\nu^4), \quad (7.27)$$

where $y \equiv \nu^{-1}\operatorname{sech}(\phi/\xi_0)\cos(\xi/\xi_0)$ and the neglected terms are $\mathcal{O}(\nu^4)$ for all $y \in (0, 1/\nu)$. It then follows that if $e^{-\pi/2\xi_0} \ll \nu$, the right-hand side of (7.26) – and hence also $\delta\tilde{\sigma}$ – is $\mathcal{O}(\nu^{-4}e^{-2\pi/\xi_0})$. We thus conclude that

$$\tilde{\sigma}_{\text{approx}}(\xi,\phi) = \sum_{k=-\infty}^{\infty} f(\xi,\phi-k\pi) \quad (7.28)$$

with f given by (7.27)[7] is a solution to the Liouville equation up to corrections of or-

[7] In principle we can obtain f to any desired order in ν, but for constructing (7.28) there is no point computing f to higher order than the correction term $\delta\tilde{\sigma}$, i.e. to $\mathcal{O}(\nu^{-4}e^{-2\pi/\xi_0})$ or higher.

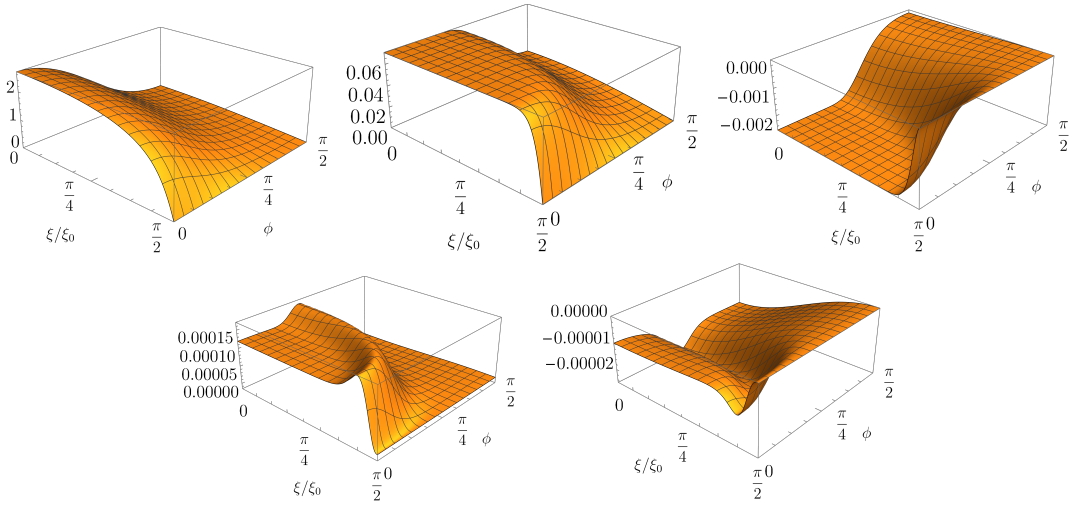


Figure 7.7: A comparison between the approximation $\tilde{\sigma}_{\text{approx}}$ and $\tilde{\sigma}_{\text{num}}$. The top left shows $\tilde{\sigma}_{\text{num}}$, while from left to right and top to bottom the subsequent plots show $\tilde{\sigma}_{\text{num}} - \tilde{\sigma}_{\text{approx}}$ with successively higher-order terms in ν included in (7.27). Here we take $\xi_0 = 0.3$ and $\nu = e^{-\pi/4\xi_0} \approx 0.07$, so that $\tilde{\sigma}_{\text{approx}}$ should agree with $\tilde{\sigma}_{\text{num}}$ up to corrections of order $\mathcal{O}(\nu^{-4}e^{-2\pi/\xi_0}) = \mathcal{O}(\nu^4)$.

der $\mathcal{O}(\nu^{-4}e^{-2\pi/\xi_0})$. To verify the validity of this approximation, in Figure 7.7 we compare $\tilde{\sigma}_{\text{approx}}$ to the numerical solution $\tilde{\sigma}_{\text{num}}$ of (7.20a) for $\nu = e^{-\pi/4\xi_0}$ and $\xi_0 = 0.3$, in which case we should expect these solutions to agree to order $\mathcal{O}(\nu^3)$, as we find that they do.

7.4 Pure JT for General m

Section 7.3.1 described how to construct the two-dimensional quotient geometry of the replica trick by identifying two copies of the Poincaré disk with defects, yielding a geometry that depends on m and on the distance D between the defects. This construction will allow us to compute the effective classical action of pure JT gravity as a function of D at arbitrary m. As noted above, this effective action should not exhibit any saddles for D except for $m = 1$, and indeed as expected we do not find any classical wormhole

solutions away from $m = 1$. However, many of the surprising nontrivial features of the continuation to small m manifest already in pure JT, and will thus be illuminating for the subsequent more complicated model in Section 7.5. In particular, the equations of motion are critical for obtaining a single-valued analytic continuation to $m < 1$.

7.4.1 JT Gravity in the Boundary Formalism

The tractability of pure JT stems largely from the fact that it can be reduced to the dynamics of a Schwarzian theory describing the boundary of nearly-AdS_2 space. This same feature is what permits us to compute the action at general m. To do so, recall that the Euclidean JT action on a manifold M is

$$I_{\rm JT} = -\frac{S_0}{4\pi}\left(\int_M R + 2\int_{\partial M} K\right) - \frac{1}{2}\int_M \varphi(R+2) - \int_{\partial M}\varphi(K-1), \tag{7.29}$$

where K is the extrinsic curvature of the boundary curve ∂M and we have set the AdS length to unity. The boundary conditions at ∂M are that

$$\text{Length}(\partial M) = \frac{\beta}{\delta}, \qquad \varphi|_{\partial M} = \frac{1}{\delta}, \qquad \delta \to 0. \tag{7.30}$$

These boundary conditions allow us to define a renormalized proper length coordinate u along ∂M through the condition that u/δ be a proper distance on ∂M. Integrating out the dilaton φ as usual forces M to have constant negative curvature, reducing the partition function to an integral over the moduli space of M and an integral over the shape of ∂M – hereafter referred to as the boundary "wiggle".

To evaluate the effective action for general m, we now take M to be the quotient geometry $\widehat{M}_m$. The action will depend on the modulus D and on the wiggle, which

is encoded in the embedding of ∂M in $\widehat{M}_m$. To write this embedding explicitly, we proceed as follows. Recall from Figure 7.4 that $\widehat{M}_m$ can be constructed from two identical pieces $\widehat{M}_m^{(2)}$ joined along a geodesic γ. Since we expect an on-shell solution for the wiggle to be symmetric about γ, we may restrict our attention to just a single copy of $\widehat{M}_m^{(2)}$. In terms of the complex coordinate z defined by (7.12), we embed ∂M in $\widehat{M}_m^{(2)}$ through the embedding $z = R(u)e^{i\Theta(u)}$, where $R(u)$ is related to $\Theta(u)$ by the condition that u/δ be a proper distance along ∂M[8]. Without loss of generality we also take ∂M to intersect γ at $u = \pm\beta/4$. Then because $\widehat{M}_m^{(2)}$ is just a subregion of the Poincaré disk, the action on $\widehat{M}_m$ is the usual Schwarzian action:

$$\widehat{I}_m = \frac{m-2}{m} S_0 + \int_{-\beta/4}^{\beta/4} du \left[\frac{\Theta''^2}{\Theta'^2} - \Theta'^2 - 2\left(\frac{\Theta''}{\Theta'}\right)' \right], \tag{7.31a}$$

$$= \frac{m-2}{m} S_0 - 2\int_{-\beta/4}^{\beta/4} du \, \mathrm{Sch}\left(\tan(\Theta/2), u\right), \tag{7.31b}$$

where $\mathrm{Sch}(f,u) \equiv (f''/f')' - f''^2/(2f'^2)$ is the Schwarzian derivative. However, our expectation that ∂M be symmetric about γ requires γ and ∂M to intersect orthogonally, which imposes nontrivial boundary conditions. A straightforward computation finds these to be

$$\Theta\left(\pm\beta/4\right) = \pm\left(\frac{\pi}{m} - \alpha\right), \qquad \Theta''\left(\pm\beta/4\right) = \mp\cot\alpha\, \Theta'\left(\pm\beta/4\right)^2, \tag{7.32}$$

where we recall that α is related to D via (7.14). Importantly, recall that in order to properly accommodate the case $m < 1$, we must *not* identify $\Theta(u)$ and $\Theta(u) + 2\pi$.

We now wish to put the wiggle $\Theta(u)$ on shell to obtain an effective action for α at

[8] Explicitly,

$$R = 1 - \Theta'\delta + \frac{\Theta'^2\delta^2}{2} - \frac{\Theta''^2}{2\Theta'}\delta^3 + \mathcal{O}(\delta^4).$$

general m. Taking a variation of (7.31a) yields the familiar equation of motion

$$\left(\frac{1}{\Theta'}\left(\frac{\Theta''}{\Theta'}\right)' + \Theta'\right)' = 0. \tag{7.33}$$

From the structure of the equation of motion and boundary conditions, we expect any solutions for $\Theta(u)$ to be odd in u. The general such odd solution is

$$\tan\left(\frac{\Theta(u)}{2}\right) = a\tan\left(\frac{bu}{\beta}\right) \tag{7.34}$$

where a and b are arbitrary and will be fixed by the boundary conditions in (7.32). There are two qualitatively different classes of solutions depending on whether a and b are both real or both imaginary, so we discuss them separately.

Exponential Solutions

First consider the case that a and b are imaginary: take $a = ia_i$ and $b = -ib_i$ with a_i and b_i real, giving

$$\tan\left(\frac{\Theta(u)}{2}\right) = a_i\tanh\left(\frac{b_i u}{\beta}\right). \tag{7.35}$$

Since $\Theta(u)$ must be monotonically increasing with u, a_i and b_i must have the same sign; without loss of generality we take both to be positive. Then imposing the boundary conditions (7.32), we find that when a solution for a_i and b_i exists it is always unique and given by

$$b_i = 2\,\mathrm{arccosh}\left(\frac{\sin(\pi/m)}{\sin\alpha}\right), \qquad a_i = \coth\left(\frac{b_i}{4}\right)\tan\left(\frac{\pi - m\alpha}{2m}\right). \tag{7.36}$$

Thus these solutions only exist when $\sin(\pi/m)/\sin\alpha \geq 1$. Moreover, note that the right-hand side of (7.35) is a regular function of $u \in (-\infty, \infty)$, while the left-hand side is singular when $\Theta = \pi \pmod{2\pi}$. Hence in order for solutions of this exponential type to be smooth in u, we also require that $\pi/m - \alpha < \pi$. From these two constraints on m and α, we find that these classes of solutions always exist for $m \geq 2$, never exist for $m \leq 1$, and exist for certain values of α but not others when $1 < m < 2$.

Oscillatory Solutions

Now take $a = a_r$ and $b = b_r$ with a_r and b_r real, giving the general solution

$$\tan\left(\frac{\Theta(u)}{2}\right) = a_r \tan\left(\frac{b_r u}{\beta}\right). \tag{7.37}$$

Again, monotonicity of $\Theta(u)$ allows us to restrict to positive a_r and b_r. Now, note that as Θ runs from zero to $\pi/m - \alpha$, the left-hand side of this expression goes through N poles, where

$$N \equiv \left\lfloor \frac{1}{2m} + \frac{\pi - \alpha}{2\pi} \right\rfloor. \tag{7.38}$$

To obtain a monotonic and smooth $\Theta(u)$, the right-hand side must go through the same number of poles as u goes from zero to $\beta/4$, yielding a constraint on b_r:

$$\frac{b_r}{2\pi} \in (2N - 1, 2N + 1]. \tag{7.39}$$

With this constraint in mind, we impose the boundary conditions (7.32); again we find that when a solution exists it is unique and given by

$$b_r = 2(-1)^{\lfloor 1/m - \alpha/\pi \rfloor} \arccos\left(\frac{\sin(\pi/m)}{\sin\alpha}\right) + 4\pi N, \tag{7.40a}$$

$$a_r = \cot\left(\frac{b_r}{4}\right) \tan\left(\frac{\pi - m\alpha}{2m}\right), \tag{7.40b}$$

where the principal branch of the inverse cosine is understood. Consequently we conclude that there is always a unique solution of this form whenever $|\sin(\pi/m)|/\sin\alpha < 1$.

7.4.2 Effective Action

Using the above solutions, it is straightforward to put the wiggle on shell and thereby obtain an effective action for the modulus α. We find

$$\widehat{I}_m[\alpha] = \frac{m-2}{m} S_0 + \begin{cases} \dfrac{2b_i^2}{\beta}, & m > 1 \text{ and } \dfrac{\sin(\pi/m)}{\sin\alpha} \geq 1, \\ -\dfrac{2b_r^2}{\beta}, & \dfrac{|\sin(\pi/m)|}{\sin\alpha} < 1, \end{cases} \tag{7.41}$$

with b_i and b_r given by (7.36) and (7.40). As a check, note that for $m = 1$ we recover $\widehat{I}_1 = S_0 - 2\pi^2/\beta$, which is the classical Schwarzian action of the disk [211, 169]. On the other hand, for $m = 2$ and using (7.14) we obtain $\widehat{I}_2 = 2D^2/\beta$, which is half the classical action of the double-trumpet once we recognize D as half the length of the trumpet's throat [24].

Note that when $m > 1$ we found saddles for the wiggle for all allowed values of α, but for $m < 1$ we have found no solutions whenever $|\sin(\pi/m)|/\sin\alpha > 1$. Hence $\widehat{I}_m[\alpha]$ is not defined for all m and α. In Figure 7.8 we show $\widehat{I}_m[\alpha]$ as a function of m, giving some indication of why this is the case: at a given value of α, as we decrease m we

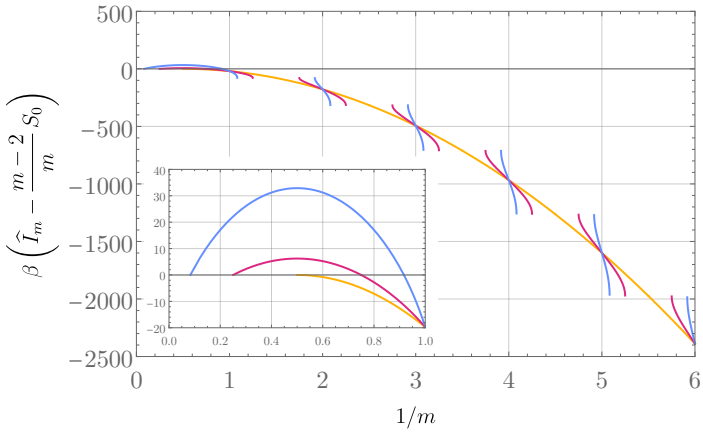


Figure 7.8: The on-shell quotient space action $\widehat{I}_m[\alpha]$ for pure JT, given in (7.41), as a function of $1/m$ for various values of α: the blue, red, and orange curves (steepest to least steep) correspond to $\alpha = \pi/12$, $\pi/4$, and $\pi/2$; since α is required to be strictly less than $\pi/2$, the orange curve can be thought of as a limiting case. Note that when $m < 1$, classical solutions for the wiggle do not always exist, and hence there are "gaps" in which the on-shell action is not defined. Solutions always exist whenever $1/m$ is an integer, in which case they yield an action independent of α, while solutions never exist for any value of $\alpha < \pi/2$ when $1/m + 1/2$ is an integer. The inset zooms in on the region $m > 1$ and shows the action to the smallest value of $1/m$ consistent with the constraint $\alpha < \min(\pi/2, \pi/m)$.

eventually reach a branch point of the inverse cosine, after which a real solution ceases to exist. Decreasing m further we reach new branch points at which solutions reappear. The locations of these branch points depend on α, but it is clear from (7.41) that a solution exists for all α whenever $1/m$ is an integer, and no solution exists for *any* values of α whenever $1/m + 1/2$ is an integer. It is also clear that at a fixed value of m, the action $\widehat{I}_m[\alpha]$ is a monotonic function of α, and hence exhibits no saddles in α.

Nevertheless, the analytic structure of $\widehat{I}_m[\alpha]$ highlights an important lesson: in the spirit of the replica trick, one might have expected that knowledge of $\widehat{I}_m[\alpha]$ for any arbitrarily small interval in m should have allowed us to analytically continue to all m. In a sense, this is indeed the case: for instance, knowing $\widehat{I}_m[\alpha]$ for $m > 1$ does allow us to analytically continue to $m < 1$. However, as shown in Figure 7.9, this analytic

continuation does not give a single-valued function of m. Instead, it yields a Riemann surface with infinitely many branches. In order to identify which, if any, of these branches correspond to the "correct" value of the action, we needed to make use of the equations of motion. The "other" branches appearing in Figure 7.9 are unphysical: they violate the constraint (7.39). We may therefore interpret them as arising from having $\Theta(u)$ wrap around the circle too many or too few times, corresponding to ∂M ending at different images of γ on the covering space of the disk, as shown in Figure 7.10. The upshot is that even working in a saddle-point approximation, a mere analytic continuation from the on-shell action at $m > 1$ is not sufficient to determine its correct behavior at $m < 1$: one needs to analytically continue the *equations of motion* themselves to $m < 1$ in order to determine the correct analytic continuation[9].

The attentive reader will note a potential concern here: even though there are no saddles for the modulus α (except when $1/m$ is an integer), $\widehat{I}_m[\alpha]$ becomes arbitrarily negative as $m \to 0$ whenever solutions exist. If we were to perform a saddle-point approximation for the path integral over the wiggle but perform a full integral over α, we might conclude that the partition function on the quotient manifold should be dominated by small-m geometries and hence exhibit a divergence in the $m \to 0$ limit of the replica trick. This would indicate a complete failure of the replica trick (or a deep pathology of pure JT gravity). However, it turns out that the saddles we have found for the wiggle are in fact unstable at sufficiently small m, and hence they should not contribute to any saddle-point approximation, which resolves this tension. We now discuss this stability

[9] A noteworthy exception is $\alpha = \pi/2$. Although a regular wormhole geometry requires $\alpha < \pi/2$ as a strict inequality, the case $\alpha = \pi/2$ can be understood as a limit in which the conical defects merge together, analogous to the "double-cone" limit of the double-trumpet. In this case, classical solutions for the wiggle exist for any $m \leq 2$ and have the on-shell action $\widehat{I}_m[\alpha = \pi/2] = (1-2/m)S_0 - 2\pi^2(1-2/m)^2/\beta$; clearly the analytic continuation of this function from any interval in m to all m is just itself, with no additional structure appearing. As can be seen in Figure 7.8, the correct branch of the Riemann surface for $\widehat{I}_m[\alpha]$ that appears for $\alpha < \pi/2$ is the one obtained via a smooth deformation away from $\widehat{I}_m[\alpha = \pi/2]$.

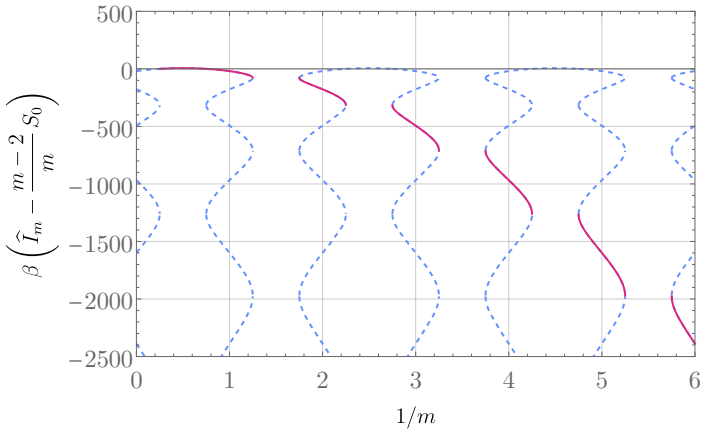


Figure 7.9: The blue dashed lines show a cross-section through the Riemann surface obtained by analytically continuing the effective action $\widehat{I}_m[\alpha]$ to all m; here we take $\alpha = \pi/4$. We only plot this Riemann surface where it takes on real values: in the "gaps" it is complex. This analytic continuation of $\widehat{I}_m[\alpha]$ exhibits infinitely many branches of which only one, shown as a solid red line, computes the correct action. The other branches correspond to violating the boundary conditions by having the wiggle $\Theta(u)$ wrap around the circle too many or too few times.

analysis.

7.4.3 Stability Analysis

To perform a stability analysis, we write $\Theta(u) = \widetilde{\Theta}(u) + \vartheta(u)$ where $\widetilde{\Theta}(u)$ is a solution to the equations of motion. We restrict our analysis here to perturbations that preserve not only the $\mathbb{Z}_2$ symmetry corresponding to reflecting across γ, but also the additional $\mathbb{Z}_2$ symmetry corresponding to reflections about the real z-axis in the left diagram of Figure 7.4. This latter symmetry amounts to taking $u \to -u$, and requires that Θ be odd in u. Then expanding the boundary conditions (7.32) to linear order in ϑ, these symmetries require that

$$\vartheta(0) = 0 = \vartheta''(0), \qquad \vartheta(\beta/4) = 0, \qquad \vartheta''(\beta/4) = -2\cot\alpha\,\widetilde{\Theta}'(\beta/4)\vartheta'(\beta/4). \tag{7.42}$$

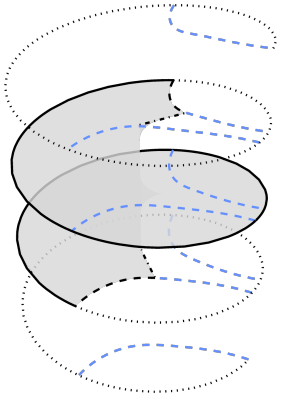


Figure 7.10: On the covering of the disk, the geodesics γ that bound $\widehat{M}_m^{(2)}$ have infinitely many images (blue dashed lines). Anchoring the boundary wiggle ∂M to these images rather than to γ gives unphysical solutions that wrap the circle too many or too few times; these additional solutions correspond to the additional branches in Figure 7.9.

Expanding the action to quadratic order in ϑ, we obtain

$$\widehat{I}_m[\Theta] = \widehat{I}_m[\widetilde{\Theta}] + 2\int_0^{\beta/4} du\, \vartheta L\vartheta + \mathcal{O}(\vartheta^3), \tag{7.43}$$

where the linear operator L is defined by

$$L\vartheta \equiv \left(\frac{\widetilde{\Theta}'\vartheta'' - 2\widetilde{\Theta}''\vartheta'}{\widetilde{\Theta}'^3}\right)'' + \left(\frac{2\widetilde{\Theta}'\widetilde{\Theta}''\vartheta'' - 3\widetilde{\Theta}''^2\vartheta'}{\widetilde{\Theta}'^4} + \vartheta'\right)'. \tag{7.44}$$

Note that L depends implicitly on m and α through its dependence on $\widetilde{\Theta}$.

It is straightforward to check that L is self-adjoint (with respect to the usual L_2 norm on $[0, \beta/4]$) on the space of functions obeying the boundary conditions (7.42). Consequently a solution $\widetilde{\Theta}$ to the equations of motion is a local minimum of the action if and only if the spectrum of L is nonnegative. Because the background solutions $\widetilde{\Theta}$ are known analytically (when they exist), it is straightforward to compute the spectrum of L

numerically using standard pseudospectral collocation methods [212]. In Figure 7.11 we show the smallest eigenvalue $\lambda_{\min}$ of L as a function of m for various values of α. As m is decreased at fixed α, $\lambda_{\min}$ remains positive until the first branch point of the action is reached and (real) classical solutions stop existing. When solutions reappear at smaller values of m, $\lambda_{\min}$ is negative. This indicates that the branch of solutions that connects continuously to $m > 1$ is stable, but the solutions that appear at smaller m past the branch points are not. So as advertised, we see that the saddles at small m that yield an arbitrarily large and negative action should not be picked up in a saddle-point approximation, and there is no worry of the replica trick giving a divergent $m \to 0$ contribution.

It is worth pausing to note the role of the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry that we imposed at the level of the stability analysis. Indeed, a careful reader might notice that $\lambda_{\min} > 0$ at $m = 1$, despite the fact that for $m = 1$ we expect there should be three zero modes associated with the $SL(2, \mathbb{R})$ symmetry of the disk. The point is that these zero modes break the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry that we have imposed and hence do not appear in our analysis. We can verify this claim by breaking, for instance, the $\mathbb{Z}_2$ corresponding to reflection symmetry about the real z-axis in the left diagram of Figure 7.4. In repeating the stability analysis with this symmetry removed, we find that L does indeed exhibit a zero mode at $m = 1$, and that a negative mode appears for all $m < 1$. Breaking the other $\mathbb{Z}_2$ corresponding to reflection symmetry across γ should recover the other two zero modes at $m = 1$. Thus we see that the additional $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry we have enforced has a stabilizing effect on the wiggle, stabilizing some of the $m < 1$ solutions that would have otherwise been unstable. This stabilizing effect will be crucial in our next model, where we will find saddles for both the wiggle *and the modulus* at $m < 1$ that are stable only if we restrict to $\mathbb{Z}_2 \times \mathbb{Z}_2$-symmetric configurations.

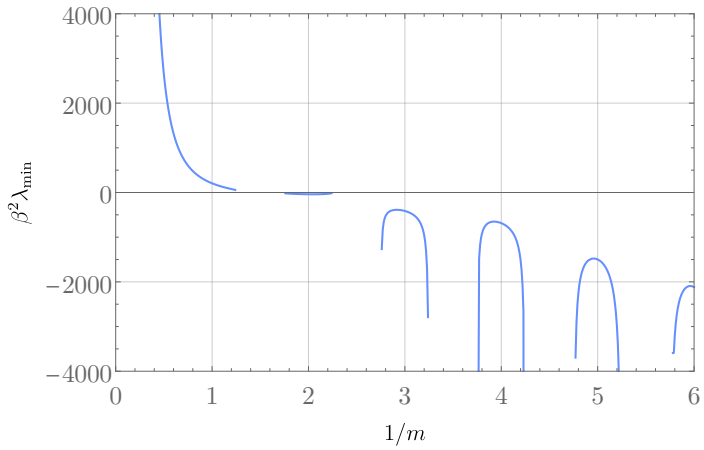


Figure 7.11: The lowest eigenvalue $\lambda_{\min}$ of L as a function of m, with $\alpha = \pi/4$. Note that $\lambda_{\min}$ is positive on the branch of solutions that continuously connects to $m > 1$, but then becomes negative on the other branches at smaller m, indicating that these branches are all unstable.

7.5 JT with a Massless Scalar

We now turn to our main model of interest: JT gravity coupled to a massless scalar. We will take the scalar ψ to be real and minimally-coupled, so that the Euclidean action of the theory is $I = I_{\text{JT}} + I_{\text{mat}}$ where

$$I_{\text{mat}} = \frac{1}{2}\int_M (\boldsymbol{\nabla}\psi)^2. \tag{7.45}$$

We can easily integrate out the scalar field by placing it on shell: since its equation of motion is $\boldsymbol{\nabla}^2\psi = 0$, the on-shell matter action just contributes a boundary term:

$$I_{\text{mat}} = \frac{1}{2}\int_{\partial M} \psi\, n^a \boldsymbol{\nabla}_a \psi, \tag{7.46}$$

where n^a is the unit outward-pointing normal to ∂M.

In order to attempt to stabilize the replica wormholes at the classical level, we will

need to turn on sources for the scalar:

$$\psi|_{\partial M} = \psi_\partial(u), \tag{7.47}$$

where the profile $\psi_\partial(u)$ can be specified arbitrarily. Note that we will require this profile to have nontrivial dependence on u: if it were a constant $\psi_\partial = c$, then the equations of motion for ψ would be solved by the constant solution $\psi = c$ everywhere. Such a solution gives a vanishing on-shell action (7.46) and reduces the total action to that of pure JT. But if $\psi_\partial(u)$ depends nontrivially on u, then the fact that the proper length u is defined by the shape of the boundary wiggle means that the on-shell action (7.46) induces a nontrivial coupling between the wiggle and the scalar. Our goal is now to understand this coupling on the quotient geometry $\widehat{I}_m$ and to show that it can provide the classical solutions for the wiggle and modulus.

7.5.1 JT + Scalar in the Boundary Formalism

Why not proceed as in pure JT by constructing the quotient manifold $\widehat{M}_m$ by taking two identical subregions of the Poincaré disk and stitching them together along a geodesic γ? The reason is that it is nontrivial to impose appropriate boundary conditions on ψ at γ, which is required for the construction of a general solution to the equations of motion and the subsequent evaluation of the on-shell matter action (7.46). Instead, it will be convenient to work with a single coordinate chart on $\widehat{M}_m$ so that we may solve for ψ by imposing only a Dirichlet condition on ∂M. We will use the elliptical coordinate chart introduced in Section 7.3.2, in terms of which the metric on $\widehat{M}_m$ can be written as

$$ds^2 = e^{2\sigma(\xi,\phi)}\left(d\xi^2 + d\phi^2\right), \tag{7.48}$$

where recall $\xi \in [0, \pi\xi_0/2)$, $\phi \in [0, 2\pi)$, the conical defects lie at $(\xi, \phi) = (0, 0)$ and $(0, \pi)$, and ∂M corresponds to $\xi = \pi\xi_0/2$. We must therefore express the JT and matter parts of the action in this coordinate chart.

JT Action

To construct the JT part of the action in the elliptical coordinate chart, we note that the near-boundary expansion of the metric takes the form

$$ds^2 = \left(\frac{1}{(\pi\xi_0/2 - \xi)^2} + g_2(\phi) + \mathcal{O}\left(\frac{\pi\xi_0}{2} - \xi\right)\right)\left(d\xi^2 + d\phi^2\right), \tag{7.49}$$

where g_2 can be extracted from the solutions to the Liouville equation constructed in Section 7.3.2:

$$g_2(\phi) = \partial_\xi^2 \tilde{\sigma}|_{\xi=\pi\xi_0/2} + \frac{1}{3\xi_0^2}, \tag{7.50}$$

where $\tilde{\sigma}$ is defined by (7.19a). In particular, in the limit of small ξ_0 and ν with $e^{-\pi/2\xi_0} \ll \nu$ (where $\nu \equiv m/(2-m)$), (7.28) gives[10]

$$g_2(\phi) = -\frac{1+2\nu}{3\nu^2\xi_0^2}\sum_{k=-\infty}^{\infty} \operatorname{sech}^2\left(\frac{\phi - k\pi}{\xi_0}\right) + \frac{1}{3\xi_0^2} + \mathcal{O}(\nu^2, \nu^{-6}e^{-2\pi/\xi_0}). \tag{7.51}$$

Interestingly, comparison with the numerical solutions to the Liouville equation shows that this expression for g_2 actually captures the ϕ-dependence of g_2 *exactly*: that is, we

[10] More compactly, we note that

$$\sum_{k=-\infty}^{\infty} \operatorname{sech}^2\left(\frac{\phi - k\pi}{\xi_0}\right) = \xi_0^2\,\partial_\phi^2 \log\left(\vartheta_2\left(i\phi/\xi_0,\, e^{\pi/\xi_0}\right)\right),$$

where $\vartheta_2(z, q) = \sum_{n=-\infty}^{\infty} q^{(n+1/2)^2} e^{(2n+1)iz}$ is the Jacobi theta function of the second kind.

find that for all ν and ξ_0,

$$g_2(\phi) = -\frac{1+2\nu}{3\nu^2\xi_0^2}\sum_{k=-\infty}^{\infty}\operatorname{sech}^2\left(\frac{\phi-k\pi}{\xi_0}\right) + C(\nu,\xi_0), \tag{7.52}$$

where $C(\nu,\xi_0)$ is independent of ϕ up to numerical resolution. We do not know of an analytic argument for why this is the case, but in practice it means that we only need to numerically solve the Liouville equation to extract the constant $C(\nu,\xi_0)$, rather than the entire functional form of $g_2(\phi)$. We discuss the computation of $C(\nu,\xi_0)$ and illustrate its behavior in Appendix D.2.1.

