The Unexpected Virtue of Ignorance:

Euclidean Wormholes and the Emergence of Spacetime



Tom Philip Yildirim

Christ Church

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2026

To see a World in a Grain of Sand
And a Heaven in a Wild Flower,
Hold Infinity in the palm of your hand
And Eternity in an hour.

— William Blake, *Auguries of Innocence*

Acknowledgements

This thesis exists because of an email that, in another version of events, would have gone unread. I had spent the first years of my PhD adrift from the questions that had drawn me to physics in the first place, and I had written to many people hoping to find a way back. Most never replied. Vijay Balasubramanian did. He had every reason not to: the sender was an unknown graduate student from a different university entirely, working in experimental physics, with no background in this field at all. He read the email, and he replied. He has been my supervisor, my mentor, and my friend ever since. Vijay's enthusiasm, intellect, and breadth of knowledge are matched only by his kindness, and to puzzle over the secrets of the universe with him has been the most intellectually fulfilling experience of my life. I hope we will go on doing so for many years to come.

That email reached him because two others first made it possible to send. Peter Davies, whose funding made this work possible in the most literal sense; without him, there would be no thesis at all. And Alan Barr, who took me on as his student and gave me the freedom to follow my curiosity wherever it led, no matter how far from the conventional course. Both of them trusted a stranger; what follows is theirs as much as it is mine.

I owe further thanks to my collaborators — Jan de Boer, Ben Craps, Marc Klinger, Juan Hernandez, Maria Knysh, Pratik Nandy, Gabriele Pascuzzi, Hanzhi Jiang, and especially James Unwin — and to Cambridge, the École Normale Supérieure, Barcelona, VUB Brussels, Utrecht, Amsterdam, Caltech, UC Santa Barbara, UC Berkeley, and the Institute for Advanced Study for hosting me for seminars and conversations throughout my DPhil.

Long before any of that, my parents fostered the curiosity that brought me here. They patiently entertained the unending stream of *... Maar waarom?* that, I have come to suspect, a physicist never quite outgrows — one only learns to ask it of stranger things. Their support has been without hesitation and without limit: the long drives to Oxford for an afternoon, the midnight pickups from the airport, the quiet certainty that someone would always come. To my siblings — Sophie, Fleur,

Floris, and little Selene — and to Quetzal, Maya, Pablo, Willem, Mickey, Nero, and Juan: thank you, simply, for being my family.

To David and Tabitha: your kindness and generosity, both during my DPhil and in the years before it, are more than I knew to expect of anyone, and humble me still.

To my friends — Matthew, Andrew, Sophie, Sean, Felix, John R, Jan, Bogdan, Tavish, Aloysius, Chris, Ben, Victoria, AC, John C, Dom, Radek, Tom D, Lucas, and the many members of Christ Church — thank you for making my time at Oxford so memorable.

And most of all, Netta. Thank you for your patience through the long hours I spent staring blankly into the distance, lost in some mathematical confusion or another; for your grace when I had to finish a paper on your birthday; for every hour you spent on a train back and forth to Oxford, and the countless other ineffable ways in which you have supported me. The last few years have been hectic, difficult and, at times, dark — and yet, because of you, they have also been the happiest I have known.

Lastly, to Oxford, and to everyone within it: thank you for the history, the camaraderie, the beauty, and for being the city where I found that which eluded me in the places I once called Home.

Whatever I have achieved, I owe to the kindness of others — more of them, by far, than these pages can hold. To all of them, named and unnamed, I dedicate this thesis.

I will end on this. A PhD is the apprenticeship of a careful mind, and most of what one is taught, rightly, is how to be careful. But the paths worth walking are rarely the ones laid out in advance, and the questions worth a life are rarely the careful ones. To follow such a path is to set yourself against two kinds of resistance: the resistance of convention, which tells you the path is not yours to take, and the resistance of circumstance, which conspires, sometimes cruelly, to keep you from walking it. Neither is easily overcome. Both must be. To anyone standing where I once stood — unsure whether the path they suspect is theirs is worth the cost of taking it — I will only say that the cost of going is real, but the cost of not going is annihilation. Go. And when the going turns hard, as it will, rage. I will leave you with the words of Dylan Thomas, which I first heard from Professor Brand in *Interstellar* — the film that, long before I knew what AdS/CFT was, sparked in me a fascination with general relativity that would result in this thesis: *Rage, rage against the dying of the light.*

Abstract

This thesis investigates what non-perturbative wormhole contributions to the Einstein-Hilbert Euclidean gravitational path integral reveal about the underlying microscopic theory of quantum gravity. Treating the saddlepoint expansion as a self-contained framework, with no holographic input, we develop a calculus of fine-grained equalities derived from coarse-grained gravitational amplitudes, and apply it to a connected sequence of questions about the non-perturbative gravity Hilbert space, its emergent geometry, and its operational content.

The principal technical tools are a *squaring trick* — which lifts a coarse-grained equality $\overline{A} = \overline{B}$ to the fine-grained statement $A = B$ via the auxiliary path integral $\overline{(A-B)^2} = 0$ — together with a basis of *shell states* defined by inserting heavy spherically symmetric dust operators into the path integral on one- and two-sided asymptotic boundaries. A Schwinger–Dyson resummation of the shell-state resolvent, controlled by a contour pole that appears once the basis is sufficiently overcomplete, establishes that these states form a semiclassically complete basis of the two-sided gravity Hilbert space $\mathcal{H}_{LR}$, with microcanonical dimension $\dim \mathcal{H}_{LR,E} = e^{S(E_L)+S(E_R)}$. A direct corollary is that bulk geometry cannot be a linear observable on the non-perturbative inner product space: microcanonical black-hole states admit expansions in shell states whose dominant saddles contain no horizon at all.

A class of mixed wormhole saddles connecting boundaries of one- and two-sided type then enforces the tensor-product factorisation of the two-boundary gravitational Hilbert space, $\mathcal{H}_{LR} = \mathcal{H}_{B_L} \otimes \mathcal{H}_{B_R}$, at the non-perturbative level. The same wormholes responsible for the apparent breakdown of amplitude factorisation are precisely those that protect Hilbert-space factorisation. Factorisation in turn yields a fine-grained gravitational proof of the partition-function-as-trace identity $\mathrm{Tr}_{\mathcal{H}_X}(e^{-\beta H}) = Z(\beta)$ — a non-trivial conjecture in gravity rather than the definition it is in quantum field theory — from which the universality of the Bekenstein–Hawking area law follows as a literal microstate count, and a purely gravitational version of the holographic replica trick for computing entropies is established with no reference to a boundary CFT. All of these results apply equally in asymptotically flat and asymptotically AdS gravity.

Drawing on these structural features of the non-perturbative Hilbert space, we then study the operational accessibility of the resulting microstates. Asymptotic shell-operator probes exhibit a sharp verification structure characteristic of the complexity class QMA: a matched probe activates detection saddles whose contribution distinguishes a specific microstate, generic probes return universal responses, and state identification without prior knowledge is exponentially hard within the saddlepoint description. For two-sided shell states, single-boundary probes can detect microstates whose interior worldvolume lies outside their entanglement wedge, accessing bulk information that no linear single-boundary operator can reconstruct. We argue on this basis that the black-hole horizon is a semiclassical manifestation of the complexity of the underlying quantum microstate, not of any genuine inaccessibility of information.

These threads are synthesised into a physical interpretation of the ensemble nature of the gravitational path integral. We identify the gravitational ensemble with the equivalence class of microstates compatible with the simple-operator data accessible to an asymptotic observer in a chaotic theory. Computations in the resulting effective state are then governed by a restricted sum over topology — a so-called *observer rule* — for which the underlying *principle of maximal ignorance* provides a clear epistemic origin: the observer retains operator contractions but suppresses state cross-contractions, which on the gravity side translates into a microscopically motivated restriction on which wormhole topologies contribute. The probe-induced shift δE in the dominant saddle energy partitions the analysis into three regimes, $\delta E = 0$, $\delta E \sim O(G_N)$, and $\delta E \sim O(1)$, that map precisely onto the nested causal $\subset$ simple $\subset$ entanglement wedge hierarchy and reproduce the Chandra–Hartman generalised entropy of the apparent horizon. The KSGNS construction on an operator system $\mathcal{S}$ and the coarse-grained state $\overline{\rho}$ defines an enlarged effective Hilbert space, with the enlargement traced to the shrinking of a null ideal of fine-tuned phase cancellations. Applied to AdS closed universes — where the fine-grained two-boundary bilinear is rank one, a direct consequence of the factorisation and trace identity established earlier — the same construction nevertheless yields an effective Hilbert space of rank as large as d_{AdS}^2. Closed-universe physics is in this sense inherently emergent from coarse-graining and admits no fine-grained analogue: semiclassical geometry, observer, and effective Hilbert space arise together as features of the same coarse-grained description — the unexpected virtue of ignorance.

Statement of Originality

I declare that the work presented in this thesis is my own, except where explicit reference is made to the work of others. No part of this thesis has been submitted for any other degree or qualification at this or any other institution. The research described here was carried out in the Particle Physics sub-department of the University of Oxford between October 2022 and May 2026, under the joint supervision of Professor Alan Barr (University of Oxford) and Professor Vijay Balasubramanian (University of Pennsylvania).

The body of this thesis is based on four collaborative research projects. While each of these works is formally co-authored, I was the primary contributor at every stage. I originated each of the four projects; performed the technical analysis, including all of the calculations, derivations and figures; constructed the principal results; and wrote the initial drafts of the manuscripts. My co-authors contributed through periodic technical discussions and editorial feedback on the drafts. Professor Balasubramanian served as my external co-supervisor as well as a co-author on each of the works listed below, and his contributions fell under the same description.

Chapter 1.
This chapter collects introductory and background material written specifically for this thesis. Where individual paragraphs have been adapted from the works listed below, this is indicated at the start of the relevant section.

Chapter 2.
This chapter is based on *A Nonperturbative Toolkit for Quantum Gravity*, written with Vijay Balasubramanian [1] and published in *Physical Review D*.

Chapter 3.
This chapter is based on *The Nonperturbative Hilbert Space of Quantum Gravity With One Boundary*, written with Vijay Balasubramanian [2] and published in the *Journal of High Energy Physics*.

Chapter 4.
This chapter is based on *How to Count States in Gravity*, written with Vijay Balasubramanian [3] and published in *Physical Review D*.

Chapter 5.
This chapter is based *Observing Spacetime*, written with Vijay Balasubramanian [4] and published in *Physical Review D*.

Chapter 6.
This chapter is based on work in progress with Vijay Balasubramanian and Jan de Boer, provisionally entitled *The Unexpected Virtue of Ignorance: On the Epistemic Origin of Observer Rules and Beyond* — to appear.

The figures appearing in Chapters 1–6 are adapted from the corresponding works listed above.

Two further pieces of work were completed during the period covered by this thesis but are not included, as they fall outside its scope. In chronological order, these are:

HLT MET Regression Using Convolutional Neural Networks.
ATLAS internal note (CERN, 2023), prepared during my time with the ATLAS collaboration under supervision of Holly Pacey.

A QCD R-Axion.
Written with James Unwin [5].

Tom Yildirim

Oxford, May 2026

Contents

1 Introduction **1**

1.1 The gravitational path integral . . . 3

1.1.1 Wormholes and coarse-graining . . . 5

1.1.2 Action of the Hamiltonian . . . 6

1.1.3 The two-sided Hilbert space . . . 8

1.2 Shell states . . . 9

1.2.1 Two-sided shell states . . . 9

1.2.1.1 Overlaps and orthogonality . . . 11

1.2.1.2 The norm and the junction conditions . . . 13

1.2.1.3 The gravitational action of the norm . . . 14

1.2.1.4 The large shell mass limit . . . 15

1.2.1.5 Shell geometry . . . 17

1.2.1.6 Multi-shell states . . . 18

1.2.1.7 Shell wormholes . . . 19

1.2.2 Single-boundary shell states . . . 21

1.2.2.1 Norm, wormholes, and the large mass limit . . . 22

1.2.2.2 Saddle geometry . . . 24

2 A Nonperturbative Toolkit for Quantum Gravity **27**

2.1 From coarse-grained to fine-grained equalities . . . 30

2.2 Traces in an overcomplete basis . . . 31

2.3 Shell states resolve the identity . . . 32

2.3.1 The folded wormhole . . . 33

2.3.2 Resummation in the planar limit . . . 35

2.3.3 The microcanonical equality and the Hilbert-space dimension . 37
2.3.4 Beyond the planar limit: the canonical equality 38
2.4 A semiclassically complete basis 41
2.4.1 Examples 46
2.5 Discussion 48
2.5.1 Macroscopic superpositions of semiclassical geometries 48
2.5.2 Other possible bases and the spectral form factor 50

3 Resolving the Factorisation Problem **53**
3.1 Single-boundary shell states span $\mathcal{H}_{\mathcal{B}}$ 54
3.2 Factorisation of the two-sided Hilbert space 57
3.2.1 Tensor-product states lie in $\mathcal{H}_{2s}$ 59
3.2.2 Two-sided shell states lie in $\mathcal{H}_L \otimes \mathcal{H}_R$ 61
3.3 Disentangling shell states and factorised geometries 64
3.4 Discussion 68

4 On the Universality of Black Hole Entropy **71**
4.1 The Gibbons-Hawking puzzle 73
4.2 The single-boundary partition function 74
4.2.1 Argument from factorisation 74
4.2.2 Explicit calculation in the shell basis 75
4.2.3 The area law counts black-hole microstates 77
4.3 Entanglement entropy and the gravitational replica trick 79
4.3.1 The trace glues density matrices 81
4.3.2 Explicit calculation in the shell basis 81
4.4 Discussion 82

5 Decoding the Black Hole Interior **87**
5.1 Observing a single-boundary universe 88
5.1.1 Universal saddles 89
5.1.2 Detection saddles 91
5.1.3 Detection-to-universal ratio 93
5.2 Observing a two-boundary universe 96
5.2.1 Coordinated two-boundary probe 96
5.2.2 Single-boundary probe of a two-boundary state 99
5.3 Verification, superpositions, and the search problem 103
5.4 Discussion 105

5.4.1 Detection vs entanglement-wedge reconstruction 105
5.4.2 Horizons and baby universes geometrize complexity 105

6 On the Unexpected Virtue of Ignorance 107
6.1 Epistemic coarse-graining 111
6.1.1 Simple observables and dephasing 111
6.1.2 Complex operators and state ensembles 112
6.1.3 Capturing an observer's ignorance 114
6.1.4 Predictions 116
6.2 Coarse-graining the black hole interior 117
6.2.1 Heavy shells and the gravitational state ensemble 117
6.2.2 Topological restriction and the simple wedge 120
6.2.3 The simple wedge observer and observer rule A1 122
6.2.3.1 Operational requirements and probe backreaction . . 123
6.2.3.2 The causal wedge observer and observer rule B1 . . . 126
6.2.4 Two-boundary gravity states 128
6.3 The observer's effective Hilbert space 131
6.3.1 The effective Hilbert space from operator overlaps 131
6.3.2 Setup and path integral methods 135
6.3.3 Global case 136
6.3.4 The ignorant Gram matrix 138
6.3.5 Two-boundary states 141
6.3.5.1 Two-boundary probes 141
6.3.5.2 Single-boundary probes 144
6.4 Maximal ignorance ensembles 146
6.4.1 The maximal ignorance framework 148
6.4.2 Ensemble of pure states 150
6.4.2.1 Type A: functionally complete simple measurements 150
6.4.2.2 Type B: Hamiltonian and low-backreaction probes . 152
6.4.3 Ensemble of mixed states 153
6.4.3.1 Type A: functionally complete simple measurements 154
6.4.3.2 Type B: thermal measurements 156
6.4.3.3 Effective dimension from the mixed-state ensemble . 156
6.4.4 Comparison of observer rules 157
6.5 A non-trivial closed universe emerges through coarse-graining 158
6.5.1 The coarse-grained bilinear form 159

6.5.2 The closed-universe Gram matrix and observer rule A1 160
6.5.3 Triviality of the fine-grained Hilbert space and its enlargement 161
6.5.4 Predictability in the closed universe 162
6.6 Discussion 164
6.6.1 Future directions 165

A Gram matrix of overlaps **169**

B Ensemble statistics and random matrix calculations **173**
B.1 Universality of ensemble statistics 173
B.1.1 Random-phase ensemble 173
B.1.2 Haar ensemble on the microcanonical shell 174
B.1.3 Ginibre ensemble 175
B.1.4 Maximum-entropy ensemble 176
B.1.5 Summary 177
B.2 Generalised Wishart ensembles 177
B.2.1 Rank-one construction: scale-mixture Wishart 177
B.2.2 Trace construction: matrix-valued Wishart mixture 179
B.2.3 Bi-ensemble construction: two-sided Gram–Wishart 180
B.3 Ensemble normalisation and the Gaussian approximation 183
B.3.1 Contour-integral derivation of the partition function 183
B.3.2 Exact finite-D result and moments 183
B.3.3 Saddle-point evaluation 184
B.3.4 Conditions for the Gaussian approximation 184
B.3.5 Two illustrative examples 185
B.4 Thermal saddle-point analysis 186
B.4.1 Continuum limit and saddle-point evaluation 186
B.4.2 Evaluation of the mean probability 186
B.4.3 Asymptotic behaviour 187
B.4.4 Comparison with the Gibbs state 187

Bibliography **189**

List of Figures

1.1 Slicing a gravity path integral boundary condition along $\mathcal{X}$ to define a state (adapted from [1]). 4

1.2 Action of $e^{-\beta H}$ on states. (**a**) Action of the Hamiltonian; from a different perspective this is preparing the state $|2\beta\rangle$. (**b**) Example state in $\mathcal{H}_{\mathcal{X}}$. (**c**) Action of the Hamiltonian on this state. (**d**) Survival amplitude after Euclidean time evolution, alternatively the overlap $\langle 2\beta||s^*\rangle_L \otimes |s\rangle_R$ (adapted from [1]). 6

1.3 Two distinct ways of preparing states in $\mathcal{H}_{\mathcal{X}_L \cup \mathcal{X}_R}$. (**a**) Construction of an example state in $\mathcal{H}_{LR}$. (**b**) Construction of an example state in $\mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$. 8

1.4 Asymptotic boundary condition for the gravity path integral defining the shell state. (**a**) Cut-open Euclidean boundary with topology $\mathbb{I}_{\frac{\beta_L+\beta_R}{2}} \times \mathbb{S}^{d-1}$ for preparation of the shell states. The shell operator $\mathcal{O}_i$ is pictured in red. In AdS/CFT we can also perform the path integral in the boundary CFT with insertion of a $\mathbb{S}^{d-1}$ symmetric operator dual to the shell. (**b**) Euclidean boundary with the $\mathbb{S}^{d-1}$ suppressed. We adopt this convention for the rest of the paper. Here $\beta_{L,R}/2$ are Euclidean "preparation times". 10

1.5 Shell asymptotic boundary condition for $\overline{\langle j|i\rangle}$, consisting of the operator insertions $\mathcal{O}_i$ and $\mathcal{O}_j^\dagger$ separated by asymptotic time extent β_L and β_R respectively. The red lines represent the shells propagating into the bulk. 12

1.6 The saddlepoints for the shell norm path integral $\overline{\langle i|i\rangle}$ are constructed by gluing together two disks. In particular the shell homology regions (purple) on each disk are discarded and the resulting geometries glued together along the shell worldvolume. Note the propagation temperatures $\Delta T_{L,R}$ are denoted as $\Delta\tau_{\mp}$ in the main text. 12

1.7 Diagrams showing the analytic continuation of the type 1-3 shell states to Lorentzian signature. (**a**) Type 1 shell state with $\beta_L < \beta_{HP}$ and $\beta_R < \beta_{HP}$ corresponding to a two-sided black hole. (**b**) Type 2 shell state with $\beta_L > \beta_{HP}$ and $\beta_R > \beta_{HP}$ consisting of two copies of thermal AdS with the addition of a compact Big-Crunch AdS cosmology. (**c**) Type 3 shell state with $\beta_L > \beta_{HP}$ and $\beta_R < \beta_{HP}$ corresponding to a disconnected geometry with an R single sided BH and an L thermal AdS factor. 18

1.8 Construction of shell wormhole saddles Z_2. We glue two disks together by discarding the regions indicated in purple and attaching the remainder along the shell worldvolumes (red, black dashes indicate the glued shells) while obeying the Israel junction conditions. Saddlepoints, i.e., solutions to the equation of motion, exist with this topology [6]. . . . 19

1.9 Fully connected wormhole geometry for $\overline{(G^n)_{ii}}$ consisting of n shell boundaries connected by a single bulk, pictured here for $n = 4$. 20

1.10 Schematic of the pinwheel wormhole construction. (**a**) Sheet diagrams obtained by cutting Fig. 1.9 along the shell worldvolumes. (**b**) Each sheet can be regarded as a portion of a disk; the pinwheel saddle is constructed by gluing two disks together along the corresponding shell worldvolumes via the junction conditions. 21

1.11 Cut-open path integral boundary condition defining the single-sided shell states. (**a**) Euclidean boundary with topology $\mathbb{R}^{<0} \times \mathbb{S}^{d-1}$. The shell operator $\mathcal{O}_i$ is pictured in red and $\beta/2$ is the Euclidean preparation time. (**b**) Euclidean boundary with the $\mathbb{S}^{d-1}$ suppressed; we sometimes depict this boundary with a curve or kink to clarify diagrams. 22

1.12 Shell-strip asymptotic boundary condition for the overlap $\langle j|i\rangle$ consisting of the line $\lim_{\alpha\to\infty}[-\alpha, \beta+\alpha]$ on which $\mathcal{O}_i$ and $\mathcal{O}_j^\dagger$ are inserted at $\tau = 0$ and $\tau = \beta$ respectively. 22

1.13 The saddle for the norm $\overline{\langle i|i\rangle}$ is constructed by gluing a disk and a strip together along the shell worldvolume, after discarding the shell homology regions (purple). The junction conditions dynamically determine the propagation times $\Delta T_{S,D}$. 23
1.14 The fully connected single-sided shell wormhole saddles for $\overline{(G^n)_{ii}}$ are constructed by gluing n strips into a disk along the shell worldvolumes, depicted here for $n = 4$. 25
1.15 Lorentzian continuation of the single-sided shell states. (**a**) Type A: a single-sided black hole with a shell behind the horizon. (**b**) Type B: thermal AdS with a disconnected compact Big-Crunch cosmology. . 25

2.1 The *tadpole* boundary corresponding to $\langle\beta|i\rangle$: a periodic asymptotic boundary of length $\frac{1}{2}(\beta+\beta_L+\beta_R)$ with a single shell operator insertion $\mathcal{O}_i$ (red). 33
2.2 (**a**) Folded wormhole geometry $\tilde{Z}_{n+2}$, depicted for $n = 2$: the $n+1$ shell-overlap boundaries and two tadpole boundaries are joined through a single bulk component. (**b**) Construction by self-folding: a Z-partition-function disk with $2(n + 1)$ shell insertions has its shell homology regions (pink) removed, and the remaining geometry is glued back along corresponding worldvolume pairs. 34
2.3 Graphical representation of the two tadpole boundaries $\langle\beta|i\rangle\langle j|\beta\rangle$ in (2.3.6). 36
2.4 Schematic of the Schwinger-Dyson resummation (2.3.7). The planar sum over partially-connected saddles is reorganised by which shell-overlap boundaries are connected to the two tadpole shells. Insertions of R correspond to the trace of the shell-state resolvent. 36
2.5 Operator boundary for $\langle i|\Psi\rangle$, with the $\mathcal{O}_\Psi$ insertion (whose stress-energy is depicted as the dotted yellow line) and $\mathcal{O}_i^\dagger$ separated by Euclidean times $(\beta_{L,R} + \tilde{\beta}_{L,R})/2$ on the two sides. 43
2.6 Construction of the $\hat{Z}_{n+2}[\Psi]$ wormhole saddles for $n = 2$. Starting from a disk saddle for the path integral with boundary Σ_n (left), the shell homology regions (purple) are discarded to produce a sheet (centre); the sheet is then glued back to itself along the corresponding shell worldvolumes to form the wormhole (right). 43

2.7 (**a**) The connected contribution to $\overline{G_{ij}^n \langle\Psi|i\rangle\langle j|\Psi\rangle \langle\Psi|\Psi\rangle}$ for $n = 1$. (**b**) Construction by gluing along the shell worldvolumes. (**c**) The $n \to -1$ limit recovers the connected wormhole contribution to $\overline{\langle\Psi|\Psi\rangle^2}$ at the original boundary length. Red denotes shell worldvolumes; chartreuse denotes $\mathcal{O}_\Psi$ matter sources. 46

3.1 Resolving the factorisation puzzle. (**a**) The two-sided Hilbert space $\mathcal{H}_{LR}$, the tensor product Hilbert space $\mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$, and the total two-sided Hilbert space $\mathcal{H}_{L\cup R}$. The shaded regions $\mathcal{H}_{2s}$ and $\mathcal{H}_L \otimes \mathcal{H}_R$ denote the spans of two-sided and tensor-product single-sided shell bases respectively. (**b**) The two- and single-sided shell states span $\mathcal{H}_{LR}$ (Chapter 2) and $\mathcal{H}_{\mathcal{B}_{L,R}}$ (Sec. 3.1) respectively in the limit of infinitely many basis states. (**c**) Sec. 3.2 shows that the two-sided shell basis coincides with the tensor-product basis, establishing $\mathcal{H}_{LR} = \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$. 55

3.2 The fully connected wormhole saddle entering the single-boundary spanning argument, depicted for $n = 1$. It is constructed by gluing $n + 1$ shell strips into a central strip along the shell worldvolumes, replacing the folded wormhole of Sec. 2.3.1. See [2] for details. 56

3.3 Saddles for $\overline{\langle i|i\rangle_L \langle j|j\rangle_R}$. (**a**) Disconnected contribution. (**b**) Connected wormhole, present only when $i = j$. (**c**) Construction of the wormhole by gluing L, R shell strips into a disk. 58

3.4 Boundary condition for the mixed overlap $\langle i|_L\langle j|_R|\mathbf{p}\rangle_{2s}$ between a tensor product of single-sided shell states and a two-sided shell state. . . 58

3.5 (**a**) Universal wormhole contribution to $\overline{|\langle i|_L\langle j|_R|\mathbf{p}\rangle_{2s}|^2}$. (**b**) Construction by gluing L- and R-strips into the corresponding disks along the single-sided shell worldvolumes, then gluing the disks together along the two-sided shell worldvolume. 59

3.6 (**a**) Fully-connected saddle contributing to (3.2.5), depicted for $n = 2$. (**b**) Construction by gluing the L- and R-disks along $n + 1$ two-sided shell worldvolumes and gluing an L- and R-strip into the respective disks along the single-sided worldvolumes. 60

3.7 (**a**) Connected saddle contributing to the cross term in (3.2.8), depicted for $n = 2$. (**b**) Construction: as in Fig. 3.6b, with an additional L- and R-strip glued in along an extra pair of single-sided shell insertions separated by $\beta_{L,R}$ from the original. 62

3.8 (**a**) Fully-connected saddle contributing to (3.2.12), depicted for $n = 2$, $m = 3$. (**b**) Construction by gluing an L- and an R-disk along the **p**-shell worldvolume, and gluing $n+1$ ($m+1$) single-sided L-strips (R-strips) into the respective disks along the single-sided shell worldvolumes. 63
3.9 (**a**) Connected saddle contributing to the cross term in the fine-grained extension of (3.2.14). (**b**) Construction: as in Fig. 3.8b, with an additional pair of **p**-shell insertions on each disk separated by $\tilde{\beta}_{L,R}$ from the original. 64
3.10 Saddle for the two-shell norm $\overline{\langle \mathbf{i},\mathbf{j}|\mathbf{i},\mathbf{j}\rangle_{2s}}$: the L- and R-disks are glued into a central disk by identifying the **i**- and **j**-shell worldvolumes. . . . 65
3.11 Lorentzian geometries of two-shell states with $\tilde{\beta}_L, \tilde{\beta}_R < \beta_{HP}$. (**a**) For $\beta_m < \beta_{HP}$, a single two-sided AdS black hole with a long interior supporting both shell trajectories. (**b**) For $\beta_m > \beta_{HP}$, two disconnected AdS black holes, with the two shells propagating in disconnected "Big Crunch" interior universes. 66
3.12 Saddle for the direct overlap (3.3.3): two disks glued together along the **i**- and **j**-shell worldvolumes via a strip of length $\beta_m/2$. 67

4.1 (**a**) The thermal partition function in gravity computed by summing path integrals with boundaries defined by a basis of states. The coloured crosses denote possible matter insertions specifying the basis element. (**b**) The Gibbons-Hawking path integral $\overline{Z(\beta)}$, defined on a periodic Euclidean boundary of length β with no insertions. 73
4.2 (**a**) Boundary condition for the path integral computing (4.2.6), depicted for $n = 3$. The n shell-overlap boundaries from G^n are joined to the $\langle j|e^{-\beta H}|i\rangle$ boundary, and the shell index loop closes around the chain. (**b**) Construction of the maximally-connected saddle by gluing $n+1$ shell strips into a disk along the corresponding shell worldvolumes. 76
4.3 The QFT replica prescription. (**a**) The boundary condition for $|\Psi\rangle$, viewable as defining a map $M_\Psi : \mathcal{H}_{\mathcal{X}_L} \to \mathcal{H}_{\mathcal{X}_R}$. (**b**) The reduced density matrix ρ_R is obtained by sewing the $|\Psi\rangle$ ket to the bra $\langle\Psi|$ along $\mathcal{X}_L$, equivalently the operator $M_\Psi M_\Psi^\dagger$. (**c**) $\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n)$ is obtained by sewing n copies of ρ_R together along $\mathcal{X}_R$ into a periodic boundary. 80
4.4 The path integral $\overline{Z_\circ(\rho_\Psi^n)(\tilde{\beta}_L,\tilde{\beta}_R)}$ rewritten as an overlap $\overline{\langle\Sigma|\Phi\rangle}$ by cutting it along $\mathcal{X}_L$ and $\mathcal{X}_R$. 82

4.5 (**a**) Boundary condition for $\rho_R = \mathrm{Tr}_{\mathcal{H}_{\mathfrak{X}_L}}|\Psi\rangle\langle\Psi|$. (**b**) Boundary condition for $\mathrm{Tr}_{\mathcal{H}_{\mathfrak{X}_R}}(\rho_R^n)$ in the shell basis, depicted for $n = 2$. 84
4.6 (**a**) Maximally-connected saddle for (4.3.2). (**b**) Construction by gluing shell strips into a central disk along the shell worldvolumes (one L and one R shell shown for clarity). 85

5.1 Asymptotic boundary condition for $\overline{\langle i|\mathcal{O}_P|i\rangle\langle i|\mathcal{O}_P^\dagger|i\rangle}$. 89
5.2 The six classes of wormhole saddle contributing to $\overline{\langle i|\mathcal{O}_P|i\rangle\langle i|\mathcal{O}_P^\dagger|i\rangle}$. Top row: the two universal saddles $U_{L,R}$ that contribute for any $\mathcal{O}_P$. Lower two rows: the four detection saddles $D_{L,R}^{\uparrow,\downarrow}$ that contribute only when $\mathcal{O}_P = \mathcal{O}_i$. 90
5.3 Construction of the universal saddles. (**a**) U_L by gluing two single-boundary norm saddles along the probe shell. (**b**) U_R by folding a four-shell disk along the probe worldvolumes and attaching shell strips at the $\mathcal{O}_i$ worldvolumes. 91
5.4 Construction of the detection saddles. (**a**) $D_L^\uparrow$. (**b**) $D_R^\uparrow$. (**c**) $D_R^\downarrow$ (and analogously $D_L^\downarrow$). 95
5.5 Asymptotic boundary condition for the coordinated two-boundary probe correlator (5.2.1). 96
5.6 Saddle contributions to (5.2.1). (**a**) The universal saddle, of disk-with-handle topology. (**b**) The detection saddle, present only when $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$. 97
5.7 Asymptotic boundary condition for the single-boundary probe (5.2.12) of a two-boundary state. 99
5.8 The two universal saddles for the single-boundary probe of a two-boundary state. (**a**) U_1: all three shells cross between boundaries through a cylinder. (**b**) U_2: only the probe shell crosses; the $\mathcal{O}_\mathbf{i}$ shells annihilate within each boundary. 100
5.9 The first three classes of detection saddle for the single-boundary probe of a two-boundary state, $D_{1,2,3}$. 101
5.10 The fourth class of detection saddle, D_4, of twice-punctured-torus topology. 101

6.1 Observer rule B1, corresponding to the coarse-graining (6.2.16). The right-hand side is the thermofield-double at the effective temperature β_{eff}. 128
6.2 Observer rule A1, corresponding to the coarse-graining (6.2.6). 128

6.3 Two classes of replica-symmetric saddles contributing to (6.3.20), depicted for $n = 4$. (**a**) n subsidiary disks of length β_{obs}, each carrying one strip, glued into a central disk of length $n\beta_R$. (**b**) A single disk of length $n\beta_{\text{obs}}$ with n strips, glued to the central disk. 137

"... an infinite series of times, in a growing, dizzying net of divergent, convergent and parallel times."
— *Jorge Luis Borges,* The Garden of Forking Paths

Chapter 1

Introduction

One of the deepest open problems in theoretical physics is to reconcile quantum mechanics with gravity. General relativity, our best classical theory of gravitation, predicts that spacetime is a dynamical entity whose geometry is determined by its matter content. Quantum mechanics, meanwhile, requires that all dynamical variables be subject to the uncertainty principle and that physical processes be described by probability amplitudes. The tension between these two frameworks becomes acute in regimes where both are important — near spacetime singularities, in the very early universe, and at the horizons of black holes.