We may now construct the boundary JT action: we embed ∂M in $\widehat{M}_m$ through the embeddings $(\xi,\phi) = (X(u),\Phi(u))$, where as before the requirement that u/δ be a proper length along ∂M relates the embeddings through

$$X(u) = \frac{\pi\xi_0}{2} - \Phi'(u)\,\delta - \frac{g_2(\Phi(u))\Phi'(u)^4 + \Phi''(u)^2}{2\Phi'(u)}\,\delta^3 + \mathcal{O}(\delta^4). \tag{7.53}$$

In terms of these embeddings, the dynamical part of the JT action becomes

$$-\int_{\partial M}\varphi(K-1) = \frac{1}{2}\int_0^\beta du\left(\frac{\Phi''(u)^2}{\Phi'(u)^2} + 3g_2(\Phi(u))\Phi'(u)^2\right), \tag{7.54a}$$

$$= \frac{1}{2}\int_0^{2\pi} d\phi\left(\frac{u''(\phi)^2}{u'(\phi)^3} + \frac{3g_2(\phi)}{u'(\phi)}\right), \tag{7.54b}$$

where we have written the second line in terms of the inverse function $u(\phi)$ defined by $\Phi(u(\phi)) = \phi$, the existence of which is guaranteed by the requirement that $\Phi(u)$ be monotonic in u. Note that Φ now wraps once around the ellipse: $\Phi(u+\beta) = \Phi(u) + 2\pi$, or $u(\phi+2\pi) = u(\phi) + \beta$.

Matter Action

To compute the contribution of the scalar to the action for the wiggle, we make use of the fact that the scalar field is a CFT, allowing us to compute the classical action (7.46) in any choice of conformal frame. A natural such choice is given by taking the conformal factor $\sigma = 0$, thereby putting the scalar on a strip:

$$ds^2 = d\xi^2 + d\phi^2, \qquad \xi \in [0, \pi\xi_0/2), \qquad \phi \in [0, 2\pi). \tag{7.55}$$

Because this domain is conformal to half of the double-trumpet, in principle we could proceed as in e.g. [204, 213] and express ψ as an integral against a bulk-to-boundary propagator on the double-trumpet with appropriate replica symmetry imposed to relate the boundary conditions on the two ends. In practice, for implementing the numerical methods described in Appendix D.2, it is more convenient to directly solve the Laplace equation for ψ via Fourier series by constructing a solution subject to appropriate boundary conditions:

- Periodicity in ϕ: $\psi(\xi, \phi + 2\pi) = \psi(\xi, \phi)$.
- Smoothness at $\xi = 0$ when conformally mapped back to the ellipse[11]:

$$\partial_\xi^p \psi(0, \phi) = (-1)^p \partial_\xi^p \psi(0, -\phi) \qquad \forall p \in \mathbb{Z}_{\geq 0}$$

[11] The reader may be concerned that the demand that ψ be regular on the ellipse excludes solutions where ψ is regular on the full wormhole geometry M_m but singular at the conical defects in $\widehat{M}_m$. But this is not the case: because classical solutions for ψ are harmonic, smooth solutions on M_m must also be smooth on $\widehat{M}_m$. One way to see this is to consider an arbitrary $\mathbb{Z}_m$-invariant holomorphic function f on the Poincaré disk (7.12), which must have an expansion in powers of z^m. Taking a quotient $z \to z^{1/m}$ to go to the Poincaré disk with a defect, one finds that f has a standard expansion in integer powers of z, which is regular. This analysis applies locally near any defect, and since the real and imaginary parts of f are harmonic, we conclude that any $\mathbb{Z}_m$-invariant harmonic function on M_m must be regular on $\widehat{M}_m$, including at the defects.

- Scalar sources: $\psi(\pi\xi_0/2, \phi) = \psi_\partial(u(\phi))$.

The general solution to the Laplace equation on this domain satisfying the first two boundary conditions is given by

$$\psi(\xi, \phi) = \sum_{k=-\infty}^{\infty} \left(a_k \cosh(k\xi) + i b_k \sinh(k\xi)\right) e^{ik\phi}, \tag{7.56}$$

where the coefficients a_k and b_k are real and obey $a_k = a_{-k}$, $b_k = b_{-k}$. These coefficients can be determined by imposing the Dirichlet condition at $\xi = \pi\xi_0/2$:

$$a_k \cosh(k\pi\xi_0/2) + i b_k \sinh(k\pi\xi_0/2) = \frac{1}{2\pi} \int_0^{2\pi} d\phi\, \psi_\partial(u(\phi)) e^{-ik\phi}. \tag{7.57}$$

Using these relations to express (7.56) in terms of $\psi_\partial(u(\phi))$, we ultimately find that the on-shell action (7.46) takes the form of a smearing of $\psi_\partial(u(\phi))$:

$$\widehat{I}_{\text{mat}} = \frac{1}{2} \int_0^{2\pi} d\phi\, d\tilde{\phi}\ \psi_\partial(u(\phi))\, \psi_\partial(u(\tilde{\phi}))\, G(\phi, \tilde{\phi}), \tag{7.58}$$

where

$$G(\phi, \tilde{\phi}) \equiv \frac{1}{\pi} \sum_{k=1}^{\infty} k \left[\tanh\left(\frac{k\pi\xi_0}{2}\right) \cos(k\phi) \cos\left(k\tilde{\phi}\right) + \coth\left(\frac{k\pi\xi_0}{2}\right) \sin(k\phi) \sin\left(k\tilde{\phi}\right) \right]. \tag{7.59}$$

Note that $G(\phi, \tilde{\phi})$ is essentially the boundary-to-boundary propagator for ψ. As written, the sum defining $G(\phi, \tilde{\phi})$ is not convergent for all ϕ and $\tilde{\phi}$ due to contact terms; for the purposes of evaluating the action, $G(\phi, \tilde{\phi})$ should be understood distributionally. For computing the matter two-point function, $G(\phi, \tilde{\phi})$ should instead be renormalized

appropriately.

Boundary Action and Equations of Motion

Putting these results together, we find that integrating out the scalar field ψ leaves us with the boundary action

$$\widehat{I}_m = -\frac{S_0}{\nu} + \frac{1}{2}\int_0^{2\pi} d\phi \left(\frac{u''(\phi)^2}{u'(\phi)^3} + \frac{3g_2(\phi)}{u'(\phi)} + \int_0^{2\pi} d\tilde{\phi}\, \psi_\partial(u(\phi))\psi_\partial(u(\tilde{\phi}))G(\phi,\tilde{\phi}) \right), \quad (7.60)$$

which we have expressed entirely in terms of the inverse wiggle $u(\phi)$. Due to the non-locality of the matter part of the action, the resulting equation of motion for $u(\phi)$ is an integro-differential equation:

$$\left(\left(\frac{u''(\phi)}{u'(\phi)^3} \right)' + \frac{3(u''(\phi)^2 + u'(\phi)^2 g_2(\phi))}{2u'(\phi)^4} \right)' + \dot{\psi}_\partial(u(\phi)) \int_0^{2\pi} d\tilde{\phi}\, \psi_\partial(u(\tilde{\phi}))G(\phi,\tilde{\phi}) = 0, \quad (7.61)$$

where $\dot{\psi}_\partial \equiv d\psi_\partial/du$.

We will also need to perform a stability analysis of the solutions to (7.61), which is done as in pure JT: we write the wiggle as $u(\phi) = \tilde{u}(\phi) + \kappa(\phi)$, where $\tilde{u}(\phi)$ solves the equation of motion. Expanding the action to quadratic order in κ, we obtain

$$\widehat{I}_m[u] = \widehat{I}_m[\tilde{u}] + \frac{1}{2}\int_0^{2\pi} d\phi\, \kappa L \kappa + \mathcal{O}(\kappa^3), \quad (7.62)$$

where now the fluctuation operator L is

$$\begin{aligned} L\kappa(\phi) \equiv & \left(\frac{\kappa''(\phi)}{\tilde{u}'(\phi)^3} - \frac{3\tilde{u}''(\phi)\kappa'(\phi)}{\tilde{u}'(\phi)^4} \right)'' + 3\left(\frac{\tilde{u}''(\phi)\kappa''(\phi)}{\tilde{u}'(\phi)^4} - \left(\frac{g_2(\phi)}{\tilde{u}'(\phi)^3} + \frac{2\tilde{u}''(\phi)^2}{\tilde{u}'(\phi)^5} \right)\kappa'(\phi) \right)' \\ & + \int_0^{2\pi} d\tilde{\phi} \left[\ddot{\psi}_\partial(\tilde{u}(\phi))\psi_\partial(\tilde{u}(\tilde{\phi}))\kappa(\phi) + \dot{\psi}_\partial(\tilde{u}(\phi))\dot{\psi}_\partial(\tilde{u}(\tilde{\phi}))\kappa(\tilde{\phi}) \right] G(\phi,\tilde{\phi}), \quad (7.63) \end{aligned}$$

where $\ddot{\psi}_\partial \equiv d^2\psi_\partial/du^2$. It is straightforward to verify that L is self-adjoint (with respect to the usual L_2 norm on $[0, 2\pi]$) on the space of functions periodic in ϕ with period 2π, and hence a saddle $\tilde{u}$ is a local minimum of the action if and only if the spectrum of L is nonnegative.

7.5.2 Stabilizing the Double-Trumpet

For general ψ_∂, (7.61) cannot be solved analytically; we will require a numerical solution. However, for $m = 1$ and $m = 2$, it is possible to obtain analytic solutions by looking for *bulk* solutions for the dilaton Φ and reconstructing the corresponding behavior of the boundary wiggle $u(\phi)$. We discuss the construction of these bulk solutions in Appendix D.3; here we simply exhibit them in order to study the effect of turning on the CFT sources. In particular, we will show that taking the amplitude of the boundary source ψ_∂ large enough allows us to stabilize the double-trumpet, and moreover that the double-trumpet dominates over the disk at sufficiently small temperatures.

The solutions constructed in Appendix D.3 correspond to the family of boundary profiles

$$\psi_\partial(u) = J \cos\left(\frac{2\pi n u}{\beta}\right) \sqrt{\frac{1+A}{1+A\cos(4\pi n u/\beta)}}, \tag{7.64}$$

with boundary wiggle given by

$$\tan\left(\frac{2\pi n u(\phi)}{\beta}\right) = \sqrt{\frac{1+A}{1-A}} \tan(n\phi), \tag{7.65}$$

where J and $A \in (-1, 1)$ are arbitrary constants and n is an arbitrary positive integer. This class of profiles and form of the wiggle is tractable because $\psi_\partial(u(\phi)) = J\cos(n\phi)$, so the smearing of ψ_∂ against $G(\phi, \tilde{\phi})$ is straightforward to compute. Saddles (for both

the wiggle and the modulus ξ_0) are then obtained through an appropriate choice of A. For instance, for $m = 2$, we have $g_2(\phi) = 1/(3\xi_0^2)$, and hence the action of the wiggle profile (7.65) is

$$\widehat{I}_2 = \frac{n^2}{\beta}\left[\frac{2\pi^2(4+1/(n\xi_0)^2)}{\sqrt{1-A^2}} - 8\pi^2 + \frac{\pi\beta J^2}{2n}\tanh\left(\frac{n\pi\xi_0}{2}\right)\right]. \tag{7.66}$$

For a given choice of A, the wiggle equation of motion is only satisfied for particular values ξ_* of the modulus ξ_0: evaluating (7.61) on (7.64) and (7.65), we obtain the constraint

$$\frac{1}{\sqrt{1-A^2}} = \sqrt{1+\left(\frac{\beta J^2/n}{8\pi(4+1/(n\xi_*)^2)\coth(n\pi\xi_*/2)}\right)^2}. \tag{7.67}$$

However, if we wish to find a saddle to the *full* path integral rather than just for the integral over the boundary wiggle, we must further require that ξ_* be a stationary point with respect to ξ_0 of the action (7.66). This requirement gives

$$\frac{1}{\sqrt{1-A^2}} = \frac{n^2\xi_*^3\beta J^2}{16\cosh^2(n\pi\xi_*/2)}. \tag{7.68}$$

Hence by simultaneously solving (7.67) and (7.68) for A and ξ_*, we can obtain a boundary profile that gives rise to a classical saddle for both the wiggle $u(\phi)$ and the modulus ξ_0. It is straightforward to see that such a saddle can only exist when $\beta J^2/n$ is sufficiently large: any ξ_* that simultaneously satisfies (7.67) and (7.68) obeys $f(n\xi_*) = n^2/(\beta^2 J^4)$, where

$$f(x) \equiv \left(\frac{x^3}{16\cosh^2(\pi x/2)}\right)^2 - \frac{1}{(8\pi(4+1/x^2)\coth(\pi x/2))^2}. \tag{7.69}$$

$f(x)$ has a global maximum at $x_{\max} \approx 0.93$, where it attains the value $f_{\max} \approx 4.7\times 10^{-5}$, so saddles for ξ_0 exist if and only if $\sqrt{\beta}J \geq f_{\max}^{-1/4}\sqrt{n} \approx 12\sqrt{n}$. So turning on matter

sources with sufficiently large amplitude, or taking the temperature sufficiently small, can give rise to classical saddles for the double-trumpet.

While the approach just described allows us to find simultaneous saddles for both the wiggle and the modulus, for the purposes of a stability analysis it is illuminating to construct the effective action $\widehat{I}_2[\xi_0]$ for the modulus obtained by only putting the wiggle on shell. To do so, we consider a family of boundary sources $\psi_\partial(u)$ of the form (7.64) with $A = A_*(\beta J^2/n)$, where

$$A_*(y) \equiv \left[1 + \left(\frac{8\pi\left(4 + \frac{1}{x_*(y)^2}\right)\coth\left(\frac{\pi x_*(y)}{2}\right)}{y}\right)^2\right]^{-1/2}, \tag{7.70a}$$

$$x_*(y) \equiv \begin{cases} x_{\max}, & y < f_{\max}^{-1/2}, \\ \text{smallest positive solution of } y^2 f(x) = 1, & y \geq f_{\max}^{-1/2}. \end{cases} \tag{7.70b}$$

This form of A ensures that (7.67) is satisfied (and hence (7.65) is a solution to the equation of motion) when $n\xi_0 = x_*(\beta J^2/n)$, which for $\sqrt{\beta}J \geq f_{\max}^{-1/4}\sqrt{n}$ corresponds to the location ξ_* of a saddle in ξ_0. We will also take $n = 2$ in order to ensure that $\psi_\partial(u)$ is symmetric about $u = 0$ and $u = \beta/4$.

With such a profile, the equation of motion (7.61) cannot be solved analytically for general ξ_0 (except, by construction, for the special case $n\xi_0 = x_*(\beta J^2/n)$), so we must proceed numerically. The details of the numerical computation are presented in Appendix D.2, and the resulting action $\widehat{I}_2[\xi_0]$ is shown in Figure 7.12. When $J = 0$, we recover the pure JT trumpet action $\widehat{I}_2[\xi_0] = 2\pi^2/(\xi_0^2\beta) = b^2/(2\beta)$, where $b = 2\pi/\xi_0$ is the circumference of the trumpet throat. At sufficiently small values of $\sqrt{\beta}J$, the action remains a monotonic function of ξ_0, exhibiting no saddles in ξ_0. When $\sqrt{\beta}J$ becomes sufficiently large, two saddles in ξ_0 appear, with one stable and the other unstable with

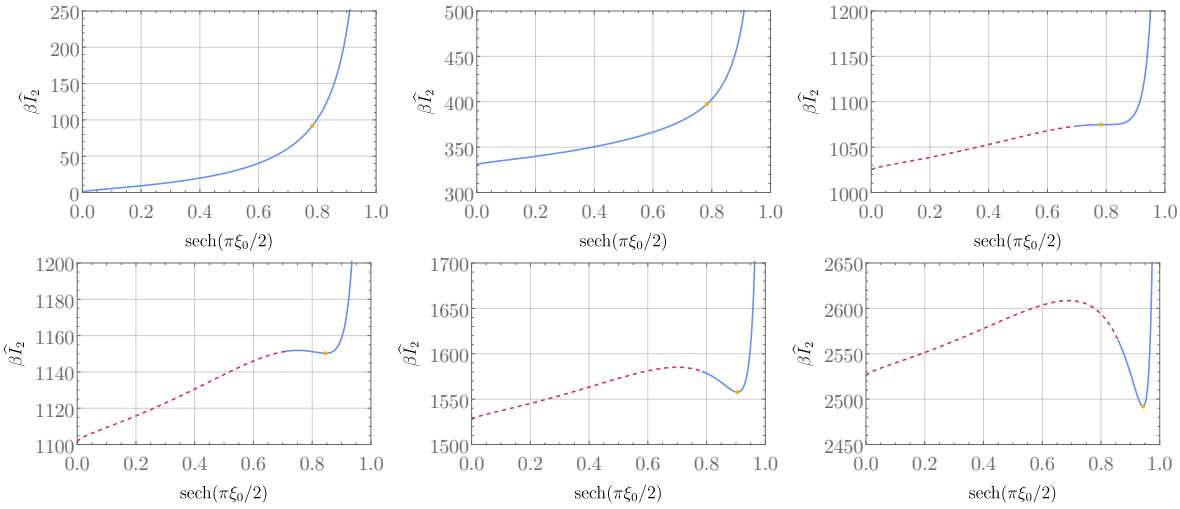


Figure 7.12: The effective action $\widehat{I}_2[\xi_0]$ obtained by placing the wiggle on shell, with boundary matter profile (7.64) with $n = 2$ and A given by (7.70); note that we plot the action as a function of the eccentricity $\mathrm{sech}(\pi\xi_0/2)$ of the elliptical domain. From left to right and top to bottom, we show $\sqrt{\beta}J = 0$, 10, 17, 18, 22, and 30. We plot the action as a solid blue curve where the spectrum of L (on the space of perturbations that preserve the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry) is nonnegative, and as a dashed red curve when L has a negative eigenvalue. The orange dots show the action (7.66) when $\xi_0 = x_*(\beta J^2/2)/2$, for which the wiggle (7.65) with A given by (7.70) satisfies the equation of motion (7.61).

respect to perturbations in ξ_0. At these intermediate values of $\sqrt{\beta}J$, the stable saddle does not *globally* minimize the action: it corresponds to a metastable solution. Further increasing $\sqrt{\beta}J$, however, decreases the action of the stable saddle until it becomes a global minimum in ξ_0.

We might expect that this global minimum of $\widehat{I}_2[\xi_0]$ should dominate the path integral. This well may be the case, but a stability analysis indicates a subtlety: not all of the saddles obtained for the wiggle are stable. Restricting our consideration to perturbations that preserve the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry, the solid blue curves in Figure 7.12 correspond to solutions for which the fluctuation operator L defined in (7.63) has a nonnegative spectrum, while the dashed red curves correspond to solutions for which L exhibits a negative eigenvalue. So although for large $\sqrt{\beta}J$ the stable saddles are global minima

of the effective action $\widehat{I}_2[\xi_0]$ obtained by keeping the wiggle on shell, it is conceivable that off-shell configurations of the wiggle could decrease the action below that of these putative global minima. We have not investigated this possibility further, but for now we assume that the $m = 2$ path integral can be approximated by this new global minimum of the effective action $\widehat{I}_2[\xi_0]$, when the minimum exists.

If the double-trumpet can be stabilized, can it ever dominate over the disk in a computation of, say, $\overline{Z^2}$? Such dominance could only ever occur in a classical limit if J and β scale appropriately with S_0, since otherwise the topological part of the action will trivially cause the disk to dominate. This is analogous to the need for the matter partition function to be of order e^{S_0} in models of black hole evaporation before wormholes can start to dominate after the Page time [49, 50]. Indeed, we find that with an appropriate scaling of J and β with S_0, an exchange of dominance between the disk and the double-trumpet can occur: in Figure 7.13 we show the difference $\widehat{I}_1 - \widehat{I}_2$ between the actions of the disk and the trumpet for various values of $J/\sqrt{S_0}$. At values of $J/\sqrt{S_0}$ around order unity or smaller, this difference is everywhere-negative, so the disk always dominates. But at larger values of $J/\sqrt{S_0}$, this difference becomes positive at sufficiently large values of $\sqrt{\beta}J$, indicating that the double-trumpet dominates at sufficiently low temperatures. This transition requires the amplitude J to scale with S_0 like $J \gtrsim \sqrt{S_0}$, so at fixed $J/\sqrt{S_0}$ the temperatures at which the double-trumpet dominates (when it does at all) scale with S_0 like $T \lesssim S_0$. This behavior is analogous to the results of [207, 213], where a phase transition between the disk and the double-trumpet was induced by turning on constant but complex sources for a massless scalar. Here we see that an analogous transition can be supported with real, replica-symmetric sources with nontrivial Euclidean time dependence. It is this nontrivial Euclidean time dependence that induces the stress tensor necessary to stabilize the wormhole.

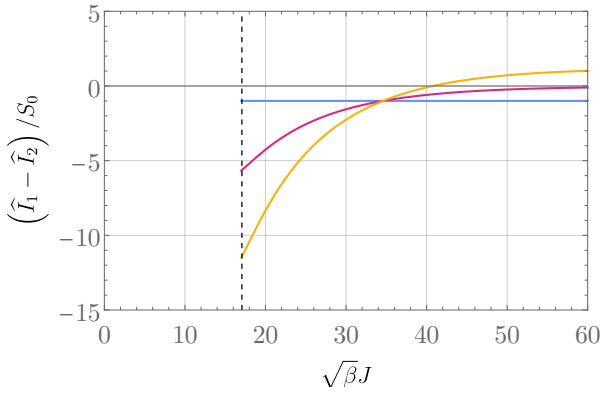


Figure 7.13: The difference $\widehat{I}_1 - \widehat{I}_2$ between the on-shell actions of the disk and the trumpet obtained by putting both the wiggle and the modulus ξ_0 on shell. The boundary matter profile is (7.64) with $n = 2$ and A given by (7.70); the corresponding trumpet action $\widehat{I}_2$ is given by (7.66) with $\xi_0 = x_*(\beta J^2/2)/2$, while the disk action $\widehat{I}_1$ is computed numerically. The blue, red, and orange curves correspond to $J/\sqrt{S_0} = 0$, 4, and 6; the orange curve becomes positive at large enough $\sqrt{\beta}J$, indicating dominance of the double-trumpet over the disk. No saddles for $\widehat{I}_2$ exist to the left of the dashed black line.

7.5.3 Wormholes for $m < 1$

Having discussed the special cases $m = 1$ and $m = 2$, we now turn our investigation to the saddles at $m < 1$ that appear in the replica trick. The first distinction to note with the $m > 1$ case is that the topological part of the action is monotonically decreasing with m. Consequently, taking S_0 large at fixed β and ψ_∂ will always cause the disk to win out over the $m > 1$ wormholes; this was the reason that we needed to scale J and β with S_0 in the previous section to get the double-trumpet to dominate over the disk. However, for this same reason, taking S_0 large will always cause the disk to be *subdominant* to any saddles at $m < 1$. Hence if there are any saddles at $m < 1$ at all, they will always dominate over the disk in a classical limit $S_0 \to \infty$ with β and ψ_∂ kept fixed. Thus we only need to look for stable saddles at $m < 1$, without needing to worry about their dominance over the disk.

Our analysis will be entirely numerical, so for simplicity we now fix the matter sources

to be the lowest nontrivial Fourier mode on the thermal circle compatible with our assumed $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry:

$$\psi_\partial(u) = J \cos\left(\frac{4\pi u}{\beta}\right). \tag{7.71}$$

Again we leave the numerical details to Appendix D.2. At relatively small values of $\sqrt{\beta}J$, we do not find any stable saddles at any m. In Figure 7.14 we show the effective action $\widehat{I}_m[\xi_0]$ with $\sqrt{\beta}J = 10$ for several values of m. This effective action is either monotonic in ξ_0, exhibiting no saddles for the modulus, or may exhibit a stable saddle for ξ_0 which is however unstable to perturbations of the wiggle, as in the fourth plot in the figure. Note that solutions do not exist for all ξ_0: the wiggle becomes singular at sufficiently small $\xi_0 = \xi_{\text{end}}$ below which we found no more solutions. This behavior is qualitatively analogous to what we observed in pure JT in Section 7.4: on-shell solutions for the wiggle did not exist for all α. In that case, the end points at which solutions stopped existing corresponded to branch points of the Riemann surface for the analytic continuation of $\widehat{I}_m[\alpha]$ to complex m (and α). The endpoints ξ_{end} shown in Figure 7.14 may play the same role: they may be branch points of the analytic continuation of $\widehat{I}_m[\xi_0]$ to complex m and ξ_0.

Increasing $\sqrt{\beta}J$ turns out to change the story substantially, however: in Figure 7.15 we show $\widehat{I}_m[\xi_0]$ with $\sqrt{\beta}J = 20$. We are still unable to find on-shell solutions for the wiggle for all ξ_0, but we also find two independent branches of solutions, with one unstable and the other stable. For $\nu \gtrsim 1/3$ (or $m \gtrsim 1/2$), these branches meet at a zero mode at which the wiggle is regular. As ν is decreased below $1/3$, the zero mode becomes singular and the two branches separate, with each one terminating at a singular endpoint analogous to those in Figure 7.14. The key feature of these two branches is that because one is stable,

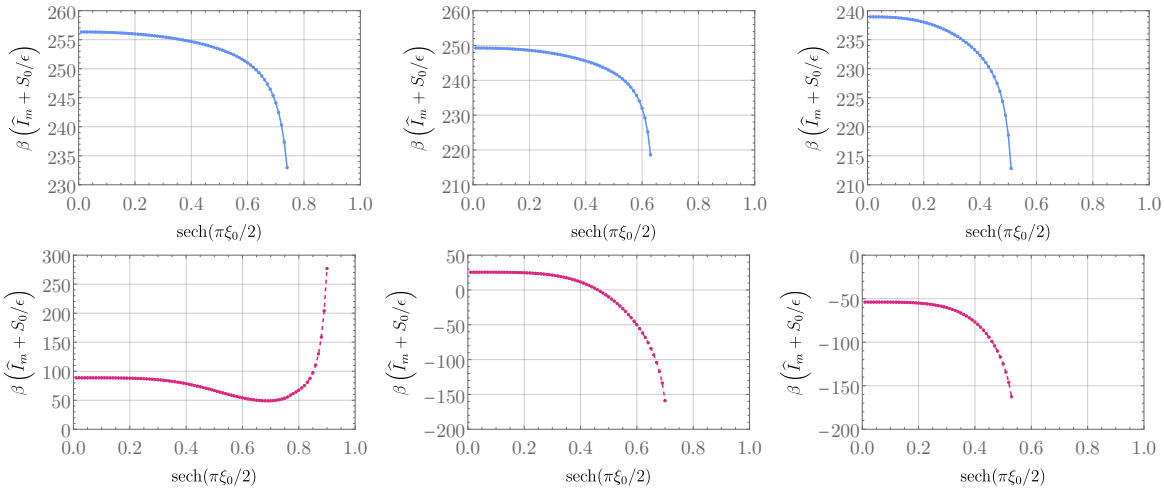


Figure 7.14: The effective action $\widehat{I}_m[\xi_0]$ of JT + CFT with boundary source given by (7.71) with $\sqrt{\beta}J = 10$. From left to right and top to bottom we show $\nu = 0.9$, 0.8, 0.7, 0.34, 0.3, 0.27, corresponding to $m \approx 0.95$, 0.89, 0.82, 0.51, 0.46, 0.43. Points correspond to numerical data; curves are drawn to guide the eye. Blue points connected by solid lines indicate that the spectrum of L is nonnegative; red points connected by dashed lines indicate that L has a negative eigenvalue. No stable saddles for both ξ_0 and the wiggle are present.

we need only find a stable saddle in ξ_0 to deduce the existence of a stable wormhole. And indeed, such a saddle exists, as seen in the fourth plot in the figure. Interestingly, this saddle does not persist to smaller m: as can be seen in the last two plots of the figure, when ν is decreased below 1/3 and the branches separate, the saddle vanishes. We are unable to find additional stable saddles by further decreasing ν. This qualitative behavior is independent of the value of $\sqrt{\beta}J$ up to the largest value $\sqrt{\beta}J = 25$ for which we have constructed solutions. The upshot is that at sufficiently large $\sqrt{\beta}J$, stable classical replica wormholes exist at $m < 1$, but only down to around $m \approx 1/2$.

Before examining the on-shell action of these wormholes in more detail, let us pause to note that requiring the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry was crucial to finding stable saddles. For reference, in Figure 7.16 we show the lowest eigenvalue of L when the parity of the perturbations about $\phi = 0$ and $\phi = \pi/2$ is modified. Only for perturbations odd about $\phi = 0$

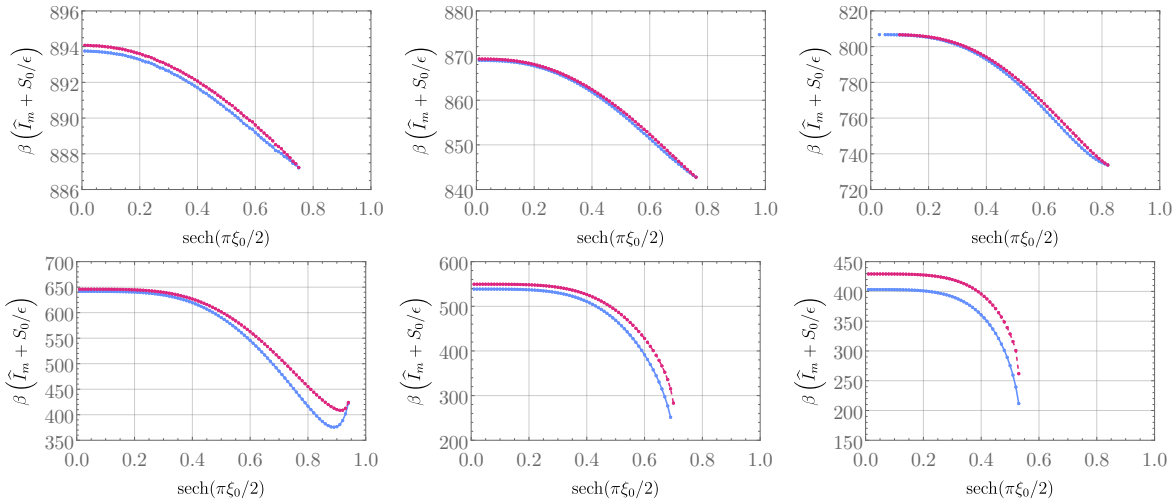


Figure 7.15: The effective action $\widehat{I}_m[\xi_0]$ of JT + CFT with boundary source given by (7.71) with $\sqrt{\beta}J = 20$. From left to right and top to bottom we show $\nu = 0.9$, 0.7, 0.5, 0.34, 0.3, 0.27, corresponding to $m \approx 0.95$, 0.82, 0.67, 0.51, 0.46, 0.43. Points correspond to numerical data; curves are drawn to guide the eye. Blue points connected by solid lines (lowermost curves) indicate that the spectrum of L is nonnegative; red points connected by dashed lines (uppermost curves) indicate that L has a negative eigenvalue. For all ν shown here there are two branches of solutions, with one stable and the other unstable; for $\nu > 1/3$ ($m > 1/2$) these branches meet at a zero mode, but for $\nu < 1/3$ these branches cease to join. Note that there is a stable saddle in fourth plot.