The most concrete evidence that quantum effects are essential to gravity comes from black hole thermodynamics. The laws of black hole mechanics, established in the early 1970s [7], bear a striking formal resemblance to the laws of thermodynamics. This analogy was elevated to physical reality by Bekenstein's proposal [8] that black holes carry an entropy proportional to their horizon area, and by Hawking's subsequent discovery [9] that quantum fields in a black hole background radiate thermally at a temperature proportional to the surface gravity. The resulting Bekenstein-Hawking entropy,

$$S_{BH} = \frac{A}{4G_N} \,, \tag{1.0.1}$$

is a profoundly non-perturbative quantity: it is proportional to $1/G_N$ and thus invisible in any finite order of perturbation theory around a fixed background. It suggests that a black hole of area A has $e^{S_{BH}}$ microstates, but identifying these microstates within a framework of quantum gravity remains a central challenge.

A powerful approach to these questions is the Euclidean gravitational path integral, pioneered by Gibbons and Hawking [10]. The idea is to evaluate a formal sum over Riemannian (positive-definite) metrics with periodic boundary conditions in (asymptotic) time, weighted by the exponential of the Euclidean gravitational action, in close analogy with the partition function of statistical mechanics. This programme immediately bore fruit: evaluating the path integral in the saddlepoint approximation on the Euclidean Schwarzschild solution reproduces the Bekenstein-Hawking entropy (1.0.1), and for asymptotically anti-de Sitter (AdS) spacetimes, the Hawking-Page phase transition [11] between thermal AdS and the AdS black hole emerges as a change in dominance between saddlepoints of the Euclidean path integral.

The Euclidean path integral is not without its difficulties. The gravitational action is unbounded below due to the conformal factor problem — conformal rescalings of the metric can make the action arbitrarily negative [12]. Moreover, a complete definition of the measure over metrics and topologies remains elusive, and it is not known whether the Euclidean approach captures all non-perturbative effects of a complete theory of quantum gravity. Despite these issues, the saddlepoint approximation to the Euclidean path integral has proven remarkably effective as a tool for extracting non-perturbative gravitational physics. A striking modern example is the role of *replica wormholes* — connected saddlepoints in the path integral computing Rényi entropies — which resolve the tension between the perturbative description of Hawking radiation and the unitarity of the underlying theory, recovering the expected Page curve for black hole evaporation [13, 14]. For these reasons, we will take the saddlepoint approximation to the gravitational path integral at face value throughout this thesis, and ask what can be learned from it.

More broadly, the role of Euclidean wormholes — connected geometries with multiple disconnected boundaries — has been a recurring theme since the earliest days of quantum gravity. In the late 1980s it was recognised that such configurations could lead to violations of cluster decomposition and apparent losses of quantum coherence [15–17]. A modern resolution of these puzzles, building on the exact solvability of two-dimensional Jackiw-Teitelboim (JT) gravity as a matrix integral [18], is that the gravitational path integral computes not the amplitudes of a single quantum theory, but rather a coarse-grained average over an ensemble of theories — or more

precisely, an average over the fine-grained data of an underlying microscopic theory. This interpretation extends to higher dimensions, where wormhole contributions have been used to argue for a microscopic origin of black hole entropy [6, 19].

This thesis investigates what non-perturbative wormhole effects reveal about the underlying microscopic theory of gravity. Chapter 2 establishes a convenient basis of states for the gravity Hilbert space that allow for particularly simple wormhole amplitudes at the saddlepoint level. In Chapter 3, we demonstrate that these effects account for the factorisation of the non-perturbative gravitational Hilbert space. Building on this result, Chapter 4 establishes the universality of the black hole entropy formula. We then draw on these structural features of the gravitational Hilbert space in Chapter 5 to argue that the black hole horizon is a semiclassical manifestation of the complexity of the underlying quantum state. Finally, Chapter 6 synthesises these findings into a physical interpretation of the ensemble nature of the gravitational path integral, identifying the additional coarse-graining imposed by limited gravitational observers as a restriction on the sum over topologies. This framework resolves the apparent tension between wormhole amplitudes and the triviality of non-perturbative gravity in closed universes: physics in the closed universe is inherently emergent from coarse-graining, and admits no fine-grained analogue.

We do so by analyzing a particular set of states defined by cutting open the Euclidean gravitational path integral. In the remainder of this introduction we develop the necessary framework. We begin in Sec. 1.1 by reviewing how states are constructed from the gravitational path integral and how wormholes lead to the coarse-graining interpretation. We then introduce, in Sec. 1.2, the *shell states* — an explicit family of states constructed by inserting heavy dust shell operators into the path integral — which serve as the concrete basis for the computations in the subsequent chapters.

1.1 The gravitational path integral

The gravitational path integral is a map $\zeta[\Sigma_b] \to \mathbb{C}$ from a set of boundary conditions for spacetime – a boundary topology, metric, and matter sources, collectively denoted by Σ_b – to the complex numbers. We are going to analyze the Euclidean gravitational path integral, and focus for concreteness on asymptotically locally anti-de Sitter (AlAdS) gravity, although essentially identical analyses follow in the asymptotically flat case. To calculate the path integral we must sum over all topologies and geometries (g), modulo diffeomorphisms, and matter fields (ϕ), that are compatible with the

asymptotic boundary conditions:

$$\zeta[\Sigma_b] = \int_{g,\phi\to\Sigma_b} \mathcal{D}g\,\mathcal{D}\phi\, e^{-I_{bulk}[g,\phi]}$$
$$I_{bulk} = -\frac{1}{16\pi G_N}\int_{\mathcal{M}}\sqrt{g}(R-2\Lambda+\mathcal{L}_{matter}) - \frac{1}{8\pi G_N}\int_{\partial\mathcal{M}=\mathcal{M}_b}\sqrt{h}K \tag{1.1.1}$$

where $\mathcal{M}$ is a bulk manifold filling in the boundary. We have used the Einstein-Hilbert action above, but could have instead used any self-consistent effective theory in any dimension, such as two-dimensional Jackiw-Teitelboim (JT) gravity. We understand this path integral as an effective description below some cutoff, completed at high energies by some ultraviolet finite theory. We can construct states of the theory by partitioning the boundary $\mathcal{M}_b = \mathcal{M}_1 \cup \mathcal{M}_2$ so that $\partial\mathcal{M}_1 = \partial\mathcal{M}_2 = \mathcal{X}$ (see Fig. 1.1). The path integral then produces a functional of the fields and geometry on the cut $\mathcal{X}$, and formally defines states $|\mathcal{M}_{1,2}\rangle \in \mathcal{H}_\mathcal{X}$ in a Hilbert space $\mathcal{H}_\mathcal{X}$ such that $\langle\mathcal{M}_2|\mathcal{M}_1\rangle = \zeta[\Sigma_b]$. In general, the path integral cannot be performed exactly, and so we resort to (perturbative expansions around) saddlepoint approximations to evaluate overlaps of states constructed in this way. Note that defining a state by partitioning boundary data is a standard way of producing states in the *bulk* quantum gravitational Hilbert space. For instance, such methods were used in [20] and [21] to construct states in JT gravity and a topological toy model respectively, and in this thesis we will apply this method in higher dimensional Einstein-Hilbert gravity. Crucially, we do not assume any kind of dual boundary CFT construction. It is clear our states live in the bulk Hilbert space as we can, in principle, compute the wavefunctionals $\Psi_{\mathcal{M}_1}(h,\phi) \equiv \langle h,\phi|\mathcal{M}_1\rangle$ between our states and states defined by some bulk time-slice data $h_{\mu\nu},\phi$ by performing the path integral subject to these joint boundary conditions.

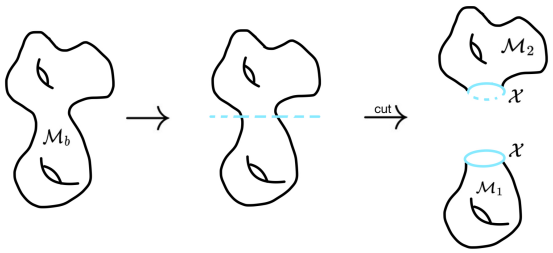


Figure 1.1: Slicing a gravity path integral boundary condition along $\mathcal{X}$ to define a state (adapted from [1]).

For a given cut $\mathcal{X}$ one option for producing different states is to fix the boundary manifold $\mathcal{M}_b$ and to create states $|i\rangle$ by inserting matter operators $\mathcal{O}_i$ on $\mathcal{M}_b$ (see,

e.g., [21–23]). The overlap of such states, $\langle i|j\rangle$, is computed by evaluating the path integral with $\mathcal{O}_{i,j}$ inserted on $\mathcal{M}_{1,2}$. The bra $\langle i|$ associated to the ket $|i\rangle$ defined by inserting $\mathcal{O}_i$ on $\mathcal{M}_1$ is obtained by inserting $\mathcal{O}_i^\dagger$ on $\mathcal{M}_1$. The ket $|i^*\rangle$ defined by inserting $\mathcal{O}_i^\dagger$ on $\mathcal{M}_1$ therefore defines the same boundary condition as $\langle i|$, and so $\langle i|j\rangle = \langle j^*|i^*\rangle$. Finally, the leading saddlepoint of the path integral for the norm of a state can be sliced open at a moment of time-reflection symmetry, provided this exists, and analytically continued to Lorentzian signature. This procedure associates a Lorentzian geometry to the state.

1.1.1 Wormholes and coarse-graining

To compute the squared overlap $\langle i|j\rangle\,\langle j|i\rangle$ we insert $\mathcal{O}_{i,j}^{(\dagger)}$ on two copies of the boundary manifold and sum over bulk geometries and topologies that fill in these boundaries. This path integral has two kinds of contributions: (a) disconnected topologies coming from products of terms contributing to $\langle i|j\rangle$, and (b) connected topologies, i.e., wormholes, that interpolate between the two boundaries. These non-perturbative contributions, which are akin to instantons in quantum field theory, can produce an apparently paradoxical result where

$$\langle i|j\rangle = \delta_{ij} \quad \text{while} \quad \langle i|j\rangle\,\langle j|i\rangle = \delta_{ij}^2 + Z_{WH}\,, \tag{1.1.2}$$

where Z_{WH} is the wormhole contribution, at least in the semiclassical saddlepoint approximation of the gravitational path integral. Following [1,2,6,13,14,18,19,21,24–30] we interpret this situation as arising because the gravitational path integral, treated as an effective theory below some cutoff, and certainly after taking a saddlepoint approximation, calculates coarse-grained observables that compute some kind of average over corresponding quantities in an underlying more fundamental theory that we refer to as the fine-grained theory.

To reflect this interpretation we will write an overbar over quantities computed using the gravitational path integral, for example,

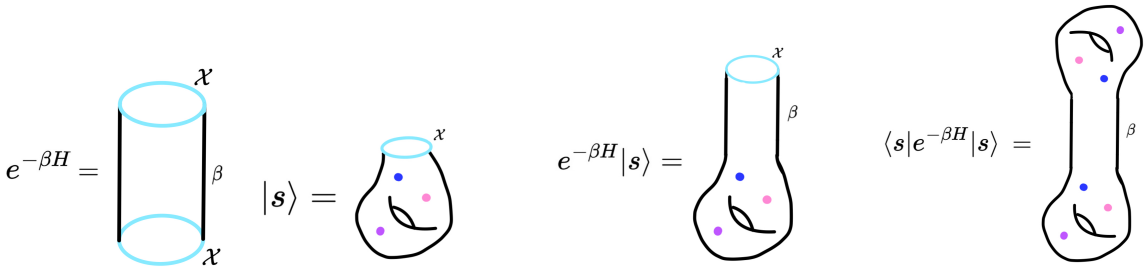
$$\overline{\langle i_1|j_1\rangle\,\langle i_2|j_2\rangle\cdots\langle i_n|j_n\rangle}\,. \tag{1.1.3}$$

Thus, it can transpire that $\overline{\langle i|j\rangle}$ vanishes because of cancellation of random phases in the coarse-grained computation of overlaps of states in the fine-grained theory, while these phases cancel in the magnitude squared $\overline{\langle i|j\rangle\,\langle j|i\rangle}$. If so, the square root of the latter quantity quantifies the actual magnitude of the overlap in the fine-grained theory.

Sharpening the precise relationship between wormhole saddlepoints and coarse-graining in the fine-grained theory — and exploiting it to extract results about the underlying quantum gravity Hilbert space — is one of the central goals of this thesis.

1.1.2 Action of the Hamiltonian

Next, we define the action of the Hamiltonian on our states. The action of $e^{-\beta H}$ is to glue an asymptotic boundary "cylinder" $\mathcal{X} \times \mathbb{I}_\beta$ to the cut (Figs. 1.2a, 1.2b and 1.2c). In other words, it lengthens the cut open boundary by a Euclidean boundary time β. This is motivated by ADM formalism for Einstein gravity, where the bulk Hamiltonian generates boundary time evolution. The so-called survival amplitude $\langle \mathbf{s}|e^{-\beta H}|\mathbf{s}\rangle$ is therefore constructed by sewing the $|\mathbf{s}\rangle$ bra and ket together via a boundary cylinder of length β resulting in the closed boundary manifold $\mathcal{M}_{\mathbf{s}}(\beta)$ (Fig. 1.2d). The



(a) (b) (c) (d)

Figure 1.2: Action of $e^{-\beta H}$ on states. (**a**) Action of the Hamiltonian; from a different perspective this is preparing the state $|2\beta\rangle$. (**b**) Example state in $\mathcal{H}_\mathcal{X}$. (**c**) Action of the Hamiltonian on this state. (**d**) Survival amplitude after Euclidean time evolution, alternatively the overlap $\langle 2\beta||s^*\rangle_L \otimes |s\rangle_R$ (adapted from [1]).

Hamiltonian can alternatively be viewed as creating a state on two copies of $\mathcal{X}$ – pictorially, we imagine bending the cylinder in Fig. 1.2a around into a half-torus. We denote the state constructed in this manner by the operator $e^{-\frac{\beta H}{2}}$ as $|\beta\rangle$ (note the 1/2 in the exponent). From this point of view, we can also think of overlaps between time-evolved states in $\mathcal{H}_\mathcal{X}$, for example $\langle i|\, e^{-\frac{\beta H}{2}}\, |j\rangle$, as overlaps between $|\beta\rangle$ and a tensor product state, $\langle \beta\, |i^*\rangle_L \otimes |j\rangle_R$. In QFT, a similar half-torus path integral prepares the Thermofield Double (TFD) state $|TFD(\beta)\rangle = \sum_n e^{\frac{-\beta E_n}{2}} |E_n\rangle^*_L |E_n\rangle_R$ defined on two copies of the theory. In what follows we refer to the norm

$$\langle \beta|\beta\rangle \equiv Z(\beta) \tag{1.1.4}$$

as the "Z partition function". We introduce this terminology as it is not clear from the gravitational Hilbert space perspective presented here why this path integral should

compute the thermal partition sum of the gravity theory. Chapter 4 is concerned with showing that the naive equality

$$Z(\beta) = \mathrm{Tr}_{\mathcal{H}_{grav}}(e^{-\beta H}) \tag{1.1.5}$$

indeed holds, but only once non-perturbative wormhole corrections to the thermal trace are taken into account.

The leading saddlepoints contributing to $Z(\beta)$ in asymptotically AdS_{d+1} gravity are well known. They are the spherically symmetric, static, Euclidean solutions

$$\mathrm{d}s^2 = f(r)\,\mathrm{d}\tau^2 + \frac{\mathrm{d}r^2}{f(r)} + r^2\,\mathrm{d}\Omega_{d-1}^2\,, \tag{1.1.6}$$

with $\tau \sim \tau + \beta$, where the blackening factor $f(r)$ takes one of three forms:

- *Thermal AdS*: $f(r) = r^2/\ell^2 + 1$, with no horizon. The Euclidean geometry has the topology of a solid cylinder $\mathbb{D}^2 \times \mathbb{S}^{d-1}$ in which the thermal circle is non-contractible.

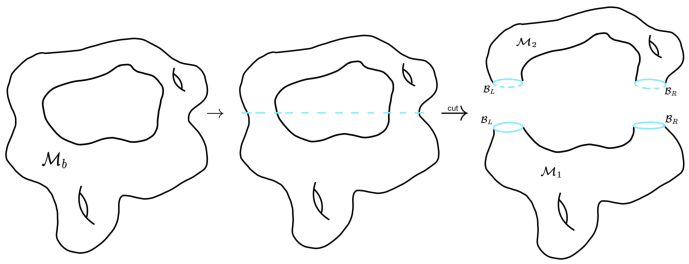
- *Euclidean Schwarzschild-AdS black hole*: $f(r) = r^2/\ell^2 + 1 - 16\pi G_N M/(d-1)V_\Omega r^{d-2}$, with a horizon at r_+ where $f(r_+) = 0$. Regularity at r_+ fixes the inverse temperature $\beta = 4\pi/f'(r_+)$. For $d \geq 2$ there are two branches of solutions at fixed β — a *small* black hole with $r_+ < r_+^{\min}$ and a *large* black hole with $r_+ > r_+^{\min}$ — corresponding to the two roots of $\beta(r_+)$. The Euclidean geometry has the topology of a disk in the (r, τ) plane fibred over $\mathbb{S}^{d-1}$, with the thermal circle contractible at the horizon.

The dominant saddle depends on the temperature. At low temperature ($\beta > \beta_{HP}$), thermal AdS dominates. At high temperature ($\beta < \beta_{HP}$), the large black hole dominates. The crossover is the Hawking-Page transition [11], occurring at an inverse temperature β_{HP} that depends on the dimension and the AdS radius. The small black hole saddle is always subdominant and unstable (it has negative specific heat), but contributes at subleading order in the path integral, depending on the contour.

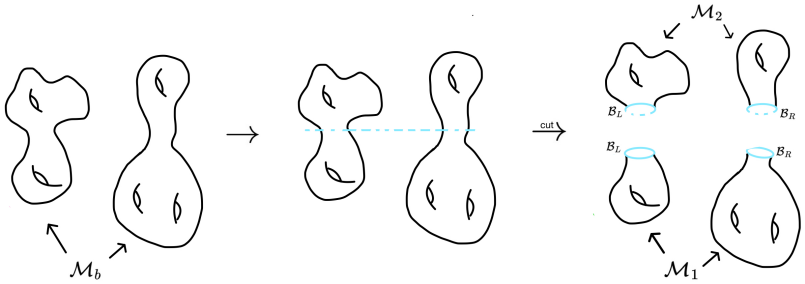
For ease of exposition we will refer to any saddlepoint geometry contributing to $Z(\beta)$ — thermal AdS, small black hole, or large black hole — collectively as a *disk*, even though only the black hole saddles literally have disk topology in the (r, τ) plane. This terminology, which is standard in the shell state literature, will allow us to formulate constructions involving multiple Z-partition function saddles uniformly without committing to which saddle dominates in a given regime. The metric (1.1.6) and the form of $f(r)$ then provide the local geometry on either side of any shell insertion in the constructions of Sec. 1.2.

1.1.3 The two-sided Hilbert space

Consider a Lorentzian quantum gravity theory for universes with a spatial asymptotic boundary having two disconnected components. We can build states for this theory by cutting open the Euclidean path integral on a slice with two disconnected boundary components $\mathcal{X} = \mathcal{X}_L \cup \mathcal{X}_R$, as after Lorentzian continuation each of these components continues to a separate asymptotic Lorentzian boundary. We will call this Hilbert space $\mathcal{H}_{\mathcal{X}_L \cup \mathcal{X}_R}$. There are two ways of preparing states in this space. First, we could cut open a single connected boundary manifold, as in Fig. 1.3a. For example, the state $|\beta\rangle$ is of this kind. Let us call the resulting Hilbert space $\mathcal{H}_{LR}$. Note that while the two components of the *boundary* cut are disconnected, there can exist bulk geometries satisfying these boundary conditions which connect the boundaries via a single bulk — the AdS-Schwarzschild black hole for example. Alternatively we could cut open two disconnected closed boundary manifolds $\mathcal{M}_L$ and $\mathcal{M}_R$ with cuts $\mathcal{X}_L$ and $\mathcal{X}_R$, naturally producing states in a tensor product Hilbert space $\mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$ (Fig. 1.3b).



(a) Construction of an example state in $\mathcal{H}_{LR}$.



(b) Construction of an example state in $\mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$.

Figure 1.3: Two distinct ways of preparing states in $\mathcal{H}_{\mathcal{X}_L \cup \mathcal{X}_R}$. (**a**) Construction of an example state in $\mathcal{H}_{LR}$. (**b**) Construction of an example state in $\mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$.

One might expect the full two-boudary Hilbert space to be an internal direct sum $\mathcal{H}_{LR} \oplus \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$, or more generally the vector-space sum $\mathcal{H}_{L \cup R} = \{u + v \mid u \in \mathcal{H}_{LR},\, v \in \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}\}$. This is puzzling from the holographic point of view:

AdS/CFT with two disconnected boundaries requires a tensor product structure, because the dual CFT lives independently on each boundary and so its Hilbert space factorises. One is therefore led to expect

$$\mathcal{H}_{L\cup R} = \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R} = \mathcal{H}_{LR}\,, \tag{1.1.7}$$

yet it is entirely non-obvious from the bulk why this should be the case — the states in $\mathcal{H}_{LR}$ are prepared by connected boundary manifolds and can correspond to geometries that connect the two boundaries through the bulk, while the tensor product states are prepared independently on each side. Reconciling these two descriptions is a non-trivial dynamical question about quantum gravity. In Chapter 3 we show that this factorisation does hold, and that it is a direct consequence of novel Euclidean wormhole saddles connecting the various boundary conditions. The existence of these saddles is itself a non-perturbative effect of the kind discussed in Sec. 1.1.1, and their discovery is one of the main results of this thesis.

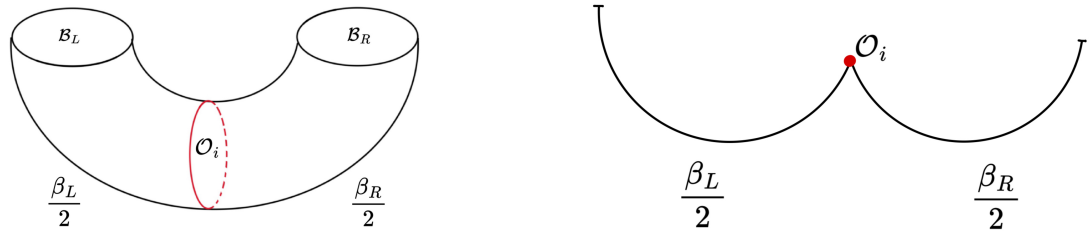
1.2 Shell states

The state construction from the path integral and the coarse-graining interpretation of wormholes reviewed above provide the general framework within which we will work. To make this framework concrete and computationally tractable, we now introduce a specific family of states for which wormhole saddlepoints can be explicitly constructed and their actions computed: the shell states. The shell states were first studied in [6, 19, 25, 31]; see [32] for a comprehensive discussion. In Chapter 2 we argue that these states form a basis of the gravity Hilbert space, thereby providing a convenient computation tool for analyzing wormhole contributions to the path integral.

1.2.1 Two-sided shell states

We start with a candidate basis for the two-sided Hilbert space $\mathcal{H}_{\mathcal{X}_L\cup\mathcal{X}_R}$, as these states are natural extensions of the familiar thermofield-double construction. The basis consists of states defined by cutting open the Euclidean gravitational path integral with a heavy spherically symmetric dust shell inserted on a periodic asymptotic boundary of topology $\mathbb{S}\times\mathbb{S}^{d-1}$. The cut consists of two disconnected components $\mathcal{B}_L = \mathcal{B}_R$, each of topology $\mathbb{S}^{d-1}$, such that the resulting cut-open boundary is of topology $\mathbb{I}\times\mathbb{S}^{d-1}$ (Fig. 1.4). We also insert a $\mathbb{S}^{d-1}$ symmetric heavy ($m\sim O(1/G_N)$) dust shell operator $\mathcal{O}_S$ separated by a boundary length $\frac{\beta_L}{2}$ and $\frac{\beta_R}{2}$ from the $\mathcal{B}_{L,R}$ cuts.

Following [6] we can obtain an infinite family of shell states in this way by varying the mass m_i of the shell operators. We call the resulting state $|i\rangle$.



(a) (b)

Figure 1.4: Asymptotic boundary condition for the gravity path integral defining the shell state. (**a**) Cut-open Euclidean boundary with topology $\mathbb{I}_{\frac{\beta_L+\beta_R}{2}} \times \mathbb{S}^{d-1}$ for preparation of the shell states. The shell operator $\mathcal{O}_i$ is pictured in red. In AdS/CFT we can also perform the path integral in the boundary CFT with insertion of a $\mathbb{S}^{d-1}$ symmetric operator dual to the shell. (**b**) Euclidean boundary with the $\mathbb{S}^{d-1}$ suppressed. We adopt this convention for the rest of the paper. Here $\beta_{L,R}/2$ are Euclidean "preparation times".

The construction of these states is intuitive within $\mathrm{AdS}_{d+1}/\mathrm{CFT}_d$ but applies to any dimension and cosmology with a time-like asymptotic boundary. In the AdS/CFT setting, the CFT lives on the boundary geometry. We define states on two copies of the CFT with Hamiltonians $H_L^{CFT} = H_R^{CFT}$ by inserting an operator $\mathcal{O}$ in the CFT path integral and evolving with the L and R Hamiltonian. Via the usual state-operator map this yields a state $|\mathcal{O}\rangle = |e^{-\frac{\beta_L H_L}{2}}\mathcal{O}e^{-\frac{\beta_R H_R}{2}}\rangle$ on $\mathcal{H}_L^{CFT} \otimes \mathcal{H}_R^{CFT}$ (Fig. 1.4). Taking $|E_n\rangle$ to label energy eigenstates for one copy of the CFT, such a state in the energy basis is given by:

$$|\mathcal{O}\rangle = \sum_{n,m} e^{-\frac{(\beta_L E_n + \beta_R E_m}{2})}\mathcal{O}_{mn}|E_n\rangle_L|E_m\rangle_R \tag{1.2.1}$$

and thus corresponds to a Partially Entangled Thermal State (PETS) [33]. The CFT dual states are defined by a choice of operator $\mathcal{O}$ that is holographically dual to the shell operator insertion on the asymptotic boundary of AdS_{d+1}. In the CFT this is achieved by inserting products of a scalar operator $\mathcal{O}_\Delta$ of $O(1)$ scaling dimension Δ. In order to make the states sufficiently heavy $m \sim O(1/G_N)$, we must insert $n \sim \ell^{d-1}/\mathrm{G_N}$ such operators (ℓ is the AdS radius) at spherically symmetric points on the $\mathbb{S}^{d-1}$, resulting in the operator $\mathcal{O}_S$.[1] In the CFT, each constituent $\mathcal{O}_\Delta$ is dual, via

[1] See [34, 35] for recent discussion regarding the existence of the large N limit of these states in the CFT dual.

the HKLL construction [36], to a massive scalar field insertion near the asymptotic boundary. The combined operator $\mathcal{O}_S$ can be chosen to create an approximately spherical thin cloud of dust particles which backreacts on the geometry at leading order in the semiclassical $G_N \to 0$ expansion. In the bulk, the cloud is effectively described as a pressureless perfect fluid localised at the worldvolume $\mathscr{W}$ of a thin shell, with energy-momentum tensor $T_{\mu\nu}|_{\mathscr{W}} = \sigma\, u_\mu\, u_\nu$, where σ is the surface density and u^μ the proper velocity of the fluid [6, 25].

To be clear, a particular microstate of the gravity theory is constructed by specifying the microscopic details of the operator insertions that make up the heavy shell. There are expected to be many such exact microstates which share the same semiclassical description as a spherically symmetric heavy dust shell in the bulk. As discussed in Sec. 1.1.1, the semiclassical gravitational path integral in the manner used here appears to compute a coarse-graining over these fine-grained details. Thus while we have some particular microscopic realisation in mind for each of our shell states, we can, remarkably, learn much about the underlying theory without specifying these details. This feature has a natural interpretation through the lens of the Eigenstate Thermalisation Hypothesis (ETH). Assuming dynamical chaos in the high-energy spectrum of the CFT, the matrix elements of $\mathcal{O}_S$ in the energy basis admit the form [25]

$$(\mathcal{O}_S)_{nm} = e^{-S(\bar{E})/2}\, g(\bar{E}, \omega)^{1/2}\, R_{nm}\ , \tag{1.2.2}$$

where $\bar{E} = (E_n + E_m)/2$, $\omega = E_n - E_m$, $g(\bar{E}, \omega)$ is a smooth envelope function encoding the microcanonical two-point function, and R_{nm} are erratically varying $O(1)$ complex coefficients. Because $\mathcal{O}_S$ is heavy, the envelope function $g(\bar{E}, \omega)$ scales non-trivially with the microcanonical entropy $S(\bar{E})$ and the shell mass m, modifying the large-N saddlepoint equations — this is dual to the classical backreaction of the thin shell on the geometry [25]. When the coefficients R_{nm} are promoted to actual random variables of zero mean and unit variance, they define an ensemble of operators consistent with the semiclassical description. We show in Chapter 6 that cumulants of correlation functions over this ensemble are captured by the wormhole contributions to the gravitational path integral described below.

1.2.1.1 Overlaps and orthogonality

We will want to compute the Gram matrix $G_{ij} \equiv \langle i|j\rangle$ of overlaps between the shell states. We will say that the path integrals calculating these quantities have *shell*

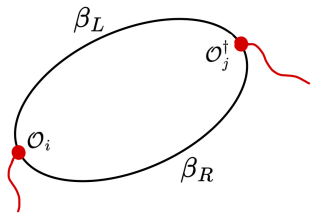


Figure 1.5: Shell asymptotic boundary condition for $\overline{\langle j|i\rangle}$, consisting of the operator insertions $\mathcal{O}_i$ and $\mathcal{O}_j^\dagger$ separated by asymptotic time extent β_L and β_R respectively. The red lines represent the shells propagating into the bulk.