(corresponding to the shape of ∂M being symmetric about the major axis of the ellipse) is one of the branches of solutions stable at the saddle for ξ_0, giving a stable wormhole. This symmetry about $\phi = 0$ is necessary for a real Lorentzian continuation obtained by cutting the ellipse across its major axis. So the stability of these wormholes – and consequently whether the quenched and annealed generating functionals Γ_Q and Γ_A differ – depends crucially on whether we demand that perturbations about the saddle admit a real Lorentzian section that contains the defects. Note that it is crucial that the Lorentzian section contain the defects: a real Lorentzian geometry could also be generated by cutting the ellipse about its minor axis, but requiring reality of such a section is not sufficient to stabilize the wormholes.

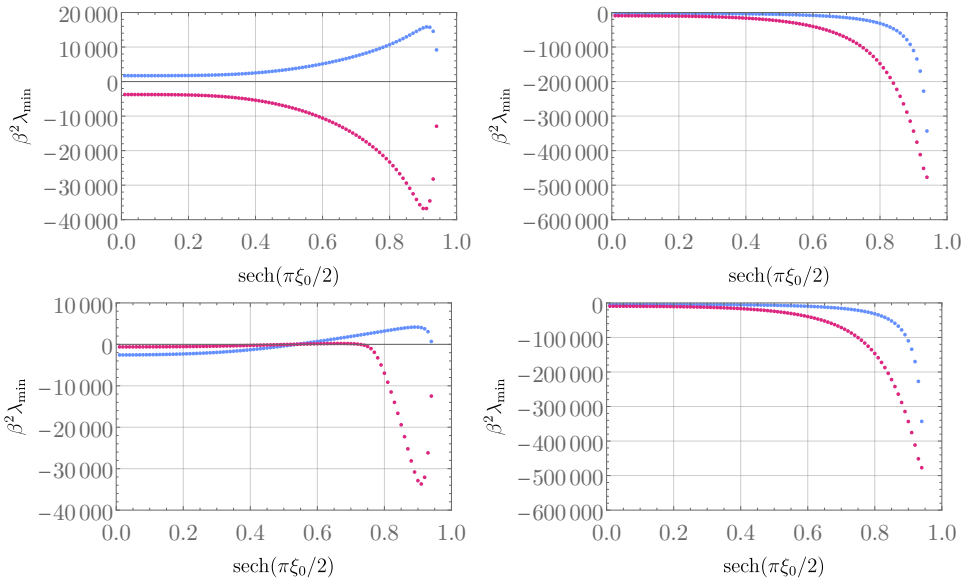


Figure 7.16: The lowest eigenvalue $\lambda_{\min}$ of the fluctuation operator L under different parities of the perturbation $\nu(\phi)$ defined in (7.63); here we show $\sqrt{\beta}J = 20$ and $\nu = 0.34$, i.e. the fourth plot in Figure 7.15. The left (right) plots take $\nu(\phi)$ to be odd (even) about $\phi = 0$, while the top (bottom) plots take $\nu(\phi)$ to be odd (even) about $\phi = \pi/2$. The colors label the two branches shown in Figure 7.15; note that $\nu(\phi)$ needs to be odd about $\phi = 0$ for at least one of the branches to be stable at the location of the saddle in ξ_0. This corresponds to the boundary curve ∂M being even about the major axis of the ellipse.

How should the need for such a real Lorentzian section be interpreted? If the path integral is to be understood as a purely Euclidean object completely removed from any Lorentzian underpinning, then there is no reason to impose any such condition. In this case, our saddles are simply unstable and never contribute to the quenched generating functional. But this interpretation is rather odd: after all, we are ultimately interested in theories with a Lorentzian counterpart, and moreover to even make the JT Euclidean path integral well-defined in the first place the dilaton φ needs to be Wick rotated to an imaginary contour: a *strictly* Euclidean definition of the JT path integral is manifestly divergent. In addition, there is the question of which $\mathbb{Z}_2$ symmetry to preserve; that is, which of the principal axes of the ellipse should correspond to the "$t = 0$" slice of the

Lorentzian section. Comparison with other replica tricks make it natural to require the conical defects to live on this $t = 0$ slice: for example, in the LM construction of von Neumann entropy, it is the $m \to 1$ limit of the conical defects that turns into the minimal or QES surfaces in the RT or QES formulas. In order for these surfaces to live in the Lorentzian section, the $t = 0$ Lorentzian slice of the quotient geometry must therefore contain the conical defects. Ultimately, whether or not the new saddles should genuinely contribute to Γ_Q will depends on the desired properties of the theory; we will revisit this question in Section 7.7.

As a final note, the need to study the $m < 1$ wormholes completely numerically somewhat obstructs the origin of the new branch of solutions and renders it difficult to completely scan the parameter space in a controlled way. To shed some light into the qualitative features exhibited by Figures 7.14 and 7.15, in Appendix D.4 we study a simpler model of JT gravity coupled to end-of-the-world branes, similar to that considered in [50]. This model can be studied analytically (up to a single transcendental equation), and we find that turning on a brane tension gives rise to stable wormholes for $m > 1$ and to two branches of solutions for the wiggle when $m < 1$, as we have seen in JT coupled to a massless scalar. However, it does not exhibit the stable wormholes at $m < 1$ we have found, so it is not sufficiently rich to exhibit our desired behavior. Nevertheless, it is instructive in showing explicitly why two branches of solutions for the wiggle can exist when $m < 1$.

7.5.4 It was the Best of Saddles, it was the Worst of Saddles

Assuming that we restrict to perturbations with a real Lorentzian section in the sense discussed above (in particular, with the defects contained on the moment of time symmetry of the Lorentzian section), we may now compute the saddle-point approximation

of the effective action $\widehat{I}_m$ by putting the modulus on-shell, and consequently obtain Γ_Q. The action $\widehat{I}_m$ as a function of m is shown in Figure 7.17. Besides the aforementioned fact that the saddles, when they exist, do not appear to persist below $m \approx 1/2$ (at least in the parameter space we have been able to probe numerically), an additional noteworthy feature is that for intermediate values of $\sqrt{\beta}J$ these saddles also do not extend all the way to $m = 1$: there can be an isolated interval of saddles for $1/2 < m < m_{\max} < 1$. Regardless, the upshot is that in the classical $S_0 \to \infty$ limit in which the dominance of geometries is controlled solely by the topological term in the action, the replica trick (7.10) gives the quenched generating functional

$$\Gamma_Q = -\widehat{I}_{1/2}. \tag{7.72}$$

We should of course be clear that our numerics are unable to determine that solutions stop existing *precisely* at $m = 1/2$, so the above equality should really be understood as taking the limit towards the leftmost endpoint of the curves shown in Figure 7.17.

An outstanding question worth considering is whether the minimization involved in the one-step RSB prescription is justified in the present context, given that the minimum of $\widehat{I}_m$ at $m = 1/2$ is not a local minimum, but rather a global extremum at the boundary of the space of solutions: that is, $m_1 = 1/2$ is not a saddle of $\widehat{I}_{m_1}$. To some extent an answer to this question requires a more comprehensive understanding of the replica trick, but there is no obvious need for the minimization over m_1 in the replica trick to be treated on the same footing as the other saddles. Some intuition can be obtained from going back to the original wormhole geometries with integer $m \geq 1$. In these geometries, m_1 is only allowed to take on a discrete set of values (namely, the divisors of m), and a dominant solution is found by minimizing m_1 over this discrete set; there is no sense in which we

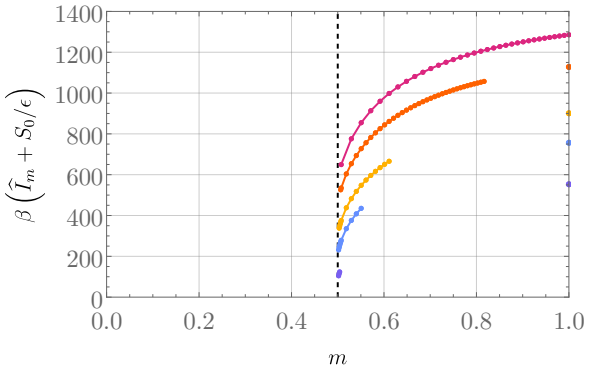


Figure 7.17: The on-shell action $\widehat{I}_m$ obtained by putting both the wiggle and the modulus on shell; from uppermost to lowermost, the curves correspond to $\sqrt{\beta}J = 25$, 23, 20, 18, and 15. Note that the stable saddles we have found do not persist below $m = 1/2$, marked as a dashed black line, and we are unable to find further saddles at smaller values of m. For sufficiently large values of $\sqrt{\beta}J$ saddles exist in the entire region $m \in (1/2, 1]$; decreasing $\sqrt{\beta}J$ shrinks the region in m in which saddles exist, until for $\sqrt{\beta}J \lesssim 15$ we find no saddles at all for $m < 1$ (the isolated points on the right-hand side of the plot are the $m = 1$ solutions, which exist for any value of $\sqrt{\beta}J$).

can look for "saddles" of m_1. When m, and consequently m_1, are continued away from the integers in the replica trick, it is reasonable to expect that the minimization over m_1 should remain a mere global minimization with no requirement that m_1 be a saddle. (Put differently: we must look for saddles in the wiggle and the modulus ξ_0 because these are degrees of freedom we integrate over in the path integral, and we approximate this integral by a saddle-point approximation. On the other hand, m_1 represents degrees of freedom that are *summed* over in the path integral, i.e. different topologies, so we only need to minimize with respect to m_1 with no need for a saddle.)

With this interpretation understood, we can now compare the quenched and annealed generating functionals using these new saddles at $m = 1/2$: we show Γ_Q and Γ_A in Figure 7.18 as functions of the "temperature" $T \equiv 1/\beta$ (though recall that these are not thermal states). Since our data is consistent with stable saddles for $m < 1$ existing to arbitrarily low temperatures, we expect Γ_Q and Γ_A to continue to be distinct down

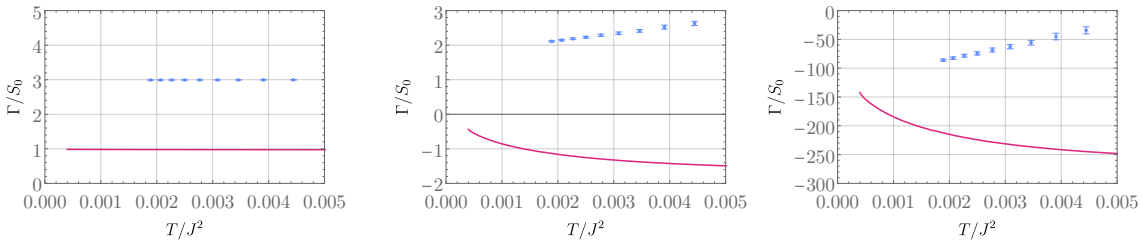


Figure 7.18: The low-temperature quenched and annealed generating functionals Γ_Q (blue points) and Γ_A (red line). From left to right, the plots correspond to $J/\sqrt{S_0} = 0.1$, 1, and 10. The error bars on the blue points are a rough estimate of the uncertainty introduced in extrapolating the data to $m = 1/2$ as required in (7.72).

to $T = 0$; Figure 7.18 merely displays our results to the lowest temperatures for which we have generated data. On the other hand, the apparent lack of stable saddles with $m < 1$ at high temperatures implies that the only saddle that can contribute to Γ_Q is the disk, and so we would expect that at temperatures higher than those shown in Figure 7.18, $\Gamma_Q = \Gamma_A$. But if this were the case, it is clear from the figure that Γ_Q would be discontinuous at this transition temperature, which is pathological behavior! This discontinuity stems from the fact that the $m = 1/2$ saddles that contribute to Γ_Q do not smoothly exchange dominance with the $m = 1$ saddle that defines Γ_A, but rather they dominate *immediately* as soon as they start existing. What are we to make of this apparent discontinuity?

One possibility is to object to the $m < 1$ saddles in the first place: after all, if we modify our stability criterion to require that saddles be stable under any real Euclidean perturbation, rather than just perturbations that admit a real Lorentzian section, then the $m < 1$ saddles are unstable. Without this stability condition, we would thus trivially find that $\Gamma_Q = \Gamma_A$ at all temperatures. However, we have discussed above our reasoning for taking the requirement of stability under real Lorentzian perturbations seriously, so we do not find this objection compelling. What we deem more likely is that the story so far is incomplete: as alluded to earlier, the structure of solutions to the equations of motion of the JT + CFT model at $m < 1$ is richer than one might have otherwise

expected, and unfortunately a numerical analysis cannot ensure that we have found all relevant saddles. It may be that there are additional saddles we have missed at smaller values of $\sqrt{\beta}J$ that allow F_Q to transition *continuously* to F_A at high temperatures. In fact, there could even be additional saddles that are not captured by our one-step RSB ansatz. As we will see in Section 7.6, it may even be that quantum effects are needed to give rise to new semiclassical saddles, which may then yield a continuous Γ_Q. Or perhaps, since our states are far from thermal there is no reason to require Γ_Q to be continuous after all, and the discontinuity is simply an interesting quirk of the theory.

Because of this incompleteness, we remain agnostic on what the "correct" answer for the quenched generating functional is. We claim that our main finding is not the particular functional form of Γ_Q shown in Figure 7.18, but rather the existence of new saddles at $m < 1$ whose stability properties depend crucially on whether or not perturbations about them are required to admit a real Lorentzian continuation. Additional investigation, either into the space of saddles, the structure of the replica trick, or non-gravitational models of JT + matter, is needed to determine what the right answer for Γ_Q actually is.

7.6 Quantum Corrections: The Adventures of Operator Twist

To assess whether our purely classical results are robust to semiclassical corrections, we now investigate the effect of turning on quantum corrections to the scalar field. Because the scalar only couples to the wiggle through the boundary sources ψ_∂, and because these sources have already been incorporated into the classical analysis, excitations of the scalar about its classical background will obey homogeneous Dirichlet boundary con-

ditions. Such quantum excitations will not couple to the wiggle, so we need only concern ourselves with the coupling to the modulus. To do so, we construct the effective matter action

$$\widehat{I}^{\text{quant}}[g_m] = -\frac{1}{m} \ln Z^{\text{quant}}[g_m], \tag{7.73}$$

where $Z^{\text{quant}}[g_m]$ is the scalar partition function on the wormhole geometry g_m, which will depend on the modulus D. Note that we have included the prefactor of $1/m$ to match conventions with the classical quotient action $\widehat{I}_m$.

After going to the quotient space, this partition function is computed in a standard way [214] by inserting appropriate twist operators at the locations of the conical defects:

$$Z^{\text{quant}}[g_m] = \int \mathcal{D}\psi\, \Sigma(p_1)\Sigma^*(p_2) e^{-I_{\text{mat}}[\widehat{g}_m]} \equiv Z^{\text{quant}}[\widehat{g}_m]\langle \Sigma(p_1)\Sigma^*(p_2)\rangle, \tag{7.74}$$

where p_1 and p_2 are the locations of the conical defects, Σ and Σ^* are the twist and anti-twist operators, which identify fields in the clockwise and anti-clockwise directions between the different Riemann sheets of the m-copy replica manifold, and the normalization factor $Z^{\text{quant}}[\widehat{g}_m]$ is the matter partition function on the quotient geometry (with no twist operator insertions). Because the massless scalar is a CFT, the partition function $Z^{\text{quant}}[\widehat{g}_m]$ can be computed straightforwardly from the conformal anomaly $\int R[\widehat{g}_m]$, which is topological and hence independent of the modulus D (though it will depend on m). Thus the normalization factor only adds an m-dependent additive constant to $\widehat{I}^{\text{quant}}$, so it can safely be ignored. Instead, all nontrivial dependence on D is contained in the correlator $\langle \Sigma(p_1)\Sigma^*(p_2)\rangle$.

In order to compute Z^{quant}, we will make two simplifications. First, we will work perturbatively about $m = 1$, since this is the regime in which the correlator $\langle \Sigma\Sigma^*\rangle$ is most tractable. Second, we will work in the limit of a CFT with large central charge.

This limit is compatible with the regime in which our classical $m < 1$ saddles exist in the following sense. In Section 7.5 we worked with a single scalar field (central charge $c = 1$ in appropriate conventions) with boundary sources with amplitude J, giving an on-shell classical action proportional to J^2. But we could equally well have worked with, say, many independent scalar fields with total central charge c, giving an on-shell classical action proportional to cJ^2. Consequently, the results of Section 7.5.3 can be interpreted as showing that $m < 1$ saddles can be supported by a large-c CFT as long as cJ^2 is sufficiently large relative to $1/\beta$.

With these assumptions understood, we turn to computing $\langle \Sigma\Sigma^* \rangle$. To do so, we first take the quotient geometry to be the unit disk with the defects located at $z_\pm = \pm a$, as in the left diagram of Figure 7.4. Here a is the modulus that sets the proper distance D between the defects. The metric takes the form

$$ds^2 = e^{2\sigma(z,\bar{z})} dz\, d\bar{z} = \frac{4}{(1 - z\bar{z})^2} \left(1 + \mathcal{O}(\varepsilon)\right) dz\, d\bar{z}, \tag{7.75}$$

where we have defined $\varepsilon \equiv 1 - m$. We then map the disk to the upper half-plane via the Möbius transformation $z = (1 + iw)/(i + w)$, giving the metric

$$ds^2 = \frac{1}{4} e^{2\sigma(z,\bar{z})} |z - i|^4 dw\, d\bar{w}. \tag{7.76}$$

Since the twist operators are primaries with scaling dimension [214][12]

$$\Delta_m = \frac{c}{12}\left(m - \frac{1}{m}\right) = -\frac{c}{6}\,\varepsilon + \mathcal{O}(\varepsilon^2), \tag{7.77}$$

[12]Note that we are analytically continuing the scaling dimension from $m > 1$ down to $m < 1$, which should be valid since we are only interested in the behavior in a neighborhood of $m = 1$ and we expect the scaling dimension to be analytic in m at $m = 1$. Analogous arguments apply for the rest of this section.

the two-point function of twist operators on the quotient geometry $\widehat{g}_m$ can be related to that on the flat upper half-plane by a standard scaling under Weyl and conformal transformations:

$$\langle \Sigma(a)\Sigma^*(-a)\rangle = \left(\frac{1}{4}\, e^{\sigma(a,a)+\sigma(-a,-a)}(1+a^2)^2\right)^{-\Delta_m} \langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\mathrm{UHP}}, \tag{7.78a}$$

$$= \left[1 + \frac{c}{3}\ln\left(\frac{1+a^2}{1-a^2}\right)\varepsilon + \mathcal{O}(\varepsilon^2)\right] \langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\mathrm{UHP}}, \tag{7.78b}$$

where $w_\pm = (1 \mp ia)/(\pm a - i)$ are the locations of the twist operator insertions in the upper half-plane.

We thus need only to compute the correlator $\langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\mathrm{UHP}}$ on the (flat) upper half-plane. This setting corresponds to working in a boundary CFT (BCFT); this will typically induce excitations on the boundary due to the breaking of full conformal symmetry, as realized by the presence of a bulk to boundary OPE. To evaluate $\langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\mathrm{UHP}}$, we can therefore follow [215]. The basic idea is that the product $\Sigma(w_+)\Sigma^*(w_-)$ can be expanded in terms of an OPE. Because we have a BCFT, there are two channels the OPE can be expanded in. The first of these is the boundary channel, where we expand each bulk operator in terms of a bulk to boundary OPE, corresponding to the boundary operator spectrum of the bulk operator. This would convert the twist two-point function into a sum over boundary two-point functions. There is also a bulk channel in which we expand the product of bulk operators in terms of a bulk OPE; in this channel the two-point function becomes a sum over bulk one-point functions. In general, these two channels are only valid in their respective OPE limits: the boundary channel is valid when the twist operators are brought closer to the boundary than they are to each other, while the bulk channel is valid when the twist operators are brought closer to each other than they are to the boundary. However, if we work in the $c \to \infty$ limit

then we can use the channels for all values of $\eta \in [0,1]$, with the two channels exchanging dominance at some η_*, where η is the cross-ratio:

$$\eta \equiv \frac{(w_+ - \overline{w}_+)(w_- - \overline{w}_-)}{(w_+ - \overline{w}_-)(w_- - \overline{w}_+)} = \left(\frac{1-a^2}{1+a^2}\right)^2. \tag{7.79}$$

Moreover, we can approximate the OPE expansion in a given channel by the identity block of that channel.

This calculation of $\langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\text{UHP}}$ in the $c \to \infty$ limit was done explicitly in [215] to linear order in ε, and yields

$$\langle \Sigma(w_+)\Sigma^*(w_-)\rangle_{\text{UHP}} = 1 - \frac{\varepsilon c}{6}\left(2\ln \epsilon_{\text{UV}} - \ln \min(\eta, 1-\eta)\right) + \mathcal{O}(\varepsilon^2), \tag{7.80}$$

where ϵ_{UV} is a cutoff that regulates the region around the operator insertions. Consequently, using (7.78b), (7.73), and (7.74) we obtain the matter effective action

$$\widehat{I}_m^{\text{quant}}[a] = -\frac{c(m-1)}{3}\max\left(0, \ln\left(\frac{1-a^2}{2a}\right)\right) + \mathcal{O}(m-1)^2, \tag{7.81}$$

up to overall additive constants independent of a. This action is shown in Figure 7.19. Note that for $a > \sqrt{2}-1$, $\widehat{I}_m^{\text{quant}}[a]$ is independent of a: this is to be expected from the fact that $a > \sqrt{2}-1$ corresponds to the boundary OPE channel $\eta < 1/2$ in which the two-point function is dominated by the proximity of the twist operators to the boundary rather than to each other. When mapped back to hyperbolic space the distance from the twist operators to the boundary is renormalized to a constant, and hence we expect that correlators dominated by this channel should be independent of a.

For $m > 1$, $\widehat{I}_m^{\text{quant}}[a]$ is negatively divergent at $a = 0$. For the double-trumpet case $m = 2$, this behavior (exhibited by bosonic matter) is well-understood and is due to

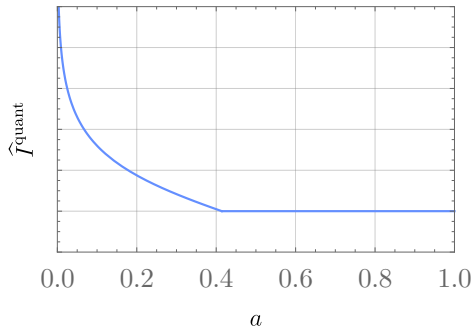


Figure 7.19: The large-c, leading-order in $m-1$ behavior of the quantum effective action $I_m^{\text{quant}}[a]$ for $m < 1$ (up to various additive constants that we have ignored). Note the divergence at $a = 0$ due to a Casimir effect.

a divergent Casimir energy as the throat of the wormhole pinches off. This divergence is extremely destabilizing: the quantum effective action is unbounded below, suggesting that a classically-stable double-trumpet cannot remain globally stable under the inclusion of semiclassical effects; see e.g. [216, 208]. The action (7.81) suggests that this destabilizing Casimir energy persists for wormholes with $m > 1$, and presumably means that any classical saddles with $m > 1$ do not remain well-defined saddles of a semiclassical theory (though they may remain metastable).

However, for the $m < 1$ wormholes relevant to the quenched generating functional Γ_Q, $\widehat{I}_m^{\text{quant}}$ is *positively* divergent at $a = 0$. Hence quantum effects do not render these wormholes pathologically unstable. Whether or not a classical saddle remains stable under the inclusion of quantum corrections then depends on the details of the combined action $\widehat{I}_m[a] + \widehat{I}_m^{\text{quant}}[a]$. For instance, a classical saddle at some $a > \sqrt{2} - 1$ will remain a semiclassical saddle, but a classical saddle at $a < \sqrt{2} - 1$ may or may not remain a semiclassical saddle. Conversely, it is possible for semiclassical saddles to appear even when no classical ones existed: for example, if $d\widehat{I}_m/da > 0$ for all a, then no classical saddles exist, but the fact that $d\widehat{I}_m^{\text{quant}}/da \to -\infty$ at $a = 0$ (and that $d\widehat{I}_m^{\text{quant}}/da = 0$ for $a > \sqrt{2}-1$) is sufficient to ensure that a semiclassical saddle will exist. Indeed, using

the fact that a is roughly inversely related to the parameter ξ_0 used in earlier sections, it is tantalizing to note that of the plots shown in Figures 7.14 and 7.15, $d\widehat{I}_m/da > 0$ on all branches with no stable modulus but with a stable wiggle. Hence it is entirely possible that semiclassical effects will stabilize some of the classically unstable wormholes we have looked at – this may even be sufficient to remedy the apparent discontinuity of Γ_Q in Figure 7.18.

Investigating these possibilities in more detail will require us to construct the quantum effective action away from a neighborhood of $m = 1$, which we leave to future work. The upshot is that the same quantum effects that manifest pathologies in $m > 1$ wormholes seem to be benign, or even beneficial, in the $m < 1$ wormholes necessary to construct the quenched generating functional.

7.7 Discussion: Great Expectations

We have investigated the potential contributions of connected saddles to the quenched generating functional Γ_Q in the semiclassical approximation by developing an LM-inspired procedure for replica symmetry breaking that admits continuation to zero replicas. Using this technology, we have shown that in a model of JT gravity coupled to a massless scalar, a computation of Γ_Q reveals the existence of a new on-shell wormhole that dominates over the usual disconnected contribution that gives the annealed result Γ_A. This new wormhole is unstable to arbitrary Euclidean perturbations, but it is stabilized by restricting to perturbations that admit a real Lorentzian section with a moment of time symmetry on which the conical defects live. Moreover, we have shown that quantum effects do not destabilize this wormhole in the same way that they destabilize conventional wormholes with $m > 1$, at least when working perturbatively about $m = 1$. In fact, they may even

have a stabilizing influence, potentially making semiclassical saddles appear where no classical ones existed.

It is not clear whether the (in)stability of the new saddle is a feature or a problem. On the one hand, the Euclidean gravitational path integral is notorious for its superior intelligence compared with its Lorentzian counterpart, as manifested in e.g. computations of the black hole entropy by Gibbons and Hawking [217]. Perhaps, then, we should take a purely Euclidean perspective, excluding saddles like the connected one we have found that are unstable under some Euclidean perturbations. On the other hand, due to the conformal mode problem, a strictly Euclidean gravitational path integral is in fact divergent (and in JT gravity, inconsistent with canonical quantization): the contour of integration of the conformal mode (or, in JT gravity, the dilaton) needs to be Wick rotated to the imaginary axis to give sensible results. Perhaps this is a hint that a Lorentzian treatment is more fundamental after all, as suggested in [218]. If so, then only behavior under Lorentzian perturbations is relevant, and the quenched generating functional should be dominated by the connected saddle after all.

Ultimately, whether or not the new saddle should contribute to Γ_Q should be diagnosed by whether its inclusion yields the desired physics. In prior discussions of replica wormholes, arguments against the inclusion of Euclidean wormholes were countered by appeal to unitarity. Is there an analogous fundamental physical guiding principle that justifies (or excludes) these wormholes from contributing to the replica trick (7.10) for Γ_Q? Unfortunately, any such principle is not immediately evident. Because the states that we consider are not thermal due to matter sources that break the $U(1)$ Euclidean time-translation symmetry, the generating functional is no longer required to obey typical constraints of thermal states. Some other constraints do exist of course: for instance, correlation functions computed from the generating functional must satisfy a number of

properties, from large-distance (or late-time) behavior to triangle inequalities. Some of these are manifest already semiclassically, while others (e.g. very late-time recurrences) are expected to be inherited from the microscopic, nonperturbative description. An investigation of whether or not these constraints are satisfied by certain saddles requires an analysis of observables computed from the generating functional. For instance, [173] found that a full nonperturbative calculation of correlation functions in JT gravity reproduces the correct late time recurrences predicted in [219]; the problem – often dubbed the Maldacena information paradox – is that the standard semiclassical gravity calculation predicts no late time recurrences. It would be interesting to investigate whether replica wormhole contributions to the generating functional can resolve this tension in a similar way to how connected topologies resolved the tension between the Page and Hawking calculations of the entropy of Hawking radiation.

Alternatively, some guidance could be provided by appealing to dual models. For example, pure JT gravity is dual to a double-scaling limit of a matrix model; such models give nonperturbative completions of JT that can be used to explicitly compute the quenched generating functional and confirm that it differs from its annealed counterpart at low temperatures [191, 192, 193, 194, 195, 196, 197, 198, 199]. Is there an analogous dual to JT gravity coupled to matter that would allow us to compute the quenched free energy directly – without reference to replica tricks or wormholes – in a semiclassical limit? For example, are there sources we could turn on in the SYK model that would approximately reproduce, in an appropriate low-energy limit, the Schwarzian theory of JT coupled to matter? If so, then an explicit computation of the quenched generating functional should reveal whether or not our prescription for computing Γ_Q via the replica trick should include contributions from the new saddles we have found.

Identifying the contributing saddle to the generating functional is part of a larger

quest to understand the rigorous underpinning of the gravitational path integral in general. Without a guiding principle such as unitarity for the von Neumann entropy to help us in picking the correct saddle for Γ_Q, for now we simply raise this question and bring up the possibility of future investigations into stability of saddles as a potential avenue for furthering our understanding of the gravitational path integral.

Appendix A

Appendix to Chapter 2

A.1 Extremal structure of H_n for $1 \leq n \leq 5$

Here we summarize the extremal structure of H_n for all $1 \leq n \leq 5$ by showing representatives of every orbit of both facets and extreme rays. Representatives of each orbit are picked as their lexicographical minimum.[1] For extreme rays, we also present their minimum realizations, exhibiting graphs where only edges of nonzero weight are shown. At every n, we only include elements which are genuinely new and not coming from zero-liftings. This is because these should always be included – by Proposition 2.5 the zero-lift of rays preserves all extreme rays, while by Theorem 2.5 the zero-lift of inequalities preserves all superbalanced facets. As for SA, Proposition 2.8 guarantees that precisely only instances involving just singletons in $[n; N]$ give rise to facets. It will thus be convenient to present results in increasing order of n.

[1] The only exception to this is inequality 1 in Table A.1, which is chosen for symmetry reasons.

A.1.1 Facets

At $n = 1$ one just has one single-element orbit of a nonnegativity facet,

$$S_1 \geq 0. \tag{A.1}$$

For $n = 2$, the cone is a simplex with 3 facets in a single orbit of SA,

$$S_1 + S_2 \geq S_{12}. \tag{A.2}$$

Lifting to $n = 3$, one gets 6 facets in the orbit of the trivial zero-lift of SA. The cone becomes again a simplex due to the appearance of the new, totally symmetric facet of MMI

$$S_{12} + S_{13} + S_{23} \geq S_1 + S_2 + S_3 + S_{123}. \tag{A.3}$$

There are no genuinely new inequalities for $n = 4$. The trivial zero-lift of the SA facet gives a length-10 orbit. Every zero-lift of (A.3) in fact lands on the same MMI orbit, which consists of another 10 facets. In total, H_4 thus has 20 facets and is not simplicial anymore.

It is at $n = 5$ that H_n begins to exhibit a richer structure. The trivial zero-lift of SA now contributes an orbit of 15 facets. The trivial zero-lift of MMI gives an orbit with 20 facets. There is now another inequivalent zero-lift of MMI which gives an orbit of length 45. Besides these, there are 5 orbits of genuinely new facets, given in Table A.1.

In order, these give rise to orbits of lengths 72, 90, 10, 60 and 60. Together with the 80 facets coming from SA and MMI, there are a total of 372 inequalities in the H-representation of H_5. Other than inequality 1, usually referred to as *cyclic* due to its symmetry under $i \to i + 1 \mod n$ which is manifest in the given representative, these

inequalities are poorly understood.

1. $S_{123} + S_{234} + S_{345} + S_{145} + S_{125} \geq S_{12} + S_{23} + S_{34} + S_{45} + S_{15} + S_{12345}$
2. $S_{14} + S_{23} + S_{125} + S_{135} + S_{145} + S_{245} + S_{345}$
   $\geq S_1 + S_2 + S_3 + S_4 + S_{15} + S_{45} + S_{235} + S_{1245} + S_{1345}$
3. $S_{123} + S_{124} + S_{125} + S_{134} + S_{135} + S_{145} + S_{235} + S_{245} + S_{345}$
   $\geq S_{12} + S_{13} + S_{14} + S_{25} + S_{35} + S_{45} + S_{234} + S_{1235} + S_{1245} + S_{1345}$
4. $2S_{123} + S_{124} + S_{125} + S_{134} + S_{145} + S_{235} + S_{245}$
   $\geq S_{12} + S_{13} + S_{14} + S_{23} + S_{25} + S_{45} + S_{1234} + S_{1235} + S_{1245}$
5. $3S_{123} + 3S_{124} + S_{125} + S_{134} + 3S_{135} + S_{145} + S_{234} + S_{235} + S_{245} + S_{345}$
   $\geq 2S_{12}+2S_{13}+S_{14}+S_{15}+S_{23}+2S_{24}+2S_{35}+S_{45}+2S_{1234}+2S_{1235}+S_{1245}+S_{1345}$

Table A.1: Representative inequalities in each of the 5 new orbits of facets of H_5.

A.1.2 Extreme rays and minimum graph realizations

Extreme rays and their minimum realizations in $K_{N_{min}}$ will be provided. Extreme rays will be labelled by a tuple $(n, N_{min} - n, \sigma)$, where $\sigma \geq 1$ is just an integer counting orbits at fixed n and N_{min} by listing their representatives lexicographically. Notice that $N_{min} - n \geq 1$ counts the number of bulk vertices needed in the minimum representation, plus the sink. For clarity, S-vector entries S_I will be separated by a semicolon whenever the cardinality of I increases.

At $n = 1$ there is a single extreme ray with minimum realization the Bell pair in Figure A.1,

$$S_{(1,1,1)} = (1). \tag{A.4}$$

The $n = 2$ cone has just the length-3 orbit of zero-lifts of the Bell-pair extreme ray $(1, 1, 1)$. For $n = 3$, the Bell-pair zero-lift now gives an orbit of 6 extreme rays. A new

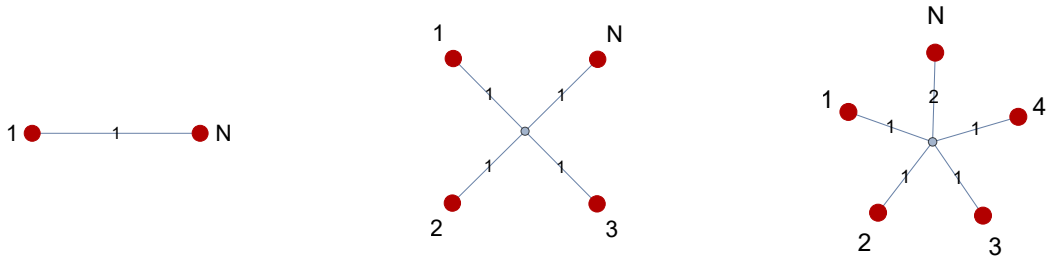
totally symmetric extreme ray appears. It has a star-graph minimum realization shown in Figure A.1 and reads

$$S_{(3,2,1)} = (1,1,1;\ 2,2,2;\ 1). \tag{A.5}$$

Lifting to $n = 4$ we get orbits of 10 extreme rays from $(1,1,1)$ and another 5 from $(3,2,1)$. A genuinely new length-5 orbit of extreme rays appears,

$$S_{(4,2,1)} = (1,1,1,1;\ 2,2,2,2,2,2;\ 3,3,3,3;\ 2), \tag{A.6}$$

which again has a star graph as minimum realization, as shown in Figure A.1.



$(1,1,1)$ $(3,2,1)$ $(4,2,1)$

Figure A.1: Minimum realizations for extreme rays in each orbit of H_n for $n \leq 4$.

At $n = 5$, lifted extreme rays become a minority. Extreme rays $(1,1,1)$, $(3,2,1)$ and $(4,2,1)$ respectively zero-lift to orbits of lengths 15, 15 and 30, totaling just 60 extreme rays. It turns out H_5 has 2267 in total, so all the others are genuinely new ones. They fall into 16 orbits, which we now present by increasing number of bulk vertices needed in their minimum realization. There are 4 distinct orbits of extreme rays realizable in a

star graph,

$$
\begin{aligned}
S_{(5,2,1)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 3,3,3,3,3,3,3,3,3,3;\ 2,2,2,2,2;\ 1),\\
S_{(5,2,2)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 3,3,3,3,3,3,3,3,3,3;\ 4,4,4,4,4;\ 3),\\
S_{(5,2,3)} &= (1,1,1,1,2;\ 2,2,2,3,2,2,3,2,3,3;\ 3,3,4,3,4,4,3,4,4,4;\ 4,3,3,3,3;\ 2),\\
S_{(5,2,4)} &= (1,1,1,2,2;\ 2,2,3,3,2,3,3,3,3,4;\ 3,4,4,4,4,5,4,4,5,5;\ 5,5,4,4,4;\ 3),
\end{aligned}
\tag{A.7}
$$

with respective orbit lengths 1, 6, 15 and 60. They can all be represented on the star graph shown in Figure A.2*(1)*, with appropriate weight assignments as specified in Table A.2. There are 6 orbits which require 2 bulk vertices,

$$
\begin{aligned}
S_{(5,3,1)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 1,3,3,3,3,3,3,3,3,3;\ 2,2,2,2,2;\ 1),\\
S_{(5,3,2)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 2,2,2,3,3,3,3,3,3,3;\ 2,2,2,2,2;\ 1),\\
S_{(5,3,3)} &= (1,1,1,1,2;\ 2,2,2,3,2,2,3,2,3,3;\ 3,3,2,3,4,4,3,4,4,4;\ 4,3,3,3,3;\ 2),\\
S_{(5,3,4)} &= (1,1,2,2,2;\ 2,3,3,3,3,3,3,4,4,4;\ 4,4,4,3,3,5,5,5,5,4;\ 4,4,4,3,3;\ 2),\\
S_{(5,3,5)} &= (2,2,2,2,3;\ 4,4,4,5,4,4,5,4,5,5;\ 4,6,5,6,5,7,6,7,7,7;\ 6,5,5,5,5;\ 3),\\
S_{(5,3,6)} &= (3,3,3,3,3;\ 6,6,6,6,6,6,6,6,6,6;\ 5,7,7,7,7,9,9,9,9,9;\ 6,6,6,6,6;\ 3),
\end{aligned}
\tag{A.8}
$$

with respective orbit lengths 10, 60, 90, 180, 360 and 90. These can be represented on graphs in Figures A.2*(2)* to A.2*(5)* following Table A.2. There are 6 orbits which require

3 bulk vertices,

$$\begin{aligned}
S_{(5,4,1)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 2,2,2,2,3,3,3,3,3,3;\ 2,2,2,2,2;\ 1),\\
S_{(5,4,2)} &= (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 2,2,3,3,2,3,3,3,3,3;\ 2,2,2,2,2;\ 1),\\
S_{(5,4,3)} &= (2,2,2,2,3;\ 4,4,4,5,4,4,5,4,5,5;\ 4,6,5,6,7,5,6,7,7,7;\ 6,5,5,5,5;\ 3),\\
S_{(5,4,4)} &= (3,3,3,3,3;\ 6,6,6,6,6,6,6,6,6,6;\ 5,7,7,7,9,9,9,7,9,9;\ 6,6,6,6,6;\ 3).
\end{aligned} \tag{A.9}$$

with respective orbit lengths 180, 60, 360 and 360. These are realizable in graphs in Figures A.2*(6)* to A.2*(8)* following Table A.2.

Finally, there is an orbit of length 360 with 4 bulk vertices,

$$S_{(5,5,1)} = (3,3,3,3,3;\ 6,6,6,6,6,6,6,6,6,6;\ 5,7,7,7,9,7,9,9,9,9;\ 6,6,6,6,6;\ 3), \tag{A.10}$$

and another one of length 15 with 5 bulk vertices,

$$S_{(5,6,1)} = (1,1,1,1,1;\ 2,2,2,2,2,2,2,2,2,2;\ 2,2,3,3,2,2,3,3,3,3;\ 2,2,2,2,2;\ 1). \tag{A.11}$$

These are respectively realizable in graphs in Figures A.2*(7)* and A.2*(8)* following Table A.2.

In summary, H_5 consists of 2267 extreme rays in 19 orbits, 2207 of which lie in 16 orbits new to $n = 5$. Note that apart from the Bell pair $(1,1,1)$ in Figure A.1, there are no edges between terminals in any of the minimum extreme-ray representations. Each is planar except for $(5,6,1)$ in Figure A.2*(8)*, which can be embedded on a torus. Terminals have degree at most 3.

A.2 Miscellaneous examples

Convexity of $H^+_{N,n}$ and $H_{N,n}$: For $n = 5$, consider extreme rays $S_{(5,2,2)}$ and the zero-lift of $S_{(3,2,1)}$, both of which are realizable in K_7. Using the ILP method in Section 2.5 we can determine that their sum is not. Therefore, neither $H^+_{7,5}$ nor $H_{7,5}$ is convex without the convex operator applied. The minimum realization of their sum is in K_8, and has edges $\{(1,6),(2,6),(3,6),(4,7),(5,7),(6,7),(7,8)\}$ with respective weights $\{2,2,2,1,1,4,4\}$.

Extreme rays of $P_{N,n}$ and $H_{N,n}$: For $n = 5$ in K_9, consider these three feasible (S, w) pairs:

$$S^1 = (3,4,3,3,3;\ 7,6,6,6,7,7,7,6,6,6;\ 6,8,6,9,9,7,8,8,8,9;\ 5,5,5,6,5;\ 2),$$

$$S^2 = (1,2,1,1,1;\ 3,2,2,2,3,3,3,2,2,2;\ 2,2,2,3,3,3,2,2,2,3;\ 1,1,1,2,1;\ 0),$$

$$S^3 = (2,2,2,2,2;\ 4,4,4,4,4,4,4,4,4,4;\ 4,6,4,6,6,4,6,6,6,6;\ 4,4,4,4,4;\ 2),$$

$$w^1 = (0,0,0,0,2,1,0,0;\ 0,0,0,3,1,0,0;\ 0,0,2,0,1,0;\ 0,1,1,1,0;\ 1,2,0,0;\ 0,1,0;\ 1,0;\ 2),$$

$$w^2 = (0,0,0,0,1,0,0,0;\ 0,0,0,2,0,0,0;\ 0,0,1,0,0,0;\ 0,1,0,0,0;\ 1,0,0,0;\ 0,0,0;\ 0,0;\ 0),$$

$$w^3 = (0,0,0,0,1,0,1,0;\ 0,0,0,0,0,2,0;\ 0,0,0,1,1,0;\ 0,1,1,0,0;\ 2,0,0,0;\ 1,1,0;\ 1,2;\ 0).$$

Here (S^1, w^1) is an extreme ray of $P_{9,5}$. However, its S coordinates have $S^1 = S^2 + S^3$, so the latter cannot be an extreme ray of $H_{9,5}$ and hence of H_5 either (cf. Theorem 2.2(b)).

Extreme Ray	Graph	Terminal Edges						Edge Weights															
$(\cdot,\cdot,\cdot)$	#	N	1	2	3	4	5	w_N	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}	w_{13}	w_{14}	w_{15}
$(5,2,1)$	1							1	1	1	1	1	1										
$(5,2,2)$	1							3	1	1	1	1	1										
$(5,2,3)$	1							2	1	1	1	1	2										
$(5,2,4)$	1							3	1	1	1	2	2										
$(5,3,1)$	2							1	1	1	1	1	1	1									
$(5,3,2)$	3							2	2	2	1	1	1	1	1	1	1						
$(5,3,3)$	2				w_5		w_3	2	1	1	2	1	1	2									
$(5,3,4)$	4							2	1	1	2	1	1	1	1	1							
$(5,3,5)$	3	w_1		w_N		w_4	w_2	2	3	2	1	1	2	1	1	1	1						
$(5,3,6)$	5							3	3	2	2	2	2	1	1	1	1	1					
$(5,4,1)$	6							2	1	2	1	1	1	1	1	1	1	1	1				
$(5,4,2)$	7							1	2	1	1	2	2	1	1	1	1	1	1				
$(5,4,3)$	7	w_1		w_N	w_3	w_4	w_2	1	3	1	1	2	3	2	1	1	1	1	1				
$(5,4,4)$	8							3	2	2	1	2	2	1	1	1	1	1	1	1	1		
$(5,5,1)$	9							3	3	3	1	1	3	2	2	1	1	1	1	1	1	1	
$(5,6,1)$	10							1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table A.2: Extreme ray vs graph cross reference table for Figure A.2. The columns under "Terminal Edges" specify which edge w_j is incident to each extended terminal $i \in [5;N]$, with a blank entry indicating $j = i$. Edge weights give all extreme rays listed throughout Appendix A.1.2, up to overall scaling.

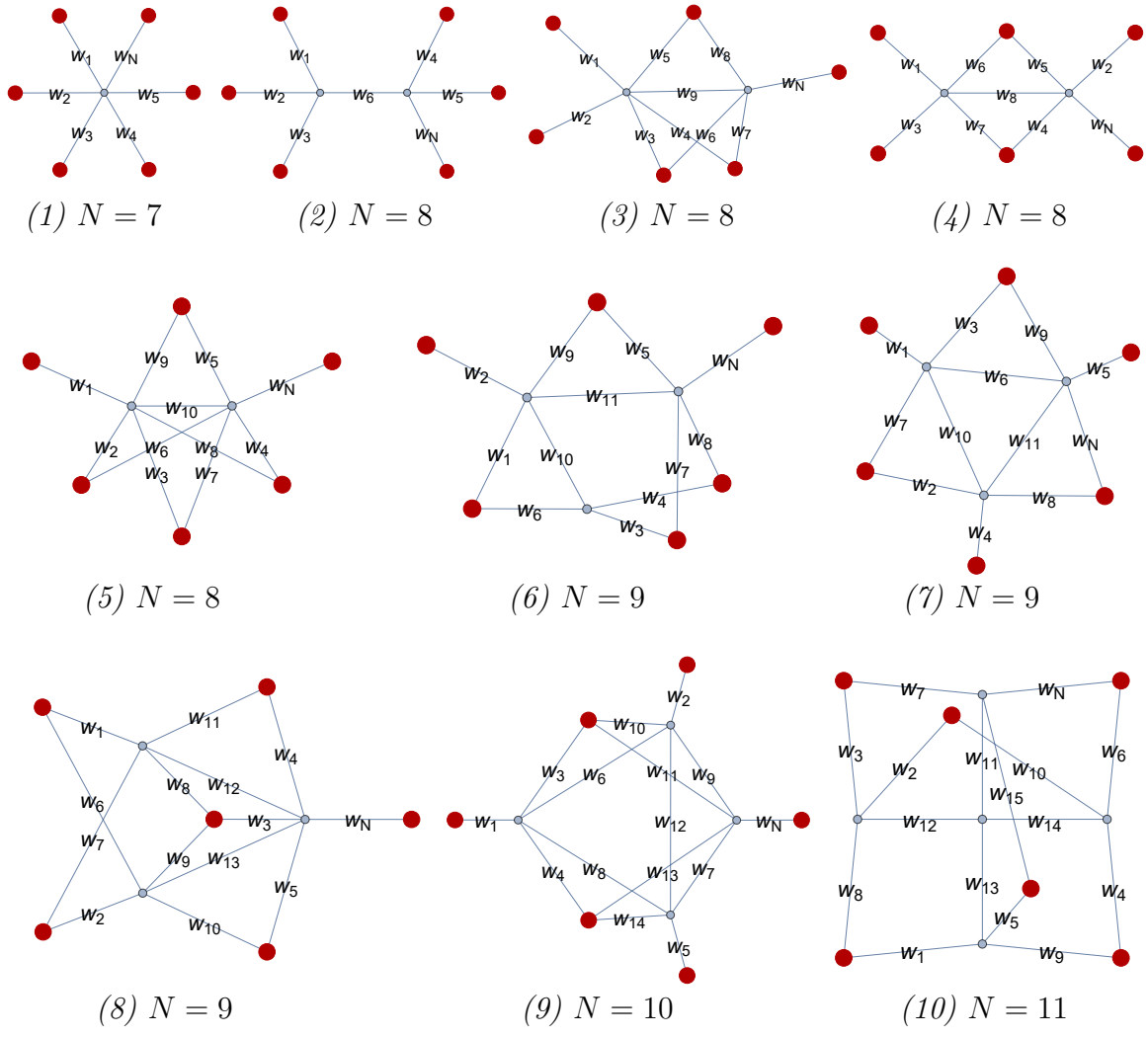


Figure A.2: Minimum realizations for extreme rays which are new for H_5.

Appendix B

Appendix to Chapter 3

B.1 Graph operations

Throughout this work we have been focusing on min-cut structures on topological graph models rather than on graph models and entropy vectors. However, as we have seen, there is a particularly important situation where the two concepts essentially coincide, namely the extreme rays of the HEC. In these cases it is useful to perform certain manipulations which simplify the graph or make manifest some underlying properties. In particular, we have used these operations to obtain realizations of all extreme rays of the HEC_5 and most of the HEC_6 by tree graphs, in order to establish a connection with the extreme rays of the $\mathrm{SAC}_{\mathsf{N}}$. In this appendix we provide a brief presentation of the operations we have used.

The kind of graph operations one is interested in are those which can be used to locally change the vertices, edges, and weights of a general graph while preserving its entropy vector. A few such basic operations were listed in Figure 6 of [42]. It turns out these can all be easily broken down into simpler operations, which we reproduce in

fig. B.1.

In the examples we discussed, however, these operations were not always sufficient to convert a graph model into a tree, and there is another graph operation which has proven remarkably useful (e.g. in obtaining the tree graphs at the bottom of fig. 3.12), the Δ-Y exchange shown in fig. B.2. Since this is a non-trivial operation, we include a proof that it indeed does not alter the entropy vector of a graph model:

Lemma B.1. *The* Δ-Y *exchange operation preserves the entropies.*

Proof. Consider two graph models $\widetilde{G_{\mathsf{N}}}$ and $\widetilde{G_{\mathsf{N}}}'$ related by the operation in fig. B.2, such that $\widetilde{G_{\mathsf{N}}}$ (on the left) contains the 3-cycle (Δ), while $\widetilde{G_{\mathsf{N}}}'$ (on the right) contains the degree-3 vertex (Y). In $\widetilde{G_{\mathsf{N}}}$, let the pertinent vertices be $\{i, j, k\}$, joined by edges with weights $\{w_{ij}, w_{ik}, w_{jk}\}$ (in fig. B.2 labeled more compactly as $\{a, b, c\}$ for simplicity of notation). To obtain $\widetilde{G_{\mathsf{N}}}'$, we preserve the same vertices, delete the edges, add a new vertex σ, and connect it to each of the former vertices by new edges of weights $\{w_{\sigma i}, w_{\sigma j}, w_{\sigma k}\}$ given by

$$w_{\sigma i} = w_{ij} + w_{ik} \qquad w_{\sigma j} = w_{ij} + w_{jk} \qquad w_{\sigma k} = w_{ik} + w_{jk}$$

An arbitrary vertex cut on $\widetilde{G_{\mathsf{N}}}$ may contain any of the 8 subsets of the vertices $\{i, j, k\}$ that make up the 3-cycle. In each case, the contribution from the edges that form the

$\widetilde{G_{\mathsf{N}}}$	weight	$\widetilde{G_{\mathsf{N}}}'$
$\varnothing, \{i, j, k\}$	0	$\varnothing, \{\sigma, i, j, k\}$
$\{i\}, \{j, k\}$	$w_{ij} + w_{ik} = w_{\sigma i}$	$\{i\}, \{\sigma, j, k\}$
$\{j\}, \{i, k\}$	$w_{ij} + w_{jk} = w_{\sigma j}$	$\{j\}, \{\sigma, i, k\}$
$\{k\}, \{i, j\}$	$w_{ik} + w_{jk} = w_{\sigma k}$	$\{k\}, \{\sigma, i, j\}$

Table B.1: Weight contributions to general candidate min-cuts on $\widetilde{G_{\mathsf{N}}}$ and $\widetilde{G_{\mathsf{N}}}'$, depending on exactly which subset of the vertices involved in the graph operation in fig. B.2 is included. The agreement between the two graphs proves lemma B.1.

cycle to the weight of the corresponding cut is given in table B.1. In $\widetilde{G_{\mathsf{N}}}'$ there are instead 16 possible subsets of $\{\sigma, i, j, k\}$ which an arbitrary cut may contain. However, 8 of them can be immediately ruled out by the fact that they cannot achieve minimum weight. To see this, consider for example a cut containing precisely $\{\sigma, i\}$. This would receive a weight $w_{\sigma j} + w_{\sigma k} = w_{ij} + w_{ik} + 2w_{jk}$, which is strictly greater than the weight that $\{i\}$ alone would give, namely, $w_{\sigma i} = w_{ij} + w_{ik}$. In general, one easily observes that for a cut to be of minimum weight, the newly added vertex σ should only participate when the cut contains at least two of $\{i, j, k\}$, thus giving only 8 possibilities, as in $\widetilde{G_{\mathsf{N}}}$. Looking at each case as in table B.1, we arrive at the desired result that min-cut weights on the two graphs indeed match. □

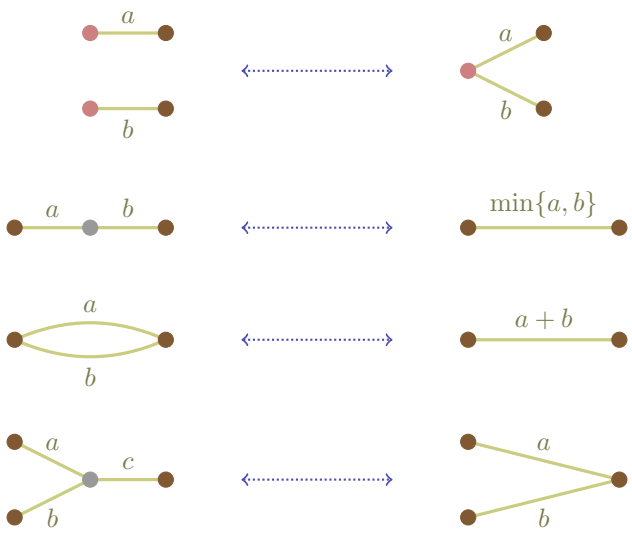


Figure B.1: Basic entropy-preserving graph operations. The last one requires $c \geq a+b$. Brown vertices are to be kept fixed under these operations. Boundary vertices are colored in red, and bulk ones in gray. In a general graph, both brown and red vertices may connect to arbitrarily many other edges, whereas gray ones should only appear as shown.

Figure B.2: The new entropy-preserving graph operation Δ-Y. Vertices are color-coded as in fig. B.1.

Appendix C

Appendix to Chapter 5

C.1 Mathematical Results

In this appendix we give the proofs of the new results stated in section 3. We will assume that the spacetime M is at least C^2, maximally extended, connected, AdS-hyperbolic and asymptotically locally $\mathrm{AdS}_n \times \mathbb{S}^k$ with $n > 2$. Similarly, ∂M is assumed to be maximally extended, connected, and globally hyperbolic. Recall that AdS-hyperbolic means that there are no closed causal curves, and for any two points p and q, the set $J^+(p) \cap J^-(q)$ is compact after conformally compactifying the AdS boundary [95].

We will assume everywhere in this section that p and q are bulk points in the domain of influence of the asymptotic boundary, so that their light-cone cuts are not empty. The results below apply to both future and past light-cone cuts which we denote $C(p)$, or $\mathcal{C}(p)$ for the extended cuts. In expressions like $\mathcal{C}(p) \cap \mathcal{C}(q)$ it will be understood that both cuts are past or both cuts are future.

Proposition 5.1. *Every light-cone cut $C(p)$ is differentiable everywhere except on a set of measure zero.*

Proof. The logic of the first part of this proof parallels that of Proposition 6.3.1 in [220].[1] Let $r \in C(p)$ and consider an open neighborhood $U_\alpha \subset \partial M$ about r. One can introduce normal coordinates $x_\alpha = \{x^\mu_\alpha : U_\alpha \to \mathbb{R} \mid \mu = 0, \ldots, n-2\}$ with ∂_0 timelike and such that the coordinate slices $\gamma_{\mathbf{c}} = \{s \in U_\alpha \mid \mathbf{x}_\alpha(s) = \mathbf{c}\}$, where $\mathbf{x}_\alpha = \{x^i_\alpha \mid i = 1, \ldots, n-2\}$, define curves intersecting both $I^-(r) \cap U_\alpha$ and $I^+(r) \cap U_\alpha$ for any constant $\mathbf{c} \in \mathbf{x}_\alpha[U_\alpha]$. By continuity and achronality, each curve $\gamma_{\mathbf{c}}$ must intersect $C(p)$ at precisely one point $s_{\mathbf{c}}$, i.e. $\{s_{\mathbf{c}}\} = \gamma_{\mathbf{c}} \cap C(p)$, and therefore the map $\mathbf{x}_\alpha : U_\alpha \cap C(p) \to \mathbb{R}^{n-2}$ is a homeomorphism onto its image.

Now define a map $\tilde{x}^0 : \mathbf{x}[U_\alpha \cap C(p)] \to x^0[U_\alpha]$ by $\tilde{x}^0(\mathbf{c}) = x^0(r_{\mathbf{c}})$, where $r_{\mathbf{c}}$ is the unique point at which $\gamma_{\mathbf{c}}$ intersects $C(p)$. Because $C(p)$ is achronal, for any two points $r, s \in U_\alpha \cap C(p)$ one has that $|\tilde{x}^0(\mathbf{x}(r)) - \tilde{x}^0(\mathbf{x}(s))| \leq K|\mathbf{x}(r) - \mathbf{x}(s)|$ for some $K \geq 1$, with $|\cdot|$ the Euclidean norm. This shows that $\tilde{x}^0$ is Lipschitz continuous. A Lipschitz continuous transition map $\varphi_{\alpha\beta} : \mathbf{x}_\alpha[U_\alpha \cap U_\beta \cap C(p)] \to \mathbf{x}_\beta[U_\alpha \cap U_\beta \cap C(p)]$ can now be constructed by direct product and composition with maps of higher differentiability class as $\varphi_{\alpha\beta} = \mathbf{x}_\alpha \circ x^{-1}_\beta \circ \{\tilde{x}^0_\beta, \mathrm{id}\}$. Thus a collection of charts $(U_\alpha \cap C(p), \mathbf{x}_\alpha)$ forms an atlas for $C(p)$ and endows it with a Lipschitz structure.

The differentiability of $C(p)$ at a point $r \in U_\alpha \cap C(p)$ is determined by the differentiability class of the transition maps $\varphi_{\alpha\beta}$ at $\mathbf{x}_\alpha(r) \in \mathbf{x}_\alpha[U_\alpha \cap U_\beta \cap C(p)]$. Because the transition map $\varphi_{\alpha\beta}$ is Lipschitz continuous, Rademacher's theorem [221] implies that the points in $\mathbf{x}_\alpha[U_\alpha \cap U_\beta \cap C(p)] \subset \mathbb{R}^{n-2}$ at which $\varphi_{\alpha\beta}$ is not differentiable form a set of Lebesgue measure zero as a subset of $\mathbb{R}^{n-2}$. Thus the set of points at which $C(p)$ fails to be differentiable has measure zero. □

Proposition 5.2. *Each point $(r, \Phi(r))$ on the extended cut $\mathcal{C}(p)$ determines the unique*

[1] There is a typo in the proof in [220]: both instances of the set $\mathcal{L}$ appearing in the penultimate sentence should be replaced by its boundary set $\dot{\mathcal{L}}$.

null geodesic from r to p.

Proof. Consider an arbitrary point $(r, \Phi(r)) \in \mathcal{C}(p)$ and let $\gamma : [0,1] \to \bar{M}$ be the unique null geodesic from $r = \gamma(0) \in C(p)$ to $p = \gamma(1) \in M$. Write $\dot{\gamma}(0) \propto V^\perp + V^\parallel$ with V some rescaled vector parallel to $\dot{\gamma}(0)$ such that $V^\perp$ has unit norm, where $V^\perp$ ($V^\parallel$) is the projection of V onto the normal (tangent) bundle of ∂M. Since r is regular, $C(p)$ is differentiable at r, and therefore there is a well-defined space tangent to $C(p)$ at r, denoted $T_r C(p)$. Because $C(p)$ is a codimension-1 spacelike subspace of ∂M, there is a unique timelike vector $T \in T_r \partial M$ normal to $C(p)$ with $T^2 = -1$. Under natural identifications of the vectors in ∂M with their inclusions in the ambient space $\bar{M}$, one can further decompose $\dot{\gamma}(0) \propto T + \cos\alpha\, V^\perp + \sin\alpha\, S$, where $S \in T_r C(p)$ is a unit spacelike vector and $\alpha \in [0, \pi/2)$. If $\alpha \neq 0$, there would be a nontrivial vector S such that one could consider a point $r_\epsilon \in C(p)$ arbitrarily close to r in the direction parallel to S. Notice that then one could deform γ infinitesimally near ∂M into a timelike piece that connects up with r_ϵ, thus making p and r_ϵ timelike-related, which contradicts the achronality of the light cone $\partial J(p)$. Hence one finds that $\alpha = 0$ necessarily, and therefore $\dot{\gamma}(0) \propto T + V^\perp$ in general. In other words, regularity of the cut point implies that the vector field $\dot{\gamma}$ tangent to γ is orthogonal to $C(p)$ at r. The dimensionality of the normal bundle of $C(p)$ in $\bar{M}$ is given by the codimension of $C(p)$, which is $k+2$ corresponding to timelike and radial bulk directions and the conformally-shrinking $\mathbb{S}^k$. The specification of $\Phi(r)$ by the extended cut point fixes the direction of $\dot{\gamma}(0)$ on $\mathbb{S}^k$, such that this vector only remains undetermined in 2 dimensions. Out of the 2 possible null directions spanning the latter, only one points inwards towards the bulk. Hence the choice of a point in $\mathcal{C}(p)$ together with the orthogonal ingoing condition fix $\dot{\gamma}(0)$ up to scaling. But because $\dot{\gamma}(0)$ is null and γ is geodesic, this suffices to determine a unique null geodesic from r to p. □

Proposition 5.5. *$\mathcal{C}(p) \cap \mathcal{C}(q)$ contains more than one point if and only if $p = q$.*

Proof.