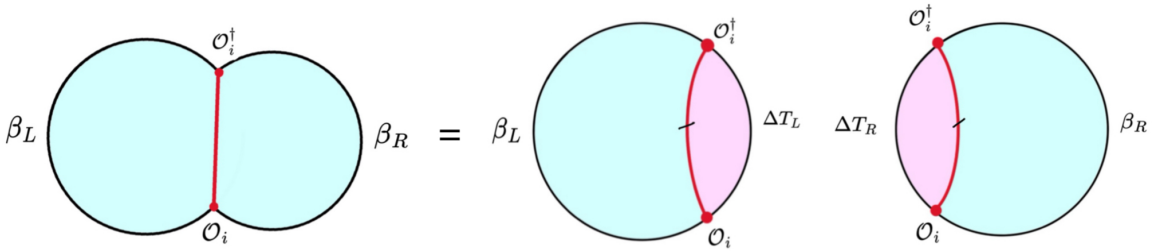


Figure 1.6: The saddlepoints for the shell norm path integral $\overline{\langle i|i\rangle}$ are constructed by gluing together two disks. In particular the shell homology regions (purple) on each disk are discarded and the resulting geometries glued together along the shell worldvolume. Note the propagation temperatures $\Delta T_{L,R}$ are denoted as $\Delta\tau_{\mp}$ in the main text.

boundaries which include operator insertions $\mathcal{O}_j$ and $\mathcal{O}_i^\dagger$ inserted on the asymptotic boundary separated by asymptotic times β_L and β_R respectively (Fig. 1.5). The overlap $\overline{\langle i|j\rangle}$ is evaluated by the gravitational path integral over geometries satisfying this boundary condition (Fig. 1.6).

States with sufficiently different shell masses have $\overline{\langle i|j\rangle} = \delta_{ij}\, Z_1$ within the saddlepoint approximation. To see this, note that shell states are not exact energy eigenstates but rather coherent-like wavepackets labelled by the classical rest mass m, so two states with $m_i \neq m_j$ will generically have overlapping support in the energy basis and are not exactly orthogonal. For the saddlepoint computation of $\overline{\langle i|j\rangle}$ to give a nonzero answer when $m_i \neq m_j$, the bulk path integral would need to mediate a transition between the two shell configurations, which requires $|m_i - m_j|$ local bulk interactions in Planck units, each suppressed by a power of the coupling. The perturbative overlap is therefore exponentially small in the mass difference. We choose the mass spacings to be large enough that this suppression is parametrically smaller than the e^{-1/G_N} wormhole corrections that are the true objects of interest, so that $\overline{\langle i|j\rangle} = \delta_{ij}\, Z_1$ to the required accuracy.

Indeed, higher topology wormhole contributions to the path integral are stabilised by the shell matter and modify the squared overlap to

$$\overline{|\langle i|j\rangle|^2} = \overline{\langle i|j\rangle\langle j|i\rangle} = Z_2 + \delta_{ij} Z_1^2 \,, \tag{1.2.3}$$

where Z_2 is the wormhole saddlepoint contribution (Fig. 1.8). The wormhole computation *is* sensitive to the non-orthogonality of states with different masses, but only through these non-perturbative moments: the connected wormhole geometry mediates precisely the kind of transition between shell configurations that is absent at the perturbative level.

In summary, we consider shell labels $\{m_i\}$ with spacings $|m_i - m_j|$ large enough that (i) any perturbative overlap suppression is subleading to e^{-1/G_N}, and (ii) each shell's local energy density ρ, curvature invariants R, $R_{\mu\nu}R^{\mu\nu}$, and typical quantum wavelengths satisfy $\rho^{1/4}$, $R^{1/2}$, $\lambda^{-1} \ll M_{\rm Pl}$. With these constraints the semiclassical saddle analysis and the resulting structure of the Gram matrix remain within the validity of the gravitational EFT.

1.2.1.2 The norm and the junction conditions

We now set $m_i = m_j$, thereby computing the norm of the shell state in the saddlepoint approximation. In such a saddle geometry, the shell worldvolume $\mathscr{W}$ bisects the Euclidean manifold into two components $X^{\pm}$. Since the shell operator $\mathcal{O}_S$ is spherically symmetric, the geometry on each side inherits the same symmetry, and Birkhoff's theorem then forces each component $X^{\pm}$ to be locally one of the disk geometries of Sec. 1.1.2, with metric of the form (1.1.6) and ADM masses $M_{\pm}$, horizon radii $r_{\pm}$, and inverse temperatures $\beta_{\pm} = 4\pi/f'_{\pm}(r_{\pm})$. We use the convention that X_- corresponds to the "left" and X_+ to the "right" region of the Euclidean geometry. The two components are glued along $\mathscr{W}$ using the Israel junction conditions [37]. Denoting the induced metric on each side by $h^{\pm}_{ab}$ and the extrinsic curvatures by $K^{\pm}_{ab}$, the junction conditions require continuity of the induced metric and relate the jump in extrinsic curvature to the shell stress-energy:

$$\Delta h_{ab} = 0 \,, \qquad \Delta K_{ab} - h_{ab}\,\Delta K = -8\pi G\, S_{ab} \,, \tag{1.2.4}$$

where $\Delta K_{ab} = K^+_{ab} - K^-_{ab}$, $\Delta K = h^{ab}\Delta K_{ab}$, and $S_{ab} = -\sigma\, u_a u_b$ for a dust shell. The total rest mass $m = \sigma\, V_\Omega\, R^{d-1}$ is conserved along $\mathscr{W}$. For a dust shell, these conditions reduce the dynamics to the motion of a non-relativistic effective particle with zero total energy [6, 25]:

$$\left(\frac{\mathrm{d}R}{\mathrm{d}T}\right)^2 + V_{\rm eff}(R) = 0 \,, \tag{1.2.5}$$

subject to the effective potential

$$V_{\text{eff}}(R) = -f_+(R) + \left(\frac{M_+ - M_-}{m} - \frac{4\pi G m}{(d-1)V_\Omega R^{d-2}} \right)^2 , \tag{1.2.6}$$

where T is the synchronous proper time of the dust particles and m is the conserved rest mass of the shell. The shell starts at the asymptotic boundary $R = r_\infty$, falls inward, bounces at a turning point $R = R_* \geq r_\pm$ on the axis of time-reflection symmetry, and returns to the boundary. The Euclidean time elapsed on each side is

$$\Delta\tau_\pm = 2 \int_{R_*}^{r_\infty} \frac{\mathrm{d}R}{f_\pm(R)} \sqrt{\frac{f_\pm(R) + V_{\text{eff}}(R)}{-V_{\text{eff}}(R)}} . \tag{1.2.7}$$

The total Euclidean periodicity on each side then fixes the physical inverse temperature to $\beta_{L,R} = \tilde{\beta}_{L,R} + \Delta\tau_\pm$, ensuring that the backreaction of the shell does not change the asymptotic mass of the geometry on either side. Note that the left and right asymptotic regions can have different temperatures $\beta_L \neq \beta_R$ because the shell acts as a domain wall.

1.2.1.3 The gravitational action of the norm

The norm $Z_1 \equiv \overline{\langle i|i\rangle}$ is computed by the gravitational path integral with the combined Euclidean action [6]

$$I[X] = -\frac{1}{16\pi G} \int_X \sqrt{g}\,(R - 2\Lambda) + \frac{1}{8\pi G} \int_{\partial X} \sqrt{h}\,K + \int_{\text{///}} \sigma + I_{\text{ct}} , \tag{1.2.8}$$

where Λ is the cosmological constant, K is the extrinsic curvature of the spacetime boundary ∂X, and I_{ct} are boundary counterterms that renormalise the on-shell action [38, 39], so that $Z_1 = e^{-I[X]}$.

It is useful to decompose the saddlepoint manifold as $X = X_- \cup X_{\text{shell}} \cup X_+$ (see Fig. 1.6), where $X_\pm$ are the portions of the Euclidean disk on either side of the shell outside the shell homology region, and X_{shell} is the region bounded by the shell worldvolume and constant Euclidean time hypersurfaces joining the horizons to the operator insertions at the boundary. By additivity of the on-shell action,

$$I[X] = I[X_-] + I[X_+] + I[X_{\text{shell}}] . \tag{1.2.9}$$

The first two terms depend only on the exterior black hole geometries. Since $X_\pm$ are angular sectors of Euclidean black hole disks of circumference $\beta_{L,R}$, subtending an angle corresponding to the preparation time $\tilde{\beta}_{L,R}$, we have [6]

$$I[X_\pm] = \tilde{\beta}_{L,R}\, F(\beta_{L,R}) = \tilde{\beta}_{L,R} \left(M_{L,R} - \frac{S_{L,R}}{\beta_{L,R}} \right) , \tag{1.2.10}$$

where $F(\beta) = -\beta^{-1} \log Z(\beta)$ is the free energy and $S_{L,R} = A_{L,R}/4G$ is the Bekenstein-Hawking entropy of the respective black hole.

The shell contribution $I[X_{\text{shell}}]$ arises from the bulk curvature in the shell region together with the worldvolume integral of σ. Using the on-shell Einstein equations, the Ricci scalar receives a smooth background piece and a distributional contribution from the shell stress-energy:

$$R - 2\Lambda = -2d + \frac{16\pi G}{d-1}\,\sigma\,\delta(y)\ , \tag{1.2.11}$$

where y is a normal coordinate to $\mathscr{W}$. Substituting into (1.2.8) and collecting terms gives [6]

$$I[X_{\text{shell}}] = \frac{d}{8\pi G}\,\text{Vol}(X_{\text{shell}}) + m\,\frac{d-2}{d-1}\,L[\gamma_{\mathscr{W}}]\ , \tag{1.2.12}$$

where $\text{Vol}(X_{\text{shell}})$ is the spacetime volume of the shell region and $L[\gamma_{\mathscr{W}}]$ is the proper length of the shell trajectory in the (r,τ) plane, given by

$$L[\gamma_{\mathscr{W}}] = 2\int_{R_*}^{r_\infty} \frac{\mathrm{d}R}{\sqrt{-V_{\text{eff}}(R)}}\ . \tag{1.2.13}$$

The second term in (1.2.12) has the form of the action of a heavy particle propagating in the (r,τ) plane, reflecting the spherical symmetry of the shell. After adding counterterms to remove long-distance divergences as $r_\infty \to \infty$, the full norm is

$$-\log Z_1 = \tilde{\beta}_L\,F(\beta_L) + \tilde{\beta}_R\,F(\beta_R) + I_{\text{ren}}[X_{\text{shell}}]\ , \tag{1.2.14}$$

where $I_{\text{ren}}[X_{\text{shell}}]$ denotes the renormalised shell action. The key feature of (1.2.14) is its clean separation into exterior and interior contributions: the exterior pieces $\tilde{\beta}_{L,R}F(\beta_{L,R})$ are determined entirely by the asymptotic data of the black hole, while the shell piece $I_{\text{ren}}[X_{\text{shell}}]$ encodes the backreaction and vanishes as $m \to 0$.

1.2.1.4 The large shell mass limit

A significant simplification occurs in the limit where the shell mass is much larger than the black hole mass, $m \gg M_\pm$ [6, 19, 31]. In this regime the effective potential (1.2.6) is dominated by the gravitational self-energy of the shell, and the turning point is pushed to large radius, $R_*^{d-1} \sim G\,m\,\ell/(d-1)V_\Omega \to \infty$. The shell therefore barely penetrates the bulk before bouncing back, and the Euclidean time elapsed on each side (1.2.7) vanishes, $\Delta\tau_\pm \to 0$. It follows from $\beta_{L,R} = \tilde{\beta}_{L,R} + \Delta\tau_\pm$ that the preparation temperatures approach the physical temperatures, $\tilde{\beta}_{L,R} \to \beta_{L,R}$.

Geometrically, this means the shell homology region X_{shell} in the decomposition (1.2.9) pinches off: as $R_* \to \infty$ the shell worldvolume is pushed to the asymptotic

boundary and the constant-time hypersurfaces bounding X_{shell} collapse onto one another. What remains of the saddlepoint are two complete Euclidean disks of circumference β_L and β_R respectively, joined at a single point on the asymptotic boundary where the shell operator is inserted.

The renormalised shell action (1.2.12) also simplifies. In the large mass limit $\text{Vol}(X_{\text{shell}}) \approx \frac{1}{d-1} L[\gamma_{\not\prime\prime}]$ [6], so the two terms in (1.2.12) combine into a single proper-length contribution. After renormalisation one finds

$$I_{\text{ren}}[X_{\text{shell}}] \;\approx\; 2(d-1)\log(G_N\, m) \;\equiv\; -\log Z_m \;, \tag{1.2.15}$$

where Z_m is a universal function of the shell mass alone, independent of the black hole parameters $\beta_{L,R}$. Since $\tilde{\beta}_{L,R} \to \beta_{L,R}$ and $I[X_\pm] \to \beta_{L,R}\, F(\beta_{L,R}) = -\log Z(\beta_{L,R})$, the norm (1.2.14) factorises as

$$Z_1 \;\approx\; Z(\beta_L)\, Z(\beta_R)\, Z_m \;, \tag{1.2.16}$$

where $Z(\cdots)$ denotes the Z-partition function defined by (1.1.4).

This factorisation has a transparent geometric origin: in the large mass limit the saddlepoint consists of two decoupled Euclidean black hole disks, one for each asymptotic region, connected only by a pointlike shell insertion at the boundary. Each disk contributes a factor of the Z partition function, and the shell contributes the universal factor Z_m. Crucially, (1.2.16) holds independently of which saddlepoint of the Z partition function dominates on each side — whether it is the Euclidean black hole or thermal AdS — because the shell homology region has pinched off and its contribution is insensitive to the geometry of the disk it is glued to.

This universality was demonstrated in detail for single-shell states in [6]. For multi-shell states in this limit, each shell independently contributes a factor Z_{m_i} and the geometries between shells contribute additional Z partition function factors (see Sec. 1.2.1.6). Furthermore, as explained in Sec. 5 of [1], the factorised form (1.2.16) is expected to persist even after perturbative bulk corrections are included, as such corrections simply map onto corresponding corrections to $Z(\beta_{L,R})$ and Z_m separately. This was verified explicitly in [40] by embedding the shell construction in a doubly holographic model.

For normalised states $|k\rangle/\sqrt{Z_1}$ the universal shell factor Z_m cancels in all physical quantities, and the overlaps and their moments depend only on the Z partition functions. This is the key simplification that makes the wormhole computations in this thesis particularly tractable.

1.2.1.5 Shell geometry

Focusing now on the asymptotically AdS case for concreteness, we can associate a geometry to a shell state by considering the leading gravitational saddlepoint contributing to the norm $\overline{\langle i|i\rangle} = \overline{Z}(\beta_L)\overline{Z}(\beta_R)Z_{m_i}$ (Fig. 1.6). As shown above, in the large shell mass limit this is determined by the leading saddlepoint for the Z partition functions $\overline{Z}(\beta_L), \overline{Z}(\beta_R)$. The leading Z-saddle depends on whether the inverse effective temperature, i.e. Euclidean time period $\beta_{L,R}$, is above or below the Hawking-Page transition [11] at β_{HP}. For example if $\beta_L > \beta_{HP}$ the L disk saddle is thermal AdS, whereas for $\beta_L < \beta_{HP}$ it is the Euclidean AdS black hole (similarly for R). The time reflection symmetric bulk slice of these saddles can be analytically continued to Lorentzian signature to associate a Lorentzian geometry to the state. As β_L, β_R can be tuned independently there are 4 regimes to consider:

1. $\beta_L, \beta_R < \beta_{HP}$: The leading saddle is a two-sided Schwarzschild-AdS black holes with a long wormhole interior (Fig. 1.7a). We will refer to these as *type 1* shell states and the space spanned by these states as $\mathcal{H}_1 \subseteq \mathcal{H}_{Shell}$. These are the states originally considered in [6]. In asymptotically flat space these saddles dominate in the finite-energy microcanonical ensemble for E_L, E_R both above the black hole threshold.

2. $\beta_L, \beta_R > \beta_{HP}$: The leading saddle is two copies of thermal AdS with the addition of a compact Big-Crunch AdS cosmology (Fig. 1.7b). This construction was studied in detail in [31]. We will refer to these as *type 2* shell states and the space spanned by them as $\mathcal{H}_2 \subseteq \mathcal{H}_{Shell}$. In asymptotically flat space these saddles dominate in the canonical ensemble, and in the finite-energy microcanonical ensemble if both the E_L, E_R energies are below the black hole threshold.

3. $\beta_L < \beta_{HP}$ and $\beta_R > \beta_{HP}$ or vice versa : The leading saddle corresponds to an L (R) single-sided AdS Black hole and a disconnected R (L) copy of thermal AdS (Fig. 1.7c). We refer to these as *type 3* shell states and the space spanned by them as $\mathcal{H}_3 \subseteq \mathcal{H}_{Shell}$. In asymptotically flat space these saddles dominate in the finite-energy microcanonical ensemble if only one of the E_L, E_R energies is above the black hole threshold.

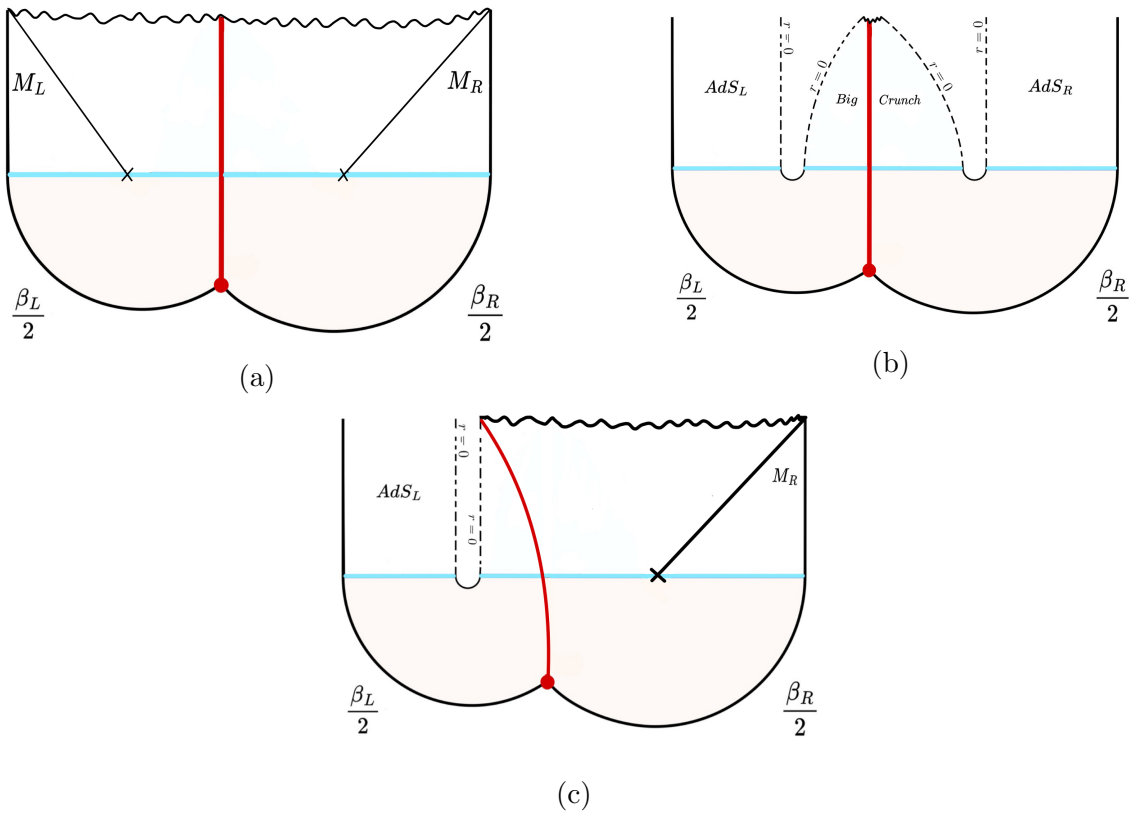


Figure 1.7: Diagrams showing the analytic continuation of the type 1-3 shell states to Lorentzian signature. (**a**) Type 1 shell state with $\beta_L < \beta_{HP}$ and $\beta_R < \beta_{HP}$ corresponding to a two-sided black hole. (**b**) Type 2 shell state with $\beta_L > \beta_{HP}$ and $\beta_R > \beta_{HP}$ consisting of two copies of thermal AdS with the addition of a compact Big-Crunch AdS cosmology. (**c**) Type 3 shell state with $\beta_L > \beta_{HP}$ and $\beta_R < \beta_{HP}$ corresponding to a disconnected geometry with an R single sided BH and an L thermal AdS factor.

1.2.1.6 Multi-shell states

More generally, we can insert multiple shells behind the horizon. Given operators $\mathcal{O}_1, \ldots, \mathcal{O}_n$ and intermediate Euclidean times $\tilde{\beta}_1, \ldots, \tilde{\beta}_{n-1}$, a multi-shell state takes the form [6]

$$|\Psi\rangle = |\rho_{\tilde{\beta}_L/2}\, \mathcal{O}_n\, \rho_{\tilde{\beta}_{n-1}} \cdots \rho_{\tilde{\beta}_1}\, \mathcal{O}_1\, \rho_{\tilde{\beta}_R/2}\rangle \ , \tag{1.2.17}$$

where $\rho_\beta = e^{-\beta H}$. Each pair of adjacent shells is separated by a region locally described by a Euclidean black hole or thermal AdS geometry, with the consistency condition $\beta_i = 2\tilde{\beta}_i + \Delta\tau_+^i + \Delta\tau_-^{i+1}$ ensuring the correct Euclidean periodicity. The thin shell formalism applies locally to each shell trajectory.

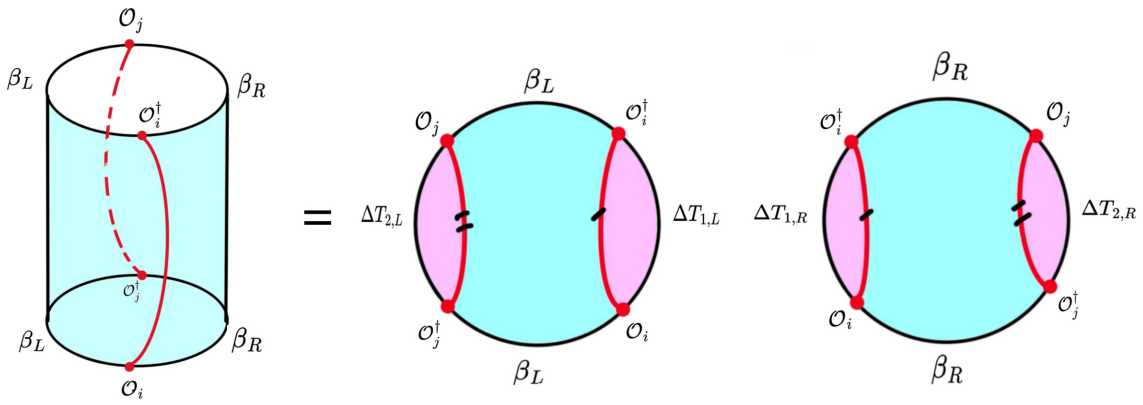


Figure 1.8: Construction of shell wormhole saddles Z_2. We glue two disks together by discarding the regions indicated in purple and attaching the remainder along the shell worldvolumes (red, black dashes indicate the glued shells) while obeying the Israel junction conditions. Saddlepoints, i.e., solutions to the equation of motion, exist with this topology [6].

1.2.1.7 Shell wormholes

The shell states are useful because they stabilise wormhole geometries whose action can be explicitly computed. Consider, for example, the gravitational path integral computing $\overline{\langle i|j\rangle \langle j|i\rangle}$, where each factor $\langle i|j\rangle$ corresponds to a periodic asymptotic boundary. While perturbatively $\overline{\langle i|j\rangle \langle j|i\rangle} \sim \delta_{ij} Z_1^2$, there exist non-perturbative wormhole corrections: saddle-points of cylinder topology in which the $\mathcal{O}_i$ insertion on one boundary propagates through the bulk and is absorbed by the $\mathcal{O}_i^\dagger$ insertion on the other, and similarly for $\mathcal{O}_j$. These wormholes are constructed by identifying shell worldvolumes across two disks subject to the Israel junction conditions as indicated in Fig. 1.8. The saddlepoint geometry on either side of each shell can be any solution to the equations of motion, so these are portions of any saddlepoint contributing to the Z partition function in (1.1.4).

Generalising this construction, we build the maximally connected wormhole saddlepoints contributing to the n-th power of the Gram matrix of overlaps $G_{ij} = \langle i|j\rangle$, which takes the form $\overline{(G^n)_{ii}} = \overline{\langle i|j_1\rangle \langle j_1|j_2\rangle \cdots \langle j_{n-1}|i\rangle}$, by connecting n asymptotic boundaries with a single multiboundary wormhole. The k-th shell propagates from $\mathcal{O}_k$ to $\mathcal{O}_k^\dagger$ on different asymptotic boundaries, so the index structure of $\overline{(G^n)_{ii}}$ manifests geometrically as the wormhole in Fig. 1.9, which we call the *pinwheel*. There are also disconnected saddlepoints connecting subsets of the boundaries, which look like collections of smaller pinwheels.

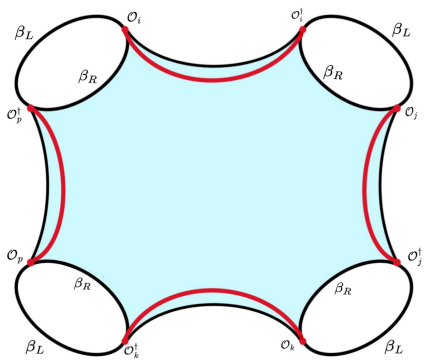


Figure 1.9: Fully connected wormhole geometry for $\overline{(G^n)_{ii}}$ consisting of n shell boundaries connected by a single bulk, pictured here for $n = 4$.

To construct a pinwheel saddlepoint, we cut Fig. 1.9 open along the shell worldvolumes to produce the *L*- and *R-sheet diagrams* in Fig. 1.10a. These sheets are obtained by removing the pink shell homology regions from the L and R disks in Fig. 1.10b. The L (R) disk boundary contains n shell operator insertions $\mathcal{O}_{1,\ldots,n}$, $\mathcal{O}^\dagger_{1,\ldots,n}$, with $\mathcal{O}_i$ and $\mathcal{O}_i^\dagger$ separated by a time $\Delta T_{L,i}$ ($\Delta T_{R,i}$) and $\mathcal{O}_i^\dagger$, $\mathcal{O}_{i+1}$ separated by a time β_L (β_R). The saddle is then assembled by removing the shell homology regions and gluing the disks together along the corresponding shell worldvolumes using the Israel junction conditions [37], which dynamically determine the propagation times ΔT_i. This procedure can be applied to any saddle of the Z partition function, so there are in general multiple saddles.

In the large mass limit of Sec. 1.2.1.4, the same geometric factorisation that simplified the norm applies to each shell in the pinwheel: the shell homology regions pinch off, each shell contributes a universal factor Z_{m_i}, and the remainder of the geometry on each side reduces to a complete disk. The fully connected n-boundary wormhole therefore decomposes as $\overline{(G^n)_{ii}}|_{\text{connected}} = \overline{Z}(n\beta_L)\,\overline{Z}(n\beta_R)\prod_{i=1}^n Z_{m_i}$, where the L and R disks now have circumference $n\beta_{L,R}$. Setting $n = 1$ recovers the norm (1.2.16). For normalised states $|k\rangle/\sqrt{Z_1}$ the shell factors cancel and we obtain

$$\overline{(G^n)_{ii}}|_{\text{connected}} = \frac{\overline{Z}(n\beta_L)\,\overline{Z}(n\beta_R)}{\overline{Z}(\beta_L)^n\,\overline{Z}(\beta_R)^n}\,, \tag{1.2.18}$$

within the saddlepoint approximation. Note that this equality is not restricted to the leading saddlepoint: each Z partition sum on the right-hand side runs over all saddlepoints, accounting for all saddles of $\overline{(G^n)_{ii}}|_{\text{connected}}$.

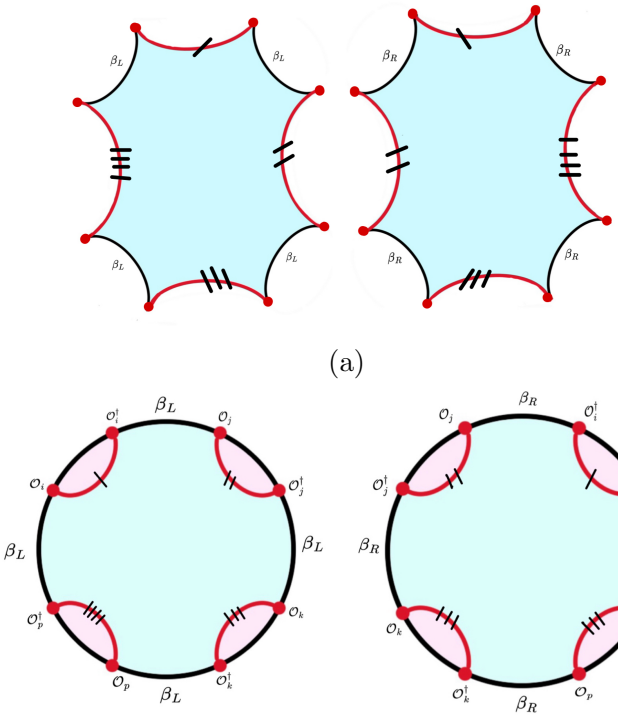


Figure 1.10: Schematic of the pinwheel wormhole construction. (**a**) Sheet diagrams obtained by cutting Fig. 1.9 along the shell worldvolumes. (**b**) Each sheet can be regarded as a portion of a disk; the pinwheel saddle is constructed by gluing two disks together along the corresponding shell worldvolumes via the junction conditions.

1.2.2 Single-boundary shell states

We build states in the single-boundary Hilbert space $\mathcal{H}_{\mathcal{B}}$ by cutting open boundary manifolds with topology $\mathbb{R}\times\mathbb{S}^{d-1}$, so that the cut $\mathcal{B}$ has topology $\mathbb{S}^{d-1}$. Our candidate basis states for $\mathcal{H}_{\mathcal{B}}$ are the single-sided thin matter shell states discussed in [2–4, 41]. To define these states we take a half-boundary with topology $\mathbb{R}^{<0}\times\mathbb{S}^{d-1}$, insert the shell operator $\mathcal{O}_S$ at $\tau=0$, and evolve for a further Euclidean time $\beta/2$ up to the cut (Fig. 1.11). In the CFT, this prepares the state

$$|\Psi_S\rangle = e^{-\frac{\beta}{2}H_{CFT}}\mathcal{O}_S|0\rangle \ , \tag{1.2.19}$$

where $\mathcal{O}_S$ is the same heavy dust shell operator described in Sec. 1.2.1. Varying the shell mass m_i produces an infinite family of states $|i\rangle$.

The single-boundary shell states are constructed in close analogy with the two-sided states of Sec. 1.2.1, and can in fact be obtained as a limit of the latter by sending one preparation temperature to infinity, $\beta_L\to\infty$. Since the thin shell formalism, the



(a) (b)

Figure 1.11: Cut-open path integral boundary condition defining the single-sided shell states. (**a**) Euclidean boundary with topology $\mathbb{R}^{<0} \times \mathbb{S}^{d-1}$. The shell operator $\mathcal{O}_i$ is pictured in red and $\beta/2$ is the Euclidean preparation time. (**b**) Euclidean boundary with the $\mathbb{S}^{d-1}$ suppressed; we sometimes depict this boundary with a curve or kink to clarify diagrams.

large mass factorisation, and the wormhole construction all carry over from the two-sided case, we focus here on the features that are genuinely new.