($\Leftarrow$) If $p = q$, $\mathcal{C}(p) \cap \mathcal{C}(q) = \mathcal{C}(p)$, which always contains more than one point.

($\Rightarrow$) Consider an arbitrary point $(r, \Phi(r)) \in \mathcal{C}(p) \cap \mathcal{C}(q)$. According to Proposition 5.2, the pair $(r, \Phi(r))$ determine a unique ingoing null geodesic γ_r. If there were two distinct such points in the intersection of the two extended cuts, their associated γ_r geodesics would pass through both p and q, which would then be either equal or conjugate to each other. But since there cannot be any conjugate points along any γ_r strictly between either of these points and their cuts, it must be the case that $p = q$. □

Theorem C.1. *$\mathcal{C}(p) \cap \mathcal{C}(q)$ contains exactly one point $(r, \Phi(r))$ if and only if $q \neq p$ and q belongs to an achronal extension of a null geodesic γ from p to a regular point $r \in C(p)$.*[2]

Proof.

($\Rightarrow$)[3] Since $\mathcal{C}(p) \cap \mathcal{C}(q)$ contains exactly one point $(r, \Phi(r))$, Proposition 5.5 already implies $p \neq q$. Then Proposition 5.2 shows that $(r, \Phi(r))$ defines the unique null geodesic γ_r associated to the regular point r. Since $(r, \Phi(r))$ belongs to the intersection of the two cuts, γ_r passes through both p and q and stays on the union of their light cones $\partial J(p) \cup \partial J(q)$. Thus the two points are null-related by an achronal geodesic through both that ends at r.

($\Leftarrow$) If p and q both lie on an achronal null geodesic γ that reaches a regular point $r \in C(p)$ and $\Phi(r)$ on $\mathbb{S}^k$, then γ lies on both $\partial J(p)$ and $\partial J(q)$. So $(r, \Phi(r))$ is clearly in both extended cuts $\mathcal{C}(p)$ and $\mathcal{C}(q)$. □

[2] The statement of an analogous result in [46] is not quite correct. In particular, q need not belong to the null geodesic from p to $r \in C^{\pm}(p)$, but instead could lie on an extension of this geodesic beyond p (i.e. p itself would lie in a null geodesic connecting q to r). This explains the qualification of the statement to an *achronal extension* of the null geodesic from p to r.

[3] This direction proves Theorem 5.1 in Section 5.3.3.

C.2 Higher-dimensional scalar propagators in global $\mathbf{AdS}_n \times \mathbb{S}^k$

Consider a free bulk scalar field of mass m with Euclidean action

$$S_\varphi = \frac{1}{2}\int \epsilon \left(|d\varphi|_g^2 + m^2\varphi^2 \right), \tag{C.1}$$

where ϵ is the volume element on all $D = n+k$ dimensions of $\mathrm{AdS}_n \times \mathbb{S}^k$. Using Poincaré coordinates in Euclidean signature,

$$g = \frac{\ell^2}{z^2}\left(dz^2 + \delta_{ij}dx^i dx^j\right) + \ell^2 d\Omega^2, \tag{C.2}$$

where Latin indices run over the $d = n-1$ spatial dimensions of AdS_n.

C.2.1 Bulk-to-bulk propagator

The higher-dimensional bulk-to-bulk scalar propagator $\mathcal{G}$ is defined as the Green function of the Klein-Gordon operator,

$$\left(-\Box_g + m^2\right)\mathcal{G}(z,x,\Omega;\,\tilde{z},\tilde{x},\tilde{\Omega}) = \frac{1}{\sqrt{\det g}}\delta^n(z-\tilde{z},x-\tilde{x})\delta^k(\Omega-\tilde{\Omega}). \tag{C.3}$$

where $\Box_g$ denotes the d'Alembertian built from the D-dimensional metric g. Because $\mathrm{AdS}_n \times \mathbb{S}^k$ is a product spacetime, this operator is diagonal and decomposes as

$$\Box_g = \Box_{\mathrm{AdS}_n} + \ell^{-2}\Delta_{\mathbb{S}^k}, \tag{C.4}$$

where $\Box_{\mathrm{AdS}_n}$ only acts on AdS coordinates (z, x) and the unit k-sphere Laplacian $\Delta_{\mathbb{S}^k}$ only acts on coordinates Ω. Explicitly,

$$\Box_{\mathrm{AdS}_n} = \frac{z^2}{\ell^2}\left(\partial_z^2 - (d-1)z^{-1}\partial_z + \partial_x^2\right), \tag{C.5}$$

and, using Cartesian coordinates on $\mathbb{R}^{k+1} \supset \mathbb{S}^k$, one can write

$$\Delta_{\mathbb{S}^k} = \sum_{\alpha>\beta}^{k} \left(x_\alpha \partial_\beta - x_\beta \partial_\alpha\right)^2. \tag{C.6}$$

Consider first the propagator G_Δ of a free scalar in AdS_n of mass μ, defined by

$$\left(-\Box_{\mathrm{AdS}_n} + \mu^2\right) G_\Delta(z, x;\, \tilde{z}, \tilde{x}) = \frac{1}{\sqrt{\det g_{\mathrm{AdS}_n}}}\delta^n(z - \tilde{z}, x - \tilde{x}). \tag{C.7}$$

This Green function is well-known and can be written in terms of the hypergeometric function ${}_2F_1$ as [75, 19]

$$G_\Delta(z, x;\, \tilde{z}, \tilde{x}) = \frac{2^{-\Delta} C_\Delta}{2\Delta - d}\, \xi^\Delta\, {}_2F_1\left(\frac{\Delta}{2}, \frac{\Delta}{2} + \frac{1}{2};\, \Delta - \frac{d}{2} + 1;\, \xi^2\right), \tag{C.8}$$

where the conformal ratio ξ is defined in terms of the coordinates of the two points by

$$\xi \equiv \frac{2z\tilde{z}}{z^2 + \tilde{z}^2 + (x - \tilde{x})^2}, \tag{C.9}$$

and the conformal dimension Δ and normalization constant C_Δ are

$$\mu^2 = \frac{\Delta(\Delta - d)}{\ell^2} \quad \text{and} \quad C_\Delta = \frac{\Gamma(\Delta)}{\pi^{d/2}\Gamma(\Delta - d/2)}. \tag{C.10}$$

The two solutions of the quadratic equation obeyed by Δ correspond to the usual two

branches

$$\Delta = \frac{d}{2} \pm \sqrt{\frac{d^2}{4} + \ell^2 \mu^2}, \tag{C.11}$$

with the positive (negative) sign giving the normalizable (non-normalizable) one.

Consider now the $\mathbb{S}^k$ term in (C.4). The eigenfunctions of $\Delta_{\mathbb{S}^k}$ are called hyperspherical harmonics $Y_L^I(\Omega)$ and labeled by their scaling degree $L \in \mathbb{Z}_{\geq 0}$ and a tuple $I_L = (i_1, \ldots, i_{k+1}) \in \mathbb{Z}^{k+1}$ with $\sum_{l=1}^{k+1} i_l = L$, which specifies an element of the representation of $SO(k+1)$ in terms of traceless symmetric tensors of degree L in $k+1$ dimensions. They are defined by the eigenvalue problem

$$\Delta_{\mathbb{S}^k} Y_L^{I_L}(\Omega) = -L(L+k-1) Y_L^{I_L}(\Omega). \tag{C.12}$$

and conventionally orthonormalized to satisfy

$$\int_{\mathbb{S}^k} d\Omega \, Y_L^{I_L\,*}(\Omega) \, Y_{\tilde{L}}^{\tilde{I}_{\tilde{L}}}(\Omega) = \delta_{L\tilde{L}} \delta^{I_L \tilde{I}_{\tilde{L}}}, \tag{C.13}$$

where $d\Omega$ is the volume element of $\mathbb{S}^k$. Additionally, as a basis for functions on $\mathbb{S}^k$, hyperspherical harmonics obey the completeness relation

$$\sum_{L=0}^{\infty} \sum_{I_L} Y_L^{I_L\,*}(\tilde{\Omega}) \, Y_L^{I_L}(\Omega) = \frac{1}{\sqrt{\det g_{\mathbb{S}^k}}} \delta^k(\Omega - \tilde{\Omega}). \tag{C.14}$$

The sum over $SO(k+1)$ representation indices I_L for fixed L can be performed explicitly and leads to [222]

$$\sum_{I_L} Y_L^{I_L\,*}(\Omega) \, Y_L^{I_L}(\tilde{\Omega}) = N_L \, C_L^{(k-1)/2}(\cos\theta), \tag{C.15}$$

where $\cos(\theta) \equiv \mathbf{n} \cdot \tilde{\mathbf{n}}$ for unit vectors $\mathbf{n}, \tilde{\mathbf{n}} \in \mathbb{R}^{k+1}$ oriented on $\mathbb{S}^k$ as specified by Ω and $\tilde{\Omega}$,

respectively, and N_L is a normalization constant given by

$$N_L = \frac{2L+k-1}{(k-1)\,\mathrm{vol}\,\mathbb{S}^k} \quad \text{where} \quad \mathrm{vol}\,\mathbb{S}^{2l-1} = \frac{2\pi^l}{\Gamma(l)}, \tag{C.16}$$

The symbol $C_l^\alpha(x)$ is a Gegenbauer polynomial, which can be written as

$$C_l^\alpha(x) = \frac{\Gamma(2\alpha+l)}{\Gamma(2\alpha)}\,{}_2F_1\left(-l,\,2\alpha+l;\,\alpha+\frac{1}{2};\,\frac{1-x}{2}\right). \tag{C.17}$$

It is now a simple matter to construct the desired propagator:

Proposition C.1. *The higher-dimensional bulk-to-bulk propagator $\mathcal{G}$ for a free scalar of mass m in global $AdS_n \times \mathbb{S}^k$ given as an infinite series by*

$$\mathcal{G}(z,x,\Omega;\,\tilde{z},\tilde{x},\tilde{\Omega}) = \sum_{L=0}^{\infty} N_L C_L^{(k-1)/2}(\cos\theta)\,G_{\Delta_L}(z,x;\,\tilde{z},\tilde{x}), \tag{C.18}$$

where G_{Δ_L} is the propagator of a free scalar in AdS_n of scaling dimension Δ_L defined to be

$$\Delta_L = \frac{d}{2} \pm \sqrt{\frac{d^2}{4} + \ell^2 M_L^2} \quad \textit{where} \quad M_L^2 = m^2 + \frac{L(L+k-1)}{\ell^2}. \tag{C.19}$$

Proof. Applying the right-hand side of (C.3) to (C.18) leads to

$$\begin{aligned}\left(-\Box_g + m^2\right)\mathcal{G}(z,x,\Omega;\,\tilde{z},\tilde{x},\tilde{\Omega}) &= \sum_{L=0}^{\infty} N_L C_L^{(k-1)/2}(\cos\theta)\left(-\Box_{\mathrm{AdS}_n} + M_L^2\right)G_{\Delta_L}(z,x;\,\tilde{z},\tilde{x}) \\ &= \frac{1}{\sqrt{\det g}}\delta^n(z-\tilde{z},x-\tilde{x})\delta^k(\Omega-\tilde{\Omega})\end{aligned} \tag{C.20}$$

where (C.4), (C.15), (C.12) and (C.19) have been used in the first equality, and (C.7), (C.15) and (C.14) in the second one. The result thus agrees with the right-hand side of (C.3). □

The series form of (C.18) may be understood as a Kaluza-Klein series expansion of the higher-dimensional bulk-to-bulk propagator. This expression reduces to a very compact form for conformally flat $\mathrm{AdS}_n \times \mathbb{S}^k$, as is the case of (C.2),[4] if one chooses the scalar to be coupled to the metric in a Weyl invariant manner [151]. This is accomplished in (C.1) by choosing the mass of the scalar to be precisely

$$m^2 = \frac{(k-1)^2 - (n-1)^2}{4\ell^2}. \tag{C.21}$$

The resulting propagator is simply a power-law in the total chordal distance along both AdS_n and $\mathbb{S}^k$, viz. (see [151] for more details)

$$\mathcal{G}(z, x, \Omega;\, \tilde{z}, \tilde{x}, \tilde{\Omega}) = \frac{\Gamma(h)}{2(2\pi)^{h+1}} \frac{1}{(\xi^{-1} - \cos\theta)^h} \quad \text{where} \quad h = \frac{n+k-2}{2}. \tag{C.22}$$

C.2.2 Bulk-to-boundary propagator

One would naively hope to be able to derive a simple expression for the bulk-to-boundary propagator starting from (C.22) and using some version of the extrapolate dictionary [160]. Unfortunately, it is not at all clear in this case how one would take the $z \to 0$ limit of (C.22). A naively reasonable guess would be to Taylor expand this object in ξ, kill off the z^l power in ξ^l of the l^{th} term with a factor of $(2l-d)z^{-l}$, take the $z \to 0$ limit and hope to be able to perform the summation of the resulting series to obtain a compact expression. However, this would neither be a kernel as defined in (5.16) nor obey the desired boundary condition in (5.18).

Instead, our approach will be to perform the summation in (5.17) directly. The terms in the summation can be obtained by applying the extrapolate dictionary to every term

[4] Recall that global $\mathrm{AdS}_n \times \mathbb{S}^k$ is conformally flat if and only if the radius of the $\mathbb{S}^k$ matches that of AdS_n.

in the series (C.18) that defines $\mathcal{G}$. These will involve the usual dimension-Δ_L bulk-to-boundary propagator [17, 160]

$$K_{\Delta_L}(z,x;\,\tilde{x}) = \lim_{\tilde{z}\to 0}(2\Delta_L - d)\tilde{z}^{-\hat{\Delta_L}}\, G_{\Delta_L}(z,x;\,\tilde{z},\tilde{x}) = C_{\Delta_L}\chi^{\Delta_L}, \tag{C.23}$$

where χ is given by

$$\chi = \frac{z}{z^2 + (x-\tilde{x})^2}. \tag{C.24}$$

The upshot is the following definition of the higher-dimensional bulk-to-boundary propagator as an infinite Kaluza-Klein series:

$$\mathcal{K}(z,x,\Omega;\,\tilde{x},\tilde{\Omega}) = \sum_{L=0}^{\infty} N_L C_{\Delta_L} C_L^{(k-1)/2}(\cos\theta)\,\chi^{\Delta_L}. \tag{C.25}$$

Note that in Lorentz signature, the limit that the bulk and boundary point become null separated corresponds to $\chi \to \infty$. Each term in this series then develops a singularity with a coefficient that is a smooth function on $\mathbb{S}^k$ peaked at the location of the bulk point.

The computation of this sum becomes tractable for Weyl invariant matter, which fixes the mass of the scalar to be given by (C.21). The resulting L^{th} term in (C.25) is

$$\mathcal{K}_{\Delta_L}(z,x,\Omega;\,\tilde{x},\tilde{\Omega}) = \frac{\Gamma\left(\frac{k-1}{2}\right)}{2\pi^{h+1}}\left(L + \frac{k-1}{2}\right)\frac{\Gamma(L+h)}{\Gamma\left(L+\frac{k-1}{2}\right)}C_L^{(k-1)/2}(\cos\theta)\,\chi^{L+h}, \tag{C.26}$$

where h was defined in (C.22). For convenience, focus on the odd-n case, for which the ratio of Γ functions may be expanded as a finite product. Using the Pochhammer symbol

$(a)_n = a(a+1)\cdots(a+n-1)$, this is

$$\frac{\Gamma(L+h)}{\Gamma\left(L+\frac{k-1}{2}\right)} = \left(L+\frac{k-1}{2}\right)_{\frac{n-1}{2}}. \tag{C.27}$$

The goal will be to manipulate (C.26) so as to be able to utilize the identity of Gegenbauer polynomials that gives their defining generating function, namely[5]

$$\sum_{L=0}^{\infty} C_L^{\alpha}(y)\,\chi^L = \frac{1}{(1-2\chi y+\chi^2)^{\alpha}}. \tag{C.28}$$

To do this, note that the right-hand side of (C.27) can be realized via differentiation in χ in the following way

$$\frac{\Gamma(L+h)}{\Gamma\left(L+\frac{k-1}{2}\right)} = \frac{1}{\chi^{L+\frac{k-1}{2}-1}}\partial_{\chi}^{\frac{n-1}{2}}\left(\chi^{L+h-1}\right), \tag{C.29}$$

and similarly one can write

$$L+\frac{k-1}{2} = \frac{1}{\chi^{L+\frac{k-1}{2}-1}}\partial_{\chi}\left(\chi^{L+\frac{k-1}{2}}\right). \tag{C.30}$$

Putting (C.29) and (C.30) together with χ^{L+h}, consider the following manipulations:

$$\begin{aligned}\left(L+\frac{k-1}{2}\right)\frac{\Gamma(L+h)}{\Gamma\left(L+\frac{k-1}{2}\right)}\chi^{L+h} &= \frac{1}{\chi^{L+\frac{k-1}{2}-1}}\partial_{\chi}\left(\chi^{L+\frac{k-1}{2}}\frac{\Gamma(L+h)}{\Gamma\left(L+\frac{k-1}{2}\right)}\right)\chi^{L+h} \\ &= \chi^{\frac{n+1}{2}}\partial_{\chi}\left(\chi\,\partial_{\chi}^{\frac{n-1}{2}}\left(\chi^{L+h-1}\right)\right).\end{aligned} \tag{C.31}$$

[5]Note that this identity holds as an equality between power series in χ. However, as an infinite series, the left-hand side is only convergent for $|\chi|<1$. While this should be kept in mind, in practice in will not be a problem: physically, one is interested in looking at each L mode independently. Every relation derived henceforth using this identity should thus be understood as an equality between power series in χ.

With this expression at hand, the infinite series that defines $\mathcal{K}$ may now be rewritten as

$$\sum_{L=0}^{\infty}\mathcal{K}_{\Delta_L}(z,x,\Omega;\,\tilde{x},\tilde{\Omega})=\frac{\Gamma\left(\frac{k-1}{2}\right)}{2\pi^{h+1}}\chi^{\frac{n+1}{2}}\partial_\chi\left(\chi\,\partial_\chi^{\frac{n-1}{2}}\left(\chi^{h-1}\sum_{L=0}^{\infty}C_L^{(k-1)/2}(\cos\theta)\,\chi^L\right)\right). \tag{C.32}$$

At this point it only remains to employ (C.28) to obtain the desired explicit form of the higher-dimensional bulk-to-boundary propagator:

$$\mathcal{K}(z,x,\Omega;\,\tilde{x},\tilde{\Omega})=\frac{\Gamma\left(\frac{k-1}{2}\right)}{2\pi^{h+1}}\chi^{\frac{n+1}{2}}\partial_\chi\left(\chi\,\partial_\chi^{\frac{n-1}{2}}\left(\frac{\chi^{h-1}}{(1-2\chi\cos\theta+\chi^2)^{\frac{k-1}{2}}}\right)\right). \tag{C.33}$$

This result is valid for any odd $n\geq 3$ and any integer $k\geq 2$. Note also that this expression only holds as an equality between coefficients in a power series in χ, the reason being that the radius of convergence of the infinite series is $|\chi|<1$. This is not a problem in Euclidean signature because $|\chi|<1$ always, but should be kept in mind for Lorentzian signature where e.g. null separation corresponds to $\chi\to\infty$.

The spacetime with one of the simplest evaluations of (C.33) is $\mathrm{AdS}_3\times\mathbb{S}^3$, for which

$$\mathcal{K}(z,x,\Omega;\,\tilde{x},\tilde{\Omega})=\frac{\chi^2\left(\chi^4+2\chi\left(\chi^2+1\right)\cos\theta-6\chi^2+1\right)}{2\pi^3\left(\chi^2-2\chi\cos\theta+1\right)^3}. \tag{C.34}$$

For the usual case of interest of $\mathrm{AdS}^5\times\mathbb{S}^5$ one gets the following:

$$\mathcal{K}(z,x,\Omega;\,\tilde{x},\tilde{\Omega})=-\frac{2\chi^4\left(\chi^6+2\chi^4\cos 2\theta-17\chi^4+25\chi^2+2\left(4\chi^4-5\chi^2-3\right)\chi\cos\theta-3\right)}{\pi^5\left(\chi^2-2\chi\cos\theta+1\right)^5}. \tag{C.35}$$

Appendix D

Appendix to Chapter 7

D.1 Liouville Equation

In this Appendix we fill in some of the details omitted from Section 7.3.2. On the rectangular domain shown in Figure 7.6, the conformal factor σ obeys the Liouville equation (7.18) subject to the requirements that σ diverge at the conformal boundary $\xi = \pi\xi_0/2$; that $\phi = 0$ and $\phi = 2\pi$ be identified; and that σ be continuous in the interior of the ellipse in the z plane. This last condition imposes boundary conditions at the $\xi = 0$ edge of the coordinate rectangle, which is mapped to a double-cover of the line segment connecting $z = \pm 1$: it requires that for all nonnegative integer p,

$$\partial_\xi^p \sigma(\xi = 0, \phi) = (-1)^p \partial_\xi^p \sigma(\xi = 0, -\phi). \tag{D.1}$$

These various boundary conditions make it convenient to extend σ from the coordinate rectangle to the infinite strip $\xi \in (-\pi\xi_0/2, \pi\xi_0/2)$, $\phi \in (-\infty, \infty)$. The point is that the $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry of the ellipse implies that on the strip σ is symmetric about

the lines $\xi = 0$ and $\phi = k\pi/2$ for all integer k, which is sufficient to ensure that σ is appropriately periodic in ϕ and that it obeys (D.1). Since the strip contains infinitely many images of the foci at $z = \pm 1$ under the map (7.16), the Liouville equation takes the form[1]

$$-4\partial_u\partial_{\bar{u}}\sigma + e^{2\sigma} = -\frac{2\pi}{\nu}\sum_{k=-\infty}^{\infty}\delta_{ik\pi}(u,\bar{u}), \tag{D.2}$$

where recall that $\nu \equiv m/(2-m)$. It is convenient to absorb the singularities, both from the delta functions and the boundary condition at the conformal boundary, into σ by defining

$$\sigma \equiv \tilde{\sigma} + \ln\left(\frac{1}{\xi_0}\sec\left(\frac{u+\bar{u}}{2\xi_0}\right)\right) + \frac{1}{\nu}\sum_{k=-\infty}^{\infty}G_{ik\pi}(u,\bar{u}), \tag{D.3}$$

where $G_{u_*}(u,\bar{u})$ is a Green's function of the Laplacian. The Green's function term absorbs the delta functions in (D.2), while the logarithmic term ensures that σ diverges at the conformal boundary if $\tilde{\sigma}$ is regular there. Its particular form is chosen to be the conformal factor of the double-trumpet, so that for $m = 2$ (or $\nu \to \infty$) the Liouville equation is solved by $\tilde{\sigma} = 0$.

It is in fact convenient to take G_{u_*} to be the Dirichlet Green's function on the strip:

$$4\partial_u\partial_{\bar{u}}G_{u_*}(u,\bar{u}) = 2\pi\delta_{u_*}(u,\bar{u}), \qquad G_{u_*}\left(\pm\frac{\pi\xi_0}{2} + i\phi, \pm\frac{\pi\xi_0}{2} - i\phi\right) = 0. \tag{D.4}$$

An explicit form of G_{u_*} can be obtained by starting with the Dirichlet Green's function on the upper half-plane, $G_{w_*}(w,\overline{w}) = \ln|(w-w_*)/(w-\overline{w}_*)|$, and then mapping the upper half-plane to the strip via $w = ie^{-iu/\xi_0}$, yielding

$$G_{u_*}(u,\bar{u}) = \ln\left|\frac{\sin\left(\frac{u-u_*}{2\xi_0}\right)}{\cos\left(\frac{u+\bar{u}_*}{2\xi_0}\right)}\right|. \tag{D.5}$$

[1]The infinite delta functions effectively ensure that the problem has the desired periodicity in ϕ.

The Liouville equation consequently yields (7.20a). Moreover, since by construction H vanishes at the conformal boundary $\xi = \pi\xi_0/2$, regularity of $\tilde{\sigma}$ in fact requires that $\tilde{\sigma}$ *vanish* there, giving the boundary conditions (7.20b).

D.1.1 Approximate Solutions

To construct the function f defined by (7.22), first note that (7.22) can be reduced to an ODE: exchanging ξ for a new variable $x \equiv \operatorname{sech}(\phi/\xi_0)\cos(\xi/\xi_0) \in (0,1)$, we obtain

$$\left[x^2(x^2-1)\partial_x^2 + 2x^3\partial_x + \xi_0 x^2 \cosh^2\left(\frac{\phi}{\xi_0}\right)\left(2x\tanh\left(\frac{\phi}{\xi_0}\right)\partial_x\partial_\phi - \partial_\phi^2\right)\right] f + \left(\frac{1-x}{1+x}\right)^{1/\nu} e^{2f} - 1 = 0, \quad \text{(D.6)}$$

while the boundary conditions require $f(x=0,\phi) = 0$ and that f be regular at $x = 1$. Consequently it is consistent to take f to be a function of x only, satisfying

$$\left(x^2(x^2-1)\partial_x^2 + 2x^3\partial_x\right) f + \left(\frac{1-x}{1+x}\right)^{1/\nu} e^{2f} - 1 = 0. \quad \text{(D.7)}$$

We now take ν to be small to construct a solution to this equation. In such a case, two scaling regimes emerge depending on the behavior of $((1-x)/(1+x))^{1/\nu}$: when x/ν is large, $((1-x)/(1+x))^{1/\nu}$ is nonperturbatively small in ν, while when x/ν is order unity or smaller, $((1-x)/(1+x))^{1/\nu}$ can be expanded perturbatively in ν. We can therefore obtain a solution for f perturbatively in ν by performing a matched asymptotic expansion: we solve for f in these two regimes and match the solutions together.

Large x/ν corresponds to the interior of the quotient geometry; in this regime, the

penultimate term in (D.7) can be ignored, leaving simply

$$\left(x^2(x^2-1)\partial_x^2 + 2x^3\partial_x\right) f_{\rm int} - 1 = 0. \tag{D.8}$$

Requiring that $f_{\rm int}$ be regular at $x = 1$ gives the family of solutions

$$f_{\rm int}(x) = C + \ln\left(\frac{x}{1+x}\right), \tag{D.9}$$

where C is an arbitrary constant (that may depend on ν). On the other hand, x/ν of order unity or smaller corresponds to a thin layer near the conformal boundary. Here we define $y \equiv x/\nu$ and expand (D.7) in ν at fixed y:

$$\left(-y^2\partial_y^2 + \nu^2 y^3\left(y\partial_y^2 + 2\partial_y\right)\right) f_{\rm bndry} + \left(1 - \frac{2y^3}{3}\,\nu^2 + \mathcal{O}(\nu^4)\right) e^{2(f_{\rm bndry}-y)} - 1 = 0. \tag{D.10}$$

We then look for solutions order-by-order in ν by writing

$$f_{\rm bndry}(y) = \sum_{n=0}^{\infty} \nu^n f_n(y). \tag{D.11}$$

Plugging this expansion into (D.10) and imposing the boundary condition $f_{\rm bndry}(0) = 0$ allows us to solve for the f_n order-by-order. For example, to $\mathcal{O}(\nu)$ we have

$$f_0 = y - \ln\left(\frac{\sinh(\alpha_0 y)}{\alpha_0 y}\right), \qquad f_1 = \alpha_1\left(\alpha_0 y\coth(\alpha_0 y) - 1\right), \tag{D.12}$$

where $\alpha_0 > 0$ and α_1 are arbitrary constants (independent of ν). One obtains a new constant α_n at each order in ν.

The constants of integration C and α_n are fixed by matching the solutions $f_{\rm int}$ and $f_{\rm bndry}$ in the overlap region: that is, we require that the expansion of $f_{\rm int}(x)$ at

small $x = \nu y$ agree with the expansion of $f_{\text{bndry}}(y)$ at large y. To $\mathcal{O}(\nu)$, the relevant expansions are

$$f_{\text{int}} = C + \ln \nu + \ln y - \nu y + \mathcal{O}(\nu^2), \tag{D.13a}$$

$$f_{\text{bndry}} = y(1-\alpha_0) + \ln y + \ln(2\alpha_0) + \nu\alpha_1 (\alpha_0 y - 1) + \mathcal{O}(\nu^2, e^{-\alpha_0 y}), \tag{D.13b}$$

and hence matching order-by-order in y and ν gives $\alpha_0 = 1$, $\alpha_1 = -1$, and $C = \ln(2/\nu) + \nu + \mathcal{O}(\nu^2)$. We can then form a composite solution by superimposing the interior and near-boundary solutions:

$$f = f_{\text{int}} + f_{\text{bndry}} - f_{\text{overlap}}, \tag{D.14}$$

where f_{overlap} is the common behavior shared by f_{int} and f_{bndry} in the matching region. Proceeding in this manner up to order ν^3 ultimately yields (7.27) in the main text. To then compute the magnitude of $\delta\tilde{\sigma}$, we use (7.27) to quantify the falloff of F outside the interval $\phi \in (-\pi/2, \pi/2)$: for ϕ outside this interval, y is $\mathcal{O}(\nu^{-1} e^{-|\phi|/\xi_0})$, which is small if $e^{-\pi/2\xi_0} \ll \nu$. Expanding F at small y gives $F = \mathcal{O}(y^2)$, which leads to the right-hand side of (7.26) being $\mathcal{O}(\nu^{-4} e^{-2\pi/\xi_0})$, as claimed.

D.2 Numerical Details

In this Appendix we briefly go into some details on the numerical approaches we use both in solving the Liouville equation for the conformal factor σ as in solving the equation of motion (7.61) for the boundary wiggle in JT coupled to a massless scalar field. For a more detailed description of some of the pseudospectral discretization methods we implement, as well as of the Newton-Raphson method for solving nonlinear problems, see e.g. [212, 223].

D.2.1 Liouville Equation

We are interested in solving the Liouville equation (7.20a), which we reproduce here for convenience:

$$-\left(\partial_\xi^2+\partial_\phi^2\right)\tilde{\sigma}+\frac{1}{\xi_0^2}\sec^2\left(\frac{\xi}{\xi_0}\right)\left(e^{2H(\xi,\phi)/\nu}e^{2\tilde{\sigma}}-1\right)=0, \tag{D.15}$$

where $H(\xi,\phi)$ is defined in (7.19b). The coordinate domain corresponding to the ellipse is given by $\xi\in[0,\pi\xi_0/2)$, $\phi\in[0,2\pi)$, and we require $\tilde{\sigma}|_{\xi=\pi\xi_0/2}=0$ and for $\tilde{\sigma}$ to be regular everywhere in the interior of the ellipse except potentially at the foci $(\xi,\phi)=(0,0)$ and $(0,\pi)$. In fact, since we expect $\tilde{\sigma}$ to be symmetric about the major and minor axes of the ellipse, it suffices to instead solve the equation on just the quarter-ellipse $\xi\in[0,\pi\xi_0/2)$, $\phi\in[0,\pi/2]$. Symmetry about the major axis of the ellipse is imposed by Neumann boundary conditions at $\phi=0$ and $\xi=0$, while symmetry about the minor axis is impose by a Neumann condition at $\phi=\pi/2$, so the boundary conditions on the computational domain are

$$\partial_\xi\tilde{\sigma}|_{\xi=0}=0,\quad \tilde{\sigma}|_{\xi=\pi\xi_0/2}=0,\quad \partial_\phi\tilde{\sigma}|_{\phi=0}=0,\quad \partial_\phi\tilde{\sigma}|_{\phi=\pi/2}=0. \tag{D.16}$$

We are thus left with a nonlinear elliptic boundary-value problem which can be solved by standard methods; we implement a Newton-Raphson nonlinear problem solver after discretization using pseudospectral methods with Chebyshev grids in both the ξ and ϕ directions. For the data discussed below, we used a grid size of 100 points in ξ and 101 points in ϕ, which is sufficient to reach machine precision except at very small values of ν and ξ_0, where large gradients make the numerics poorly-behaved.

Once a solution for $\tilde{\sigma}$ has been obtained, we can extract the near-boundary metric

function $g_2(\phi)$ appearing in (7.49):

$$g_2(\phi) = \partial_\xi^2 \tilde{\sigma}|_{\xi=\pi\xi_0/2} + \frac{1}{3\xi_0^2}. \tag{D.17}$$

As discussed in Section 7.5.1, comparison between our numerical results and analytic approximations reveals that

$$g_2(\phi) = -\frac{1+2\nu}{3\nu^2\xi_0^2} \sum_{k=-\infty}^{\infty} \operatorname{sech}^2\left(\frac{\phi - k\pi}{\xi_0}\right) + C(\nu, \xi_0), \tag{D.18}$$

where $C(\nu, \xi_0)$ is independent of ϕ to within the accuracy of our numerics. In particular, the difference between the analytic approximation (7.28) and the numerical solution of (7.20b) *does* exhibit nontrivial dependence on ϕ, as can be seen in the final plot of Figure 7.7; it is only in $g_2(\phi)$ that the difference becomes independent of ϕ. In Figure D.1 we show the ϕ-independence of $C(\nu, \xi_0)$ by plotting the relative difference

$$\Delta \equiv 1 - \frac{\min\limits_{\phi}\left[g_2^{\text{num}}(\phi) + \dfrac{1+2\nu}{3\nu^2\xi_0^2}\displaystyle\sum_{k=-\infty}^{\infty} \operatorname{sech}^2\left(\dfrac{\phi - k\pi}{\xi_0}\right)\right]}{\max\limits_{\phi}\left[g_2^{\text{num}}(\phi) + \dfrac{1+2\nu}{3\nu^2\xi_0^2}\displaystyle\sum_{k=-\infty}^{\infty} \operatorname{sech}^2\left(\dfrac{\phi - k\pi}{\xi_0}\right)\right]}, \tag{D.19}$$

where $g_2^{\text{num}}(\phi)$ is extracted from the numerical solutions of the Liouville equation via (D.17). Δ is order 10^{-7} or smaller in the entire parameter range (except for a narrow region around where $C(\nu, \xi_0)$ vanishes, where Δ is consequently a poor diagnostic of the ϕ-independence of $C(\nu, \xi_0)$). This is a consistent with a value of $\Delta = 0$ to the resolution of our numerics, and hence with $C(\nu, \xi_0)$ being independent of ϕ. Because numerical errors in our computation of $\tilde{\sigma}$ (and hence g_2^{num}) are largest near $\phi = 0$, in

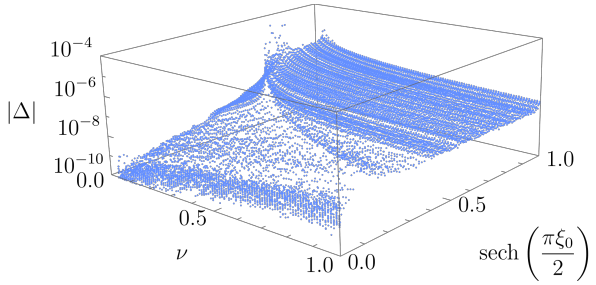


Figure D.1: The quantity Δ illustrating the ϕ-independence of $C(\nu, \xi_0)$; $\Delta \lesssim 10^{-7}$ in almost the entire parameter range, consistent with $\Delta = 0$ within the resolution of our numerics. The apparent spike in Δ appears around a region where $C(\nu, \xi_0)$ vanishes, and where Δ is consequently a poor diagnostic of the ϕ-independence of $C(\nu, \xi_0)$. Nevertheless, even in that region Δ never exceeds 10^{-2} or so.

practice we determine $C(\nu, \xi_0)$ by evaluating at $\phi = \pi/2$:

$$C(\nu, \xi_0) = g_2^{\text{num}}(\pi/2) + \frac{1+2\nu}{3\nu^2\xi_0^2} \sum_{k=-\infty}^{\infty} \text{sech}^2\left(\frac{(2k-1)\pi}{2\xi_0}\right). \tag{D.20}$$

As a check, we may compare the behavior of $C(\nu, \xi_0)$ extracted from the numerics with the analytic approximations in various limiting cases. For example, when ξ_0 and ν are small with $e^{-\pi/4\xi_0} \lesssim \nu$, we have that $C(\nu, \xi_0) = 1/(3\xi_0^2) + \mathcal{O}(\nu^2)$; likewise we have $C(\nu, \xi_0) = 1/(3\xi_0^2)$ exactly when $m = 2$ (corresponding to $\nu = \infty$); and in the limit $\xi_0 \to \infty$ we have $C(\nu, \xi_0) \to -1/(3\nu^2)$. Figure D.2 shows the difference $C(\nu, \xi_0) - 1/(3\xi_0^2)$; as expected, it is very small (in fact, limited by machine precision) at small ξ_0. In fact, the quality of the approximation is much better than would be expected from the analytic arguments: even at values of ν of order unity, $C(\nu, \xi_0) - 1/(3\xi_0^2)$ becomes arbitrarily small at small ξ_0. Evidently the analytic approximation for g_2 is much better than that for $\bar{\sigma}$.

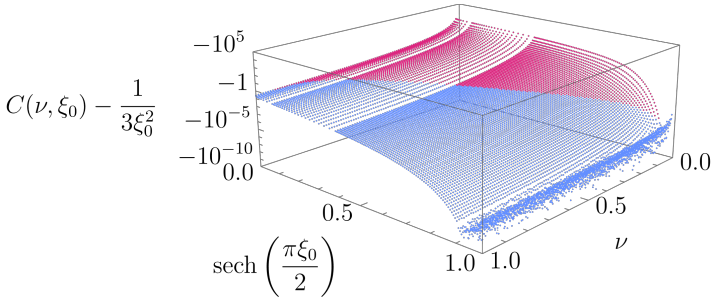


Figure D.2: The difference $C(\nu,\xi_0)-1/(3\xi_0^2)$, expected to be small when ν is large, and also when ξ_0 and ν are small with $e^{-\pi/4\xi_0}\lesssim\nu$. This difference becomes very small at small ξ_0, where it is eventually limited by machine precision. Points colored in blue (red) lie within (outside) the region $e^{-\pi/4\xi_0}<\nu$.

D.2.2 Boundary Wiggle

In our study of JT gravity coupled to a CFT, we must solve the equation of motion (7.61) and find the lowest eigenvalue of the fluctuation operator L defined in (7.63). The equation of motion for $u(\phi)$ is nonlinear, so we solve it using a standard Newton-Raphson method. We will not discuss this method further, except to note that the basin of attraction for solutions of (7.61) is relatively small, so in practice we need to implement a *damped* Newton-Raphson iterative scheme in order to achieve convergence. Likewise, the discretization of the "local" parts of the equation of motion and fluctuation operator can be performed using standard methods, which we will also avoid discussing in any detail. Instead, here we will focus on discretization of the more uncommon "nonlocal" parts of (7.61) and (7.63) defined by a smearing against the kernel $G(\phi,\tilde{\phi})$: that is, the discretization of the linear operator $\mathcal{G}$ defined by

$$(\mathcal{G}f)(\phi)=\int_0^{2\pi}d\tilde{\phi}\,G(\phi,\tilde{\phi})f(\tilde{\phi}), \tag{D.21}$$

where $G(\phi,\tilde{\phi})$ is defined in (7.59) and $f(\phi)$ is a periodic function of ϕ. We implement two approaches using pseudospectral methods on either a Fourier or a Chebyshev grid.

Fourier Discretization

Because the functions in which we are interested are periodic in ϕ, it is perhaps most natural to work on an evenly-spaced grid: that is, we discretize the domain into N grid points $\phi_j = 2\pi j/N$ for $j = 1, \ldots, N$ and work with the discrete values $f_j = f(\phi_j)$. We take N to be even. We then consider the discrete Fourier and inverse Fourier transforms of f_j, constructed by restricting wave numbers to just a single copy of the Brillouin zone:

$$\hat{f}_k = \frac{1}{N} \sum_{j=1}^{N} f_j \, e^{-ik\phi_j}, \qquad f_j = \sum_{k=-N/2}^{N/2\,\prime} \hat{f}_k \, e^{ik\phi_j}, \qquad \forall j, k \in \mathbb{Z}, \tag{D.22}$$

where the prime on the second sum indicates that the terms at $k = \pm N/2$ are to be multiplied by one half to compensate for the fact that the wave numbers $k = -N/2$ and $k = N/2$ should be identified. To compute the discretization of $\mathcal{G}$, we first consider the approximant

$$p(\phi) = \sum_{k=-N/2}^{N/2\,\prime} \hat{f}_k \, e^{ik\phi}, \tag{D.23}$$

which by construction satisfies $p(\phi_j) = f_j$. We then compute the action of $\mathcal{G}$ on $p(\phi)$, evaluate the result at the grid points $\phi = \phi_j$, and express the results in terms of f_j to obtain a matrix representation of $\mathcal{G}$. This procedure automatically truncates the Fourier series that defines $G(\phi, \tilde{\phi})$, and we find $(\mathcal{G}f)(\phi_i) = \sum_{j=1}^{N} \mathcal{G}_{ij} f_j$ where

$$\mathcal{G}_{ij} \equiv \frac{2}{N} \sum_{k=1}^{N/2\,\prime} k \left[\tanh\left(\frac{k\pi\xi_0}{2}\right) \cos(k\phi_i)\cos(k\phi_j) + \coth\left(\frac{k\pi\xi_0}{2}\right) \sin(k\phi_i)\sin(k\phi_j) \right], \tag{D.24}$$

with the prime still denoting that the term $k = N/2$ is multiplied by a factor of one half. By construction, this discretization of $\mathcal{G}$ (as well as the analogous discretizations of the derivative operators $d/d\phi$, $d^2/d\phi^2$, etc.) is *exact* when acting on any Fourier mode with

wave number up to $N/2$.

In practice, we expect solutions for the boundary wiggle to be symmetric about the axes of the ellipse, which in terms of $u(\phi)$ corresponds to $u'(\phi)$ being even in ϕ at $\phi = 0$ and $\pi/2$, assuming $u(\phi = 0) = 0$. Likewise, we expect eigenfunctions of the fluctuation operator L to have definite parity about $\phi = 0$ and $\pi/2$. It is numerically economical to exploit these symmetries by working on only the quarter-ellipse $\phi \in [0, \pi/2]$, which effectively improves our numerical resolution at fixed grid size by a factor of four: a grid size of n_ϕ (including $\phi = 0$) on the quarter-ellipse corresponds to a grid size of $N = 4(n_\phi - 1)$ around the full ellipse. Doing so is straightforward: for example, if a function $f(\phi)$ is symmetric about $\phi = 0$ and $\pi/2$, we have that $f_{N-i} = f_{N/2-i} = f_i$. Under the assumption of such a symmetry, any $N \times N$ matrix acting on f_i for $i = 1, \ldots, N$ can be expressed as an $n_\phi \times n_\phi$ matrix acting on f_i for $i = 0, \ldots, n_\phi - 1$ (where $f_0 \equiv f_N$). Consequently we restrict our attention to the quarter-ellipse.

We work with grid sizes up to $n_\phi = 500$ (corresponding to keeping Fourier modes up to wave number $k = 998$). However, in certain regions of parameter space, the wiggle function $u(\phi)$ exhibits large gradients near the computational boundaries $\phi = 0$ and $\phi = \pi/2$ which even such a large grid has difficulty resolving. Rather than continuing to increase the grid size, in such cases we implement discretization based on a Chebyshev grid, which naturally clusters more grid points near the computational boundaries and is much more effective at resolving large gradients there with smaller grid sizes.

Chebyshev Discretization

To implement pseudospectral methods on the quarter-ellipse with a Chebyshev grid, we could proceed directly by truncating the sum that defines $G(\phi, \tilde{\phi})$ and then using Clenshaw-Curtis quadrature to compute the integral that defines the operator $\mathcal{G}$. How-

ever, unlike the case of Fourier discretization just described, with a Chebyshev grid there is no natural rule for how to truncate the sum defining $G(\phi, \tilde{\phi})$, so one would need to introduce an additional parameter to capture this order of truncation. Consequently, we find it more convenient to rewrite $G(\phi, \tilde{\phi})$ as a much more rapidly-converging sum using Poisson summation.

To do so, we write $G(\phi, \tilde{\phi})$ in the form

$$G(\phi,\tilde{\phi}) \equiv \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\left[\left(k\tanh\left(\frac{k\pi\xi_0}{2}\right)-|k|\right)\cos(k\phi)\cos\Big(k\tilde{\phi}\Big)\right.$$
$$\left.+\left(k\coth\left(\frac{k\pi\xi_0}{2}\right)-|k|\right)\sin(k\phi)\sin\Big(k\tilde{\phi}\Big)\right]+\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}|k|e^{ik(\phi-\tilde{\phi})}. \quad \text{(D.25)}$$

The first sum converges exponentially in k, so we may freely use the Poisson summation formula on it. The second sum is not convergent but is still a well-defined distribution: noting that the Fourier transform of $|x|$ is $\mathcal{D}_\omega^2 \ln|\omega|/(2\pi^2)$ with $\mathcal{D}_\omega$ a distributional derivative[2], the Poisson summation formula gives the distributional relation

$$\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}|k|e^{ik(\phi-\tilde{\phi})} = \frac{1}{\pi}\sum_{k=-\infty}^{\infty}\mathcal{D}_{\tilde{\phi}}^2\ln\left|\tilde{\phi}-\phi+2k\pi\right|; \quad \text{(D.26)}$$

the right-hand sum is convergent (it is $(1/\pi)\mathcal{D}_{\tilde{\phi}}^2\ln|\sin\Big((\tilde{\phi}-\phi)/2\Big)|$). Consequently we

[2]That is, $\mathcal{D}_\omega^n \ln|\omega|$ is a homogeneous distribution of degree $-n$. Essentially, when integrated against a test function one integrates by parts "ignoring the singularity": e.g. for $a, b \neq 0$,

$$\int_a^b d\omega\, f(\omega)\mathcal{D}_\omega \ln|\omega| \equiv [f(\omega)\ln|\omega|]_a^b - \int_a^b d\omega\, f'(\omega)\ln|\omega| = \mathrm{PV}\int_a^b d\omega\,\frac{f(\omega)}{\omega},$$
$$\int_a^b d\omega\, f(\omega)\mathcal{D}_\omega^2 \ln|\omega| \equiv \left[\frac{f(\omega)}{\omega}\right]_a^b - [f'(\omega)\ln|\omega|]_a^b + \int_a^b d\omega\, f''(\omega)\ln|\omega|,$$

where PV denotes the Cauchy principal value.

find

$$G(\phi,\tilde{\phi}) = \frac{1}{\pi}\sum_{k=-\infty}^{\infty} \mathcal{D}_{\tilde{\phi}}^2 \ln\left|\frac{1-e^{-(\tilde{\phi}-\phi+2k\pi)/\xi_0}}{1+e^{-(\tilde{\phi}+\phi+2k\pi)/\xi_0}}\right|, \tag{D.27}$$

so the action of $\mathcal{G}$ on a test function is

$$(\mathcal{G}f)(\phi) = \frac{2\phi}{\pi\xi_0}f'(0) + \frac{1}{\pi}\sum_{k=-\infty}^{\infty}\int_0^{2\pi} d\tilde{\phi}\, f''(\tilde{\phi}) \ln\left|\frac{1-e^{-(\tilde{\phi}-\phi+2k\pi)/\xi_0}}{1+e^{-(\tilde{\phi}+\phi+2k\pi)/\xi_0}}\right|. \tag{D.28}$$

The terms in the sum are $\mathcal{O}(e^{-2\pi|k|/\xi_0})$ at large $|k|$, so the sum is rapidly convergent. However, the integrand of the $k = 0$ term is singular at $\tilde{\phi} = \phi$, so we cannot yet straightforwardly discretize this expression for $\mathcal{G}$. Instead we integrate by parts to obtain

$$\begin{aligned}(\mathcal{G}f)(\phi) = \frac{2\phi}{\pi\xi_0}f'(0) - \frac{\pi\xi_0}{2}f''(0) \\ + \sum_{k=-\infty}^{\infty}\left[\frac{4\phi}{\xi_0}f''(0)H(-k) - \int_0^{2\pi} d\tilde{\phi}\, f'''(\tilde{\phi})\widehat{G}_3(\phi,\tilde{\phi}+2k\pi)\right],\end{aligned} \tag{D.29}$$

where $H(x)$ is the Heaviside step function (with the convention $H(0) = 0$) and we have defined

$$\widehat{G}_3(\phi,\tilde{\phi}) \equiv \frac{\xi_0}{\pi}\,\mathrm{Re}\left[\mathrm{Li}_2\left(e^{-(\tilde{\phi}-\phi)/\xi_0}\right) - \mathrm{Li}_2\left(-e^{-(\tilde{\phi}+\phi)/\xi_0}\right)\right], \tag{D.30}$$

with $\mathrm{Li}_2(z)$ the dilogarithm. The dilogarithm is a continuous function on the real line, and hence the integrand is everywhere-finite, as desired. Moreover, note that in integrating by parts, we included a nonzero constant of integration to ensure that the sum is rapidly convergent: as can be verified using the asymptotics of the dilogarithm, the terms of this sum also decay like $e^{-2\pi|k|/\xi_0}$ at large $|k|$. In practice we find that truncating to $|k| \leq 15$ is more than sufficient for obtaining accurate results for all the values of ξ_0 we consider (i.e. up to $\xi_0 = 10$). To discretize $\mathcal{G}$ on the quarter-ellipse, we impose appropriate parity of $f(\phi)$ across $\phi = 0$ and $\phi = \pi/2$, which allows us to evaluate all integrals on the

reduced domain $\phi \in [0, \pi/2]$. We then discretize the integrals in (D.29) via Clenshaw-Curtis quadrature on a Chebyshev grid on this domain.

It should be noted that the price we pay for reexpressing $\mathcal{G}$ in the form (D.29) is the loss of spectral accuracy due to the need to integrate the dilogarithm $\mathrm{Li}_2(z)$ through its non-analytic behavior at $z = 1$. Nevertheless, in practice we find that in this new formulation, a Chebyshev grid of size 201 on the quarter-ellipse is sufficient to comfortably resolve the large gradients in $u(\phi)$ near $\phi = 0$ and $\phi = \pi/2$.

D.3 Bulk Solutions to JT + Scalar for $m = 1$ and 2

In this Appendix we explain how to obtain the analytic solutions (7.65) for the wiggle which satisfy the equations of motion with the matter sources given by (7.64) (under an appropriate choice of A) on the Poincaré disk an on the double-trumpet. The idea is to solve the *bulk* problem by constructing a solution for the dilaton Φ and the scalar ψ on the disk or the double-trumpet, and from this bulk solution to then extract the boundary quantities $u(\phi)$ and $\psi_\partial(u)$.

D.3.1 General Dilaton Solution

We begin by constructing a general solution to the JT + matter equations of motion on any portion of a Riemann surface of constant negative curvature that can be covered with a single coordinate chart, and with arbitrary conserved matter stress tensor. With bulk action $I = I_{\mathrm{JT}} + I_{\mathrm{mat}}$, the equation of motion for the metric gives

$$-\boldsymbol{\nabla}_a \boldsymbol{\nabla}_b \Phi + \Phi g_{ab} = T_{ab} - T g_{ab}, \tag{D.31}$$

where T_{ab} is the stress tensor obtained from the matter action I_{mat} (and T is its trace).

We solve (D.31) in conformal gauge, in which the metric takes the form

$$ds^2 = e^{2\sigma} dz\, d\bar{z} \tag{D.32}$$

where $\sigma(z, \bar{z})$ solves the Liouville equation (except for potentially at isolated conical defects). Conservation of the stress tensor yields

$$\partial_{\bar{z}} T_{zz} + e^{2\sigma} \partial_z \left(e^{-2\sigma} T_{z\bar{z}} \right) = 0, \tag{D.33}$$

and likewise with $z \leftrightarrow \bar{z}$. In this conformal gauge, (D.31) gives

$$-\partial_z \partial_z \Phi + 2\partial_z \sigma \partial_z \Phi = T_{zz}, \tag{D.34a}$$

$$-\partial_z \partial_{\bar{z}} \Phi + \frac{1}{2} e^{2\sigma} \Phi = -T_{z\bar{z}}, \tag{D.34b}$$

$$-\partial_{\bar{z}} \partial_{\bar{z}} \Phi + 2\partial_{\bar{z}} \sigma \partial_{\bar{z}} \Phi = T_{\bar{z}\bar{z}}, \tag{D.34c}$$

which can be integrated explicitly. To do so, we first integrate (D.34a) to get

$$\Phi(z, \bar{z}) = A(\bar{z}) + \int_{z_0}^{z} dw\, e^{2\sigma(w,\bar{z})} \left[B(\bar{z}) - \int_{z_0}^{w} du\, e^{-2\sigma(u,\bar{z})} T_{zz}(u, \bar{z}) \right], \tag{D.35}$$

where A and B are arbitrary antiholomorphic functions and the integrals are contour integrals that start at an arbitrary point z_0 and end at z (avoiding conical defects if there are any). We can relate A and B by inserting this solution into (D.34b), which gives

$$A(\bar{z}) = 2B'(\bar{z}) - s_1(\bar{z}) B(\bar{z}) - \mathcal{T}_1(\bar{z}), \tag{D.36}$$

where we have defined the antiholomorphic functions

$$s_1(\bar{z}) \equiv -4\partial_{\bar{z}}\sigma + \int_{z_0}^{z} dw\, e^{2\sigma(w,\bar{z})}, \tag{D.37a}$$

$$\begin{aligned}\mathcal{T}_1(\bar{z}) \equiv\ & 2e^{-2\sigma}T_{z\bar{z}} + 4\partial_{\bar{z}}\sigma \int_{z_0}^{z} dw\, e^{-2\sigma(w,\bar{z})}T_{zz}(w,\bar{z}) \\ & + \int_{z_0}^{z} dw \left[2\partial_{\bar{z}}\left(e^{-2\sigma(w,\bar{z})}T_{zz}(w,\bar{z})\right) - e^{2\sigma(w,\bar{z})}\int_{z_0}^{w} du\, e^{-2\sigma(u,\bar{z})}T_{zz}(u,\bar{z})\right];\end{aligned} \tag{D.37b}$$

the fact that these are antiholomorphic follows from the Liouville equation and the conservation of the stress tensor. Consequently we may evaluate them at any value of z; choosing $z = z_0$ gives the simpler expressions

$$s_1(\bar{z}) = -4\partial_{\bar{z}}\sigma(z_0,\bar{z}), \qquad \mathcal{T}_1(\bar{z}) = 2e^{-2\sigma(z_0,\bar{z})}T_{z\bar{z}}(z_0,\bar{z}). \tag{D.38}$$

Finally, using (D.36) in (D.34c) yields a third-order differential equation for B:

$$-2B'''(\bar{z}) + s_2(\bar{z})B'(\bar{z}) + \frac{1}{2}s_2'(\bar{z})B(\bar{z}) + \mathcal{T}_2(\bar{z}) = 0, \tag{D.39}$$

where we have introduced the additional antiholomorphic functions

$$s_2(\bar{z}) \equiv 8\left((\partial_{\bar{z}}\sigma)^2 - \partial_{\bar{z}}^2\sigma\right), \tag{D.40a}$$

$$\mathcal{T}_2(\bar{z}) \equiv e^{2\sigma}\partial_{\bar{z}}\left[e^{-2\sigma}\left(\mathcal{T}_1' + \int_{z_0}^{z} dw\, \partial_{\bar{z}}\left(e^{2\sigma(w,\bar{z})}\int_{z_0}^{w} du\, e^{-2\sigma(u,\bar{z})}T_{zz}(u,\bar{z})\right)\right)\right] - T_{\bar{z}\bar{z}}; \tag{D.40b}$$

as with s_1 and $\mathcal{T}_1$, one can verify that s_2 and $\mathcal{T}_2$ are antiholomorphic using the Liouville equation and the conservation of the stress tensor. So we may again evaluate these

functions at $z = z_0$, which gives the relations

$$s_2 = 2s_1' + \frac{s_1^2}{2}, \qquad \mathcal{T}_2 = \mathcal{T}_1'' + \frac{1}{2}\, s_1 \mathcal{T}_1' - T_{\bar{z}\bar{z}}(z_0, \bar{z}). \tag{D.41}$$

Using these and (D.36) we may express (D.39) entirely in terms of A:

$$-A'' - \frac{s_1}{2}\, A' - T_{\bar{z}\bar{z}}(z_0, \bar{z}) = 0. \tag{D.42}$$

Using (D.38) this can be integrated to give

$$A(\bar{z}) = c_1 + \int_{\bar{z}_0}^{\bar{z}} d\bar{w}\, e^{2\sigma(z_0,\bar{w})} \left[c_2 - \int_{\bar{z}_0}^{\bar{w}} d\bar{u}\, e^{-2\sigma(z_0,\bar{u})} T_{\bar{z}\bar{z}}(z_0, \bar{u}) \right], \tag{D.43}$$

where c_1, c_2, and $\bar{z}_0$ are arbitrary (complex) constants. Finally, we then integrate (D.36) to obtain

$$B(\bar{z}) = e^{-2\sigma(z_0,\bar{z})} \left[c_3 + \frac{1}{2} \int_{\bar{z}_0}^{\bar{z}} d\bar{w}\, e^{2\sigma(z_0,\bar{w})} \left(A(\bar{w}) + \mathcal{T}_1(\bar{w}) \right) \right], \tag{D.44}$$

where c_3 is another arbitrary complex constant. Consequently, equation (D.35) with A and B given by (D.43) and (D.44) is the general solution for the dilaton for arbitrary σ and matter stress tensor. Note that any changes to z_0 and $\bar{z}_0$ can be reabsorbed into the c_i, so there are only three independent degrees of freedom in this solution.

When the matter is classical and conformal, $T_{z\bar{z}} = 0$, and hence T_{zz} and $T_{\bar{z}\bar{z}}$ are holomorphic and antiholomorphic, respectively. This simplification, along with use of the Liouville equation, allows us to integrate (D.35) to obtain

$$\begin{aligned} \Phi = 4e^{-2\sigma(z_0,\bar{z})} & \left[\tilde{c}_3 + 2\tilde{c}_1 \partial_z \sigma + 4\tilde{c}_2 \partial_z^2 \sigma - 2\partial_z (e^{-2\sigma} \partial_z (e^{2\sigma} \overline{F})) \right]\Big|_{z=z_0} \left(\partial_{\bar{z}} \sigma(z, \bar{z}) - \partial_{\bar{z}} \sigma(z_0, \bar{z}) \right) \\ & + \tilde{c}_1 + 4\tilde{c}_2 \partial_z \sigma(z_0, \bar{z}) - 2e^{-2\sigma(z_0,\bar{z})}\, \partial_z \left(e^{2\sigma} \overline{F} \right)\Big|_{z=z_0} - 2e^{-2\sigma(z,\bar{z})} \partial_{\bar{z}} \left(e^{2\sigma} F \right), \end{aligned} \tag{D.45}$$

where the $\tilde{c}_i$ are constants and we have defined

$$F(z,\bar{z}) \equiv \int_{z_0}^{z} dw\, e^{-2\sigma(w,\bar{z})} T_{zz}(w), \qquad \overline{F}(z,\bar{z}) \equiv \int_{\bar{z}_0}^{\bar{z}} d\bar{w}\, e^{-2\sigma(z,\bar{w})} T_{\bar{z}\bar{z}}(\bar{w}). \tag{D.46}$$

We'd now like to use this expression to solve for the dilaton in the presence of conformal matter on the Poincaré disk and the double-trumpet. For the real massless scalar field ψ, the equation of motion $\boldsymbol{\nabla}^2\psi = 0$ is solved by

$$\psi(z,\bar{z}) = f(z) + \bar{f}(\bar{z}) \tag{D.47}$$

for an arbitrary holomorphic function f, and the corresponding components of the stress tensor are $T_{zz} = (f')^2$ and $T_{\bar{z}\bar{z}} = (\bar{f}')^2$. In principle we should determine f by imposing Dirichlet boundary conditions at ∂M, but for our purposes in this section we will instead choose f and then determine the corresponding boundary conditions from it.

D.3.2 Poincaré Disk

The Poincaré disk is given by conformal factor $\sigma = \ln(2/(1-z\bar{z}))$, with z covering the unit disk. We take the scalar field to be given by (D.47) with $f(z) = Jz^n/2$ for positive integer n. Then (D.45) with $z_0 = \bar{z}_0 = 0$ gives the general solution

$$\Phi = \frac{\alpha_1(1+r^2) + 2r(\alpha_2\cos\theta + \alpha_3\sin\theta) + \frac{nJ^2r^{2n}}{4}\left(\frac{r^2}{2n+1} - \frac{1}{2n-1}\right)\cos(2n\theta)}{1-r^2}, \tag{D.48}$$

where the α_i are real and we have converted to the usual polar coordinates using $z = re^{i\theta}$. We expect the boundary conditions (7.30) to constrain the α_i, but we also expect there to be residual freedom in these constants due to the $SL(2,\mathbb{R})$ symmetry of the Poincaré

disk. So for the purposes of constructing a particular solution, we set $\alpha_2 = 0 = \alpha_3$. Then the boundary condition $\Phi|_{\partial M} = 1/\delta$ gives an embedding of the boundary ∂M:

$$r = 1 - \Phi_0(\theta)\,\delta + \mathcal{O}(\delta)^2, \text{ where } \Phi_0(\theta) = \alpha_1 - \frac{nJ^2}{4(4n^2-1)}\cos(2n\theta). \tag{D.49}$$

With this embedding we may fix α_1 by imposing the requirement that the length of ∂M be β/δ:

$$\beta = \int_0^{2\pi} \frac{d\theta}{\Phi_0(\theta)} \quad \Rightarrow \quad \alpha_1 = \frac{2\pi}{\beta}\sqrt{1 + \left(\frac{n\beta J^2}{8\pi(4n^2-1)}\right)^2}. \tag{D.50}$$

Finally, we may define the proper length coordinate u along ∂M using $du = d\theta/\Phi_0(\theta)$, which gives the wiggle function $u(\theta)$ (7.65) with

$$A = \left[1 + \left(\frac{8\pi(4n^2-1)}{n\beta J^2}\right)^2\right]^{-1/2}. \tag{D.51}$$

Likewise, the boundary profile of the scalar is given simply by $\psi|_{\partial M} = J\cos(n\theta)$, which written in terms of u takes the form (7.64).