The key difference from the two-sided construction is that the boundary now extends to infinite Euclidean time on one side. Many path integrals of interest will have boundary conditions consisting of an infinite half-line $[-\infty, 0]$, a series of operator insertions and Euclidean time evolutions spanning a total time T_E, and another infinite half-line $[T_E, \infty]$. We refer to T_E as the "length of the strip" and denote such a boundary condition by $S(T_E)$. The Gram matrix $G_{ij} \equiv \langle i|j\rangle$ is then defined by a *shell strip* boundary of length β, with $\mathcal{O}_i$ and $\mathcal{O}_j^\dagger$ inserted at the two ends (Fig. 1.12). Orthogonality for $m_i \neq m_j$ follows by the same reasoning as in Sec. 1.2.1.1.

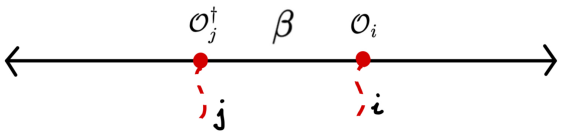


Figure 1.12: Shell-strip asymptotic boundary condition for the overlap $\langle j|i\rangle$ consisting of the line $\lim_{\alpha\to\infty}[-\alpha, \beta + \alpha]$ on which $\mathcal{O}_i$ and $\mathcal{O}_j^\dagger$ are inserted at $\tau = 0$ and $\tau = \beta$ respectively.

1.2.2.1 Norm, wormholes, and the large mass limit

The saddlepoints computing the norm $\overline{\langle i|i\rangle}$ are constructed by gluing a Euclidean disk (any Z partition function saddle) to a Euclidean strip — that is Euclidean vaccum AdS with non-compact time — along the shell worldvolume via the Israel junction

conditions (Fig. 1.13). The gravitational action decomposes into disk, strip, and shell contributions in exact analogy with the two-sided case (Sec. 1.2.1.3), with the strip replacing one of the two disks.

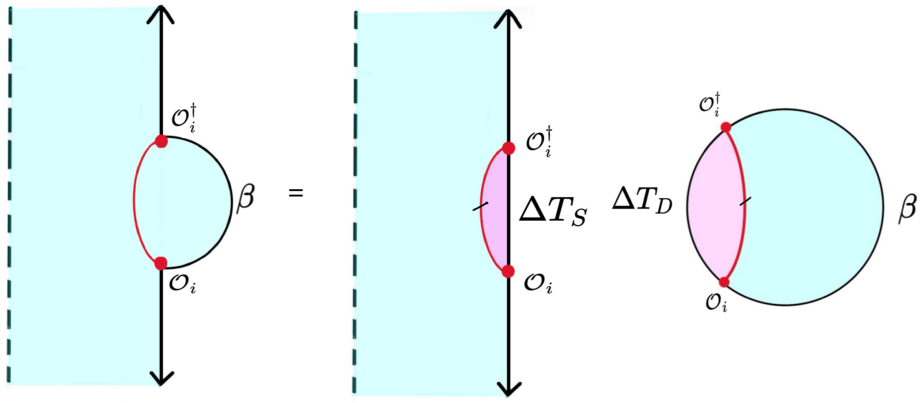


Figure 1.13: The saddle for the norm $\overline{\langle i|i\rangle}$ is constructed by gluing a disk and a strip together along the shell worldvolume, after discarding the shell homology regions (purple). The junction conditions dynamically determine the propagation times $\Delta T_{S,D}$.

In the large mass limit of Sec. 1.2.1.4, the shell homology regions pinch off just as in the two-sided case. Since the factorisation is independent of $\beta_{L,R}$ and the junction conditions are local, the two-sided results carry over directly upon taking $\beta_L \to \infty$.[2] The norm factorises as

$$\overline{\langle i|j\rangle} = \delta_{ij}\, Z_{m_i} \times \overline{Z(\beta)} \times \overline{S}(0)\ , \tag{1.2.20}$$

where $\overline{Z(\beta)}$ is the Z partition function from the disk and $\overline{S}(0)$ is the zero-length strip path integral.[3] This mirrors the two-sided factorisation (1.2.16), with $Z(\beta_L)$ replaced by $S(0)$.

Wormhole saddlepoints for the higher moments $\overline{(G^n)_{ii}}$ are constructed by gluing n strips into a single disk of circumference $n\beta$ along the shell worldvolumes (Fig. 1.14b), in direct analogy with the pinwheel construction of Sec. 1.2.1.7. In the large mass limit the saddle factorises as $\overline{G^n_{ii}} = \overline{Z}(n\beta) \times (\overline{S}(0))^n \times \prod_{j=1}^n Z_{m_j}$. Normalising the shell states to unit norm gives the single-sided analogue of (1.2.18):

$$|k\rangle \to \frac{|k\rangle}{\sqrt{Z_{m_k} \times \overline{Z(\beta)} \times \overline{S}(0)}} \quad ; \quad \overline{G^n_{ii}} = \frac{\overline{Z}(n\beta)}{\overline{Z}(\beta)^n}\ . \tag{1.2.21}$$

[2]The solution to the junction conditions for the single-sided shell states can be found in the appendix of [41], from which this behaviour is confirmed.

[3]The regularisation of the strip drops out of all calculations upon normalising the shell states, so we do not specify it.

Note that the strip factors cancel upon normalisation, so the final expression for the moments has the same form as the two-sided result (1.2.18) with a single Z factor rather than a product of L and R factors.

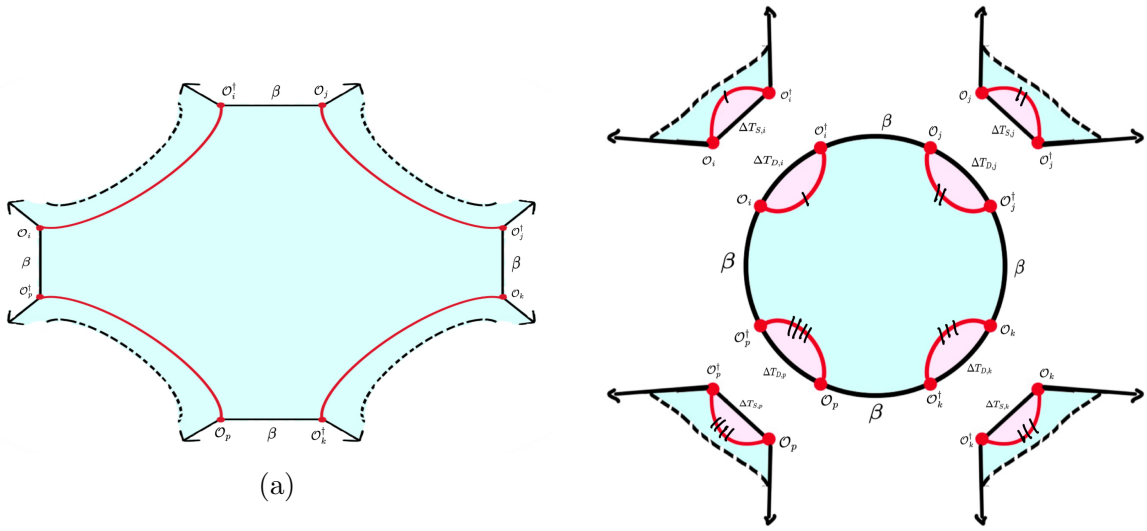
As in the two-sided case (Sec. 1.2.1.3), the factorised form (1.2.20) is expected to persist after perturbative bulk corrections are included [1], but this will not be important for the considerations in this thesis, as these are restricted to the saddle-point regime.

1.2.2.2 Saddle geometry

The Lorentzian geometry associated to a single-sided shell state is determined, as in the two-sided case, by the leading Z partition function saddle. There are two classes:

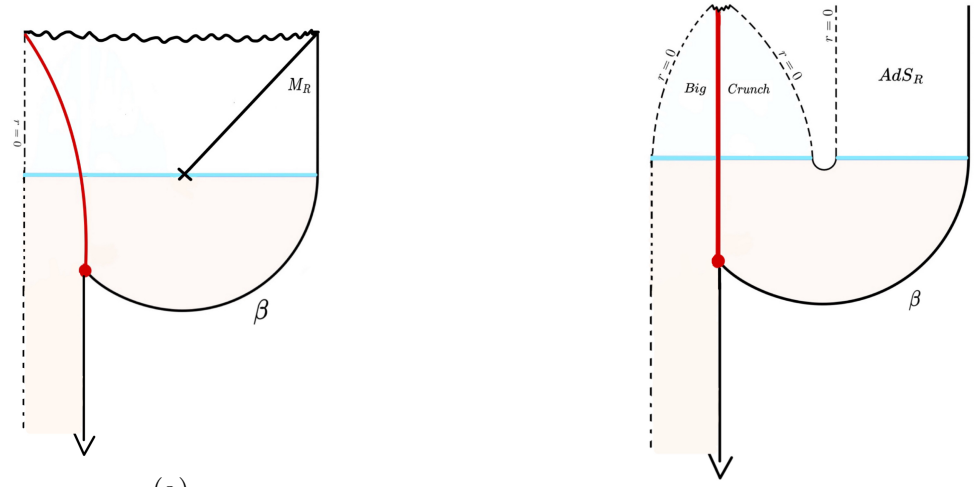
- **Type A** ($\beta < \beta_{HP}$): The leading disk saddle is the Euclidean black hole, and the Lorentzian geometry is a single-sided black hole with a shell in the interior, capped off by the vacuum behind the shell (Fig. 1.15a). In asymptotically flat space these saddles arise in the finite-energy microcanonical ensemble for energies above the black hole threshold.

- **Type B** ($\beta > \beta_{HP}$): The leading disk saddle has a non-contractible thermal circle, and the Lorentzian geometry consists of thermal AdS together with a disconnected compact "Big Crunch" universe in which the shell propagates (Fig. 1.15b). In asymptotically flat space these saddles arise for any β in the canonical ensemble.



(a) (b)

Figure 1.14: The fully connected single-sided shell wormhole saddles for $\overline{(G^n)_{ii}}$ are constructed by gluing n strips into a disk along the shell worldvolumes, depicted here for $n = 4$.



(a) (b)

Figure 1.15: Lorentzian continuation of the single-sided shell states. (**a**) Type A: a single-sided black hole with a shell behind the horizon. (**b**) Type B: thermal AdS with a disconnected compact Big-Crunch cosmology.

"Everything should be made as simple as possible, but no simpler."
— Albert Einstein

Chapter 2

A Nonperturbative Toolkit for Quantum Gravity

The framework reviewed in Chapter 1 provides only an indirect window onto quantum gravity. The wormhole contributions discussed in Sec. 1.1.1, in particular, show that the gravitational path integral cannot compute amplitudes of a single quantum theory: there exist quantities $\langle i|j\rangle$ which vanish on average while $|\langle i|j\rangle|^2$ does not. The natural interpretation is that the path integral computes a coarse-grained average over underlying fine-grained data, with wormhole saddles encoding the moments of the associated distribution. These moments are precisely the gravitational quantities that carry non-trivial information about the fine-grained theory — its operator norms, its level statistics, the dimension of its Hilbert space. The basic question this chapter addresses is how to systematically extract that information.

Two ingredients are needed. The first is a procedure for converting coarse-grained data into fine-grained statements. The second is a basis of states in which the relevant fine-grained observables are most naturally resolved — one for which the informative wormhole moments organise themselves transparently at the level of the saddlepoint approximation. This chapter develops both, building on [20, 29]. The first is the simple but powerful observation that if we want to establish $\mathcal{A} = \mathcal{B}$ in the fine-grained theory and the path integral gives us only coarse-grained data, computing $\overline{\mathcal{A}} = \overline{\mathcal{B}}$ is not enough: this is consistent with $\mathcal{A} \neq \mathcal{B}$ on individual configurations whose differences average to zero. To rule this out, it suffices to show in addition

that $\overline{(\mathcal{A}-\mathcal{B})^2}=0$, since the average of a non-negative quantity vanishes only if the quantity itself vanishes on every configuration in the support of the distribution. Both ingredients of this "squaring trick" are themselves gravitational path integrals — but together they constrain the fine-grained theory. The strength of the resulting fine-grained statement reflects the level of approximation at which the path integrals on either side are evaluated; in this chapter we work in the saddlepoint approximation, supplemented where appropriate by perturbative corrections around saddles. Even within this approximation, the sum-over-topologies computing $\overline{(\mathcal{A}-\mathcal{B})^2}$ are often intractable.

The second ingredient is a basis that resolves this difficulty, and is the principal technical contribution of the chapter. We argue that the shell states defined in Sec. 1.2 — those used in [6, 19, 28, 29] to account for black-hole microstate counts within microcanonical windows — form a *semiclassically complete* basis of the two-sided gravity Hilbert space $\mathcal{H}_{LR}$, with the formal projector onto their span

$$\Pi_{\mathcal{H}_{shell}} \;\equiv\; G^{-1}_{ij}\,|i\rangle\langle j| \tag{2.0.1}$$

acting as the identity on every state in the non-perturbative, semiclassical theory. As a corollary, the dimension of the microcanonical sector of the gravity Hilbert space is

$$\dim\mathcal{H}_{LR,E} \;=\; e^{\mathbf{S}(E_L)+\mathbf{S}(E_R)}\,, \tag{2.0.2}$$

identifying the Bekenstein-Hawking entropy with the dimension of the full microcanonical Hilbert space rather than of a restricted "shell superselection sector". All results are derived using only the rules of the gravitational path integral and make no assumptions about holographic duality.

The shell states are far from orthogonal: their wormhole-stabilised overlaps are precisely the non-perturbative effects that motivated the coarse-graining picture in the first place, and this is what makes them well suited to the present purpose. Standard linear algebra allows us to write the formal projector (2.0.1) onto the closed linear span of any κ shell states, with the inverse Gram matrix understood by analytic continuation from positive integer powers of G. The crucial observation is that inserting (2.0.1) into a gravitational path integral, and then taking a limit in which the shell basis is infinitely overcomplete ($\kappa\to\infty$), drastically simplifies the resulting sum over topologies: only geometries that preserve every maximal index loop in the boundary conditions survive, since any topology that breaks an index loop is suppressed by powers of $1/\kappa$. The wormhole correlations that encode the fine-grained data can thus become manifest at the saddlepoint level once observables are resolved

in the shell basis, and the formal expression (2.0.1) becomes a tractable computational device. We will exploit it repeatedly in subsequent chapters — to establish the factorisation of the two-boundary gravity Hilbert space (Chapter 3) and to give a microscopic account of the universality of black-hole entropy (Chapter 4).

These two ingredients together deliver the chapter's main results. The argument that $\Pi_{\mathcal{H}_{shell}}$ acts as the identity on $\mathcal{H}_{LR}$ proceeds by first establishing the equality $\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle = \langle\beta|\beta\rangle$ for the half-torus state $|\beta\rangle$ at any inverse temperature β. We give two derivations. The first is a microcanonical argument at finite κ in which a Schwinger-Dyson resummation of the shell-state resolvent, followed by a contour integral around a pole that emerges once $\kappa \geq e^{\mathbf{S}_L(E)+\mathbf{S}_R(E)}$, yields the dimension formula (2.0.2) as an immediate corollary. The second is a canonical argument in the strict $\kappa \to \infty$ limit, technically simpler than the first, that lifts via the squaring trick to a fine-grained equality at the saddlepoint level. We then extend the argument from $|\beta\rangle$ to a broad class of states whose norms admit a saddlepoint description, providing strong evidence that $\Pi_{\mathcal{H}_{shell}}$ acts as the identity on all of $\mathcal{H}_{LR}$. This extension rests on a structural assumption about the bijection between certain bulk saddle geometries, which we verify explicitly for representative classes of states; the spanning result is therefore strongly supported by the evidence but not proven in full generality.

A striking consequence of the spanning result is that geometries with horizons can be written as superpositions of states whose individual semiclassical descriptions involve no horizons but instead contain disconnected closed universes. This in turn implies that there can be no linear "geometry operator" on the gravity Hilbert space, consistent with the conjecture that bulk geometry emerges from quantum entanglement [42, 43].

The chapter is organised as follows. In Sec. 2.1 we make the squaring trick precise and discuss the levels of approximation at which we will work. In Sec. 2.2 we adapt the standard Gram matrix machinery (reviewed in Appendix A) to the gravity setting. In Sec. 2.3 we establish the shell resolution of the identity on the diagonal subspace of $\mathcal{H}_{LR}$ via the two derivations described above, recovering (2.0.2) as a corollary of the microcanonical analysis. Sec. 2.4 extends the argument to states with saddlepoint-dominated norms and works through illustrative examples, providing evidence that the shell states form a semiclassically complete basis of $\mathcal{H}_{LR}$. We conclude in Sec. 2.5 with a discussion of the implications for emergent geometry and of alternative bases that may probe the late-time spectral form factor.

2.1 From coarse-grained to fine-grained equalities

The gravitational path integral computes coarse-grained averages, denoted with an overbar, of fine-grained quantities (Sec. 1.1.1). Suppose we want to establish that two quantities $\mathcal{A}$ and $\mathcal{B}$ in the fine-grained theory are equal. Computing $\overline{\mathcal{A}} = \overline{\mathcal{B}}$ from the path integral does not suffice: this is consistent with $\mathcal{A} \neq \mathcal{B}$ on individual configurations whose differences average to zero. To rule this out we use the elementary observation that

$$\overline{\mathcal{A}} = \overline{\mathcal{B}} \quad \text{and} \quad \overline{(\mathcal{A}-\mathcal{B})^2} = 0 \quad \Longrightarrow \quad \mathcal{A} = \mathcal{B}\,. \tag{2.1.1}$$

The implication holds because $(\mathcal{A}-\mathcal{B})^2$ is non-negative on every configuration, so its average vanishes only if it itself vanishes pointwise.[1] Both ingredients on the left-hand side are coarse-grained gravitational path integrals — but together they constrain the fine-grained theory.

The squaring trick reduces fine-grained equality to two path-integral computations. These cannot in general be performed exactly, and we must resort to approximations. Three natural levels suggest themselves:

(I) Saddlepoint approximation. Sum over all topologies admitting a metric that solves the equations of motion, evaluating each at its saddlepoint action. We sometimes approximate further by retaining only the dominant saddle.

(II) Perturbative corrections. Include the Gaussian and higher fluctuations around each level-(I) saddle.

(III) Off-shell geometries. Include topologies that admit no saddle, and off-shell configurations that are not connected to any saddle by perturbative deformation.

The fine-grained statement derived via (2.1.1) is only as strong as the level of approximation at which both sides of (2.1.1) are evaluated. Throughout this chapter we work at level (I), and indicate where the arguments naturally extend to level (II): the heavy shells used in our construction localize to the asymptotic boundary and contribute a universal factor independent of the bulk geometry, so perturbative corrections on either side of (2.1.1) simply match (Sec. 2.4). We do not attempt to access level (III). An ultraviolet-finite completion of the theory — in something like string

[1] For complex-valued quantities the second condition should be read as $\overline{|\mathcal{A}-\mathcal{B}|^2} = 0$. All quantities to which we apply the trick in this chapter are manifestly real (traces of Hermitian operators, norms of states, partition functions), so we keep the simpler notation throughout.

theory — could require additional ingredients that are not captured by the effective gravitational description; those also lie beyond our scope. Some heuristic arguments regarding an extension to level (III) can be found in [1].

We emphasise that the saddlepoint approximation is itself non-perturbative in G_N: the wormhole saddles that play the central role below contribute at order e^{-1/G_N} and are invisible at any finite order of perturbation theory around a fixed background. The "nonperturbative toolkit" of the chapter title refers to this sense of non-perturbative, not to a claim of access to level (III).

2.2 Traces in an overcomplete basis

Consider a set of κ states $\{|i\rangle\}$ in the gravity Hilbert space constructed by cutting open the Euclidean path integral with operator insertions, as in Sec. 1.1. We characterise their overlaps by the Gram matrix

$$G_{ij} \equiv \langle i|j\rangle \ . \tag{2.2.1}$$

Although these states may be orthogonal in the low-energy effective theory, non-perturbative wormhole contributions to the path integral can render them non-orthogonal in the fine-grained theory. The span $\mathcal{H}_\kappa \equiv \text{span}\{|i\rangle\}$ may therefore have dimension less than κ, and the states $\{|i\rangle\}$ can be overcomplete as a basis for $\mathcal{H}_\kappa$.

The standard linear algebra of overcomplete bases (reviewed in Appendix A, following [29, 44]) supplies an orthonormal basis for $\mathcal{H}_\kappa$ together with formulas for the resolution of the identity, the trace of an operator, and the dimension of the span:

$$\mathbb{1}_{\mathcal{H}_\kappa} = G^{-1}_{ij}\, |i\rangle\langle j| \ , \tag{2.2.2}$$

$$\text{Tr}_{\mathcal{H}_\kappa}(\mathcal{O}) = G^{-1}_{ij}\, \langle j|\mathcal{O}|i\rangle \ , \tag{2.2.3}$$

$$\dim(\mathcal{H}_\kappa) = \lim_{n\to 0} \text{Tr}(G^n) \ , \tag{2.2.4}$$

where summation over repeated indices is understood throughout. The generalised inverse G^{-1} is defined by the analytic continuation $G^{-1} \equiv \lim_{n\to -1} G^n$ from positive integer powers of G, which is well-defined because G^n inverts the nonzero eigenvalues of G and leaves the zero eigenvalues untouched (Appendix A).

The advantage of this formalism in the gravity setting is that we never need to diagonalise G explicitly. Positive integer powers

$$G^n_{ij} = \langle i|j_1\rangle\langle j_1|j_2\rangle \cdots \langle j_{n-1}|j\rangle \tag{2.2.5}$$

are themselves overlaps, computed (in the coarse-grained sense) by the gravitational path integral over geometries with n asymptotic boundaries. The right-hand sides of (2.2.3) and (2.2.4) are therefore directly accessible in the path integral by analytically continuing such multi-boundary amplitudes:

$$\overline{\dim(\mathcal{H}_\kappa)} = \lim_{n\to 0} \overline{\mathrm{Tr}(G^n)} \,, \tag{2.2.6}$$

and similarly for traces of operators and matrix elements of resolutions of the identity. The remaining sections of this chapter exploit these formal expressions repeatedly to manipulate the gravitational path integral.

2.3 Shell states resolve the identity

Equipped with the formalism of Sec. 2.2, we now turn to the chapter's central technical claim: *the shell states defined in Sec. 1.2.1 resolve the identity on the two-sided Hilbert space $\mathcal{H}_{LR}$ in the large-κ limit.* Concretely, we will argue that

$$\Pi_{\mathcal{H}_{shell}} \;\equiv\; G^{-1}_{ij}\,|i\rangle\langle j| \;=\; \mathbb{1}_{\mathcal{H}_{LR}} \,, \tag{2.3.1}$$

where the equality is between quantities in the *fine-grained* theory, albeit at the saddlepoint level. Recall that $\mathcal{H}_{LR}$ denotes the gravity Hilbert space spanned by intrinsically two-boundary states, per the discussion around (1.1.7).

It is convenient to test (2.3.1) against the states $|\beta\rangle$ obtained by cutting the half-torus path integral (Sec. 1.1.2). The standard Gibbons-Hawking interpretation of $Z(\beta)$ on a periodic boundary $\mathbb{S}^1_\beta \times \mathbb{S}^{d-1}$ as a thermal trace — which we revisit in Chapter 4 — requires that $|\beta\rangle$ has support on many states in $\mathcal{H}_{LR}$. The proposed equality (2.3.1) therefore predicts

$$\overline{\langle\beta|\beta\rangle} \;\overset{?}{=}\; \overline{\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle} \;\equiv\; \lim_{n\to -1} \overline{G^n_{ij}\,\langle\beta|i\rangle\langle j|\beta\rangle} \,. \tag{2.3.2}$$

The left-hand side is a path integral subject to a single asymptotic boundary condition of length β. The right-hand side is the analytic continuation in n of an $(n+2)$-boundary path integral. The bulk of the section is devoted to evaluating the right-hand side and confirming the equality.

Our strategy proceeds in four steps. First, we identify the maximally-connected saddle contributing to $\overline{G^n_{ij}\langle\beta|i\rangle\langle j|\beta\rangle}$ at fixed positive integer n — the *folded wormhole* (Sec. 2.3.1). Second, we observe that the sum over saddles at fixed n contains many less-connected contributions, but that in the planar limit it can be reorganised by a Schwinger-Dyson resummation in terms of the maximally-connected pieces alone

(Sec. 2.3.2). Third, the resulting series cannot be summed analytically in the canonical ensemble; but its microcanonical projection can, and the equality (2.3.1) follows on each microcanonical band individually after a contour argument for the resolvent (Sec. 2.3.3). The microcanonical analysis yields the dimension of $\mathcal{H}_{LR,E}$ as a corollary. Finally, we observe that taking $\kappa \to \infty$ beyond the planar limit suppresses every less-than-fully-connected saddle altogether, so the full sum over topologies trivialises and the canonical equality follows directly; this canonical argument additionally lifts via the squaring trick to a fine-grained equality (Sec. 2.3.4).

2.3.1 The folded wormhole

The amplitude $\langle\beta|i\rangle$ corresponds to a periodic asymptotic boundary of length $\frac{1}{2}(\beta_L + \beta_R + \beta)$ with a single shell operator $\mathcal{O}_i$ inserted; here $\beta_{L,R}$ are the shell preparation temperatures from Sec. 1.2.1. We refer to this boundary topology as a *tadpole* (Fig. 2.1). The factor $\langle\beta|i\rangle\langle j|\beta\rangle$ in (2.3.2) is two tadpole boundaries, joined to the n shell-overlap boundaries of G^n_{ij} to give $n+2$ asymptotic components in total.

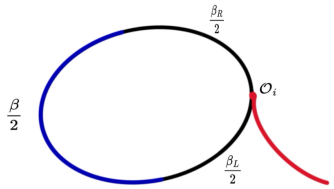
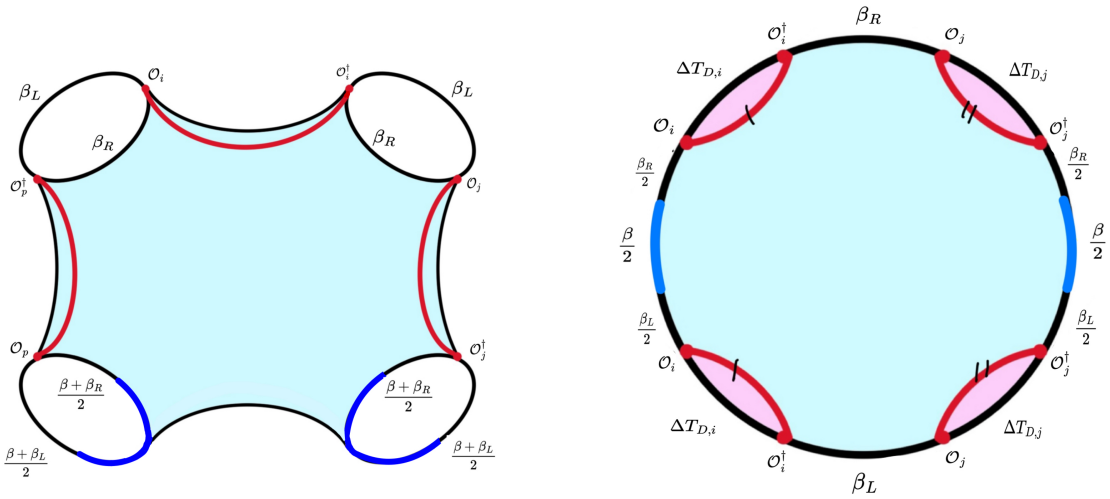


Figure 2.1: The *tadpole* boundary corresponding to $\langle\beta|i\rangle$: a periodic asymptotic boundary of length $\frac{1}{2}(\beta + \beta_L + \beta_R)$ with a single shell operator insertion $\mathcal{O}_i$ (red).

Saddles contributing to $\overline{G^n_{ij}\langle\beta|i\rangle\langle j|\beta\rangle}$ at fixed positive integer n are constructed by joining shell operator insertions across the various asymptotic boundaries via the Israel junction conditions, in direct analogy with the pinwheel construction of Sec. 1.2.1.7. A shell sourced by $\mathcal{O}_i$ can only annihilate against $\mathcal{O}_j^\dagger$ if $i = j$, so the bulk connectivity must be compatible with the shell index loops in the boundary condition. In this section we focus on the maximally-connected saddle: a fully-connected geometry in which both tadpole shells are absorbed by the chain coming from G^n_{ij}, so all $n+2$ asymptotic boundaries are joined through a single bulk component (Fig. 2.2a). We call it the *folded wormhole* and denote the sum over all such saddles — there is one for each saddle of the underlying Z partition function — by $\tilde{Z}_{n+2}$. Recall that we refer to these saddles as "disks" for ease, even though the bulk topology of thermal

AdS is an annulus. The folded wormhole differs from the pinwheel saddles considered in [6, 19, 28, 29], and is a generalisation of the $n = 0$ case discussed in Sec. 3 of [25].

A folded wormhole is constructed by "folding" a single Z-partition function disk onto itself (Fig. 2.2b). Consider a Euclidean disk with $2(n+1)$ pairs $(\mathcal{O}_{i,j,\dots}, \mathcal{O}^\dagger_{i,j\dots})$ of shell insertions on its asymptotic boundary, separated by Euclidean preparation times $\Delta T_{D,i,j\dots}$, with consecutive pairs separated by β_R in the upper half of the disk circle and β_L in the lower. In other words, each shell $i, j, \cdots$ appears in pairs on opposite sides of the disk boundary. The folded wormhole is obtained by discarding the shell homology regions and identifying the two copies of each shell's worldvolume, the propagation times $\Delta T_{D,i}$ being dynamically determined by the junction conditions. This construction can be performed for any saddle of the Z partition function, so there are in general multiple folded-wormhole saddles at each n.



(a) (b)

Figure 2.2: (**a**) Folded wormhole geometry $\tilde{Z}_{n+2}$, depicted for $n = 2$: the $n + 1$ shell-overlap boundaries and two tadpole boundaries are joined through a single bulk component. (**b**) Construction by self-folding: a Z-partition-function disk with $2(n+1)$ shell insertions has its shell homology regions (pink) removed, and the remaining geometry is glued back along corresponding worldvolume pairs.

In the large mass limit (Sec. 1.2.1.4) the homology regions pinch off, each $\Delta T_{D,i} \to 0$, and each shell contributes a universal factor Z_{m_i}. The pre-gluing disk has total asymptotic length $(n + 1)(\beta_L + \beta_R) + \beta$, so the folded wormhole reduces to a Z-partition function disk of this circumference together with the $n + 1$ universal shell

factors. After normalising the shell states to unit norm, these factors cancel and

$$\tilde{Z}_{n+2} \;=\; \frac{\overline{Z}((n+1)(\beta_L+\beta_R)+\beta)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}}\,, \tag{2.3.3}$$

where each $\overline{Z}$ is the full sum over Z-partition saddles, not just the dominant one. The shell preparation temperatures enter only through the normalisation in the denominator. Equation (2.3.3) is exact within the saddlepoint approximation: the right-hand side is itself a coarse-grained quantity computed by the gravitational path integral.

2.3.2 Resummation in the planar limit

The folded wormhole is one saddle topology among many. In addition to it, $\overline{G^n_{ij}\langle\beta|i\rangle\langle j|\beta\rangle}$ receives contributions from saddles in which the bulk is partially connected — two tadpoles joined to one another but not to the chain, or a chain split into multiple bulk components, and so on. We organise these contributions by working in the *planar limit* [6, 13]

$$\kappa,\, e^{1/G_N} \;\to\; \infty\,, \qquad \kappa\, e^{-1/G_N} \;\sim\; \mathcal{O}(1)\,, \tag{2.3.4}$$

in which planar fully-connected and planar partially-connected geometries contribute at leading order in $1/\kappa$ and e^{-1/G_N}. As a representative example, $\overline{\mathrm{Tr}\,G^2} = \kappa^2\,\overline{Z}_2 + \kappa\,\overline{Z}_1^2$, where the disconnected piece costs a factor of $1/\kappa$ from breaking an index loop and the higher-topology piece Z_2 costs a factor of e^{-1/G_N}. Non-planar geometries are subleading: they require breaking at least two index loops and are dominated by a planar geometry with the same boundary connectivity (see Sec. 5.3 of [44]).

In this regime the sum over topologies for $\overline{G^n_{ij}\langle\beta|i\rangle\langle j|\beta\rangle}$ admits a Schwinger-Dyson resummation [6, 13, 29, 44] that re-organises the full planar sum in terms of the maximally-connected pieces $\tilde{Z}_{n+2}$ alone. Define the resolvent matrix, a generating function of the moments of G, as

$$R_{ij}(\lambda) \;\equiv\; \frac{\delta_{ij}}{\lambda} + \sum_{n=1}^{\infty} \frac{G^n_{ij}}{\lambda^{n+1}}\,, \tag{2.3.5}$$

so that

$$\overline{G^n_{ij}\,\langle\beta|i\rangle\langle j|\beta\rangle} \;=\; \oint_{c_0} \frac{d\lambda}{2\pi i}\,\lambda^n\,\overline{R_{ij}(\lambda)\,\langle\beta|i\rangle\langle j|\beta\rangle}\,, \tag{2.3.6}$$

where c_0 encircles the spectrum of the resolvent (see [44] for details on the contour). Diagrammatically, the two tadpole shells must be matched against shells of the resolvent (Fig. 2.3), and the planar combinatorics of this matching organise the sum over

partially-connected planar saddles into geometric powers of R (Fig. 2.4):

$$\overline{R_{ij}(\lambda)\,\langle\beta|i\rangle\langle j|\beta\rangle} \;=\; \sum_{n=0}^{\infty} R^{n+1}\,\tilde{Z}_{n+2} \;=\; \sum_{n=0}^{\infty} R^{n+1}\,\frac{\overline{Z}((n+1)(\beta_L+\beta_R)+\beta)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}}\,, \tag{2.3.7}$$

where $R \equiv \mathrm{Tr}\, R_{ij}$ is the trace of the resolvent. The resummation is possible because the action (2.3.3) of the maximally-connected saddle is independent of the masses of the shells in the index loop — the universal shell factors having cancelled against the normalisation.

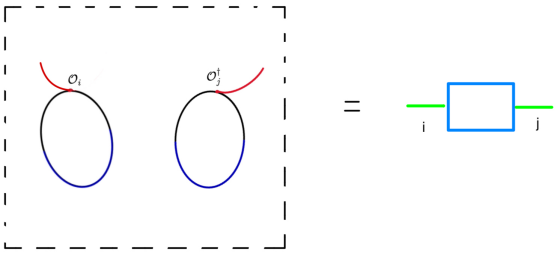


Figure 2.3: Graphical representation of the two tadpole boundaries $\langle\beta|i\rangle\langle j|\beta\rangle$ in (2.3.6).

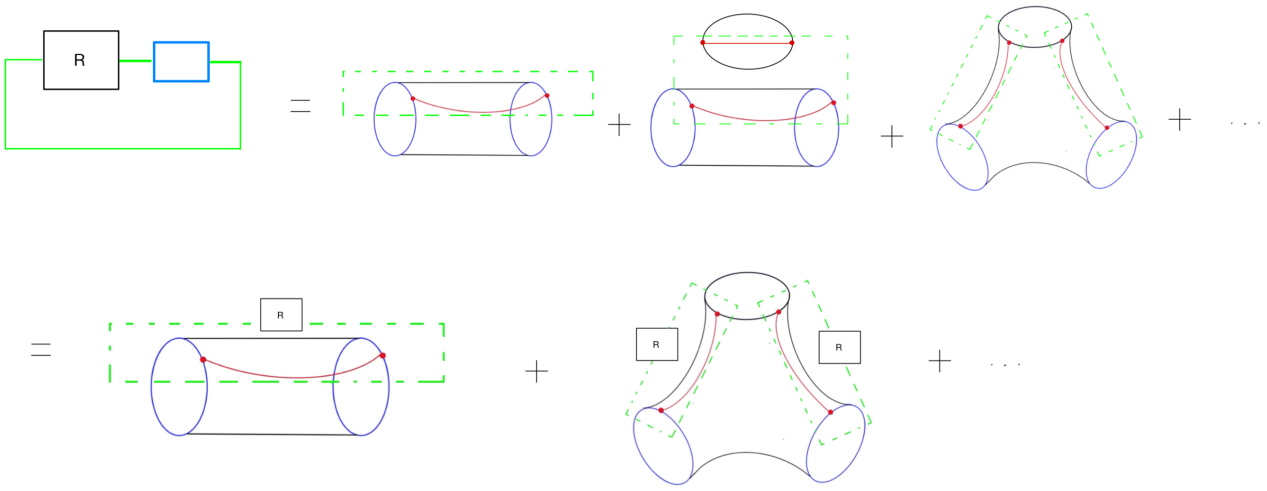


Figure 2.4: Schematic of the Schwinger-Dyson resummation (2.3.7). The planar sum over partially-connected saddles is reorganised by which shell-overlap boundaries are connected to the two tadpole shells. Insertions of R correspond to the trace of the shell-state resolvent.

The geometric series (2.3.7) cannot be evaluated in closed form: the dominant Z-partition function saddle in the numerator $\overline{Z}((n+1)(\beta_L+\beta_R)+\beta)$ shifts as n varies

through the Hawking-Page transition, so even within the leading-saddle approximation the summand has a piecewise structure that resists analytic resummation. The way forward is to project the equality (2.3.2) onto a microcanonical band, in which $\overline{Z}$ collapses to a single exponential and the sum can be performed.

2.3.3 The microcanonical equality and the Hilbert-space dimension

We project both sides of (2.3.2) onto microcanonical bands of width $\Delta E_{L,R}$ centred on energies E_L, E_R. Microcanonical shell states are obtained by Laplace transform of the canonical shell states [29],

$$|i, \beta_{L,R}, E_{L,R}\rangle \;=\; \int_{E_{L,R}-\Delta E_{L,R}/2}^{E_{L,R}+\Delta E_{L,R}/2} d\tilde{E}^{L,R}\, e^{-\beta_{L,R}\tilde{E}^{L,R}/2} \int d\tilde{\beta}_{L,R}\, e^{\tilde{E}^{L,R}\tilde{\beta}_{L,R}/2}\, |i, \tilde{\beta}_{L,R}\rangle \;. \tag{2.3.8}$$

Since $|\beta\rangle$ is diagonal in the energy basis, its microcanonical projection enforces $E_L = E_R \equiv E$. Within the saddlepoint approximation, $\overline{Z}(\beta)|_E = e^{-\beta E+\mathbf{S}(E)}$, where $\mathbf{S}(E)$ is the Bekenstein-Hawking entropy of the dominant Z-saddle at energy E (or $\mathbf{S}(E) = 0$ below the black hole threshold, where thermal AdS dominates). We denote the microcanonical resolvent as $R_E(\lambda)$. Substituting into (2.3.7) collapses the geometric series to a closed form:

$$\sum_{n=0}^{\infty} R_E^{n+1}\, \frac{\overline{Z}((n+1)(\beta_L+\beta_R)+\beta)\,|_E}{\overline{Z}(\beta_L)^{n+1}|_E\, \overline{Z}(\beta_R)^{n+1}|_E} \;=\; e^{-\beta E+\mathbf{S}(E)}\, \frac{R_E\, e^{-2\mathbf{S}(E)}}{1-R_E\, e^{-2\mathbf{S}(E)}} \;, \tag{2.3.9}$$

the canonical $(n+1)(\beta_L+\beta_R)$ dependence of the numerator having cancelled against that of the denominator term by term. Substituting back into (2.3.6) and (2.3.2),

$$\overline{\langle\beta|\Pi_{\mathcal{H}_{shell,E}}|\beta\rangle}|_E \;=\; e^{-\beta E+\mathbf{S}(E)} \lim_{n\to -1} \oint_{c_0} \frac{d\lambda}{2\pi i}\, \lambda^n\, \frac{R_E(\lambda)\, e^{-2\mathbf{S}(E)}}{1-R_E(\lambda)\, e^{-2\mathbf{S}(E)}} \;. \tag{2.3.10}$$

The analytic structure of $R_E(\lambda)$ was worked out in detail in [29]. The relevant feature is that once $\kappa \geq e^{\mathbf{S}_L(E)+\mathbf{S}_R(E)}$, the resolvent develops an isolated pole at $\lambda = 0$, and the contour c_0 can be deformed to a contour c_2 encircling the origin clockwise. The residue of the integrand at $\lambda = 0$ is exactly 1, so

$$\lim_{n\to -1} \oint_{c_2} \frac{d\lambda}{2\pi i}\, \lambda^n\, \frac{R_E(\lambda)\, e^{-2\mathbf{S}(E)}}{1-R_E(\lambda)\, e^{-2\mathbf{S}(E)}} \;=\; 1 \;, \tag{2.3.11}$$

and we obtain

$$\overline{\langle\beta|\Pi_{\mathcal{H}_{shell,E}}|\beta\rangle}|_E \;=\; e^{-\beta E+\mathbf{S}(E)} \;=\; \overline{Z(\beta)}|_E \;=\; \overline{\langle\beta|\beta\rangle}|_E \;. \tag{2.3.12}$$

Below the threshold $\kappa < e^{\mathbf{S}_L(E)+\mathbf{S}_R(E)}$ the resolvent does not develop the pole, the contour deformation fails, and the projector $\Pi_{\mathcal{H}_{shell,E}}$ does not act as the identity on the microcanonical band; the analysis there requires a detailed knowledge of the regular part of $R_E(\lambda)$. The insertion of $G^n_{ij}|i\rangle\langle j|$ followed by analytic continuation $n \to -1$ is therefore a non-trivial operation, in contrast to a simpler insertion such as $e^{-\hat{\mathcal{O}}/a}$ followed by the limit $a \to \infty$.

We have found that for any $\kappa \geq e^{\mathbf{S}_L(E)+\mathbf{S}_R(E)}$, the shell states resolve the identity on the diagonal microcanonical subspaces centered around $E_L = E_R = E$, as (2.3.1) holds at the saddlepoint level after such a microcanonical projection. As a corollary, the dimension of this microcanonical Hilbert sub-space coincides with the trace of this projector on $\mathcal{H}_{shell,E}$, giving

$$\dim \mathcal{H}_{LR,E} \;=\; e^{\mathbf{S}(E_L=E)+\mathbf{S}(E_R=E)}\,, \tag{2.3.13}$$

where for general unequal energies the entropy factorises as written. This puts the state counting of [6,19,28,32] on firmer footing, showing it measures the dimension of the full microcanonical Hilbert space of $\mathcal{H}_{LR}$, rather than that of a restricted "shell superselection sector".

In summary: the shell states span $\mathcal{H}_{LR}$, yet carry subtle non-perturbative overlaps that make them overcomplete rather than orthogonal. A careful non-perturbative accounting of their span — via the resolvent argument above — yields precisely the Bekenstein-Hawking entropy as the dimension of the microcanonical Hilbert space. This provides a microscopic origin of that entropy, at least within the coarse-grained description accessible to the gravitational path integral. We return to the state-counting perspective in Chapter 4, where we aim to explain the universality of this entropy law.

Still, the equality (2.3.12) has been established only at the level of the gravitational path integral, and is therefore a coarse-grained statement. The following section shows how taking κ to infinity beyond the planar limit upgrades this to a fine-grained equality.

2.3.4 Beyond the planar limit: the canonical equality

The microcanonical analysis required projecting onto a fixed-energy band because the canonical sum (2.3.7) was not a geometric series. We now observe that taking κ to infinity *beyond* the planar limit removes the obstruction altogether: the sum over

partially-connected topologies trivialises, leaving only the maximally-connected piece, and the canonical equality (2.3.2) follows directly. Concretely, consider the limit

$$\kappa \;\to\; \infty\,, \qquad \kappa\, e^{-1/G_N} \;\to\; \infty\,. \tag{2.3.14}$$

Some interpretation of this limit is warranted. Microcanonically, sending $\kappa \to \infty$ goes far beyond what the resolvent argument of Sec. 2.3.3 required: the microcanonical Hilbert space has finite dimension $e^{\mathbf{S}_L(E)+\mathbf{S}_R(E)}$, and any κ exceeding this threshold already suffices for the projector to act as the identity on each energy band. In the canonical ensemble, however, the situation is qualitatively different: $\mathcal{H}_{LR}$ contains states at all energies and is therefore infinite-dimensional. A basis for it must itself have infinite cardinality, so the strict $\kappa \to \infty$ limit is not a pathological strengthening of the planar limit but the natural basis size in the canonical ensemble — any finite κ would underspan an infinite-dimensional space. The accompanying condition $\kappa\, e^{-1/G_N} \to \infty$ in (2.3.14) is what makes this limit useful in practice, additionally driving the partially-connected topologies of Sec. 2.3.2 to zero; we now show this in detail.

The boundary condition for $\overline{G_{ij}^n \langle\beta|i\rangle\langle j|\beta\rangle}$ contains a single *maximal index loop* closing the shell label around the entire chain, contributing a factor of κ. Any saddle that breaks this loop — by, for example, leaving one tadpole shell unconnected to the chain — forces an explicit identification of indices and loses one or more factors of κ relative to the maximally-connected saddle. Higher-topology contributions are themselves suppressed by powers of e^{-1/G_N}. Both kinds of less-than-fully-connected topology survive in the planar limit (2.3.4), where they contribute the various powers of R in (2.3.7). In the stronger limit (2.3.14), however, both are suppressed: only the folded wormhole survives in the saddle expansion, and the resolvent resummation of Sec. 2.3.2 reduces to its leading term:

$$\overline{G_{ij}^n \,\langle\beta|i\rangle\langle j|\beta\rangle} \;=\; \kappa^{n+1}\,\tilde{Z}_{n+2} \;=\; \kappa^{n+1}\,\frac{\overline{Z}((n+1)(\beta_L+\beta_R)+\beta)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}}\,. \tag{2.3.15}$$

The analytic continuation $n \to -1$ is now performed directly on (2.3.15). We treat the limits in the order: regulate the theory with a high-energy cutoff Λ, take $\kappa \to \infty$ at fixed Λ, perform the analytic continuation $n \to -1$, then remove Λ. Formally substituting $n = -1$ on the right-hand side gives

$$\overline{\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle} \;=\; \overline{Z(\beta)}\,, \tag{2.3.16}$$

recovering (2.3.2) directly in the canonical ensemble. The shell preparation temperatures $\beta_{L,R}$ have dropped out, as they must: the equality (2.3.1) cannot depend on

the choice of basis. Note that this argument uses no microcanonical projection, no Laplace transform, and no contour manipulation: the topology sum trivialises before any of these are needed. The trade-off is that the dimension count (2.3.13), which fell out of the microcanonical contour argument, is not directly accessible here.

Fine-grained extension. The canonical argument has the additional virtue of admitting a fine-grained extension via the squaring trick of Sec. 2.1. To upgrade (2.3.16) to the fine-grained statement

$$Z(\beta) \;=\; \langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle \;, \tag{2.3.17}$$

we need

$$\overline{\left(Z(\beta) - \langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle\right)^2} \;=\; 0 \;. \tag{2.3.18}$$

The expansion of the square has three terms. The cross term $\overline{Z(\beta)\,\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle}$ factorises trivially: the $Z(\beta)$ boundary contains no shell insertions and so cannot be connected to the rest of the path integral by any wormhole that requires shell stabilisation. By (2.3.16) this cross term equals $\overline{Z(\beta)} \times \overline{Z(\beta)}$. Similarly $\overline{Z(\beta)^2}$ factorises in the saddlepoint approximation[2] as $\overline{Z(\beta)} \times \overline{Z(\beta)}$. Equation (2.3.18) therefore reduces to

$$\overline{\left(\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle\right)^2} \;=\; \left(\overline{Z(\beta)}\right)^2 \;. \tag{2.3.19}$$

Expanding the projector,

$$\overline{\left(\langle\beta|\Pi_{\mathcal{H}_{shell}}|\beta\rangle\right)^2} \;=\; \lim_{n,m\to -1} \overline{G^n_{ij}\,\langle\beta|i\rangle\langle j|\beta\rangle\; G^m_{kl}\,\langle\beta|k\rangle\langle l|\beta\rangle} \;. \tag{2.3.20}$$

The boundary condition now contains *two* maximal index loops — one closing the shell label around the G^n chain, one around the G^m chain. In the limit (2.3.14) only saddles preserving both loops contribute. The two loops are independent: each must lie within a single connected component of the bulk, but those components need not be connected to one another. The only contribution that survives as $\kappa \to \infty$ is therefore the disconnected product of two folded wormholes, evaluating to $\overline{Z(\beta)} \times \overline{Z(\beta)}$ via (2.3.16) and confirming (2.3.19).

This establishes the *fine-grained* equality (2.3.17) at the saddlepoint level. The shell states resolve the identity on the diagonal subspace generated by $|\beta\rangle$ for any β, at any choice of shell preparation temperature $\beta_{L,R}$, as a fine-grained equality and

[2]For real positive β only disconnected saddles contribute. Constrained-instanton wormholes connecting the two $Z(\beta)$ boundaries with fixed length or fixed equal energies were identified in [45–47]; they do not contribute at the saddle level considered here, though it would be interesting to revisit them within the present framework.

not merely an average. The fine-grained extension required only that the $\kappa \to \infty$ limit force the two index loops to decouple — precisely the simplification that the planar-limit microcanonical argument of Sec. 2.3.3 cannot achieve, since it works at finite κ. The same observation, in the JT gravity setting, was the route by which [44] extended coarse-grained equalities to fine-grained statements there. Sec. 2.4 extends our argument to a broad class of states in $\mathcal{H}_{LR}$, providing strong evidence that the shell states span the full semiclassical Hilbert space.

Remark: small black hole instability. The folded wormhole (2.3.3) is built from saddles of the Z partition function. In the AdS case there are three such saddles — the large black hole, thermal AdS, and the small black hole — and the small black hole carries a negative mode that contributes a factor of i to its one-loop determinant [48]. This instability could in principle threaten the interpretation of $\overline{Z(\beta)}$ as a positive overlap. In the canonical ensemble the small black hole is never the dominant saddle, and after summing over saddles $\overline{Z(\beta)}$ remains positive; [48] also argues that a careful microcanonical treatment resolves the issue. We do not consider the negative mode further.

2.4 A semiclassically complete basis

The result of Sec. 2.3, combined with the factorisation $\mathcal{H}_{LR} = \mathcal{H}_L \otimes \mathcal{H}_R$ and the identification of the Hamiltonian as the generator of asymptotic boundary time evolution (Sec. 1.1.2), already implies a striking partial spanning statement: the state $|\beta\rangle$ is the thermofield double $\sum_n e^{-\beta E_n/2}|E_n\rangle_L|E_n\rangle_R$ of the gravity theory, and varying β shows the shell states span the entire diagonal subspace $\{|E_n\rangle_L|E_n\rangle_R\}$ of $\mathcal{H}_{LR}$. This is in itself surprising, since the shell states are the gravitational analogues of partially entangled thermal states (PETS),

$$|\mathcal{O}\rangle \;=\; \sum_{n,m} e^{-(\beta_L E_n + \beta_R E_m)/2}\, \mathcal{O}_{mn}\, |E_n\rangle_L |E_m\rangle_R \;, \tag{2.4.1}$$

which generically have non-zero support on many off-diagonal pairs $|E_n\rangle_L|E_m\rangle_R$ as well. To leverage this for the full spanning statement $\mathcal{H}_{shell} = \mathcal{H}_{LR}$, we now extend the identity-insertion argument of Sec. 2.3 from $|\beta\rangle$ to arbitrary states $|\Psi\rangle \in \mathcal{H}_{LR}$.

Concretely, $\mathcal{H}_{shell} = \mathcal{H}_{LR}$ follows from

$$\langle\Psi|\Psi\rangle \;=\; \langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle \tag{2.4.2}$$

for every $|\Psi\rangle \in \mathcal{H}_{LR}$, by Cauchy-Schwarz. By the path-integral construction of states (Sec. 1.1), any such $|\Psi\rangle$ is either of the form

$$|\Psi\rangle \;=\; |\, e^{-\tilde{\beta}_L H_L/2}\, \mathcal{O}_\Psi\, e^{-\tilde{\beta}_R H_R/2}\,\rangle\;, \tag{2.4.3}$$

with $\mathcal{O}_\Psi$ a (possibly non-local) operator and $\langle\Psi|\Psi\rangle \equiv \overline{Z}_{\mathcal{O}_\Psi}(\tilde{\beta}_L+\tilde{\beta}_R)$, or a superposition of such states. We will show (2.4.2) for any state of the form (2.4.3) whose norm $\overline{\langle\Psi|\Psi\rangle}$ admits a dominant saddlepoint; the conclusion then extends to superpositions of such states by linearity. If, additionally, every state in $\mathcal{H}_{LR}$ is a superposition of saddlepoint-dominated states of this form — a physically natural assumption, but one we do not prove — then $\mathcal{H}_{shell} = \mathcal{H}_{LR}$ as a fine-grained identity. The shell states are then analogous to coherent states: a non-orthogonal, overcomplete basis whose elements are each individually associated with a classical bulk geometry.

Coarse-grained span. We follow the same two-step strategy as in Sec. 2.3: first show

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle} \;=\; \overline{\langle\Psi|\Psi\rangle} \tag{2.4.4}$$

in the coarse-grained theory, then lift to (2.4.2) by the squaring trick of Sec. 2.1. Resolving the projector via the trace formula,

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle} \;=\; \lim_{n\to -1} \overline{G_{ij}^{n}\,\langle\Psi|i\rangle\langle j|\Psi\rangle}\;. \tag{2.4.5}$$

The boundary condition is the same shell-overlap chain as in Sec. 2.3.1, with the two tadpole boundaries replaced by "operator boundaries" carrying the $\mathcal{O}_\Psi$ insertions (Fig. 2.5). The single maximal index loop closing the shell labels around the chain is unchanged, so in the $\kappa \to \infty$ limit only fully-connected saddles contribute. We denote the sum of such saddles by $\hat{Z}_{n+2}[\Psi]$.

The $\hat{Z}_{n+2}[\Psi]$ saddles are constructed in direct analogy with the folded wormholes of Sec. 2.3.1. Cut the wormhole open along the $n+1$ shell worldvolumes to obtain a sheet diagram (Fig. 2.6, centre). This sheet is geometrically a portion of a saddle of the path integral with asymptotic boundary Σ_n carrying $\mathcal{O}_\Psi, \mathcal{O}_\Psi^\dagger$ insertions and $2(n+1)$ pairs of shell insertions, separated by propagation times ΔT_i on the boundary (Fig. 2.6, left). The wormhole is then recovered by gluing the sheet back to itself along the corresponding shell worldvolumes via the junction conditions, with the ΔT_i determined dynamically by the gluing.

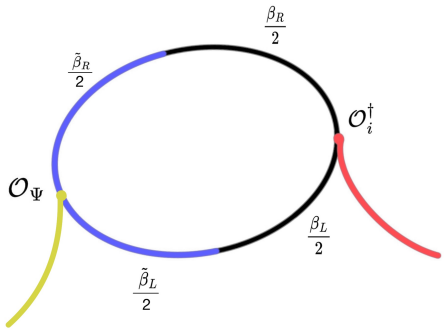


Figure 2.5: Operator boundary for $\langle i|\Psi\rangle$, with the $\mathcal{O}_\Psi$ insertion (whose stress-energy is depicted as the dotted yellow line) and $\mathcal{O}_i^\dagger$ separated by Euclidean times $(\beta_{L,R}+\tilde{\beta}_{L,R})/2$ on the two sides.

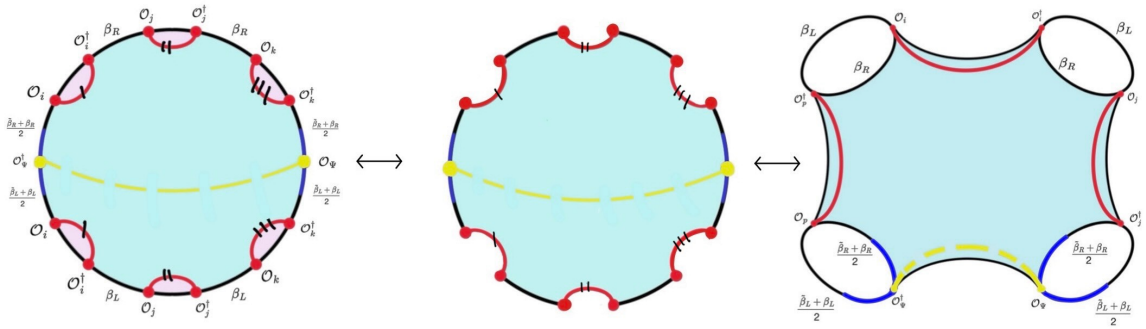


Figure 2.6: Construction of the $\hat{Z}_{n+2}[\Psi]$ wormhole saddles for $n = 2$. Starting from a disk saddle for the path integral with boundary Σ_n (left), the shell homology regions (purple) are discarded to produce a sheet (centre); the sheet is then glued back to itself along the corresponding shell worldvolumes to form the wormhole (right).

In the heavy shell mass limit (Sec. 1.2.1.4), the shell homology regions pinch off, the propagation times ΔT_i vanish, and the sheet grows to cover the entire disk. The dependence on the shell masses reduces to the universal factor $\prod_{i=1}^{n+1} Z_{m_i}$, which will be cancelled by the shell-state normalisations. The remaining bulk action is then exactly that of the path integral whose boundary is the $\mathcal{O}_\Psi$-insertion configuration with the shells removed — an elongated version of the boundary preparing $\langle\Psi|\Psi\rangle$ with total length $(n+1)(\beta_L+\beta_R)+\tilde{\beta}_L+\tilde{\beta}_R$. Each saddle of $\overline{Z}_{\mathcal{O}_\Psi}$ at this elongated length therefore produces, via the cut-and-glue construction, a saddle of $\hat{Z}_{n+2}[\Psi]$, and the sum over saddles on the two sides is in bijection.[3] Including the shell-state

[3]This bijection is exact for operators $\mathcal{O}_\Psi$ that do not back-react on the geometry, since the matter sector and the gravitational sector decouple. For back-reacting operators it amounts to assuming that the cut-and-glue construction continues to produce valid solutions of the equations of motion at the elongated boundary length, which we expect on physical grounds. The assumption is verified explicitly for the specific states considered in Sec. 2.4.1.

normalisations,

$$\hat{Z}_{n+2}[\Psi] \;=\; \frac{\overline{Z}_{\mathcal{O}_\Psi}\Big((n+1)(\beta_L+\beta_R)+\tilde{\beta}_L+\tilde{\beta}_R\Big)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}}\,, \tag{2.4.6}$$

where the Z_{m_i} factors have cancelled against the shell-state normalisations as in Sec. 2.3.1. Substituting (2.4.6) into (2.4.5) and using $\overline{G_{ij}^n\langle\Psi|i\rangle\langle j|\Psi\rangle} = \kappa^{n+1}\hat{Z}_{n+2}[\Psi]$,

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle} \;=\; \lim_{n\to-1}\kappa^{n+1}\,\frac{\overline{Z}_{\mathcal{O}_\Psi}\Big((n+1)(\beta_L+\beta_R)+\tilde{\beta}_L+\tilde{\beta}_R\Big)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}} \;=\; \overline{Z}_{\mathcal{O}_\Psi}(\tilde{\beta}_L+\tilde{\beta}_R) \;=\; \overline{\langle\Psi|\Psi\rangle}\,, \tag{2.4.7}$$

the coarse-grained version (2.4.4) of (2.4.2). The shell preparation temperatures $\beta_{L,R}$ have dropped out, as required by basis-independence: a different choice of shell preparation gives a different basis for $\mathcal{H}_{shell}$ but the same projector identity.

Fine-grained extension. The squaring trick of Sec. 2.1 lifts (2.4.4) to (2.4.2) provided

$$\overline{\big(\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle - \langle\Psi|\Psi\rangle\big)^2} \;=\; 0\,. \tag{2.4.8}$$

The expansion of the square contains two cross terms $\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle\,\langle\Psi|\Psi\rangle}$, the square $\overline{(\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle)^2}$, and the no-insertion piece $\overline{\langle\Psi|\Psi\rangle^2}$. We show that, in the $\kappa\to\infty$ limit, the first three each reduce to $\overline{\langle\Psi|\Psi\rangle^2}$, so the squared difference vanishes term by term.

Consider first the cross term. Resolving the identity insertion as in (2.4.5),

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle\,\langle\Psi|\Psi\rangle} \;=\; \lim_{n\to-1}\overline{G_{ij}^n\,\langle\Psi|i\rangle\langle j|\Psi\rangle\,\langle\Psi|\Psi\rangle}\,. \tag{2.4.9}$$

The boundary condition has a single maximal index loop closing the shell labels around the chain, exactly as in (2.4.5), plus a disjoint pair of operator boundaries from the extra $\langle\Psi|\Psi\rangle$ factor. In the $\kappa\to\infty$ limit no saddle may break this index loop, so contributing saddles are those in which all shell-carrying boundaries are joined in a single bulk component. Such saddles split into a *disconnected* sector, in which the extra $\langle\Psi|\Psi\rangle$ factor lives on its own bulk component, and a *connected* sector, in which a further wormhole links the extra $\langle\Psi|\Psi\rangle$ factor to the shell-carrying component.

The disconnected sector factorises into the $\hat{Z}_{n+2}[\Psi]$ wormhole of (2.4.6) and a separate $\overline{Z}_{\mathcal{O}_\Psi}(\tilde{\beta}_L+\tilde{\beta}_R)$; multiplying by κ^{n+1} and taking $n\to-1$ gives $\overline{\langle\Psi|\Psi\rangle}^2$. The connected sector is constructed by the same cut-and-glue procedure as in the coarse-grained argument, but now also joining the extra $\langle\Psi|\Psi\rangle$ boundary to the boundary Σ_n via a further wormhole (Fig. 2.7a). The shell worldvolumes are identified by the

gluing (Fig. 2.7b), the homology regions pinch off in the heavy shell mass limit, and the universal shell factors cancel against the normalisations. What remains is a path integral on the topology of Fig. 2.7c: an elongated $\langle\Psi|\Psi\rangle^2$ boundary condition, with the $\langle\Psi|i\rangle\langle j|\Psi\rangle$ side stretched by $(n+1)\beta_{L,R}$ relative to the unmodified $\langle\Psi|\Psi\rangle$ side. Denoting this path integral $Z_{\Psi\tilde{\Psi}}((n+1)\beta_L,(n+1)\beta_R)$,

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle\,\langle\Psi|\Psi\rangle}|_{\text{connected}} \;=\; \lim_{n\to-1}\kappa^{n+1}\,\frac{Z_{\Psi\tilde{\Psi}}((n+1)\beta_L,(n+1)\beta_R)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}} \;=\; Z_{\Psi\tilde{\Psi}}(0,0)\;, \tag{2.4.10}$$

where the second equality follows by the same bijection between saddles at elongated and unmodified boundary lengths that was used in (2.4.7). But $Z_{\Psi\tilde{\Psi}}(0,0)$ is precisely the connected wormhole contribution to $\overline{\langle\Psi|\Psi\rangle^2}$ at the original boundary length. The disconnected and connected pieces of the cross term therefore sum to $\overline{\langle\Psi|\Psi\rangle^2}$ in full.