These solutions solve the equation of motion (7.61) in the boundary formalism. This can be verified most easily by noting that when $m = 1$, the conical defects vanish and hence the on-shell action should be independent of ξ_0. For simplicity we can therefore work purely in the $\xi_0 \to \infty$ limit of the action (7.60) in which the ellipse degenerates into the Poincaré disk (with no defects). In this $\xi_0 \to \infty$ limit with $m = 1$, we have

$$g_2(\phi) \to -\frac{1}{3}, \; G(\phi, \tilde{\phi}) \to \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} |k| e^{ik(\phi-\tilde{\phi})}. \tag{D.52}$$

We then find that the equation of motion (7.61) is satisfied by (7.65) if

$$\frac{1}{\sqrt{1-A^2}} = \sqrt{1+\left(\frac{n\beta J^2}{8\pi(4n^2-1)}\right)^2}\,; \tag{D.53}$$

this agrees precisely with the expression (D.51) obtained from the bulk solutions. The action of the wiggle profile (7.65) is

$$\widehat{I}_1 = -S_0 + \frac{1}{\beta}\left[\frac{2\pi^2(4n^2-1)}{\sqrt{1-A^2}} - 8\pi^2 n^2 + \frac{\pi n\beta J^2}{2}\right], \tag{D.54}$$

which recovers the pure JT Schwarzian result $\widehat{I}_1 = -S_0 - 2\pi^2/\beta$ when $J = 0$.

D.3.3 Double Trumpet

To treat the double-trumpet, we will work on the quotient space (i.e. the single trumpet) with appropriate regularity conditions imposed to ensure smoothness of the scalar and the dilaton on the quotient geometry. The conformal factor on the trumpet is

$$\sigma = \ln\left[\frac{1}{\xi_0}\sec\left(\frac{z+\bar{z}}{2\xi_0}\right)\right], \tag{D.55}$$

with $z \equiv \xi + i\phi$ covering the strip $\mathrm{Re}\,(z) \in [0, \pi\xi_0/2)$, $\mathrm{Im}\,(z) \in [0, 2\pi)$. A wormhole solution will involve fixing not just the dilaton Φ, but also ξ_0, which sets the size of the wormhole. We take the scalar field to be given by (D.47) with

$$f(z) = \frac{J\cosh(nz)}{2\cosh(n\pi\xi_0/2)} \tag{D.56}$$

for positive integer n; this choice ensures that ψ obeys all the regularity conditions discussed in Section 7.5.1. Using (D.45) with $z_0 = \bar{z}_0 = 0$, we find that imposing these

same regularity conditions on Φ fixes the three constants $\tilde{c}_i$ uniquely, leaving the solution

$$\Phi = \frac{nJ^2\xi_0}{4\cosh^2(n\pi\xi_0/2)}\left[n\left(\xi_0 + \xi\tan\left(\frac{\xi}{\xi_0}\right)\right)\right.$$
$$\left. -\frac{\cos(2n\phi)}{1+4n^2\xi_0^2}\left(n\xi_0\cosh(2n\xi) + \frac{1}{2}\sinh(2n\xi)\tan\left(\frac{\xi}{\xi_0}\right)\right)\right]. \quad \text{(D.57)}$$

Again we impose the boundary condition $\Phi|_{\partial M} = 1/\delta$ to obtain an embedding of ∂M:

$$\xi = \frac{\pi\xi_0}{2} - \Phi_0(\phi)\,\delta + \mathcal{O}(\delta^2), \quad \text{(D.58)}$$

where

$$\Phi_0(\phi) = \frac{nJ^2\xi_0^2}{8\cosh^2(n\pi\xi_0/2)}\left(n\pi\xi_0 - \frac{\cos(2n\phi)\sinh(n\pi\xi_0)}{1+4n^2\xi_0^2}\right). \quad \text{(D.59)}$$

We must have $\Phi_0(\phi) > 0$ for all ϕ, but $\Phi_0(0)$ becomes negative at large ξ_0; hence solutions only exist when ξ_0 is sufficiently small[3]. Requiring that the length of ∂M be β/δ then fixes the allowed values ξ_* of ξ_0 in terms of J and n:

$$\beta = \int_0^{2\pi}\frac{d\phi}{\Phi_0(\phi)} \quad \Rightarrow \quad \frac{\beta J^2}{n} = \frac{16\cosh^2(n\pi\xi_*/2)}{n^3\xi_*^3\sqrt{1-\left(\frac{\sinh(n\pi\xi_*)}{n\pi\xi_*(1+4n^2\xi_*^2)}\right)^2}}. \quad \text{(D.60)}$$

Again we may define the proper length coordinate u along ∂M using $du = d\phi/\Phi_0(\phi)$, which gives the wiggle function $u(\phi)$ (7.65) with

$$A = \frac{\sinh(n\pi\xi_*)}{n\pi\xi_*(1+4n^2\xi_*^2)}. \quad \text{(D.61)}$$

The boundary profile of the scalar is given simply by $\psi|_{\partial M} = J\cos(n\phi)$, which written in terms of u takes the form (7.64).

[3]Specifically, when $n\xi_0 < x_*$, where $x_* \approx 1.36$ is the positive solution of $\pi x(1+4x^2) = \sinh(\pi x)$.

It is straightforward to check that (D.60) and (D.61) satisfy (7.67) and (7.68), so again we confirm that the solutions obtained using the bulk and boundary formalisms coincide. In particular, the need for $\sqrt{\beta}J$ to be sufficiently large in order for solutions to exist can be inferred from (D.60): in the region where the right-hand side is real, it has a global minimum at $n\xi_* \approx 0.93$ where it attains the value ≈ 146. Hence solutions for ξ_0 only exist when $\sqrt{\beta}J \geq \sqrt{\beta}J_{\text{min}} \approx 12\sqrt{n}$, just as we found in Section 7.5.2.

D.4 JT + Branes

In this Appendix we discuss a model of JT gravity coupled to end-of-the-world (EOW) branes. This model is effectively a classical version of that considered in [50], except that we do not give the branes any internal degrees of freedom. Our purpose is illustrative: though we do not find stable wormholes at $m < 1$, we will see very clearly that for $m < 1$ multiple branches of solutions for the wiggle can appear in a way analogous to the more involved model studied in Section 7.5.

D.4.1 Boundary Action

The JT + brane model has the Euclidean action

$$I_{\text{brane}} = I_{\text{JT}} + \mu \int_B ds, \tag{D.62}$$

where B is an EOW brane anchored to the boundary ∂M and $\mu > 0$ is its tension. With m boundaries, the geometry M consists of the Poincaré disk with m geodesic "bites" removed corresponding to the location of m disconnected portions of B, as shown in Figure D.3. The boundary ∂M consists of m disconnected pieces anchored to B, and

we take the length of each of these pieces of ∂M to be β/δ. After quotienting by the $\mathbb{Z}_m$ replica symmetry, it is clear that the quotient space geometry $\widehat{M}_m$ is in fact identical to (a single copy of) the geometry $\widehat{M}_m^{(2)}$ discussed in Section 7.4.1. Consequently, the JT part of the quotient action takes half its value in pure JT after the replacement $\beta \to 2\beta$:

$$\widehat{I}_m = \frac{m-2}{2m} S_0 - \int_{-\beta/2}^{\beta/2} du \, \mathrm{Sch}\left(\tan\left(\frac{\Theta}{2}\right), u\right) + \mu \, \mathrm{Length}(B). \tag{D.63}$$

(Hence the case $\mu = 0$ can be thought of as a quotient of pure JT by a $\mathbb{Z}_2$ symmetry about the geodesic B.) To put the brane on-shell, we must compute its length up to the cutoff boundary ∂M. From Figure D.3, it is clear that B is diffeomorphic to a geodesic on the Poincaré disk sweeping out an angle 2α on the boundary. Moreover, footnote 8 relating the embedding functions $\Theta(u)$ and $R(u)$ indicates that this geodesic is cut off by ∂M at the radial cutoffs $R_\pm = 1 - \Theta'(\pm\beta/2)\,\delta + \mathcal{O}(\delta^2)$. The length of B up to these cutoffs is then

$$\mathrm{Length}(B) = \ln\left(\frac{\beta^2}{\delta^2}\right) - \ln\left(\frac{4\beta^2\Theta'(\beta/2)\Theta'(-\beta/2)}{\sin^2\alpha}\right) + \mathcal{O}(\delta). \tag{D.64}$$

The first term is the expected UV divergent piece, and it can be cancelled out by adding an appropriate counterterm such as $-2\mu \ln(\beta\Phi)|_{\partial M}$ to the action. After this cancellation, the renormalized action for the wiggle is

$$\widehat{I}_m = \frac{m-2}{2m} S_0 - \int_{-\beta/2}^{\beta/2} du \, \mathrm{Sch}\left(\tan\left(\frac{\Theta}{2}\right), u\right) - \mu \ln\left(\frac{4\beta^2\Theta'(\beta/2)\Theta'(-\beta/2)}{\sin^2\alpha}\right). \tag{D.65}$$

We must next determine the boundary conditions to impose on the wiggle $\Theta(u)$. One condition is fixed as in pure JT by requiring that $\Theta(u)$ wrap around the entire boundary of $\widehat{M}_m^{(2)}$, while another will stem from fixing the angle at which ∂M intersects B. One

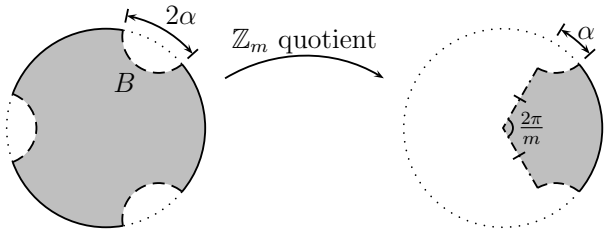


Figure D.3: For positive integer m, the genus-zero wormhole geometry corresponding to the JT + brane model consists of the Poincaré disk with m geodesic "bites" B removed, shown at left. The quotient geometry is shown at right; the dot-dashed radial lines are identified, so the two half-geodesics correspond to a single copy of one of the branes B. This quotient geometry coincides with the geometry $\widehat{M}_m^{(2)}$ shown in Figure 7.4.

way to infer this latter boundary condition is to treat the brane as a particle of mass μ which scatters with the boundary trajectory ∂M and impose conservation of the $SL(2,\mathbb{R})$ charges in this scattering process, as discussed in [224, 225]. Alternatively, we may impose the boundary condition $n^a \boldsymbol{\nabla}_a \Phi|_B = \mu$ (with n^a a unit outward-pointing normal to B) on the dilaton in the bulk [50] and convert it to the desired constraint on the intersection of ∂M and B. Ultimately we find that $n \cdot u = \mu\delta + \mathcal{O}(\delta^2)$, where n^a and u^a are outward-pointing unit normals to B and ∂M[4]. This constraint ultimately leads to the wiggle boundary conditions

$$\Theta\left(\pm\beta/2\right) = \pm\left(\frac{\pi}{m} - \alpha\right), \quad \Theta''\left(\pm\beta/2\right) = \mp\left(\cot\alpha\,\Theta'\left(\pm\beta/2\right)^2 - \mu\Theta'\left(\pm\beta/2\right)\right). \tag{D.66}$$

[4]One way of determining this condition explicitly is to work on the Poincaré disk, placing the brane on the axis $\theta = \pm\pi/2$, and considering the family of dilaton solutions

$$\Phi = \frac{(1+r_0^2)(1+r^2) - 4r_0 r\cos\theta}{(1-r_0^2)(1-r^2)}$$

which are obtained from the "standard" solution $\Phi = (1+r^2)/(1-r^2)$ by translating the origin to the right a coordinate distance r_0 using an $SL(2,\mathbb{R})$ transformation. r_0 and μ can be related using the bulk boundary condition $n^a \boldsymbol{\nabla}_a \Phi|_B = 2r_0/(1-r_0^2) = \mu$, while it is easy to verify that where the level sets $\Phi = 1/\delta$ intersect the brane, they satisfy $n \cdot u = 2r_0\delta/(1-r_0^2) + \mathcal{O}(\delta^2) = \mu\delta + \mathcal{O}(\delta^2)$. This is a local condition on the intersection of ∂M and B, and hence it must hold in any other geometry as well.

Finally, we will ultimately need to investigate the stability of the wiggle. As usual, we restrict to perturbations that exhibit a $\mathbb{Z}_2$ symmetry corresponding to reflection about $u = 0$. The stability analysis proceeds just as in the pure JT case discussed in Section 7.4.3: we write $\Theta = \widetilde{\Theta} + \vartheta$ with $\widetilde{\Theta}$ a solution to the equations of motion and expand the action to second order in ϑ. The resulting fluctuation operator L is identical to the one for pure JT (7.44), except that it acts on the space of functions obeying the boundary conditions

$$\vartheta(0) = 0 = \vartheta''(0), \quad \vartheta(\beta/2) = 0, \quad \vartheta''(\beta/2) = -\left(2\cot\alpha\,\widetilde{\Theta}'(\beta/2) - \mu\right)\vartheta'(\beta/2). \quad \text{(D.67)}$$

Hence a solution $\widetilde{\Theta}$ is stable if and only if the spectrum of L is nonnegative on the space of functions obeying these boundary conditions. We compute the spectrum of L numerically as described in Appendix D.2.

D.4.2 Saddles

We now look for saddles of the action (D.65). As in the main text, we proceed by first looking for saddles for the wiggle $\Theta(u)$ at fixed modulus α and then evaluate the action on these saddles to obtain an effective action $\widehat{I}_m[\alpha]$ for α, which we then examine to look for saddles for α.

Saddles for the Wiggle

With the boundary conditions (D.66), the variational problem for the action (D.65) is well-posed and leads to the same equation of motion (7.33) we obtained in pure JT. Since the boundary conditions (D.66) are odd in u, we may again consider the most general

odd solution:

$$\tan\left(\frac{\Theta(u)}{2}\right) = a\tan\left(\frac{bu}{2\beta}\right) \tag{D.68}$$

(the extra factor of 2 on the right-hand side relative to (7.34) is inserted for convenience to account for the relative factor of 2 in β between pure JT and the JT + brane model). The constants a and b are determined by imposing the boundary conditions (D.66) just as in pure JT. In short, we first take $a = ia_i$ and $b = -ib_i$ with a_i and b_i both real and positive. Then solutions can only exist when $\pi/m - \alpha < \pi$ in which case a_i and b_i must satisfy

$$\cosh\left(\frac{b_i}{2}\right) + \frac{\beta\mu}{b_i}\sinh\left(\frac{b_i}{2}\right) = \frac{\sin(\pi/m)}{\sin\alpha}, \tag{D.69a}$$

$$a_i = \coth\left(\frac{b_i}{4}\right)\tan\left(\frac{\pi - m\alpha}{2m}\right). \tag{D.69b}$$

On the other hand, taking $a = a_r$ and $b = b_r$ with a_r and b_r both real and positive, we find that solutions must satisfy

$$\cos\left(\frac{b_r}{2}\right) + \frac{\beta\mu}{b_r}\sin\left(\frac{b_r}{2}\right) = \frac{\sin(\pi/m)}{\sin\alpha}, \tag{D.70a}$$

$$\frac{b_r}{2\pi} \in \begin{cases} (2N, 2N+1] & \text{if } \tan\left(\frac{\pi-m\alpha}{2m}\right) > 0, \\ (2N-1, 2N) & \text{if } \tan\left(\frac{\pi-m\alpha}{2m}\right) < 0, \end{cases} \qquad N \equiv \left\lfloor \frac{1}{2m} + \frac{\pi-\alpha}{2\pi} \right\rfloor, \tag{D.70b}$$

$$a_r = \cot\left(\frac{b_r}{4}\right)\tan\left(\frac{\pi - m\alpha}{2m}\right). \tag{D.70c}$$

It is now straightforward to see how turning on a brane can give rise to new branches of solutions. For any value of μ, the left-hand side of (D.69a) is monotonic in $b_i > 0$, so when solutions of exponential type exist (i.e. when $\pi/m - \alpha < \pi$ and $\sin(\pi/m)/\sin\alpha > 1 + \mu\beta/2$), then precisely one solution exists. Similarly, for $\mu = 0$ the left-hand side of (D.70a)

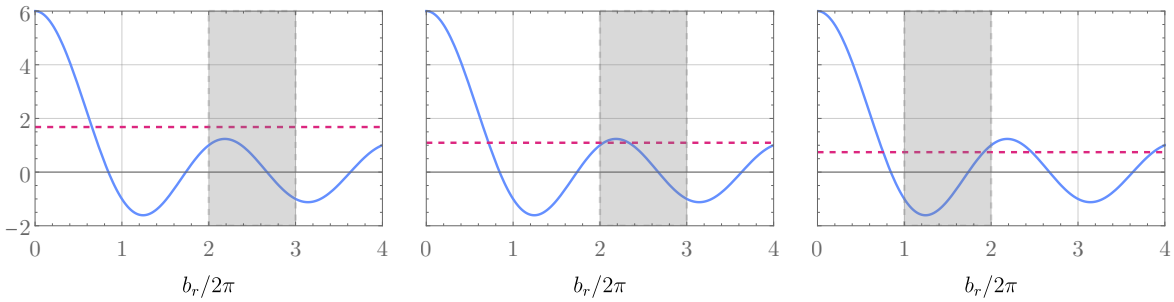


Figure D.4: Graphical solutions of (D.70) for b_r when $m < 1$. The solid blue curves and the dashed red lines show the left- and right-hand sides of (D.70a), respectively, while the shaded region is the interval satisfying (D.70b); hence solutions of (D.70) correspond to intersections of the blue and red lines in the shaded regions. All three plots take $\beta\mu = 10$ and $m = 0.45$, while from left to right we show $\alpha = \pi/8$, $\pi/5$, and $\pi/3$. At small α there are no solutions; as α grows, two solutions appear at a common value of b_r and branch apart; as α grows further, the larger of the two solutions disappears abruptly as the region satisfying (D.70b) changes.

is monotonic in b_r in the interval allowed by the constraint (D.70b), so again at most one solution of oscillatory type can exist (and it is given by the pure JT solution (7.40)). However, for $\mu \neq 0$ the left-hand side of (D.70a) is *not* in general monotonic in the interval allowed by (D.70b), and consequently may admit an additional solution. In Figure D.4 we graphically illustrate the structure of solutions for b_r for various values of α, showing how zero, one, or two solutions may exist.

Saddles for the Modulus

Putting the wiggle on-shell, we are left with an effective action $\widehat{I}_m[\alpha]$ for α:

$$\widehat{I}_m[\alpha] = \frac{m-2}{2m} S_0 + \begin{cases} \dfrac{b_i^2}{2\beta} - 2\mu \ln\left(\dfrac{2b_i}{\sinh(b_i/2)} \dfrac{\sin(\pi/m-\alpha)}{\sin\alpha}\right), & \text{exponential,} \\ -\dfrac{b_r^2}{2\beta} - 2\mu \ln\left(\dfrac{2b_r}{\sin(b_r/2)} \dfrac{\sin(\pi/m-\alpha)}{\sin\alpha}\right), & \text{oscillatory,} \end{cases} \tag{D.71}$$

with b_i and b_r implicit functions of α (as well as m and $\beta\mu$) through (D.69) and (D.70).

When $m \geq 1$, we always have $N = 0$ and $\sin(\pi/m)/\sin\alpha \geq 0$ for any allowed α, from

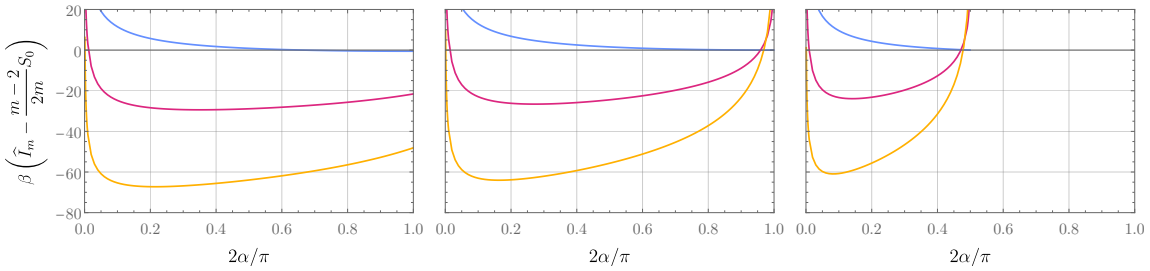


Figure D.5: The effective action $\widehat{I}_m[\alpha]$ for various values of $m > 1$ and $\beta\mu$. From left to right we show $m = 3/2$, 2, and 4, while the blue, red, and orange curves (uppermost to lowermost within each plot) correspond to $\beta\mu = 0$, 5, and 10, respectively. Note that for any $m > 1$ and $\beta\mu \neq 0$, $\widehat{I}_m[\alpha]$ exhibits a local minimum in α. The spectrum of L is also nonnegative on all of these solutions, so these minima in α correspond to stable wormholes.

which it follows that precisely one solution will exist for any allowed value of α or μ. Figure D.5 shows $\widehat{I}_m[\alpha]$ for various values of $\beta\mu$ and m. Note that for any nonzero $\beta\mu$ there is a local minimum, meaning that the modulus is stabilized. Moreover, we have verified that the spectrum of L is nonnegative for all of these solutions, so we conclude that turning on the EOW branes stabilizes the wormholes when $m > 1$ (though the wormholes do not appear to dominate over the disk: it is clear from Figure D.5 that the on-shell action $\widehat{I}_m[\alpha_{\min}]$ evaluated at the saddle $\alpha_{\min}$ grows with m). On the other hand, the behavior of $\widehat{I}_m[\alpha]$ when $m < 1$ is quite different. It is clear from (D.70) that $\widehat{I}_m[\alpha]$ is single-valued and independent of α whenever $1/m$ is an integer, just as in pure JT. For intermediate values of $1/m$, there are instead always two branches of $\widehat{I}_m[\alpha]$, as shown in Figure D.6 for $\beta\mu = 10$; see also Figure D.7 (the behavior for other nonzero values of $\beta\mu$ is analogous). These branches can exhibit either stable, unstable, or no saddles in α. However, any stable saddles for α coincide with unstable saddles for the wiggle; conversely, the branches of solutions on which the wiggle is stable (which only exist for $m \geq 1/2$) only exhibit unstable saddles for the modulus. Hence this classical JT+brane model does not admit any stable wormholes for $m < 1$.

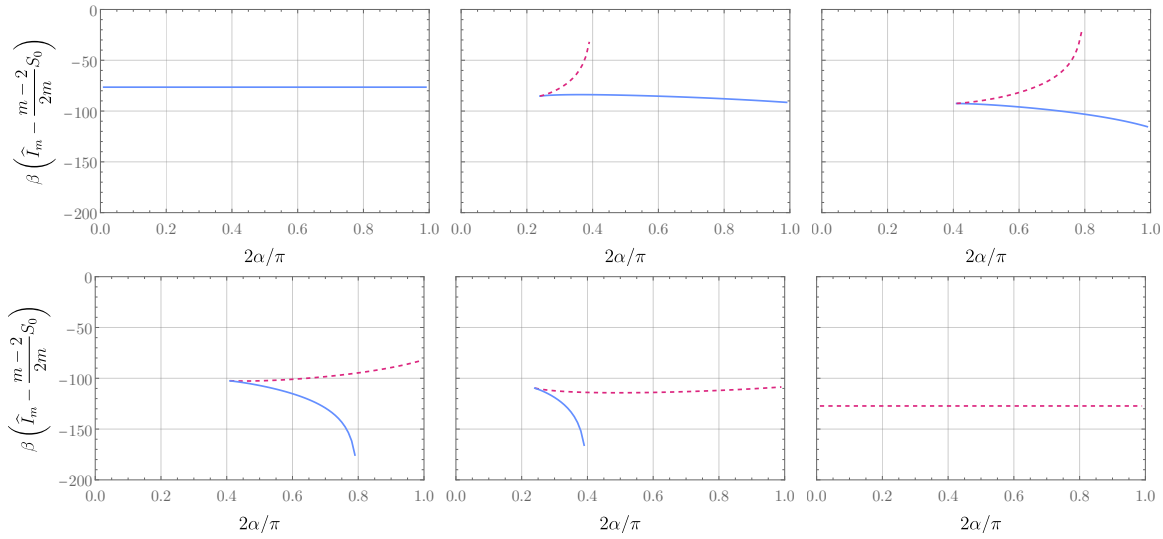


Figure D.6: The effective action $\widehat{I}_m[\alpha]$ for various values of $m \leq 1$ with $\beta\mu = 10$. From left to right and top to bottom, we show $1/m = 1$, 1.2, 1.4, 1.6, 1.8, and 2. Except for when $1/m$ is an integer, there are two branches: on the ones drawn as solid blue curves, the spectrum of L is nonnegative, while on the dashed red curves L has a negative eigenvalue. Note that although there are some saddles for α, these saddles are all unstable, either to perturbations of α (as on the solid blue branch in the second plot) or of the wiggle (as in the dashed red branch in the fifth plot). The behavior for larger $1/m$ is analogous, except that all of the solutions for the wiggle are unstable.

Just as in pure JT, the branches in Figures D.6 and D.7 that appear to simply end are indicative of additional sheets of the analytic continuation of $\widehat{I}_m[\alpha]$ to complex m (and complex α, if desired). This analytic continuation is obtained by inverting (D.70a) for b_r ignoring the constraint (D.70b); the inverse function $b_r(\sin(\pi/m)/\sin\alpha)$ is a meromorphic function of m whose Riemann surface contains infinitely many sheets, and the action (D.71) is infinitely-sheeted as well. Hence we can attribute the ends of the "branches to nowhere" in Figures D.6 and D.7 as stemming from the constraint (D.70b) that fixes the allowed branches of this Riemann surface. The upshot is that as in pure JT, imposing the equations of motions is crucial to constraining the allowed behavior of the analytic continuation of $\widehat{I}_m$: merely continuing a portion of the action (from, say, $m > 1$) to all m without invoking the equations of motion is insufficient to uniquely fix the allowed

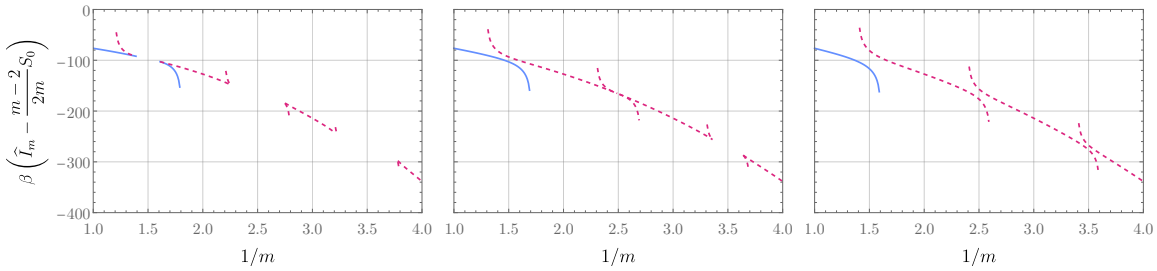


Figure D.7: The effective action $\widehat{I}_m[\alpha]$ as a function of m for $\beta\mu = 10$ and various values of α; from left to right we show $\alpha = \pi/5$, $3\pi/10$, and $2\pi/5$. On the solid blue branch(es) the spectrum of L is nonnegative; on the dashed red branches L has a negative eigenvalue.

behavior of the action at $m < 1$.

(It is also worth noting that the endpoints of the "branches to nowhere" in Figures D.6 and D.7 correspond to singular solutions for the wiggle, which can be seen as follows. As the graphical analysis in Figure D.4 indicates, as α and/or m are varied a single branch of solutions can end when the allowed interval for b_r (D.70b) changes due to a change in the sign of $\tan((\pi - m\alpha)/2m)$. On this branch, $b_r/2\pi$ is not an integer as this transition is reached, so from (D.70c) and (D.71) we conclude that a_r either vanishes or diverges there while the action remains finite. Hence these endpoints correspond to singular configurations of the wiggle with finite action.)

Bibliography

[1] V. Chandrasekaran, N. Engelhardt, S. Fischetti and S. Hernández-Cuenca, *A Tale of Two Saddles*, 2207.09472.

[2] S. Hernández-Cuenca, V.E. Hubeny and M. Rota, *The holographic entropy cone from marginal independence*, 2204.00075.

[3] M. Fadel and S. Hernández-Cuenca, *Symmetrized holographic entropy cone*, *Phys. Rev. D* **105** (2022) 086008 [2112.03862].

[4] N. Bao, N. Cheng, S. Hernández-Cuenca and V.P. Su, *Topological Link Models of Multipartite Entanglement*, *Quantum* **6** (2022) 741 [2109.01150].

[5] C. Akers, S. Hernández-Cuenca and P. Rath, *Quantum Extremal Surfaces and the Holographic Entropy Cone*, *JHEP* **11** (2021) 177 [2108.07280].

[6] Å. Folkestad and S. Hernández-Cuenca, *Conformal rigidity from focusing*, *Class. Quant. Grav.* **38** (2021) 215005 [2106.09037].

[7] S. Hernández-Cuenca, G.T. Horowitz, G. Treviño and D. Wang, *Boundary Causality Violating Metrics in Holography*, *Phys. Rev. Lett.* **127** (2021) 8 [2103.05014].

[8] D. Avis and S. Hernández-Cuenca, *On the foundations and extremal structure of the holographic entropy cone*, 2102.07535.

[9] N. Bao, N. Cheng, S. Hernández-Cuenca and V.P. Su, *A Gap Between the Hypergraph and Stabilizer Entropy Cones*, 2006.16292.

[10] S. Hernández-Cuenca and G.T. Horowitz, *Bulk reconstruction of metrics with a compact space asymptotically*, *JHEP* **08** (2020) 108 [2003.08409].

[11] N. Bao, N. Cheng, S. Hernández-Cuenca and V.P. Su, *The Quantum Entropy Cone of Hypergraphs*, *SciPost Phys.* **9** (2020) 067 [2002.05317].

[12] S. Hernández-Cuenca, V.E. Hubeny, M. Rangamani and M. Rota, *The quantum marginal independence problem*, 1912.01041.

[13] S. Hernández-Cuenca, *Holographic entropy cone for five regions*, *Phys. Rev. D* **100** (2019) 026004 [1903.09148].

[14] G. 't Hooft, *Dimensional reduction in quantum gravity*, *Conf. Proc. C* **930308** (1993) 284 [gr-qc/9310026].

[15] L. Susskind, *The World as a Hologram*, *J. Math. Phys.* **36** (1995) 6377 [hep-th/9409089].

[16] J.M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231 [hep-th/9711200].

[17] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [hep-th/9802150].

[18] O. Aharony, S.S. Gubser, J.M. Maldacena, H. Ooguri and Y. Oz, *Large N field theories, string theory and gravity*, *Phys. Rept.* **323** (2000) 183 [hep-th/9905111].

[19] E. D'Hoker and D.Z. Freedman, *Supersymmetric gauge theories and the AdS / CFT correspondence*, in *TASI 2001*, 1, 2002 [hep-th/0201253].