In short, every wormhole saddle contributing to $\overline{\langle\Psi|\Psi\rangle^2}$ induces, via the cut-and-glue construction at elongated boundary length, a corresponding fully-connected saddle of the cross term, and conversely; the two sums are therefore in one-to-one correspondence and equal term by term.

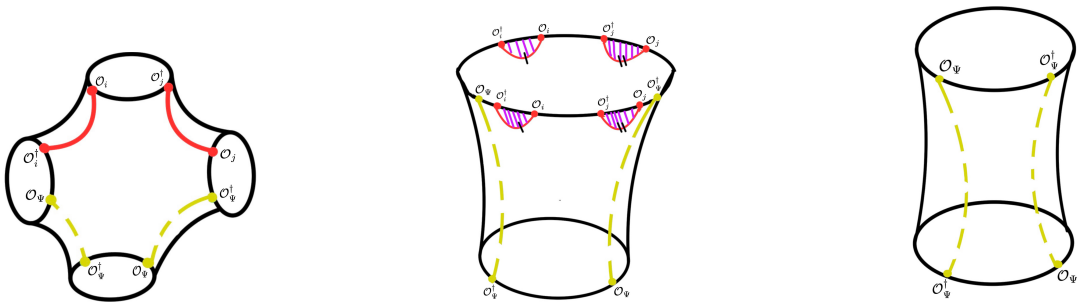
The same argument applies to $\overline{(\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle)^2}$. The boundary now has *two* maximal index loops, one from each $\Pi_{\mathcal{H}_{shell}}$ insertion. The $\kappa\to\infty$ limit forces each loop to lie in a connected bulk component, and the contributing saddles again split into disconnected (giving $\hat{Z}_{n+2}[\Psi]^2 \to \overline{\langle\Psi|\Psi\rangle}^2$) and connected (giving the cut-and-glue wormhole, which reduces by the same argument to $Z_{\Psi\tilde{\Psi}}(0,0)$). The sum is again $\overline{\langle\Psi|\Psi\rangle^2}$ in full.

The squared difference (2.4.8) therefore vanishes, and subject to the bijection assumption stated above,

$$\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle \;=\; \langle\Psi|\Psi\rangle \tag{2.4.11}$$

holds at the saddlepoint level. The same logic applied to bilinears in distinct states gives $\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Phi\rangle = \langle\Psi|\Phi\rangle$ for any $|\Psi\rangle, |\Phi\rangle$ in the class considered.

Conclusion. We have shown $\mathcal{H}_{shell} = \mathcal{H}_{LR}$ as a fine-grained equality, subject to the saddlepoint and bijection assumptions stated above. The result holds for any preparation temperature $\beta_{L,R}$, since the shell-state normalisations cancel the dependence on $\beta_{L,R}$ in the cut-and-glue construction. The argument extends to all bilinears $\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Phi\rangle$, not only norms, and so identifies $\Pi_{\mathcal{H}_{shell}}$ with the identity operator on $\mathcal{H}_{LR}$ at the saddlepoint level.



(a) (b) (c)

Figure 2.7: (**a**) The connected contribution to $\overline{G_{ij}^n \langle\Psi|i\rangle\langle j|\Psi\rangle \langle\Psi|\Psi\rangle}$ for $n = 1$. (**b**) Construction by gluing along the shell worldvolumes. (**c**) The $n \to -1$ limit recovers the connected wormhole contribution to $\overline{\langle\Psi|\Psi\rangle^2}$ at the original boundary length. Red denotes shell worldvolumes; chartreuse denotes $\mathcal{O}_\Psi$ matter sources.

Extension to perturbative corrections. The above argument operates at level (I) of the classification of Sec. 2.1. It extends to level (II) provided the bijection assumption also holds for the perturbative corrections around each saddle. In the heavy shell limit the shells become classical and their trajectories localise to the asymptotic boundary and the corrected interior bulk action equals that of the corresponding perturbative correction to the dominant saddle of $\overline{\langle\Psi|\Psi\rangle}$, plus a universal shell contribution that cancels against the normalisation. For non-back-reacting operators this is again exact; for back-reacting operators it is subject to the same caveat as at level (I). We do not extend the argument to level (III): topologies that admit no saddle, or off-shell metrics not connected to a saddle by perturbative deformation, can in principle contribute to one side of (2.4.11) without a matching contribution to the other.

After this work was completed we became aware of [49, 50], which use a similar argument in JT gravity to show that any microcanonical gravity state can be reconstructed from a basis of $e^{\mathbf{S}}$ states. The present analysis can be viewed as a higher-dimensional realisation, and a fine-grained extension, of that argument.

2.4.1 Examples

The examples below verify the bijection assumption explicitly for two representative classes of states, beyond the $|\beta\rangle$ case of Sec. 2.3. Together with the $|\beta\rangle$ result they cover the principal families of states arising in the gravitational path integral, and we take them as strong evidence that the assumption holds in general.

A shell state. Take $|\Psi\rangle$ to be a shell state created with shell operator $\mathcal{O}_S$ at preparation temperatures $\tilde{\beta}_L, \tilde{\beta}_R$, with $\mathcal{O}_S$ chosen distinct from the operators defining the basis (or, alternatively, charged under a global symmetry under which the basis is neutral). In the heavy shell limit the norm is $\overline{\langle\Psi|\Psi\rangle} = \overline{Z}(\tilde{\beta}_L)\,\overline{Z}(\tilde{\beta}_R)\,Z_{m_S}$. The path integral $\overline{G^n_{ij}\langle\Psi|i\rangle\langle j|\Psi\rangle}$ is computed by the pinwheel of Fig. 1.9 with $n+2$ asymptotic boundaries; adapting (1.2.18),

$$\overline{G^n_{ij}\,\langle\Psi|i\rangle\langle j|\Psi\rangle} \;=\; Z_{m_S}\,\frac{\overline{Z}\Big((n+1)\beta_L+\tilde{\beta}_L\Big)\;\overline{Z}\Big((n+1)\beta_R+\tilde{\beta}_R\Big)}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{n+1}}\;, \tag{2.4.12}$$

which limits to $\overline{\langle\Psi|\Psi\rangle}$ as $n \to -1$. The saddle correspondence here is exact by construction: the $\mathcal{O}_S$ shell and the basis shells propagate independently through the bulk, so the bijection between $\hat{Z}_{n+2}[\Psi]$ saddles and elongated-disk saddles holds with no additional assumptions. The cross term required for the fine-grained extension is computed by an $(n+3)$-boundary pinwheel and gives

$$\overline{\langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle\,\langle\Psi|\Psi\rangle} \;=\; Z^2_{m_S}\,\overline{Z}(2\tilde{\beta}_L)\,\overline{Z}(2\tilde{\beta}_R)\;, \tag{2.4.13}$$

which is precisely the cylinder-topology connected contribution to $\overline{\langle\Psi|\Psi\rangle\,\langle\Psi|\Psi\rangle}$ (Fig. 1.8). The other cross term works analogously, and we conclude $\langle\Psi|\Psi\rangle = \langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle$ at the fine-grained level.

When $\mathcal{O}_S$ carries a global charge, this argument shows that the symmetry is broken by non-perturbative wormhole contributions, in line with [49]. This therefore provides a concrete higher-dimensional realization of the folklore that quantum gravity breaks bulk global symmetries.

A light operator. Take $\mathcal{O}_\Psi$ heavy enough for the geodesic approximation (see, for example, [51, 52]) to be valid but light enough not to back-react on the geometry significantly. The norm takes the schematic form

$$\overline{\langle\Psi|\Psi\rangle} \;=\; \sum_X \overline{Z}_X(\tilde{\beta}_L+\tilde{\beta}_R)\sum_i e^{-m_\Psi L^i_X(\tilde{\beta}_L,\tilde{\beta}_R)}\;, \tag{2.4.14}$$

where the outer sum runs over Z-saddles X and the inner sum runs over windings i of the bulk geodesic connecting the operator insertions on background X (see e.g. [31]). Since $\mathcal{O}_\Psi$ does not back-react, the connected contribution to $\overline{G^n_{ij}\langle\Psi|i\rangle\langle j|\Psi\rangle}$ consists of Ψ-matter geodesics propagating on a folded wormhole background, which is itself a Z-saddle on a longer boundary:

$$\overline{G^n_{ij}\,\langle\Psi|i\rangle\langle j|\Psi\rangle} \;=\; \sum_X \overline{Z}_X(L_{tot})\sum_i e^{-m_\Psi L^i_X(L^{tot}_L,\,L^{tot}_R)}\;, \tag{2.4.15}$$

with $L^{tot}_{L,R} \equiv (n+1)\beta_{L,R} + \tilde{\beta}_{L,R}$ and $L_{tot} = L^{tot}_L + L^{tot}_R$. The bijection between saddles on either side is explicit: each Z-saddle X on the left corresponds to the same saddle X at elongated boundary length on the right, and the geodesic windings are in one-to-one correspondence because the matter sector decouples from the gravitational one. The $n \to -1$ limit then recovers (2.4.14) directly. The fine-grained extension is immediate: $\mathcal{O}_\Psi$ does not stabilise additional connected wormhole contributions to the cross terms, so each cross term factorises without further assumption, and we obtain $\langle\Psi|\Psi\rangle = \langle\Psi|\Pi_{\mathcal{H}_{shell}}|\Psi\rangle$ at the fine-grained level.

Note that the one-to-one correspondence relies on summing over all geodesic windings on both sides: had we retained only the dominant winding, the two sides would not have matched, and the equality (2.4.11) would have failed. This is a further illustration of the non-triviality of the shell projector insertion: the analytic continuation $n \to -1$ probes the full saddle structure of the norm, not merely its leading contribution.

These two examples, together with the $|\beta\rangle$ case of Sec. 2.3, establish (2.4.11) for states representing black holes in thermal equilibrium, arbitrary shell excitations above them, and probe-matter insertions of any mass in the geodesic regime. The bijection assumption is seen to follow automatically whenever $\mathcal{O}_\Psi$ and the basis shells decouple in the bulk; for genuinely back-reacting operators it remains an assumption, though one consistent with the general structure of the cut-and-glue construction.

2.5 Discussion

The arguments of this chapter rely only on the rules of the gravitational path integral and the Einstein-Hilbert action, treated as a universal effective field theory of gravity; no holographic duality is assumed. We phrased the discussion in terms of asymptotically AdS gravity for concreteness, but the main results carry over unchanged to the asymptotically flat case. The specific saddlepoint calculations rely on the explicit form of the action, but the general method — resolution of the identity in the shell basis at large κ combined with the squaring trick — requires only that the action consist of local bulk and boundary terms. The main results are therefore expected to be insensitive to higher-derivative corrections to the Einstein-Hilbert action.

2.5.1 Macroscopic superpositions of semiclassical geometries

A striking consequence of the spanning result of Sec. 2.4 is that the canonical shell states resolve the identity at *any* preparation temperature β_L, β_R. The basis change

between different choices of preparation temperature,

$$|i, \beta_L, \beta_R\rangle = \sum_j a_j(i, \beta_L, \beta_R)\, |j, \tilde{\beta}_L, \tilde{\beta}_R\rangle \ , \tag{2.5.1}$$

is therefore in general non-trivial. Recall from Sec. 1.2.1.5 that the norm of a shell state is $\overline{\langle i|i\rangle} = \overline{Z}(\beta_L)\,\overline{Z}(\beta_R)\,Z_{m_i}$, so each shell state is associated with up to nine possible saddle geometries: three for each of $Z(\beta_L)$ and $Z(\beta_R)$ (large black hole, small black hole, thermal AdS). For any given β_L, β_R one of these dominates, and in the strict $G_N \to 0$ limit only this one survives. In the perturbative limit we may therefore associate a single semiclassical geometry to each canonical shell state, but in general a shell state should be thought of as a quantum superposition of saddlepoint configurations with and without horizons. The mixing coefficients a_j in (2.5.1) are nonzero precisely because of $\mathcal{O}(e^{-1/G_N})$ wormhole corrections to the shell overlaps; in the strict $G_N \to 0$ limit shell states at different preparation temperatures become linearly independent.

Of particular interest are the type-2 shell states with both preparation temperatures below the black hole threshold, $\beta_{L,R} > 2\pi\ell/\sqrt{2}$ (where ℓ is the AdS radius): these are associated with horizonless thermal AdS saddles, with a disconnected closed universe behind the shell. Combining (2.5.1) with the canonical-to-microcanonical Laplace transform

$$|i\rangle_{E_L,E_R} \propto \int d\beta_L\, d\beta_R\, e^{(\beta_L E_L + \beta_R E_R)/2} \sqrt{Z(\beta_L)\, Z(\beta_R)}\, |i, \beta_L, \beta_R\rangle \ , \tag{2.5.2}$$

a microcanonical black hole state can be expanded as a superposition of type-2 shell states at any fixed preparation temperature:

$$|i\rangle_{E_L,E_R} \propto \sum_j b_j(i, E_L, E_R)\, |j, \tilde{\beta}_L, \tilde{\beta}_R\rangle \ . \tag{2.5.3}$$

By choosing $\tilde{\beta}_L, \tilde{\beta}_R > 2\pi\ell/\sqrt{2}$ in (2.5.3), a microcanonical black hole state — whose own dominant saddle has a horizon — is written as a superposition of type-2 geometries with *no* horizons but containing disconnected closed universes (Fig. 1.7b).

This conclusion appears striking. However, observers inside the compact disconnected universe in such a configuration cannot send classical signals to an asymptotic observer, much like observers inside a black hole. Unlike a black hole, the asymptotic observer cannot send a classical probe into the closed universe either, but the Euclidean state preparation entangles the closed universe with the asymptotic component, so observers in either region might use the entanglement to send signals to each other — in the spirit of recent proposals that the black hole interior lies within the

entanglement wedge of the distant Hawking radiation [13, 14]. See [31] for a detailed discussion of the encoding of compact universes into the boundary theory.

Geometry must be emergent. A direct consequence of (2.5.3) is that bulk geometry cannot be a linear quantum observable. Suppose a state $|\Psi\rangle$ has a norm dominated by a single saddle with metric $\mathbf{g}_\Psi$, so we may speak of $\mathbf{g}_\Psi$ as the semiclassical geometry of $|\Psi\rangle$ and write $|\Psi\rangle = |\mathbf{g}_\Psi\rangle$. If $\mathbf{g}_\Psi$ is to be a well-defined quantum observable, it should correspond to the eigenvalue of some Hermitian "geometry operator" $\mathcal{G}$ [53]:

$$\mathcal{G}\,|\mathbf{g}_\Psi\rangle = \mathbf{g}_\Psi\,|\mathbf{g}_\Psi\rangle\ . \tag{2.5.4}$$

Since the type-2 shell states with $\tilde{\beta}_{L,R} > 2\pi\ell/\sqrt{2}$ form a basis of $\mathcal{H}_{LR}$ and their norms are dominated by a horizonless saddle, this basis would furnish a complete set of eigenvectors of $\mathcal{G}$. But the microcanonical black hole states also have a single dominant saddle (with horizons distinct from any type-2 saddle), and (2.5.3) expands them as superpositions of type-2 states. It is therefore impossible for $\mathcal{G}$ to be Hermitian with respect to the non-perturbative inner product. A version of this claim was made in [6] for the wormhole length: long type-1 wormholes can be expanded as superpositions of short ones, so there is no linear operator that extracts wormhole length. The non-existence of a linear geometry operator is consistent with the conjecture that bulk geometry emerges from quantum entanglement [42, 43] in AdS/CFT, since entanglement itself is not measurable by a linear operator.

In the $G_N \to 0$ limit, shell states at different temperatures are linearly independent and (2.5.4) can be Hermitian with respect to the perturbative inner product. This matches the AdS/CFT expectation that the semiclassical gravity description emerges in the strict $N \to \infty$ (equivalently $G_N \to 0$) limit. Related ideas have been developed from the perspectives of state-dependent black-hole interior reconstruction [54, 55], non-isometry in the holographic code [56, 57], the non-existence of a topology operator [58], the exterior description of a black hole [59], and the algebraic description of the semiclassical limit [60–64].

2.5.2 Other possible bases and the spectral form factor

By the argument of Sec. 2.4, many types of sets of κ gravitational states satisfying $\overline{\langle i|j\rangle} \propto \delta_{ij}$ and $\overline{\langle i|j\rangle\langle j|i\rangle} \neq 0$ for distinct $|i\rangle, |j\rangle$ should form a basis of $\mathcal{H}_{LR}$ as $\kappa \to \infty$. The shell states are useful because $\overline{\langle i|j\rangle\langle j|i\rangle} \neq 0$ already at the saddlepoint level, with explicitly constructible wormhole saddles. A potential alternative is the family

of time-evolved TFD states discussed in [65], obtained by Lorentzian time-evolving $|\beta\rangle$ by varying amounts. As pointed out in [66], showing that these states span $\mathcal{H}_{LR}$ would imply the saturation of the plateau of the gravitational spectral form factor; and [67] showed that in chaotic quantum systems the time-evolved TFD states do satisfy the orthogonality and overlap conditions above. If the corresponding wormhole amplitudes can be constructed, it would follow that the gravitational spectral form factor in any dimension has a late-time plateau. The relevant wormholes have not yet been constructed, however, and may not exist as saddles since there is no matter stress-energy supporting their topology.

"All things by immortal power, Near or far, Hiddenly, To each other linked are, That thou canst not stir a flower, Without troubling of a star."
— *Francis Thompson,* The Mistress of Vision

Chapter 3

Resolving the Factorisation Problem

The factorisation puzzle introduced in Sec. 1.1.3 asks whether the two-sided Hilbert space of quantum gravity factorises:

$$\mathcal{H}_{L\cup R} \;=\; \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R} \;=\; \mathcal{H}_{LR}\,. \tag{3.0.1}$$

Holography with two disconnected AdS boundaries requires a tensor product structure — the dual CFT lives independently on each boundary — but the bulk Euclidean path integral includes geometries in which the two boundaries are connected through the bulk, which seem to threaten this factorisation [68,69]. In this chapter we resolve the puzzle, showing that (3.0.1) holds at the level of the gravitational saddlepoint expansion. The argument uses only the rules of the gravitational path integral and the toolkit of Chapter 2; no holographic duality is assumed, and the conclusion applies equally to asymptotically flat and asymptotically AdS gravity.

Earlier work has made partial progress. In two-dimensional JT gravity, [44] showed that traces in the two-sided Hilbert space factorise to all orders in e^{-1/G_N}, and [29] extended the leading-saddle version of this argument to higher dimensions. More generally, [23] used operator-algebraic methods to show that $\mathcal{H}_{LR}$ must be a sum of

products of factors. None of these results identify the factors with the single-boundary Hilbert spaces $\mathcal{H}_{\mathcal{B}_{L,R}}$ themselves, which is the genuine content of (3.0.1). Closing this gap requires explicitly constructing bases of $\mathcal{H}_{LR}$ and of $\mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$ and showing that they coincide, which is the strategy adopted here.

Two clarifying remarks are in order. First, the factorisation puzzle is sometimes posed in Lorentzian signature, by asking how the effective field theory Hilbert space in the eternal black hole background can factorise as expected from AdS/CFT. This formulation is misleading: the code subspace around any single entangled state in a large factorised Hilbert space need not itself factorise. AdS/CFT requires (3.0.1) only of the full non-perturbative theory, and this is the statement we prove. Second, the factorisation of the Hilbert space (3.0.1) is distinct from the failure of *amplitude* factorisation: the gravitational path integral generically gives $\overline{\mathcal{A}\mathcal{A}} \neq (\overline{\mathcal{A}})^2$ because of wormhole contributions, an effect attributed to coarse-graining. The two notions are not in tension. On the contrary, we will show the same wormholes responsible for non-factorisation of coarse-grained amplitudes are precisely what enforce the fine-grained Hilbert-space factorisation.

Our strategy, illustrated in Fig. 3.1, proceeds in two steps. First, we establish that the single-boundary shell states defined in Sec. 1.2.2 form a semiclassically complete basis of $\mathcal{H}_{\mathcal{B}}$. The argument is a direct adaptation of the corresponding two-sided result in Chapter 2, and we record only the result and its consequences in Sec. 3.1. Second, in Sec. 3.2 we show that the basis of two-sided shell states for $\mathcal{H}_{LR}$ established in Chapter 2 and the basis of tensor-product single-boundary shell states for $\mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$ span the same space, by establishing both inclusions. The technical input is a class of mixed wormhole saddles with both single- and two-sided shell boundaries, whose action again simplifies in the heavy shell limit. Sec. 3.3 introduces an alternative basis of *two-shell* states which makes the factorisation manifest at the level of Lorentzian geometry: tuning the entanglement between the two factors yields a basis of states whose dominant saddles are disconnected. We conclude in Sec. 3.4 with a discussion of the implications for operator algebras and emergent geometry.

3.1 Single-boundary shell states span $\mathcal{H}_{\mathcal{B}}$

We denote a single-boundary shell state of mass m_i and preparation temperature β_R by $|i\rangle_R$, and write $\mathcal{H}_R \equiv \text{span}\{|i\rangle_R, i = 1, \ldots, \kappa_R\}$ for the closed span of κ_R such states with varying mass. The states $|i\rangle_R$ are non-orthogonal: their wormhole-stabilised

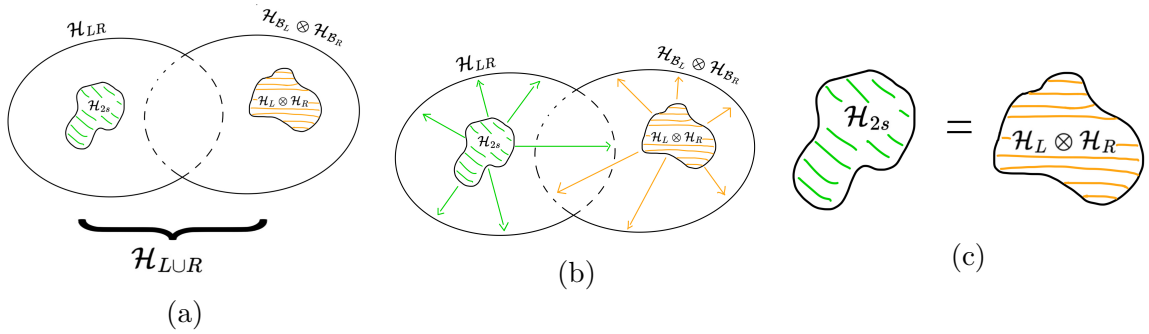


Figure 3.1: Resolving the factorisation puzzle. (**a**) The two-sided Hilbert space $\mathcal{H}_{LR}$, the tensor product Hilbert space $\mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$, and the total two-sided Hilbert space $\mathcal{H}_{L\cup R}$. The shaded regions $\mathcal{H}_{2s}$ and $\mathcal{H}_L \otimes \mathcal{H}_R$ denote the spans of two-sided and tensor-product single-sided shell bases respectively. (**b**) The two- and single-sided shell states span $\mathcal{H}_{LR}$ (Chapter 2) and $\mathcal{H}_{\mathcal{B}_{L,R}}$ (Sec. 3.1) respectively in the limit of infinitely many basis states. (**c**) Sec. 3.2 shows that the two-sided shell basis coincides with the tensor-product basis, establishing $\mathcal{H}_{LR} = \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$.

overlaps are precisely the kind of non-perturbative effects discussed in Sec. 1.1.1. Following the formalism of Chapter 2, we define the formal projector onto $\mathcal{H}_R$ via the Gram matrix $G_{ij} \equiv \langle i|j\rangle_R$,

$$\Pi_{\mathcal{H}_R} \;\equiv\; G^{-1}_{ij}\, |i\rangle_R \langle j|_R \;, \tag{3.1.1}$$

with G^{-1} understood by analytic continuation from positive integer powers of G (Sec. 2.2).

Spanning result. In the limit $\kappa_R \to \infty$ the projector (3.1.1) acts as the identity on the full single-boundary Hilbert space at the saddlepoint level:

$$\Pi_{\mathcal{H}_R} \;=\; \mathbb{1}_{\mathcal{H}_{\mathcal{B}}} \;, \qquad \text{equivalently} \qquad \mathcal{H}_R \;=\; \mathcal{H}_{\mathcal{B}} \;. \tag{3.1.2}$$

The proof is a direct adaptation of the two-sided spanning argument given in Sec. 2.4 of Chapter 2; we omit it here and refer the reader to [2] for the explicit construction. The two structural changes from the two-sided argument are as follows. First, the asymptotic boundary in the single-boundary case extends to infinite Euclidean time on either side, so the maximally-connected wormhole saddles contributing to $\overline{G^n_{ij}\,\langle\Psi|i\rangle\langle j|\Psi\rangle}$ are constructed by gluing $n+1$ shell strips into a central *strip* along the shell worldvolumes (Fig. 3.2), in place of the folded wormhole of Sec. 2.3.1 which was built from a self-folded disk. Second, the universal heavy-shell factor Z_{m_i} and the strip path integral $\overline{S}(0)$ both cancel against the shell-state normalisation, leaving the same combinatorial structure as in the two-sided case. The squaring trick of Sec. 2.1

then lifts the resulting coarse-grained equality to the fine-grained statement (3.1.2) in the limit $\kappa_R \to \infty$, exactly as in Sec. 2.3.4.

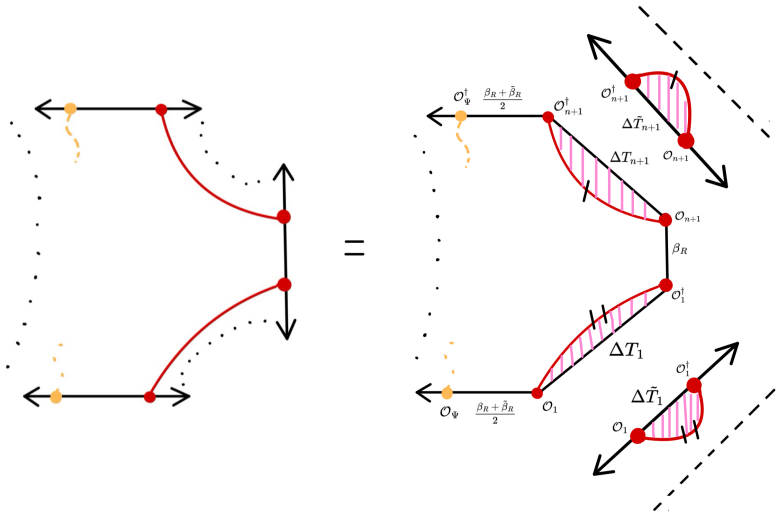


Figure 3.2: The fully connected wormhole saddle entering the single-boundary spanning argument, depicted for $n = 1$. It is constructed by gluing $n + 1$ shell strips into a central strip along the shell worldvolumes, replacing the folded wormhole of Sec. 2.3.1. See [2] for details.

Consequences. Two consequences of (3.1.2) are worth recording.

First, the spanning result holds for *any* preparation temperature β_R, since neither the structure of the strip wormhole saddle nor the squaring argument depends on the choice of β_R. Recall from Sec. 1.2.2 that the dominant saddle for the norm of $|i\rangle_R$ is a single-sided black hole with a shell behind the horizon (type A) when $\beta_R < \beta_{HP}$, and thermal AdS together with a disconnected compact "Big Crunch" universe (type B) when $\beta_R > \beta_{HP}$. We are therefore free to choose either family as a basis of $\mathcal{H}_\mathcal{B}$. In particular, *a single-sided black hole state can be expanded as a superposition of horizonless type-B states with a closed universe behind the shell, and vice versa.* As argued in Sec. 2.5 of Chapter 2 for the two-sided case, this is incompatible with the existence of a linear "geometry operator" on the non-perturbative gravitational Hilbert space: the same conclusion applies here, providing a single-boundary realisation of the principle that bulk geometry must be emergent.

Second, the spanning result may extend beyond the leading saddle approximation in the same way as its two-sided counterpart (Sec. 2.4). In the heavy-shell limit each shell trajectory localises to the asymptotic boundary, where it is unaffected by

perturbative bulk corrections; the cut-and-glue construction then produces, on each side of (3.1.2), the same perturbative correction to the underlying norm path integral. Topologies that admit no saddle, or off-shell metrics not connected to a saddle by perturbative deformation, lie beyond the scope of this argument. As discussed in [2], a surgery argument analogous to that of [44] heuristically extends the conclusion to the full path integral; we do not reproduce it here.

3.2 Factorisation of the two-sided Hilbert space

We have established two complementary spanning results: Chapter 2 showed that two-sided shell states span $\mathcal{H}_{LR}$, and Sec. 3.1 that single-sided shell states span $\mathcal{H}_{\mathcal{B}}$. Tensor products of single-sided bases therefore span $\mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}$, and the factorisation problem (3.0.1) reduces to

$$\mathcal{H}_{2s} \;=\; \mathcal{H}_L \otimes \mathcal{H}_R \,, \tag{3.2.1}$$

where $\mathcal{H}_{2s}$ denotes the span of two-sided shell states with preparation temperatures $\tilde{\beta}_{L,R}$ and $\mathcal{H}_L \otimes \mathcal{H}_R$ the span of tensor products $|i\rangle_L \otimes |j\rangle_R$ of single-sided shell states with preparation temperatures $\beta_{L,R}$. Each family of bases spans for any choice of preparation temperature, so the values of $\tilde{\beta}_{L,R}$ and $\beta_{L,R}$ are independent and we keep them general throughout. Two-sided shell states are denoted $|\mathbf{i}\rangle_{2s}$ in boldface, and the corresponding single-sided states as $|i\rangle_L$, $|j\rangle_R$.

Setup: tensor norms and mixed overlaps. Two preliminary path integrals fix conventions and establish that the two subspaces are non-trivially related.

The norm $\overline{\langle i|i\rangle_L \langle j|j\rangle_R}$ has two saddlepoint contributions: the disconnected product $\overline{\langle i|i\rangle_L}\,\overline{\langle j|j\rangle_R}$, and a connected wormhole (Fig. 3.3b) in which the L and R shells annihilate against one another. By the orthogonality argument of Sec. 1.2.1.1, the wormhole contributes only when $m_i = m_j$, giving

$$\overline{\langle i|i\rangle_L \langle j|j\rangle_R} \,/\, Z_{m_i} Z_{m_j} \;=\; \overline{Z}(\beta_L)\,\overline{Z}(\beta_R)\,\overline{S}(0)^2 \;+\; \delta_{ij}\,\overline{Z}(\beta_L+\beta_R)\,\overline{S}(0)^2 \,. \tag{3.2.2}$$

We choose the $\mathcal{H}_{L,R}$ bases to consist of shells of mutually distinct masses, so the second term vanishes and the tensor product norm factorises into single-sided norms.

The mixed overlap $\overline{\langle i|_L \langle j|_R |\mathbf{p}\rangle_{2s}}$ between a tensor product and a two-sided shell state (Fig. 3.4) is exponentially suppressed for our basis: the single-sided shells $\mathcal{O}_{i,j}$ and the two-sided shell $\mathcal{O}_{\mathbf{p}}$ can only be matched via $|m_{\mathbf{p}} - (m_i + m_j)|$ Planck-scale interactions. The squared overlap, however, has a universal connected wormhole

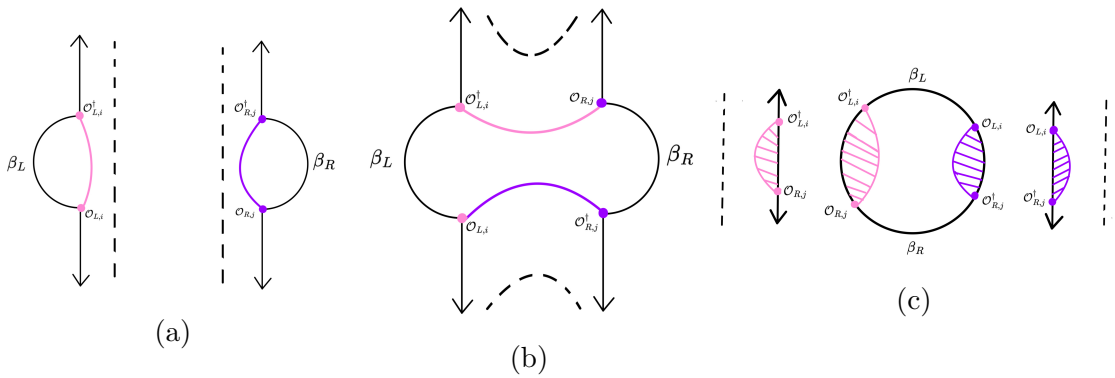


Figure 3.3: Saddles for $\overline{\langle i|i\rangle_L \langle j|j\rangle_R}$. (**a**) Disconnected contribution. (**b**) Connected wormhole, present only when $i = j$. (**c**) Construction of the wormhole by gluing L, R shell strips into a disk.

contribution (Fig. 3.5) in which the L- and R-disks are glued together along the $\mathcal{O}_{\mathbf{p}}$ insertion and shell strips are glued to the disks along the $\mathcal{O}_{i,j}$ insertions. In the heavy mass limit, after normalisation,

$$\overline{|\langle i|_L \langle j|_R |\mathbf{p}\rangle_{2s}|^2} \;=\; \frac{\overline{Z}(\beta_L + \tilde{\beta}_L)\,\overline{Z}(\beta_R + \tilde{\beta}_R)}{\overline{Z}(\beta_L)\,\overline{Z}(\tilde{\beta}_L)\,\overline{Z}(\beta_R)\,\overline{Z}(\tilde{\beta}_R)}\,. \tag{3.2.3}$$

This is non-zero, so the two-sided and tensor-product shell sets have non-trivial overlap as subsets of $\mathcal{H}_{L\cup R}$, and (3.2.1) is at least a sensible candidate equality.

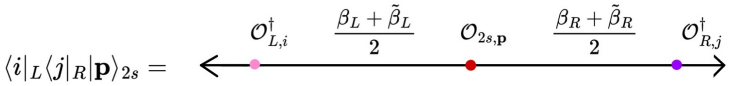
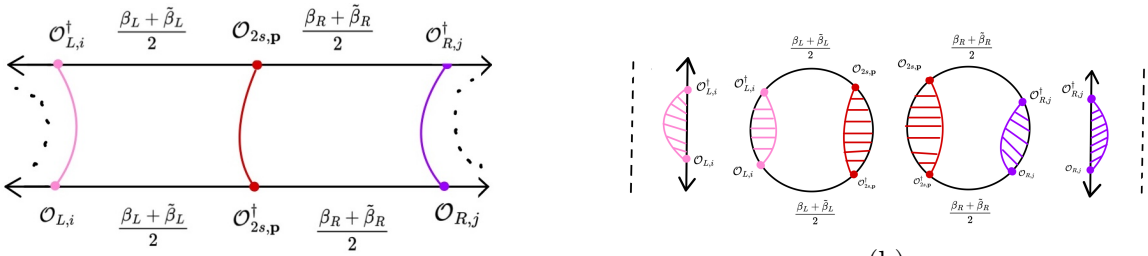


Figure 3.4: Boundary condition for the mixed overlap $\langle i|_L \langle j|_R |\mathbf{p}\rangle_{2s}$ between a tensor product of single-sided shell states and a two-sided shell state.

Strategy. We establish (3.2.1) by showing both inclusions $\mathcal{H}_L \otimes \mathcal{H}_R \subseteq \mathcal{H}_{2s}$ and $\mathcal{H}_{2s} \subseteq \mathcal{H}_L \otimes \mathcal{H}_R$. The structure of each inclusion is identical to that of the spanning arguments of Chapter 2 and Sec. 3.1: a coarse-grained projector identity follows from a maximally-connected wormhole construction, and is lifted to a fine-grained statement by the squaring trick of Sec. 2.1 in the limit of infinitely overcomplete shell bases. The two inclusions are closely parallel, and we present the first in detail and the second more briefly.



(a) (b)

Figure 3.5: (**a**) Universal wormhole contribution to $\overline{|\langle i|_L\langle j|_R|\mathbf{p}\rangle_{2s}|^2}$. (**b**) Construction by gluing L- and R-strips into the corresponding disks along the single-sided shell worldvolumes, then gluing the disks together along the two-sided shell worldvolume.

3.2.1 Tensor-product states lie in $\mathcal{H}_{2s}$

To show $\mathcal{H}_L \otimes \mathcal{H}_R \subseteq \mathcal{H}_{2s}$ it suffices, since the tensor products $|i\rangle_L|j\rangle_R$ form a basis of $\mathcal{H}_L \otimes \mathcal{H}_R$, to verify

$$\langle i|_L\langle j|_R\,\Pi_{\mathcal{H}_{2s}}\,|i\rangle_L|j\rangle_R \;=\; \langle i|i\rangle_L\,\langle j|j\rangle_R \tag{3.2.4}$$

at the saddlepoint level.

Coarse-grained. Expanding the projector and analytically continuing,

$$\overline{\langle i|_L\langle j|_R\,\Pi_{\mathcal{H}_{2s}}\,|i\rangle_L|j\rangle_R} \;=\; \lim_{n\to-1}\overline{(G_{2s})^n_{\mathbf{kl}}\,\langle i|_L\langle j|_R|\mathbf{k}\rangle_{2s}\langle\mathbf{l}|_{2s}|i\rangle_L|j\rangle_R}\,, \tag{3.2.5}$$

where G_{2s} is the two-sided Gram matrix. The boundary condition consists of n two-shell boundaries (Fig. 1.5) and two mixed boundaries (Fig. 3.4), and contains a single maximal two-sided shell index loop. In the limit $\kappa_{2s}\to\infty$ no saddle may break this loop, and only fully-connected geometries contribute (Fig. 3.6a). Such a saddle is constructed by taking an L- and an R-disk each carrying $n+1$ pairs of two-sided shell insertions and one pair of single-sided shell insertions, gluing the two disks together along the $n+1$ two-sided worldvolumes, and gluing an L-strip and R-strip into the respective disks along the single-sided worldvolumes (Fig. 3.6b). In the heavy shell mass limit the homology regions pinch off; after the universal two-sided shell factors cancel against the basis normalisation,

$$\overline{(G_{2s})^n_{\mathbf{kl}}\,\langle i|_L\langle j|_R|\mathbf{k}\rangle_{2s}\langle\mathbf{l}|_{2s}|i\rangle_L|j\rangle_R} \;=\; \kappa_{2s}^{n+1}\,\frac{\overline{S}_L(0)\,\overline{S}_R(0)\,\overline{Z}\Big(\beta_L+(n+1)\tilde\beta_L\Big)\;\overline{Z}\Big(\beta_R+(n+1)\tilde\beta_R\Big)\;Z_{m_i,L}\,Z_{m_j,R}}{\overline{Z}(\tilde\beta_L)^{n+1}\,\overline{Z}(\tilde\beta_R)^{n+1}}\,, \tag{3.2.6}$$

independent of the indices $\mathbf{k}, \mathbf{l}$ in the loop. Taking $n \to -1$,

$$\overline{\langle i|_L \langle j|_R \, \Pi_{\mathcal{H}_{2s}} \, |i\rangle_L |j\rangle_R} \;=\; \overline{S}_L(0)\, \overline{Z}(\beta_L)\, Z_{m_i,L}\; \overline{S}_R(0)\, \overline{Z}(\beta_R)\, Z_{m_j,R} \;=\; \overline{\langle i|i\rangle_L \, \langle j|j\rangle_R}\,, \tag{3.2.7}$$

the coarse-grained projector identity. The two-sided preparation temperatures $\tilde{\beta}_{L,R}$ have dropped out, as required: the basis change between $\mathcal{H}_{2s}$ at any choice of $\tilde{\beta}_{L,R}$ and $\mathcal{H}_L \otimes \mathcal{H}_R$ at any choice of $\beta_{L,R}$ is consistent.

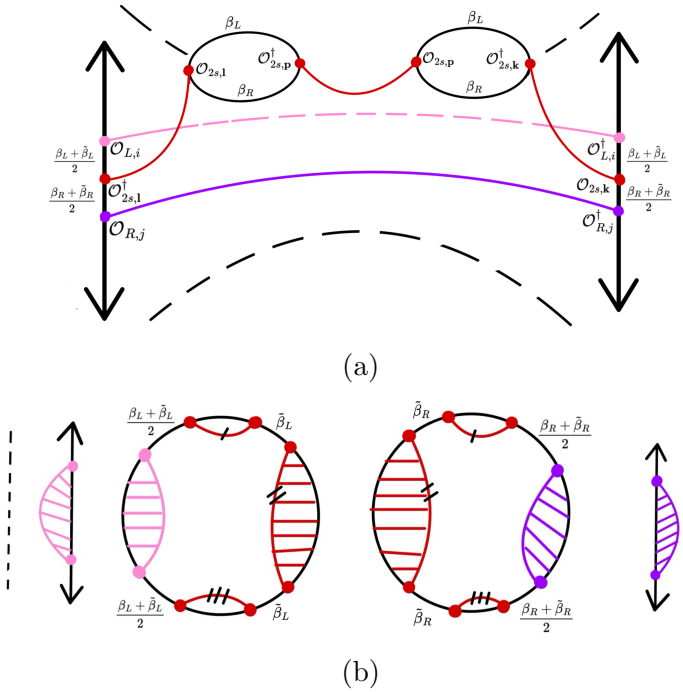


Figure 3.6: (**a**) Fully-connected saddle contributing to (3.2.5), depicted for $n = 2$. (**b**) Construction by gluing the L- and R-disks along $n+1$ two-sided shell worldvolumes and gluing an L- and R-strip into the respective disks along the single-sided worldvolumes.

Fine-grained. The squaring trick of Sec. 2.1 reduces (3.2.4) to

$$\overline{\left(\langle i|_L \langle j|_R \, \Pi_{\mathcal{H}_{2s}} \, |i\rangle_L |j\rangle_R \;-\; \langle i|i\rangle_L \, \langle j|j\rangle_R\right)^2} \;=\; 0\,. \tag{3.2.8}$$

Each term in the expansion of the square contains either two maximal two-sided shell index loops — the squared and cross terms involving $\Pi_{\mathcal{H}_{2s}}$ — or none — the term $\overline{\langle i|i\rangle_L^2 \, \langle j|j\rangle_R^2}$. By the same argument as in Sec. 2.3.4, the limit $\kappa_{2s} \to \infty$ forces the two index loops, when present, to remain seperate. Within each connected component,

the $\mathcal{O}_i$ and $\mathcal{O}_j$ insertions on the L and R disks may either be connected by a single-sided wormhole or remain disconnected, giving two further classes of saddle.

The disconnected saddles factorise into two copies of (3.2.7), giving $(\overline{\langle i|i\rangle_L \langle j|j\rangle_R})^2$ — the disconnected part of $\overline{\langle i|i\rangle_L^2 \langle j|j\rangle_R^2}$. The connected saddles for the cross term $\overline{\langle i|_L\langle j|_R \Pi_{\mathcal{H}_{2s}} |i\rangle_L|j\rangle_R \langle i|i\rangle_L \langle j|j\rangle_R}$ are constructed by augmenting (3.2.7) with one additional $\mathcal{O}_i$ pair on the L-disk separated by β_L from the original (similarly on the R side), and gluing a second L- (resp. R-) strip into the disks along the new worldvolumes (Fig. 3.7b). After normalisation, the saddle action is

$$\frac{\overline{S}_L(0)^2\, Z_{m_i,L}^2\, \overline{Z}\Big(2\beta_L + (n+1)\tilde{\beta}_L\Big)\;\; \overline{S}_R(0)^2\, Z_{m_j,R}^2\, \overline{Z}\Big(2\beta_R + (n+1)\tilde{\beta}_R\Big)}{\overline{Z}(\tilde{\beta}_L)^{n+1}\, \overline{Z}(\tilde{\beta}_R)^{n+1}}\,, \tag{3.2.9}$$

which limits as $n \to -1$ to the connected piece of $\overline{\langle i|i\rangle_L^2 \langle j|j\rangle_R^2}$, namely

$$\overline{S}_L(0)^2\, Z_{m_i,L}^2\, \overline{Z}(2\beta_L)\, \overline{S}_R(0)^2\, Z_{m_j,R}^2\, \overline{Z}(2\beta_R) \tag{3.2.10}$$

by (1.2.18) at $n = 2$. The squared term $\overline{(\langle i|_L\langle j|_R \Pi_{\mathcal{H}_{2s}} |i\rangle_L|j\rangle_R)^2}$ admits a parallel construction with two pairs of strips on each disk and gives the same result. Each term in (3.2.8) therefore reduces to $\overline{\langle i|i\rangle_L^2 \langle j|j\rangle_R^2}$, establishing (3.2.4).

We conclude $\mathcal{H}_L \otimes \mathcal{H}_R \subseteq \mathcal{H}_{2s}$ at the saddlepoint level.

3.2.2 Two-sided shell states lie in $\mathcal{H}_L \otimes \mathcal{H}_R$

For the reverse inclusion we show

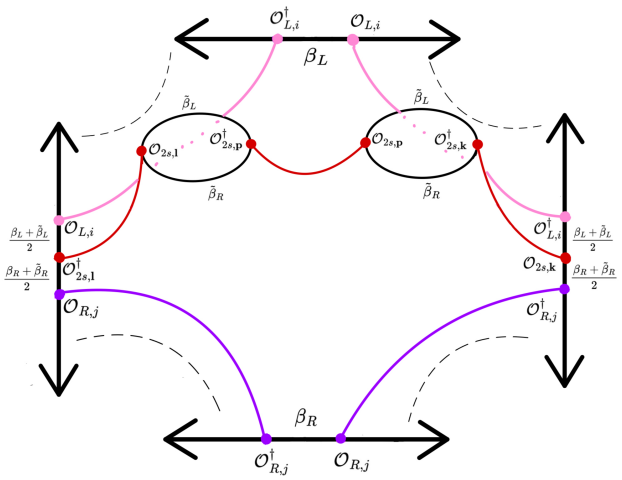
$$\langle \mathbf{p}|\, \Pi_{\mathcal{H}_L\otimes\mathcal{H}_R}\, |\mathbf{p}\rangle_{2s} \;=\; \langle \mathbf{p}|\mathbf{p}\rangle_{2s} \tag{3.2.11}$$

for any two-sided shell state $|\mathbf{p}\rangle_{2s}$. Since the Gram matrix on $\mathcal{H}_L \otimes \mathcal{H}_R$ factorises as $G_{L\otimes R} = G_L \otimes G_R$, the projector splits, $\Pi_{\mathcal{H}_L\otimes\mathcal{H}_R} = \Pi_{\mathcal{H}_L} \otimes \Pi_{\mathcal{H}_R}$.

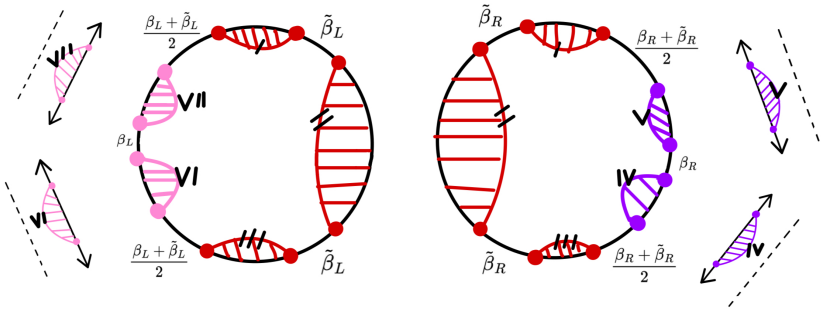
Coarse-grained. Expanding both projectors and analytically continuing,

$$\overline{\langle \mathbf{p}|\, \Pi_{\mathcal{H}_L\otimes\mathcal{H}_R}\, |\mathbf{p}\rangle_{2s}} \;=\; \lim_{n,m\to -1} \overline{(G_L)^n_{ac}\, (G_R)^m_{il}\, \langle \mathbf{p}|_{2s}|a\rangle_L|i\rangle_R\, \langle c|_L\langle l|_R|\mathbf{p}\rangle_{2s}}\,. \tag{3.2.12}$$

The boundary condition has two single-sided shell index loops — one of length n on the L side, one of length m on the R side — so in the limit $\kappa_L, \kappa_R \to \infty$ neither can be broken (Fig. 3.8a). The maximally-connected saddle is built by gluing an L- and an R-disk together along the single $\mathbf{p}$-shell worldvolume, and gluing $n+1$ single-sided



(a)



(b)

Figure 3.7: (**a**) Connected saddle contributing to the cross term in (3.2.8), depicted for $n = 2$. (**b**) Construction: as in Fig. 3.6b, with an additional L- and R-strip glued in along an extra pair of single-sided shell insertions separated by $\beta_{L,R}$ from the original.

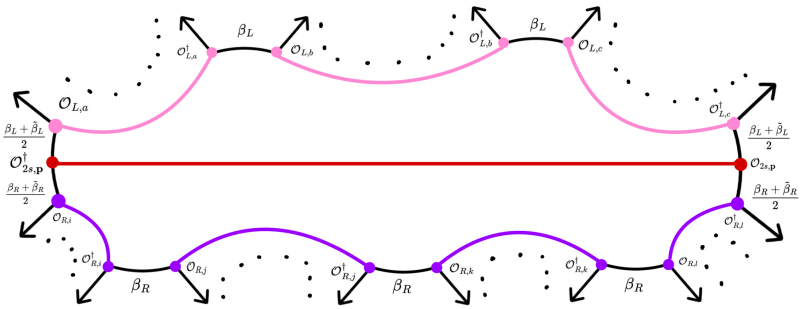
L-strips and $m+1$ single-sided R-strips into the respective disks (Fig. 3.8b). In the heavy mass limit, after normalisation of the single-sided shell states,

$$\begin{aligned}&\overline{(G_L)^n_{ac}\,(G_R)^m_{il}\,\langle\mathbf{p}|_{2s}|a\rangle_L|i\rangle_R\,\langle c|_L\langle l|_R|\mathbf{p}\rangle_{2s}}\\ &\quad=\ \kappa_L^{n+1}\kappa_R^{m+1}\,\frac{\overline{Z}\Big(\tilde\beta_L+(n+1)\beta_L\Big)\;\overline{Z}\Big(\tilde\beta_R+(m+1)\beta_R\Big)\;Z_{m_{\mathbf{p}},2s}}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{m+1}}\,,\end{aligned}\tag{3.2.13}$$

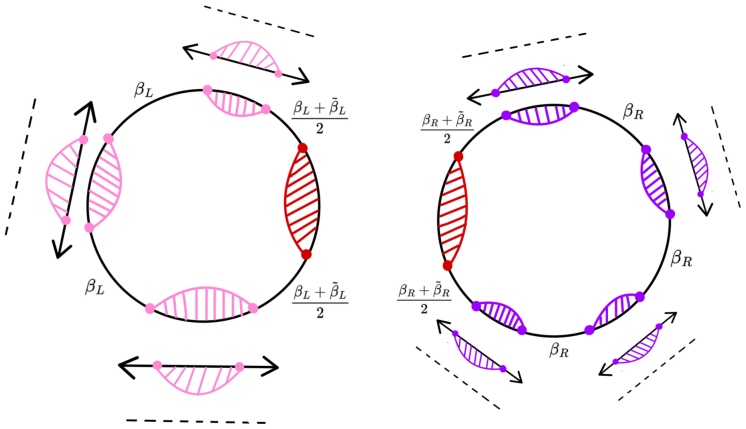
independent of the single-sided indices in either loop. Taking $n, m \to -1$ recovers the two-sided shell norm,

$$\overline{\langle\mathbf{p}|\,\Pi_{\mathcal{H}_L\otimes\mathcal{H}_R}\,|\mathbf{p}\rangle_{2s}}\ =\ \overline{Z}(\tilde\beta_L)\,\overline{Z}(\tilde\beta_R)\,Z_{m_{\mathbf{p}},2s}\ =\ \overline{\langle\mathbf{p}|\mathbf{p}\rangle_{2s}}\,.\tag{3.2.14}$$

The single-sided preparation temperatures $\beta_{L,R}$ have dropped out, as before.



(a)



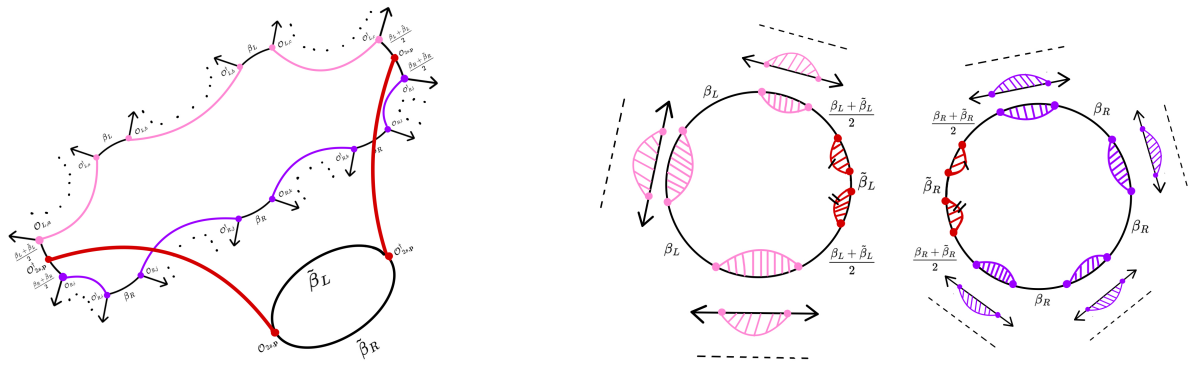
(b)

Figure 3.8: (**a**) Fully-connected saddle contributing to (3.2.12), depicted for $n = 2$, $m = 3$. (**b**) Construction by gluing an L- and an R-disk along the **p**-shell worldvolume, and gluing $n+1$ ($m+1$) single-sided L-strips (R-strips) into the respective disks along the single-sided shell worldvolumes.

Fine-grained. The argument is exactly parallel to that of Sec. 3.2.1. The cross term $\overline{\langle \mathbf{p}|\Pi_{\mathcal{H}_L\otimes\mathcal{H}_R}|\mathbf{p}\rangle_{2s}\,\langle \mathbf{p}|\mathbf{p}\rangle_{2s}}$ admits two classes of fully-connected saddle. Those in which the two $|\mathbf{p}\rangle$ insertions remain in disconnected bulk components factorise via (3.2.14) to $\overline{\langle \mathbf{p}|\mathbf{p}\rangle_{2s}}^2$, the disconnected piece of $\overline{\langle \mathbf{p}|\mathbf{p}\rangle^2_{2s}}$. Those in which the two $|\mathbf{p}\rangle$ insertions are joined by a **p**-shell wormhole are constructed by augmenting Fig. 3.8b with a second pair of **p**-shell insertions on each of the L, R disks separated by $\tilde{\beta}_{L,R}$ from the original (Fig. 3.9b). The resulting connected saddle action is

$$\frac{\overline{Z}\Big(2\tilde{\beta}_L + (n+1)\beta_L\Big)\;\overline{Z}\Big(2\tilde{\beta}_R + (m+1)\beta_R\Big)\;Z^2_{m_{\mathbf{p}},2s}}{\overline{Z}(\beta_L)^{n+1}\,\overline{Z}(\beta_R)^{m+1}}\,, \tag{3.2.15}$$

which limits as $n, m \to -1$ to $\overline{Z}(2\tilde{\beta}_L)\,\overline{Z}(2\tilde{\beta}_R)\,Z^2_{m_{\mathbf{p}},2s}$ — the connected piece of $\overline{\langle \mathbf{p}|\mathbf{p}\rangle^2_{2s}}$ by (1.2.18). The squared term works analogously, and each term in the squared



(a) (b)

Figure 3.9: (**a**) Connected saddle contributing to the cross term in the fine-grained extension of (3.2.14). (**b**) Construction: as in Fig. 3.8b, with an additional pair of **p**-shell insertions on each disk separated by $\tilde{\beta}_{L,R}$ from the original.

difference reduces to $\overline{\langle \mathbf{p}|\mathbf{p}\rangle_{2s}^2}$, establishing (3.2.11).

Conclusion. Combining the two inclusions,

$$\mathcal{H}_{LR} \;=\; \mathcal{H}_{2s} \;=\; \mathcal{H}_L \otimes \mathcal{H}_R \;=\; \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}\,, \tag{3.2.16}$$

where the outer equalities use the spanning results of Chapter 2 and Sec. 3.1. Since $\mathcal{H}_{L\cup R} = \{u+v \mid u \in \mathcal{H}_{LR},\, v \in \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R}\}$ (Sec. 1.1.3) and these two sets coincide as subspaces of $\mathcal{H}_{L\cup R}$, the factorisation problem (3.0.1) is resolved. The argument used only the rules of the gravitational path integral and is therefore equally valid for asymptotically flat and asymptotically AdS gravity, with no holographic input.

The mixed wormhole saddles connecting single- and two-sided shell boundaries (Figs. 3.6b and 3.8b) were crucial. Without them, the cross overlaps between $|\mathbf{p}\rangle_{2s}$ and $|i\rangle_L \otimes |j\rangle_R$ would vanish, the projector identities (3.2.7) and (3.2.14) would fail, and the two-sided and tensor-product shell sets would span orthogonal subspaces of $\mathcal{H}_{L\cup R}$ rather than coincide. The perturbative gravity Hilbert space obtained in the strict $G_N \to 0$ limit, in which these $\mathcal{O}(e^{-1/G_N})$ wormholes are absent, therefore does not factorise. Factorisation of the non-perturbative two-sided Hilbert space is thus a direct, and quantitative, consequence of including the wormhole sum.

3.3 Disentangling shell states and factorised geometries

The factorisation result (3.2.16) resolves the puzzle posed in Sec. 1.1.3 as a non-perturbative statement about the gravity Hilbert space. There remains, however,

a residual conceptual puzzle. The bases of $\mathcal{H}_{LR}$ produced by the two-sided shell construction generally have semiclassical geometries that look intrinsically two-sided, with no manifest factorisation structure. As reviewed in Sec. 1.2.1.5, the three classes of two-sided shell state correspond to a two-sided black hole with a long wormhole interior (type 1), thermal AdS with a disconnected closed-universe partner (type 2), or a single-sided black hole disconnected from a thermal-AdS factor (type 3). None of these states is a product of single-sided black holes, and in holographic settings they map to partially-entangled thermal states (PETS) in the dual CFT [33]. Reading off a tensor-product structure from such a basis is, a priori, difficult.

A tensor-product structure must nevertheless be present. To make it manifest at the level of Lorentzian geometry, we tune down the entanglement between the two factors by introducing an alternative basis in which two shell operators are inserted, separated by a controllable Euclidean time. Define the two-shell state

$$|\mathbf{i},\mathbf{j}\rangle_{2s} \;=\; |\, e^{-\tilde{\beta}_L H_L/2}\, \mathcal{O}_{\mathbf{i}}\, e^{-\beta_m H_R/2}\, \mathcal{O}_{\mathbf{j}}\, e^{-\tilde{\beta}_R H_R/2} \,\rangle\;, \tag{3.3.1}$$

with $\mathcal{O}_{\mathbf{i,j}}$ a pair of two-sided shell operators of varying mass; we have used H_R for the inter-shell evolution but H_L would do equally well. From the holographic perspective, varying β_m tunes the entanglement between the L and R Hilbert space factors.

The two-shell states span $\mathcal{H}_{LR}$ at any value of β_m. The argument is a direct adaptation of the spanning result of Chapter 2: in any wormhole construction where the single-shell case glues two disks together along a single shell worldvolume, the analogous two-shell construction glues both disks into a central disk containing the two shells separated by $\beta_m/2$ on either side (Fig. 3.10). In the heavy mass limit this introduces only an additional factor of $\overline{Z}(\beta_m)$ at each order, which contributes $\lim_{n\to -1}\overline{Z}(\beta_m)^{n+1} = 1$ to the analytic continuation. The spanning conclusion is unchanged.

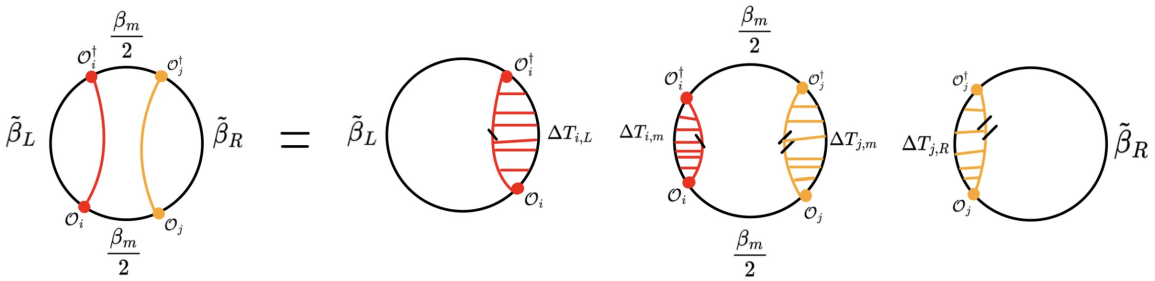


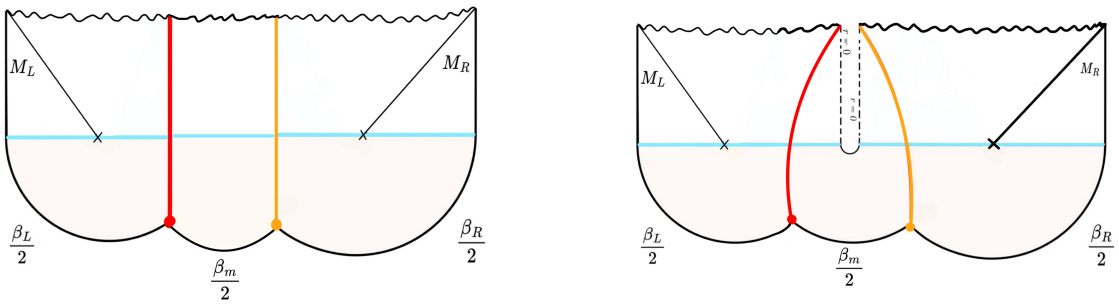
Figure 3.10: Saddle for the two-shell norm $\overline{\langle \mathbf{i},\mathbf{j}|\mathbf{i},\mathbf{j}\rangle_{2s}}$: the L- and R-disks are glued into a central disk by identifying the **i**- and **j**-shell worldvolumes.

The saddlepoint norm is, in the heavy shell limit,

$$\overline{\langle \mathbf{i},\mathbf{j}|\mathbf{i},\mathbf{j}\rangle_{2s}} \;=\; \overline{Z}(\tilde{\beta}_L)\,\overline{Z}(\beta_m)\,\overline{Z}(\tilde{\beta}_R)\,Z_{m_i}\,Z_{m_j}\;, \tag{3.3.2}$$

and the associated Lorentzian geometry is determined by the leading saddles of the three Z partition functions. Each may be above or below the Hawking-Page transition independently, yielding $2^3 = 8$ classes of two-shell states. For each of the four classes of single-shell two-sided states catalogued in Sec. 1.2.1.5, the two-shell construction inserts a thermal-AdS or AdS-black-hole patch "between" the shells.