[20] G.T. Horowitz and J. Polchinski, *Gauge/gravity duality*, gr-qc/0602037.

[21] J. Polchinski, *Introduction to Gauge/Gravity Duality*, in *TASI 2010*, 10, 2010, DOI [1010.6134].

[22] V.E. Hubeny, *The AdS/CFT Correspondence*, *Class. Quant. Grav.* **32** (2015) 124010 [1501.00007].

[23] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl. Phys. B* **72** (1974) 461.

[24] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, 1903.11115.

[25] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199.

[26] J.M. Bardeen, B. Carter and S.W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161.

[27] R.M. Wald, *The thermodynamics of black holes*, *Living Rev. Rel.* **4** (2001) 6 [gr-qc/9912119].

[28] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333.

[29] J.D. Bekenstein, *Generalized second law of thermodynamics in black hole physics*, *Phys. Rev. D* **9** (1974) 3292.

[30] G.T. Horowitz and J. Polchinski, *A Correspondence principle for black holes and strings*, *Phys. Rev. D* **55** (1997) 6189 [hep-th/9612146].

[31] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys. Lett. B* **379** (1996) 99 [hep-th/9601029].

[32] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [hep-th/0603001].

[33] T. Nishioka, S. Ryu and T. Takayanagi, *Holographic Entanglement Entropy: An Overview*, *J. Phys. A* **42** (2009) 504008 [0905.0932].

[34] M. Rangamani and T. Takayanagi, *Holographic Entanglement Entropy*, vol. 931, Springer (2017), 10.1007/978-3-319-52573-0, [1609.01287].

[35] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090 [1304.4926].

[36] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [0705.0016].

[37] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [1307.2892].

[38] N. Engelhardt and A.C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [1408.3203].

[39] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [1005.3035].

[40] P. Hayden, M. Headrick and A. Maloney, *Holographic Mutual Information is Monogamous*, *Phys. Rev. D* **87** (2013) 046003 [1107.2940].

[41] V. Balasubramanian, P. Hayden, A. Maloney, D. Marolf and S.F. Ross, *Multiboundary Wormholes and Holographic Entanglement*, *Class. Quant. Grav.* **31** (2014) 185015 [1406.2663].

[42] N. Bao, S. Nezami, H. Ooguri, B. Stoica, J. Sully and M. Walter, *The Holographic Entropy Cone*, *JHEP* **09** (2015) 130 [1505.07839].

[43] M. Taylor, *Generalized entanglement entropy*, *JHEP* **07** (2016) 040 [1507.06410].

[44] P.A.R. Jones and M. Taylor, *Entanglement entropy in top-down models*, *JHEP* **08** (2016) 158 [1602.04825].

[45] N. Engelhardt and G.T. Horowitz, *Towards a Reconstruction of General Bulk Metrics*, *Class. Quant. Grav.* **34** (2017) 015004 [1605.01070].

[46] N. Engelhardt and G.T. Horowitz, *Recovering the spacetime metric from a holographic dual*, *Adv. Theor. Math. Phys.* **21** (2017) 1635 [1612.00391].

[47] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [1905.08762].

[48] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, *JHEP* **09** (2020) 002 [1905.08255].

[49] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013 [1911.12333].

[50] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, *JHEP* **03** (2022) 205 [1911.11977].

[51] N. Engelhardt, S. Fischetti and A. Maloney, *Free energy from replica wormholes*, *Phys. Rev. D* **103** (2021) 046021 [2007.07444].

[52] N. Pippenger, *What are the laws of information theory*, in *Special Problems on Communication and Computation Conference*, pp. 3–5, 1986.

[53] Z. Zhang and R.W. Yeung, *A non-shannon-type conditional inequality of information quantities*, *IEEE Trans. Inf. Theory* **43** (1997) 1982.

[54] Z. Zhang and R.W. Yeung, *On characterization of entropy function via information inequalities*, *IEEE Trans. Inf. Theory* **44** (1998) 1440.

[55] F. Matus, *Infinitely many information inequalities*, in *2007 IEEE International Symposium on Information Theory*, pp. 41–44, IEEE, 2007.

[56] N. Pippenger, *The inequalities of quantum information theory*, *IEEE Trans. Inf. Theory* **49** (2003) 773.

[57] N. Linden and A. Winter, *A new inequality for the von neumann entropy*, *Commun. Math. Phys.* **259** (2005) 129 [quant-ph/0406162].

[58] R. Dougherty, C. Freiling and K. Zeger, *Networks, matroids, and non-shannon information inequalities*, *IEEE Trans. Inf. Theory* **53** (2007) 1949.

[59] J. Cadney, N. Linden and A. Winter, *Infinitely many constrained inequalities for the von neumann entropy*, *IEEE Trans. Inf. Theory* **58** (2012) 3657.

[60] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2010), 10.1017/CBO9780511976667.

[61] B. Ibinson, *Quantum Information and Entropy*, Ph.D. thesis, University of Bristol, 2006.

[62] N. Linden, F. Matúš, M.B. Ruskai and A. Winter, *The Quantum Entropy Cone of Stabiliser States*, *LIPIcs* **22** (2013) 270 [1302.5453].

[63] D. Gross and M. Walter, *Stabilizer information inequalities from phase space distributions*, *Journal of Mathematical Physics* **54** (2013) 082201 [1302.6990].

[64] M. Walter, *Multipartite Quantum States and their Marginals*, Ph.D. thesis, Zurich, ETH, 2014. 1410.6820. 10.3929/ethz-a-010250985.

[65] B. Czech and X. Dong, *Holographic Entropy Cone with Time Dependence in Two Dimensions*, *JHEP* **10** (2019) 177 [1905.03787].

[66] T. He, V.E. Hubeny and M. Rangamani, *Superbalance of Holographic Entropy Inequalities*, *JHEP* **07** (2020) 245 [2002.04558].

[67] V.E. Hubeny, M. Rangamani and M. Rota, *Holographic entropy relations*, *Fortsch. Phys.* **66** (2018) 1800067 [1808.07871].

[68] V.E. Hubeny, M. Rangamani and M. Rota, *The holographic entropy arrangement*, *Fortsch. Phys.* **67** (2019) 1900011 [1812.08133].

[69] M. Walter and F. Witteveen, *Hypergraph min-cuts from quantum entropies*, 2002.12397.

[70] S. Fujishige, *Submodular Functions and Optimization*, vol. 58 of *Ann. Discrete Math.*, Elsevier (2005), 10.1016/S0167-5060(13)71057-4.

[71] W.H. Cunningham, *On submodular function minimization*, *Comb.* **5** (1985) 185.

[72] S. Nezami and M. Walter, *Multipartite Entanglement in Stabilizer Tensor Networks*, *Phys. Rev. Lett.* **125** (2020) 241602 [1608.02595].

[73] D. Kleitman and G. Markowsky, *On Dedekind's Problem: The Number of Isotone Boolean Functions. II*, *Trans. Am. Math. Soc.* **213** (1975) 373.

[74] A.E. Brouwer, C. Mills, W. Mills and A. Verbeek, *Counting families of mutually intersecting sets*, *Electron. J. Comb.* **20** (2013) .

[75] S.S. Gubser, I.R. Klebanov and A.M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys. Lett.* **B428** (1998) 105 [hep-th/9802109].

[76] A. Almheiri, X. Dong and D. Harlow, *Bulk Locality and Quantum Error Correction in AdS/CFT*, *JHEP* **04** (2015) 163 [1411.7041].

[77] D.L. Jafferis, A. Lewkowycz, J. Maldacena and S.J. Suh, *Relative entropy equals bulk relative entropy*, *JHEP* **06** (2016) 004 [1512.06431].

[78] X. Dong, D. Harlow and A.C. Wall, *Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality*, *Phys. Rev. Lett.* **117** (2016) 021601 [1601.05416].

[79] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle and M. Walter, *Entanglement Wedge Reconstruction via Universal Recovery Channels*, *Phys. Rev. X* **9** (2019) 031011 [1704.05839].

[80] M. Headrick, V.E. Hubeny, A. Lawrence and M. Rangamani, *Causality & holographic entanglement entropy*, *JHEP* **12** (2014) 162 [1408.6300].

[81] V.E. Hubeny, H. Liu and M. Rangamani, *Bulk-cone singularities & signatures of horizon formation in AdS/CFT*, *JHEP* **01** (2007) 009 [hep-th/0610041].

[82] J. Hammersley, *Extracting the bulk metric from boundary information in asymptotically AdS spacetimes*, *JHEP* **12** (2006) 047 [hep-th/0609202].

[83] J. Maldacena, D. Simmons-Duffin and A. Zhiboedov, *Looking for a bulk point*, *JHEP* **01** (2017) 013 [1509.03612].

[84] V.E. Hubeny, *Extremal surfaces as bulk probes in AdS/CFT*, *JHEP* **07** (2012) 093 [1203.1044].

[85] S. Bilson, *Extracting Spacetimes using the AdS/CFT Conjecture: Part II*, *JHEP* **02** (2011) 050 [1012.1812].

[86] N. Bao, C. Cao, S. Fischetti and C. Keeler, *Towards Bulk Metric Reconstruction from Extremal Area Variations*, *Class. Quant. Grav.* **36** (2019) 185002 [1904.04834].

[87] T. Takayanagi and K. Umemoto, *Entanglement of purification through holographic duality*, *Nature Phys.* **14** (2018) 573 [1708.09393].

[88] S. Dutta and T. Faulkner, *A canonical purification for the entanglement wedge cross-section*, *JHEP* **03** (2021) 178 [1905.00577].

[89] P. Hayden, R. Jozsa, D. Petz and A. Winter, *Structure of states which satisfy strong subadditivity of quantum entropy with equality*, *Commun. Math. Phys.* **246** (2004) 359–374 [quant-ph/0304007].

[90] H. Casini, E. Teste and G. Torroba, *Modular Hamiltonians on the null plane and the Markov property of the vacuum state*, *J. Phys. A* **50** (2017) 364001 [1703.10656].

[91] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter and Z. Yang, *Holographic duality from random tensor networks*, *JHEP* **11** (2016) 009 [1601.01694].

[92] J. Sorce, *Holographic entanglement entropy is cutoff-covariant*, *JHEP* **10** (2019) 015 [1908.02297].

[93] E. Witten, *APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory*, *Rev. Mod. Phys.* **90** (2018) 045003 [1803.04993].

[94] D. Marolf, M. Rota and J. Wien, *Handlebody phases and the polyhedrality of the holographic entropy cone*, *JHEP* **10** (2017) 069 [1705.10736].

[95] A.C. Wall, *Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy*, *Class. Quant. Grav.* **31** (2014) 225007 [1211.3494].

[96] M. Rota and S.J. Weinberg, *New constraints for holographic entropy from maximin: A no-go theorem*, *Phys. Rev. D* **97** (2018) 086013 [1712.10004].

[97] N. Bao and M. Mezei, *On the Entropy Cone for Large Regions at Late Times*, 1811.00019.

[98] X. Dong, A. Lewkowycz and M. Rangamani, *Deriving covariant holographic entanglement*, *JHEP* **11** (2016) 028 [1607.07506].

[99] T. Faulkner and A. Lewkowycz, *Bulk locality from modular flow*, *JHEP* **07** (2017) 151 [1704.05464].

[100] P. Hayden and G. Penington, *Learning the Alpha-bits of Black Holes*, *JHEP* **12** (2019) 007 [1807.06041].

[101] B. Czech, L. Lamprou, S. McCandlish and J. Sully, *Integral Geometry and Holography*, *JHEP* **10** (2015) 175 [1505.05515].

[102] B. Czech, L. Lamprou, S. McCandlish, B. Mosk and J. Sully, *A Stereoscopic Look into the Bulk*, *JHEP* **07** (2016) 129 [1604.03110].

[103] N. Bao, C. Cao, S. Fischetti, J. Pollack and Y. Zhong, *More of the Bulk from Extremal Area Variations*, *Class. Quant. Grav.* **38** (2021) 047001 [2009.07850].

[104] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, *Holography from Conformal Field Theory*, *JHEP* **10** (2009) 079 [0907.0151].

[105] D. Harlow, *The Ryu–Takayanagi Formula from Quantum Error Correction*, *Commun. Math. Phys.* **354** (2017) 865 [1607.03901].

[106] T. He, M. Headrick and V.E. Hubeny, *Holographic Entropy Relations Repackaged*, *JHEP* **10** (2019) 118 [1905.06985].

[107] M. Freedman and M. Headrick, *Bit threads and holographic entanglement*, *Commun. Math. Phys.* **352** (2017) 407 [1604.00354].

[108] S.X. Cui, P. Hayden, T. He, M. Headrick, B. Stoica and M. Walter, *Bit Threads and Holographic Monogamy*, *Commun. Math. Phys.* **376** (2019) 609 [1808.05234].

[109] C. Akers and P. Rath, *Entanglement Wedge Cross Sections Require Tripartite Entanglement*, *JHEP* **04** (2020) 208 [1911.07852].

[110] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [1908.10996].

[111] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *The entropy of Hawking radiation*, *Rev. Mod. Phys.* **93** (2021) 035002 [2006.06872].

[112] S. Raju, *Lessons from the information paradox*, *Phys. Rept.* **943** (2022) 1 [2012.05770].

[113] C. Akers, N. Engelhardt, G. Penington and M. Usatyuk, *Quantum Maximin Surfaces*, *JHEP* **08** (2020) 140 [1912.02799].

[114] A. Almheiri, R. Mahajan and J.E. Santos, *Entanglement islands in higher dimensions*, *SciPost Phys.* **9** (2020) 001 [1911.09666].

[115] R. Emparan, *Black hole entropy as entanglement entropy: A Holographic derivation*, *JHEP* **06** (2006) 012 [hep-th/0603081].

[116] R.C. Myers, R. Pourhasan and M. Smolkin, *On Spacetime Entanglement*, *JHEP* **06** (2013) 013 [1304.2030].

[117] M. Rozali, J. Sully, M. Van Raamsdonk, C. Waddell and D. Wakeham, *Information radiation in BCFT models of black holes*, *JHEP* **05** (2020) 004 [1910.12836].

[118] H.Z. Chen, R.C. Myers, D. Neuenfeld, I.A. Reyes and J. Sandor, *Quantum Extremal Islands Made Easy, Part I: Entanglement on the Brane*, *JHEP* **10** (2020) 166 [2006.04851].

[119] H.Z. Chen, R.C. Myers, D. Neuenfeld, I.A. Reyes and J. Sandor, *Quantum Extremal Islands Made Easy, Part II: Black Holes on the Brane*, *JHEP* **12** (2020) 025 [2010.00018].

[120] J. Hernandez, R.C. Myers and S.-M. Ruan, *Quantum extremal islands made easy. Part III. Complexity on the brane*, *JHEP* **02** (2021) 173 [2010.16398].

[121] V. Chandrasekaran, M. Miyaji and P. Rath, *Including contributions from entanglement islands to the reflected entropy*, *Phys. Rev. D* **102** (2020) 086009 [2006.10754].

[122] R. Bousso and E. Wildenhain, *Gravity/ensemble duality*, *Phys. Rev. D* **102** (2020) 066005 [2006.16289].

[123] D. Neuenfeld, *Homology conditions for RT surfaces in double holography*, *Class. Quant. Grav.* **39** (2022) 075009 [2105.01130].

[124] H. Geng and A. Karch, *Massive islands*, *JHEP* **09** (2020) 121 [2006.02438].

[125] T. Takayanagi, *Holographic Dual of BCFT*, *Phys. Rev. Lett.* **107** (2011) 101602 [1105.5165].

[126] M. Fujita, T. Takayanagi and E. Tonni, *Aspects of AdS/BCFT*, *JHEP* **11** (2011) 043 [1108.5152].

[127] M. Headrick and T. Takayanagi, *A Holographic proof of the strong subadditivity of entanglement entropy*, *Phys. Rev. D* **76** (2007) 106013 [0704.3719].

[128] F. Sanches and S.J. Weinberg, *Holographic entanglement entropy conjecture for general spacetimes*, *Phys. Rev. D* **94** (2016) 084034 [1603.05250].

[129] Y. Nomura, P. Rath and N. Salzetta, *Pulling the Boundary into the Bulk*, *Phys. Rev. D* **98** (2018) 026010 [1805.00523].

[130] C.A. Agón and J.F. Pedraza, *Quantum bit threads and holographic entanglement*, *JHEP* **02** (2022) 180 [2105.08063].

[131] C. Akers and G. Penington, *Leading order corrections to the quantum extremal surface prescription*, *JHEP* **04** (2021) 062 [2008.03319].

[132] A.W. Ingleton, *Representation of matroids*, in *Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969)*, pp. 149–167, Academic Press, London, 1971.

[133] V. Balasubramanian, B. Czech, B.D. Chowdhury and J. de Boer, *The entropy of a hole in spacetime*, *JHEP* **10** (2013) 220 [1305.0856].

[134] V. Balasubramanian, B.D. Chowdhury, B. Czech, J. de Boer and M.P. Heller, *Bulk curves from boundary data in holography*, *Phys. Rev.* **D89** (2014) 086004 [1310.4204].

[135] R.C. Myers, J. Rao and S. Sugishita, *Holographic Holes in Higher Dimensions*, *JHEP* **06** (2014) 044 [1403.3416].

[136] B. Czech, X. Dong and J. Sully, *Holographic Reconstruction of General Bulk Surfaces*, *JHEP* **11** (2014) 015 [1406.4889].

[137] M. Headrick, R.C. Myers and J. Wien, *Holographic Holes and Differential Entropy*, *JHEP* **10** (2014) 149 [1408.4770].

[138] B. Czech and L. Lamprou, *Holographic definition of points and distances*, *Phys. Rev.* **D90** (2014) 106005 [1409.4473].

[139] C. Cao, S.M. Carroll and S. Michalakis, *Space from Hilbert Space: Recovering Geometry from Bulk Entanglement*, *Phys. Rev.* **D95** (2017) 024031 [1606.08444].

[140] N. Bao, A. Chatwin-Davies, B.E. Niehoff and M. Usatyuk, *Bulk Reconstruction Beyond the Entanglement Wedge*, *Phys. Rev. D* **101** (2020) 066011 [1911.00519].

[141] J.M. Maldacena, *Wilson loops in large N field theories*, *Phys. Rev. Lett.* **80** (1998) 4859 [hep-th/9803002].

[142] E.T. Newman, *Heaven and Its Properties*, *Gen. Rel. Grav.* **7** (1976) 107.

[143] R.O. Hansen, E.T. Newman, R. Penrose and K.P. Tod, *The Metric and Curvature Properties of H Space*, *Proc. Roy. Soc. Lond.* **A363** (1978) 445.

[144] C.N. Kozameh and E.T. Newman, *Theory of light cone cuts of null infinity*, *J. Math. Phys.* **24** (1983) 2481.

[145] G. Trevino, *Reconstruction of an AdS Radiation/Boson Star Bulk Geometry Using Light-cone Cuts*, *JHEP* **11** (2017) 111 [1708.03331].

[146] N. Engelhardt, *Into the Bulk: A Covariant Approach*, *Phys. Rev.* **D95** (2017) 066005 [1610.08516].

[147] M. Gary, S.B. Giddings and J. Penedones, *Local bulk S-matrix elements and CFT singularities*, *Phys. Rev. D* **80** (2009) 085005 [0903.4437].

[148] M. Dodelson and H. Ooguri, *High-energy behavior of Mellin amplitudes*, *Phys. Rev. D* **101** (2020) 066008 [1911.05274].

[149] M. Taylor, *Holography for degenerate boundaries*, hep-th/0001177.

[150] C. Bär, N. Ginoux and F. Pfäffle, *Wave Equations on Lorentzian Manifolds and Quantization*, ESI Lectures in Mathematics and Physics, European Mathematical Society (2007).

[151] H. Dorn, M. Salizzoni and C. Sieg, *On the propagator of a scalar field on AdS x S and on the BMN plane wave*, *JHEP* **02** (2005) 047 [hep-th/0307229].

[152] C. Fefferman and C.R. Graham, *Conformal invariants*, in *Élie Cartan et les mathématiques d'aujourd'hui*, no. S131 in Astérisque, pp. 95–116, Société mathématique de France (1985).

[153] L.F. Alday and E. Perlmutter, *Growing Extra Dimensions in AdS/CFT*, *JHEP* **08** (2019) 084 [1906.01477].

[154] A. May, G. Penington and J. Sorce, *Holographic scattering requires a connected entanglement wedge*, *JHEP* **08** (2020) 132 [1912.05649].

[155] A. Mollabashi, N. Shiba and T. Takayanagi, *Entanglement between Two Interacting CFTs and Generalized Holographic Entanglement Entropy*, *JHEP* **04** (2014) 185 [1403.1393].

[156] A. Karch and C.F. Uhlemann, *Holographic entanglement entropy and the internal space*, *Phys. Rev. D* **91** (2015) 086005 [1501.00003].

[157] S. Gao and R.M. Wald, *Theorems on gravitational time delay and related issues*, *Class. Quant. Grav.* **17** (2000) 4999 [gr-qc/0007021].

[158] C. Akers, J. Koeller, S. Leichenauer and A. Levine, *Geometric Constraints from Subregion Duality Beyond the Classical Regime*, 1610.08968.

[159] N. Engelhardt and S. Fischetti, *The Gravity Dual of Boundary Causality*, *Class. Quant. Grav.* **33** (2016) 175004 [1604.03944].

[160] D. Harlow and D. Stanford, *Operator Dictionaries and Wave Functions in AdS/CFT and dS/CFT*, 1104.2621.

[161] L. Susskind and E. Witten, *The Holographic bound in anti-de Sitter space*, hep-th/9805114.

[162] T. Banks, M.R. Douglas, G.T. Horowitz and E.J. Martinec, *AdS dynamics from conformal field theory*, hep-th/9808016.

[163] J.J. Halliwell and J.B. Hartle, *Integration Contours for the No Boundary Wave Function of the Universe*, *Phys. Rev. D* **41** (1990) 1815.

[164] D. Berenstein and D. Grabovsky, *The Tortoise and the Hare: A Causality Puzzle in AdS/CFT*, *Class. Quant. Grav.* **38** (2021) 105008 [2011.08934].

[165] S.R. Coleman, *Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence*, *Nucl. Phys. B* **307** (1988) 867.

[166] S.B. Giddings and A. Strominger, *Loss of Incoherence and Determination of Coupling Constants in Quantum Gravity*, *Nucl. Phys. B* **307** (1988) 854.

[167] J.M. Maldacena and L. Maoz, *Wormholes in AdS*, *JHEP* **02** (2004) 053 [hep-th/0401024].

[168] N. Arkani-Hamed, J. Orgera and J. Polchinski, *Euclidean wormholes in string theory*, *JHEP* **12** (2007) 018 [0705.2768].

[169] D. Harlow and D. Jafferis, *The Factorization Problem in Jackiw-Teitelboim Gravity*, *JHEP* **02** (2020) 177 [1804.01081].

[170] D. Stanford and E. Witten, *JT gravity and the ensembles of random matrix theory*, *Adv. Theor. Math. Phys.* **24** (2020) 1475 [1907.03363].

[171] L.V. Iliesiu, *On 2D gauge theories in Jackiw-Teitelboim gravity*, 1909.05253.

[172] D. Kapec, R. Mahajan and D. Stanford, *Matrix ensembles with global symmetries and 't Hooft anomalies from 2d gauge theory*, *JHEP* **04** (2020) 186 [1912.12285].

[173] P. Saad, *Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity*, 1910.10311.

[174] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, *JHEP* **08** (2020) 044 [2002.08950].

[175] D. Marolf and H. Maxfield, *Observations of Hawking radiation: the Page curve and baby universes*, *JHEP* **04** (2021) 272 [2010.06602].

[176] S.B. Giddings and G.J. Turiaci, *Wormhole calculus, replicas, and entropies*, *JHEP* **09** (2020) 194 [2004.02900].

[177] I. Aref'eva and I. Volovich, *Gas of baby universes in JT gravity and matrix models*, *Symmetry* **12** (2020) 975 [1905.08207].

[178] E. Witten, *Matrix Models and Deformations of JT Gravity*, *Proc. Roy. Soc. Lond. A* **476** (2020) 20200582 [2006.13414].

[179] N. Afkhami-Jeddi, H. Cohn, T. Hartman and A. Tajdini, *Free partition functions and an averaged holographic duality*, *JHEP* **01** (2021) 130 [2006.04839].

[180] A. Belin, J. De Boer, P. Nayak and J. Sonner, *Charged eigenstate thermalization, Euclidean wormholes and global symmetries in quantum gravity*, *SciPost Phys.* **12** (2022) 059 [2012.07875].

[181] M. Van Raamsdonk, *Comments on wormholes, ensembles, and cosmology*, *JHEP* **12** (2021) 156 [2008.02259].

[182] A. Blommaert, *Dissecting the ensemble in JT gravity*, 2006.13971.

[183] A. Maloney and E. Witten, *Averaging over Narain moduli space*, *JHEP* **10** (2020) 187 [2006.04855].

[184] J. Cotler and K. Jensen, *AdS_3 gravity and random CFT*, *JHEP* **04** (2021) 033 [2006.08648].

[185] A. Pérez and R. Troncoso, *Gravitational dual of averaged free CFT's over the Narain lattice*, *JHEP* **11** (2020) 015 [2006.08216].

[186] O. Janssen, M. Mirbabayi and P. Zograf, *Gravity as an ensemble and the moment problem*, *JHEP* **06** (2021) 184 [2103.12078].

[187] G.J. Turiaci, M. Usatyuk and W.W. Weng, *2D dilaton-gravity, deformations of the minimal string, and matrix models*, *Class. Quant. Grav.* **38** (2021) 204001 [2011.06038].

[188] H. Maxfield and G.J. Turiaci, *The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral*, *JHEP* **01** (2021) 118 [2006.11317].

[189] N. Benjamin, C.A. Keller, H. Ooguri and I.G. Zadeh, *Narain to Narnia*, *Commun. Math. Phys.* **390** (2022) 425 [2103.15826].

[190] C. Peng, J. Tian and Y. Yang, *Half-Wormholes and Ensemble Averages*, 2205.01288.

[191] C.V. Johnson, *Nonperturbative Jackiw-Teitelboim gravity*, *Phys. Rev. D* **101** (2020) 106023 [1912.03637].

[192] C.V. Johnson, *Jackiw-Teitelboim supergravity, minimal strings, and matrix models*, *Phys. Rev. D* **103** (2021) 046012 [2005.01893].

[193] C.V. Johnson, *Explorations of nonperturbative Jackiw-Teitelboim gravity and supergravity*, *Phys. Rev. D* **103** (2021) 046013 [2006.10959].

[194] C.V. Johnson, *Low Energy Thermodynamics of JT Gravity and Supergravity*, 2008.13120.

[195] C.V. Johnson, *On the Quenched Free Energy of JT Gravity and Supergravity*, 2104.02733.

[196] C.V. Johnson, *Quantum Gravity Microstates from Fredholm Determinants*, *Phys. Rev. Lett.* **127** (2021) 181602 [2106.09048].

[197] C.V. Johnson, *Consistency Conditions for Non-Perturbative Completions of JT Gravity*, 2112.00766.

[198] C.V. Johnson, *The Microstate Physics of JT Gravity and Supergravity*, 2201.11942.

[199] C.V. Johnson, *The Distribution of Ground State Energies in JT Gravity*, 2206.00692.

[200] A. Belin and J. de Boer, *Random statistics of OPE coefficients and Euclidean wormholes*, *Class. Quant. Grav.* **38** (2021) 164001 [2006.05499].

[201] J. Pollack, M. Rozali, J. Sully and D. Wakeham, *Eigenstate Thermalization and Disorder Averaging in Gravity*, *Phys. Rev. Lett.* **125** (2020) 021601 [2002.02971].

[202] L. Eberhardt, *Summing over Geometries in String Theory*, *JHEP* **05** (2021) 233 [2102.12355].

[203] P. Saad, S.H. Shenker, D. Stanford and S. Yao, *Wormholes without averaging*, 2103.16754.

[204] A.M. García-García and V. Godet, *Half-wormholes in nearly AdS_2 holography*, *SciPost Phys.* **12** (2022) 135 [2107.07720].

[205] A. Blommaert, L.V. Iliesiu and J. Kruthoff, *Gravity factorized*, 2111.07863.

[206] X. Dong, *The Gravity Dual of Renyi Entropy*, *Nature Commun.* **7** (2016) 12472 [1601.06788].

[207] A.M. García-García and V. Godet, *Euclidean wormhole in the Sachdev-Ye-Kitaev model*, *Phys. Rev. D* **103** (2021) 046014 [2010.11633].

[208] U. Moitra, S.K. Sake and S.P. Trivedi, *Jackiw-Teitelboim gravity in the second order formalism*, *JHEP* **10** (2021) 204 [2101.00596].

[209] U. Moitra, S.K. Sake and S.P. Trivedi, *Aspects of Jackiw-Teitelboim gravity in Anti-de Sitter and de Sitter spacetime*, *JHEP* **06** (2022) 138 [2202.03130].

[210] M. Mezard, G. Parisi and M. Virasoro, *Spin Glass Theory and Beyond*, World Scientific (1986), 10.1142/0271.

[211] J. Maldacena, D. Stanford and Z. Yang, *Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space*, *PTEP* **2016** (2016) 12C104 [1606.01857].

[212] L.N. Trefethen, *Spectral Methods in MATLAB*, SIAM (2000), 10.1137/1.9780898719598.

[213] A.M. García-García, V. Godet, C. Yin and J.P. Zheng, *Euclidean-to-Lorentzian wormhole transition and gravitational symmetry breaking in the Sachdev-Ye-Kitaev model*, 2204.08558.

[214] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *J.Phys.* **A42** (2009) 504005 [0905.4013].

[215] J. Sully, M.V. Raamsdonk and D. Wakeham, *BCFT entanglement entropy at large central charge and the black hole interior*, *JHEP* **03** (2021) 167 [2004.13088].

[216] U. Moitra, S.K. Sake, S.P. Trivedi and V. Vishal, *Jackiw-Teitelboim Model Coupled to Conformal Matter in the Semi-Classical Limit*, *JHEP* **04** (2020) 199 [1908.08523].

[217] G.W. Gibbons and S.W. Hawking, *Cosmological Event Horizons, Thermodynamics, and Particle Creation*, *Phys. Rev. D* **15** (1977) 2738.

[218] D. Marolf, *Gravitational thermodynamics without the conformal factor problem: Partition functions and Euclidean saddles from Lorentzian Path Integrals*, 2203.07421.

[219] J.M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021 [hep-th/0106112].

[220] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (2, 2011), 10.1017/CBO9780511524646.

[221] H. Rademacher, *Über partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale*, *Mathematische Annalen* **79** (1919) 340.

[222] Z. Wen and J. Avery, *Some properties of hyperspherical harmonics*, *J. Math. Phys.* **26** (1985) 396.

[223] O.J.C. Dias, J.E. Santos and B. Way, *Numerical Methods for Finding Stationary Gravitational Solutions*, *Class. Quant. Grav.* **33** (2016) 133001 [1510.02804].

[224] J. Maldacena, D. Stanford and Z. Yang, *Diving into traversable wormholes*, *Fortsch. Phys.* **65** (2017) 1700034 [1704.05333].

[225] I. Kourkoulou and J. Maldacena, *Pure states in the SYK model and nearly-AdS_2 gravity*, 1707.02325.