The geometrically interesting regime is $\tilde{\beta}_L, \tilde{\beta}_R < \beta_{HP}$ together with $\beta_m > \beta_{HP}$ (Fig. 3.11b). The leading saddles for $\overline{Z}(\tilde{\beta}_{L,R})$ are AdS black holes while that for $\overline{Z}(\beta_m)$ is thermal AdS, and the resulting Lorentzian geometry is two disconnected single-sided AdS black holes. If we choose $\beta_L = \tilde{\beta}_L$ and $\beta_R = \tilde{\beta}_R$, this geometry is identical to that of the tensor product state $|i\rangle_L \otimes |j\rangle_R$ with the same shell content. *Every state in $\mathcal{H}_{LR}$ can therefore be expanded in a basis of two-shell states whose dominant saddles are manifestly factorised*, which makes the tensor-product structure (3.2.16) visible at the level of geometry.



(a) (b)

Figure 3.11: Lorentzian geometries of two-shell states with $\tilde{\beta}_L, \tilde{\beta}_R < \beta_{HP}$. (**a**) For $\beta_m < \beta_{HP}$, a single two-sided AdS black hole with a long interior supporting both shell trajectories. (**b**) For $\beta_m > \beta_{HP}$, two disconnected AdS black holes, with the two shells propagating in disconnected "Big Crunch" interior universes.

Saturation of Cauchy-Schwarz. For finite β_m the two-shell state $|\mathbf{i},\mathbf{j}\rangle_{2s}$ and the tensor product $|i\rangle_L \otimes |j\rangle_R$ cannot be literally identical: there is residual entanglement between the L and R factors of (3.3.1) for any finite inter-shell separation. The resolution is that the perturbative semiclassical approximation of retaining only the leading Z-saddle does not detect the fine correlations that distinguish them.

To see this explicitly, consider the direct overlap $\overline{{}_{2s}\langle \mathbf{i},\mathbf{j}|i\rangle_L|j\rangle_R}$ when $\beta_L = \tilde{\beta}_L$ and $\beta_R = \tilde{\beta}_R$. The connecting saddle is constructed by gluing two disks together along the **i**- and **j**-shell worldvolumes via a strip of length $\beta_m/2$ (Fig. 3.12). In the heavy mass limit, after normalisation,

$$\frac{\overline{{}_{2s}\langle \mathbf{i},\mathbf{j}|i\rangle_L|j\rangle_R}}{\sqrt{\overline{\langle \mathbf{i},\mathbf{j}|\mathbf{i},\mathbf{j}\rangle_{2s}}\;\overline{\langle i|i\rangle_L}\;\overline{\langle j|j\rangle_R}}} = \frac{\overline{S}(\beta_m/2)}{\sqrt{\overline{Z}(\beta_m)}}\,, \tag{3.3.3}$$

where $\overline{S}(\beta_m/2)$ denotes the strip path integral with shells separated by Euclidean time $\beta_m/2$. When $\beta_m > \beta_{HP}$ and only the leading Z-saddle is retained, $\overline{Z}(\beta_m)$ is dominated by thermal AdS, for which $\overline{Z}(\beta_m/2)^2 = \overline{Z}(\beta_m)$ and $\overline{S}(\beta_m/2) = \overline{Z}(\beta_m/2)$.[1] The right-hand side of (3.3.3) is then exactly 1, saturating the Cauchy-Schwarz inequality and forcing $|\mathbf{i},\mathbf{j}\rangle_{2s}$ and $|i\rangle_L \otimes |j\rangle_R$ to be linearly dependent at this level of approximation.

The two states must of course be genuinely distinct, which is restored as soon as subleading saddles are included. Retaining the subdominant Euclidean black hole saddle in $\overline{Z}(\beta_m)$ corrects (3.3.3) to a value strictly less than 1, breaking the Cauchy-Schwarz saturation and re-establishing linear independence. In the alternative regime $\beta_m < \beta_{HP}$, the leading saddle for $\overline{Z}(\beta_m)$ is the AdS black hole while that for the strip $\overline{S}(\beta_m/2)$ is empty AdS, so (3.3.3) is strictly less than unity already at the leading-saddle level — consistent with the dominant geometries of $|\mathbf{i},\mathbf{j}\rangle_{2s}$ and $|i\rangle_L \otimes |j\rangle_R$ being different in this regime (Fig. 3.11a).

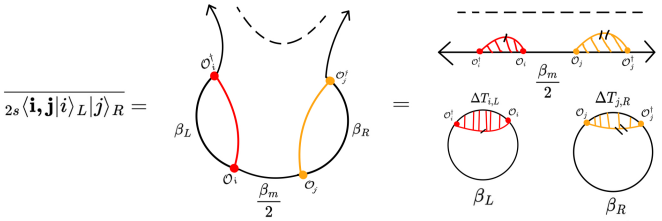


Figure 3.12: Saddle for the direct overlap (3.3.3): two disks glued together along the **i**- and **j**-shell worldvolumes via a strip of length $\beta_m/2$.

The construction provides a quantitative version of the basis-change phenomenon discussed in Sec. 2.5 of Chapter 2. There, the choice of basis at finite G_N determined whether the dominant saddle of an individual state contained a horizon, motivating the conclusion that bulk geometry cannot be a linear quantum observable. Here, the choice of β_m in the two-shell basis determines whether the dominant saddle is

[1] Thermal AdS is a section of the strip with the endpoints identified.

connected or disconnected — but the underlying state, once subleading saddles are restored, contains both pieces simultaneously. *Connectedness*, like geometry, is not a linear observable on the non-perturbative gravity Hilbert space.

3.4 Discussion

We have shown that the non-perturbative two-sided Hilbert space of quantum gravity factorises as

$$\mathcal{H}_{LR} \;=\; \mathcal{H}_{\mathcal{B}_L} \otimes \mathcal{H}_{\mathcal{B}_R} \,, \tag{3.4.1}$$

using only the rules of the gravitational path integral and the toolkit of Chapter 2. The result is driven by mixed wormhole saddles connecting boundaries of single- and two-sided types, which exist only beyond the perturbative $G_N \to 0$ limit and which simultaneously obstruct *amplitude* factorisation at the coarse-grained level. The argument applies equally to asymptotically flat and asymptotically AdS gravity, with no holographic input. We conclude with a discussion of the implications for operator algebras, the matrix-model interpretation of the gravity path integral, and emergent geometry.

Operator algebras. The factorisation (3.2.16) has direct consequences for the von Neumann algebras of observables on the two-sided Hilbert space. In the perturbative gravity limit, the two-sided black hole supports algebras of type III [60, 61], which admit no traces, no density matrices, and no pure states. Imposing gravitational constraints upgrades these to type II crossed-product algebras [62–64], which admit traces (up to constants) and density matrices, but still no pure states and no representation on a tensor-product Hilbert space. The non-perturbative tensor-product structure (3.2.16) requires the full L, R algebras to be of type I — the unique class of algebras admitting pure states, a unique trace, and a tensor-product Hilbert-space representation. Furthermore, since the two-sided Hilbert space factorises into two copies of the single-boundary Hilbert space, the non-perturbative L, R algebras must be unitarily equivalent to the non-perturbative single-sided operator algebra.

In holographic settings these statements are tautological from the dual quantum-mechanical theory on the asymptotic boundary. The non-trivial content is that the same algebraic structure can be deduced from the gravitational path integral alone, and is therefore equally a feature of asymptotically flat quantum gravity, which has no known holographic dual. Independently, [23] showed using operator-algebraic axioms that $\mathcal{H}_{LR}$ must factorise as a sum of products of factors, but did not identify the

factors with the single-boundary Hilbert spaces. Our results fill this gap, and it would be interesting to determine which additional axioms in the framework of [23] would imply (3.2.16) directly.

Universality of $\overline{Z}$. After normalising the shell states, every wormhole saddle encountered in this chapter — whether connecting two-sided boundaries (Chapter 2), single-sided boundaries (Sec. 3.1), or a mix — evaluates to a product or ratio of the Z partition function $\overline{Z}(\beta)$. This pattern is consistent with proposals, originating in lower-dimensional gravity [18, 70] and recently extended to higher dimensions [25, 26, 71, 72], in which the gravitational path integral is interpreted as a maximum-ignorance random matrix (or matrix-tensor) ensemble whose only inputs are coarse-grained observables of an underlying microscopic theory — prominently including the disorder-averaged thermal partition function — with wormhole contributions captured by moments of the ensemble. It would be valuable to verify directly from such an ensemble that the wormhole sums considered here reduce to products of $\overline{Z}(\beta)$, and to understand what additional constraints — if any — factorisation places on its construction. We return to these points in Chapter 6.

"Behind it all is surely an idea so simple, so beautiful, that when we grasp it — in a decade, a century, or a millennium — we will all say to each other, how could it have been otherwise?"
— John Archibald Wheeler

Chapter 4

On the Universality of Black Hole Entropy

The Bekenstein-Hawking entropy formula

$$\mathbf{S}_{BH} \;=\; \frac{A}{4G_N} \tag{4.0.1}$$

appears to apply to every black hole in any theory with an Einstein-gravity low-energy description, regardless of dimension, matter content, or asymptotic boundary conditions. This universality is striking, and is one of the central facts that any theory of quantum gravity is expected to explain. The most common derivation traces it back to the Gibbons-Hawking saddlepoint calculation [10], which proceeds by direct analogy with statistical field theory. There the thermal partition function $\mathrm{Tr}(e^{-\beta H})$ is, by construction, computed by the Euclidean path integral on a manifold $\mathcal{M} \times \mathbb{S}^1_\beta$ with periodic time of period β. Gibbons and Hawking carried this boundary condition over to gravity, defining

$$\overline{Z(\beta)} \;\equiv\; \zeta[\mathbb{S}^1_\beta \times \mathcal{X}] \tag{4.0.2}$$

as the gravitational path integral on a periodic asymptotic boundary, and treating it as a thermal partition function. Evaluating (4.0.2) at saddlepoint — where the dominant

Euclidean black hole has on-shell action I_{BH}, so $\overline{Z(\beta)} \approx e^{-I_{BH}}$ — and applying the standard thermodynamic identities then reproduces (4.0.1). The construction carries the formal features of statistical mechanics — a sensible β-dependence, Hawking-Page-type phase transitions between competing saddles — but the analogy on which it rests is far from evident. In QFT on a fixed background, $\mathrm{Tr}(e^{-\beta H})$ equalling the periodic-time path integral is automatic; wheras in gravity it is a conjecture

$$\mathrm{Tr}_{\mathcal{H}_\chi}\left(e^{-\beta H}\right) \stackrel{?}{=} \overline{Z(\beta)}, \tag{4.0.3}$$

where $\mathcal{H}_\chi$ is the Hilbert space of single-boundary quantum gravity. The path integral $\overline{Z(\beta)}$ is built by summing over all geometries asymptotic to a periodic time circle, with no a priori connection to a trace, and the dominant black-hole saddle has the topology of a disk rather than a cylinder (Sec. 1.1.2). Until (4.0.3) is established, the apparent universality of (4.0.1) is the universality of a thermodynamic analogy — not the universality of a microscopic state count.

This chapter establishes (4.0.3), and uses it to give a microscopic account of why the Bekenstein-Hawking formula is universal. Equation (4.0.3) is shown by two independent arguments: an abstract one that reduces it to the factorisation of the two-boundary Hilbert space of Chapter 3, and a concrete one that evaluates the trace directly in the single-sided shell basis of Chapter 3 and finds that the wormhole sums collapse to the Gibbons-Hawking path integral. Both arguments depend critically on non-perturbative wormhole saddles; in perturbative gravity, where these wormholes are absent, $\overline{Z(\beta)}$ is not a thermal trace. The arguments use no specific form of the gravitational action, no holographic input, and apply equally in asymptotically flat and asymptotically AdS gravity. With (4.0.3) in hand, we then unpack its microcanonical content: the Bekenstein-Hawking formula is the saddlepoint Laplace transform of (4.0.3); combining this with the spanning result of Chapter 3 identifies $e^{\mathbf{S}_{BH}}$ as a literal count of microstates above the black-hole threshold; and the basis of microstates being counted consists of single-sided black holes sharing the exterior of the Gibbons-Hawking saddle but distinguished by interior shell content. Universality of the area law is then a corollary of universality of the partition-function-as-trace identity (4.0.3).

A second application of the same machinery concerns gravitational entanglement. For a two-boundary gravity state $|\Psi\rangle \in \mathcal{H}_{\chi_L \cup \chi_R}$, the factorisation of Chapter 3 permits the construction of a reduced density matrix ρ_R on either boundary factor, whose Rényi entropies $\mathrm{Tr}_{\mathcal{H}_{\chi_R}}(\rho_R^n)$ admit a direct Hilbert-space evaluation. The holographic replica prescription instead computes $\mathrm{Tr}(\rho^n)$ as a path integral on an n-fold replicated

boundary, and is the input to the bulk derivation [73] of the RT and HRT formulas [74–76]. The two prescriptions look very different. We show that they agree by an argument structurally identical to the one used for (4.0.3). The conclusion provides a purely gravitational — and not specifically holographic — justification for the replica trick.

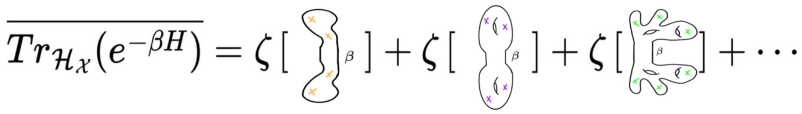
The chapter is organised as follows. Sec. 4.1 states the Gibbons-Hawking puzzle precisely. Sec. 4.2 establishes (4.0.3) and develops its microcanonical consequences. Sec. 4.3 establishes the gravitational replica equality. Sec. 4.4 concludes. Throughout, we work in the saddlepoint approximation; the squaring trick of Sec. 2.1 lifts saddlepoint equalities to fine-grained statements where appropriate.

4.1 The Gibbons-Hawking puzzle

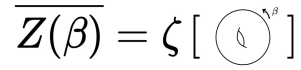
The thermal partition function on the single-boundary Hilbert space $\mathcal{H}_\chi$, evaluated using the trace formula of Sec. 2.2, is

$$\mathrm{Tr}_{\mathcal{H}_\chi}\left(e^{-\beta H}\right) \;=\; \sum_\gamma \langle v_\gamma | e^{-\beta H} | v_\gamma \rangle \;=\; \sum_\gamma \zeta[\mathcal{M}_\gamma(\beta)]\,, \tag{4.1.1}$$

where $\{|v_\gamma\rangle\}$ is any orthonormal basis of $\mathcal{H}_\chi$ that may be constructed by cutting open the Euclidean path integral, and $\mathcal{M}_\gamma(\beta)$ is the closed boundary manifold obtained by sewing each survival amplitude $\langle v_\gamma | e^{-\beta H} | v_\gamma \rangle$ along its asymptotic time circle (Fig. 4.1a). The right-hand side is a sum of distinct gravitational path integrals, one per basis element. When the basis contains states with matter excitations, the boundary conditions $\mathcal{M}_\gamma(\beta)$ carry the corresponding asymptotic operator insertions.



(a)



(b)

Figure 4.1: (**a**) The thermal partition function in gravity computed by summing path integrals with boundaries defined by a basis of states. The coloured crosses denote possible matter insertions specifying the basis element. (**b**) The Gibbons-Hawking path integral $\overline{Z(\beta)}$, defined on a periodic Euclidean boundary of length β with no insertions.

The Gibbons-Hawking path integral $\overline{Z(\beta)}$ defined in (4.0.2), by contrast, is a single gravitational path integral on a periodic boundary $\mathbb{S}^1_\beta \times \mathfrak{X}$ with no operator insertions. Cutting this boundary at any chosen Euclidean time identifies $\overline{Z(\beta)}$ with the squared norm of the half-torus state of Sec. 1.1.2,

$$\overline{Z(\beta)} \;=\; \overline{\langle\beta|\beta\rangle}\,. \tag{4.1.2}$$

The two expressions look superficially different. (4.1.1) is a basis-dependent sum of distinct path integrals with operator insertions; (4.0.2) is a single insertion-free path integral. The puzzle of this chapter is:

> *Why is the sum of path integrals on the right-hand side of (4.1.1) equal to the single path integral (4.0.2)?*

Equivalently: in what sense does the Gibbons-Hawking path integral compute a trace over the gravity Hilbert space?[1]

4.2 The single-boundary partition function

We establish (4.0.3) by two arguments before turning to its consequences for the universality of the area law.

4.2.1 Argument from factorisation

Let $\{|\alpha_i\rangle\}$ denote any (in general non-orthogonal and overcomplete) spanning set of $\mathcal{H}_\mathfrak{X}$ constructed by cutting open the path integral, with Gram matrix $G_{ij} = \langle\alpha_i|\alpha_j\rangle$. Using the trace formula of Sec. 2.2,

$$\overline{\mathrm{Tr}_{\mathcal{H}_\mathfrak{X}}(e^{-\beta H})} \;=\; \lim_{n\to-1} \overline{G^n_{ij}\,\langle\alpha_j|e^{-\beta H}|\alpha_i\rangle} \;=\; \lim_{n\to-1} \overline{G^n_{ij}\,\langle\alpha_i^*|e^{-\beta H}|\alpha_j^*\rangle}\,, \tag{4.2.1}$$

where the dual states $|\alpha^*\rangle$ are defined as in Sec. 1.1. To make contact with $\overline{Z(\beta)} = \overline{\langle\beta|\beta\rangle}$, we split $e^{-\beta H}$ into two halves and insert a resolution of the identity (2.2.2) on $\mathcal{H}_\mathfrak{X}$ between them:

$$\overline{\mathrm{Tr}_{\mathcal{H}_\mathfrak{X}}(e^{-\beta H})} \;=\; \lim_{n,m\to-1} \overline{G^n_{ij}\,G^m_{kl}\,\langle\alpha_i^*|e^{-\beta H/2}|\alpha_k\rangle\langle\alpha_l|e^{-\beta H/2}|\alpha_j^*\rangle}\,. \tag{4.2.2}$$

The action of $e^{-\beta H/2}$ admits two equivalent geometrical interpretations (Sec. 1.1.2): time evolution by $\beta/2$ on a single asymptotic boundary, or preparation of the state $|\beta\rangle$

[1] See [21, 77] for additional discussion on putting the Gibbons-Hawking path integral on firmer footing.

on two copies of $\mathcal{X}$. Reading it the second way, $\langle\alpha_i^*|e^{-\beta H/2}|\alpha_k\rangle = \langle\beta|\alpha_i\rangle_L \otimes |\alpha_k\rangle_R$ and similarly for the other matrix element. Substituting and recognising the resolutions of identity on $\mathcal{H}_{\mathcal{X}_L}$ and $\mathcal{H}_{\mathcal{X}_R}$,

$$\overline{\mathrm{Tr}_{\mathcal{H}_\mathcal{X}}(e^{-\beta H})} \;=\; \overline{\langle\beta|\,\mathbb{1}_{\mathcal{X}_L}\otimes\mathbb{1}_{\mathcal{X}_R}\,|\beta\rangle}\,. \tag{4.2.3}$$

The state $|\beta\rangle$ lives in the two-boundary Hilbert space $\mathcal{H}_{\mathcal{X}_L\cup\mathcal{X}_R}$, while $\mathbb{1}_{\mathcal{X}_L}\otimes\mathbb{1}_{\mathcal{X}_R}$ is the identity on $\mathcal{H}_{\mathcal{X}_L}\otimes\mathcal{H}_{\mathcal{X}_R}$. By the factorisation result of Chapter 3 these spaces coincide, the two identities are equal, and

$$\overline{\mathrm{Tr}_{\mathcal{H}_\mathcal{X}}(e^{-\beta H})} \;=\; \overline{\langle\beta|\beta\rangle} \;=\; \overline{Z(\beta)}\,. \tag{4.2.4}$$

The squaring trick of Sec. 2.1 lifts (4.2.4) to a fine-grained equality: the same insertion procedure applied to $\overline{(\mathrm{Tr}-Z)^2}$ reduces it to $\overline{Z^2}-\overline{Z}^2=0$ in the saddlepoint approximation, so

$$\mathrm{Tr}_{\mathcal{H}_\mathcal{X}}\big(e^{-\beta H}\big) \;=\; Z(\beta) \tag{4.2.5}$$

holds at the fine-grained level. The argument is striking in its generality: it uses no specific basis, no specific form of the gravitational action, and only two ingredients — that the gravity Hamiltonian generates evolution in asymptotic boundary time (a consequence of bulk diffeomorphism invariance), and the Hilbert space factorisation of Chapter 3.

4.2.2 Explicit calculation in the shell basis

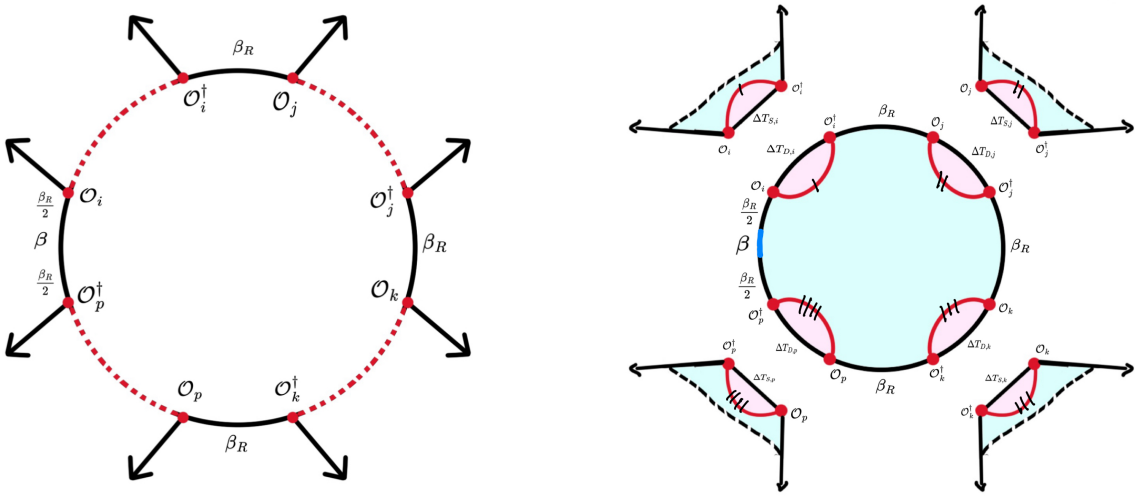
The factorisation argument is sufficient but perhaps not necessary. We now establish (4.2.5) directly by evaluating the trace in the single-sided shell basis of Chapter 3, without invoking factorisation.

Denote a single-sided shell state of mass m_i and preparation temperature β_R by $|i\rangle_R$, and the closed span of κ_R such states by $\mathcal{H}_R$. By Chapter 3, $\mathcal{H}_R=\mathcal{H}_\mathcal{X}$ as $\kappa_R\to\infty$, as the shell states span the single-boundary Hilbert space. The trace then resolves over the shell basis as

$$\overline{\mathrm{Tr}_{\mathcal{H}_\mathcal{X}}(e^{-\beta H})} \;=\; \lim_{n\to-1}\overline{G^n_{ij}\,\langle j|_R\,e^{-\beta H}\,|i\rangle_R}\,, \tag{4.2.6}$$

with summation over i,j implicit.

The boundary condition consists of n shell strips of length β_R (from G^n, see Fig. 1.12) joined to one elongated shell strip of length $\beta+\beta_R$ (from $\langle j|e^{-\beta H}|i\rangle$). The shell labels close into a single maximal index loop running around the chain



(a) (b)

Figure 4.2: (**a**) Boundary condition for the path integral computing (4.2.6), depicted for $n = 3$. The n shell-overlap boundaries from G^n are joined to the $\langle j|e^{-\beta H}|i\rangle$ boundary, and the shell index loop closes around the chain. (**b**) Construction of the maximally-connected saddle by gluing $n+1$ shell strips into a disk along the corresponding shell worldvolumes.

(Fig. 4.2a). In the limit $\kappa_R \to \infty$ no saddle may break this loop, and only fully-connected geometries contribute. These maximally-connected saddles are constructed by gluing the $n+1$ shell strips into a single Euclidean disk along the corresponding shell worldvolumes (Fig. 4.2b), in direct analogy with the pinwheel construction of Sec. 1.2.1.7, with one strip elongated by the additional β from the trace insertion. In the heavy mass limit the homology regions pinch off, the universal shell factors Z_{m_i} and the strip path integrals $\overline{S}(0)$ cancel against the shell-state normalisations as in (1.2.21), and the saddle contribution reduces to

$$\overline{G_{ij}^n \, \langle j|_R \, e^{-\beta H} \, |i\rangle_R} \;=\; \kappa_R^{n+1} \, \frac{\overline{Z}(\beta + (n+1)\beta_R)}{\overline{Z}(\beta_R)^{n+1}} \,. \tag{4.2.7}$$

This expression contains contributions from every saddle of the Z partition function, not merely the dominant one. Taking $n \to -1$,

$$\overline{\mathrm{Tr}_{\mathcal{H}_\chi}(e^{-\beta H})} \;=\; \overline{Z(\beta)} \,, \tag{4.2.8}$$

where the shell preparation temperature has dropped out as required by basis-independence. The squaring trick lifts (4.2.8) to the fine-grained statement (4.2.5): each term in $\overline{(\mathrm{Tr} - \overline{Z})^2}$ contains either two maximal index loops or none, and the $\kappa_R \to \infty$ limit

forces the two index loops to lie in disconnected bulk components, each evaluating to (4.2.8).[2]

The shell-basis derivation does not invoke factorisation. The wormhole saddle (4.2.7) connects only single-boundary states, never mixing single- and two-boundary types. This shows that the wormholes responsible for $\mathrm{Tr} = \overline{Z}$ are a strictly weaker requirement than those responsible for the factorisation $\mathcal{H}_{\mathcal{X}_L \cup \mathcal{X}_R} = \mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$ used in Sec. 4.2.1 — the latter requires also mixed wormholes between single- and two-boundary states. We return to this dichotomy in Sec. 4.4.

4.2.3 The area law counts black-hole microstates

The Bekenstein-Hawking formula makes two claims about the spectrum of quantum gravity at energies above the black-hole threshold: that the microcanonical density of states is $e^{A(E)/4G_N}$, and that the states in any such microcanonical window are black-hole microstates — geometries sharing the same exterior but with non-trivial interior content. We can now use (4.2.5) to give a microscopic account of why both claims are universal.

The argument proceeds in three steps: (i) (4.2.5) implies the Bekenstein-Hawking formula gives the leading-order microcanonical density of states; (ii) the spanning result of Chapter 2 permits the construction of an explicit basis of the microcanonical Hilbert space; (iii) above the black-hole threshold this basis consists of states whose dominant saddles are black-hole microstates.

Step 1: $e^{\mathbf{S}_{BH}}$ is the microcanonical density of states. With (4.2.5) established, the standard thermodynamic identification of the partition function with a thermal trace makes the inverse Laplace transform $\rho(E) = \int d\beta\, e^{\beta E}\, Z(\beta)$ the microcanonical density of states of $\mathcal{H}_{\mathcal{X}}$. In the saddlepoint approximation, $\overline{Z(\beta)}$ may have multiple competing saddles — in asymptotically AdS gravity these are thermal AdS, the large black hole, and the small black hole (Sec. 1.1.2) — and the dominant one determines $\rho(E)$. Above the Hawking-Page transition the black hole dominates, the saddlepoint Laplace transform yields

$$\rho(E) \;=\; e^{\mathbf{S}(E)}\,, \qquad \mathbf{S}(E) \;=\; \frac{A(E)}{4G_N}\,, \tag{4.2.9}$$

[2] Concretely: the equality $\overline{\langle\beta|\mathbb{1}_{\mathcal{X}_L} \otimes \mathbb{1}_{\mathcal{X}_R}|\beta\rangle} = \overline{\langle\beta|\beta\rangle}$ holds as a full path integral equality by Sec. 3.2.1 of Chapter 3, and the surgery argument of [44] extends the conclusion (4.2.5) beyond the saddlepoint approximation to the full path integral.

where $A(E)$ is the horizon area at energy E. The same conclusion follows for any low-energy theory described by Einstein gravity, regardless of dimension or cosmological constant. This already explains the *universality* of the area law: it is the saddlepoint Laplace transform of the universal identity (4.2.5). What remains is to identify which states (4.2.9) is counting.

Step 2: an explicit basis of $\mathcal{H}_\chi|_E$. Let Π_E project onto the microcanonical subspace $\mathcal{H}_\chi|_E$, and let $|i^\beta\rangle = e^{-\beta H/2}|\hat{\imath}\rangle$ denote a canonical shell state with preparation temperature β. In the energy basis $\{|n\rangle\}$,

$$\Pi_E|i^\beta\rangle \;=\; \sum_{E_n=E} e^{-\beta E_n/2}\langle n|\hat{\imath}\rangle|n\rangle \;, \tag{4.2.10}$$

so canonical shell states with the same operator content but different β project onto the same microcanonical state up to a β-dependent prefactor. The microcanonical projection is therefore implemented by a Laplace transform,

$$|i^E\rangle \;\propto\; \int d\beta\, e^{\beta E/2}|i^\beta\rangle \;. \tag{4.2.11}$$

Since the canonical shell states $|i^\beta\rangle$ span $\mathcal{H}_\chi$ at any preparation temperature (Chapter 3), the projected states $|i^E\rangle$ span $\mathcal{H}_\chi|_E$. This basis has $e^{\mathbf{S}(E)}$ elements at energies above the black-hole threshold, by (4.2.9) and the dimension argument of Chapter 2.

Step 3: the basis consists of black-hole microstates. The semiclassical geometry associated to $|i^E\rangle$ is the dominant saddle of its norm,

$$\overline{\langle i^E|i^E\rangle} \;\propto\; \int d\beta\, e^{\beta E}\,\overline{\langle i^{\tilde\beta}|i^\beta\rangle} \;=\; Z_{m_i}\,\overline{S}(0)\int d\beta\, e^{\beta E}\,\overline{Z}(\beta+\tilde\beta) \;, \tag{4.2.12}$$

where the second equality uses the canonical shell norm (1.2.20) at some arbitrary preparation temperature $\tilde\beta$. The Laplace transform of $\overline{Z}(\beta+\tilde\beta)$ is the microcanonical density of states by Step 1, and at energies above the black-hole threshold this is dominated by the black-hole saddle. The semiclassical norm therefore evaluates to

$$\overline{\langle i^E|i^E\rangle} \;=\; Z_{m_i}\,\overline{S}(0)\, e^{-\tilde\beta E+\mathbf{S}(E)} \;, \tag{4.2.13}$$

with the dominant geometry being a single-sided black hole at energy E with the shell behind the horizon (Fig. 1.15a) — a type-A shell state in the language of Sec. 1.2.2.

The Bekenstein-Hawking formula above the black-hole threshold therefore counts genuine black-hole microstates, all sharing the exterior of the Gibbons-Hawking saddle

but distinguished by interior shell content. Below threshold the same construction recovers a count of horizonless geometries, with $e^{\mathbf{S}(E)}$ corresponding to the dimension of the gas-of-gravitons sector.[3]

Two-boundary corollary. Applied to each factor of $\mathcal{H}_{LR} = \mathcal{H}_{\mathcal{X}_L} \otimes \mathcal{H}_{\mathcal{X}_R}$, the identity (4.2.5) yields the two-boundary trace factorisation

$$\mathrm{Tr}_{\mathcal{H}_{LR}}\left(e^{-\beta_L H_L} e^{-\beta_R H_R}\right) \;=\; Z(\beta_L)\, Z(\beta_R) \tag{4.2.14}$$

and a microcanonical entropy $\mathbf{S}(E_L) + \mathbf{S}(E_R)$, identifying the two-sided shell states of [6, 19, 28, 29, 32] as a basis of two-sided black-hole microstates and recovering the result of Chapter 2 from this angle.

4.3 Entanglement entropy and the gravitational replica trick

The same machinery yields a purely gravitational understanding of the holographic replica trick. For a two-boundary gravity state $|\Psi\rangle \in \mathcal{H}_{\mathcal{X}_L \cup \mathcal{X}_R}$, the factorisation of Chapter 3 permits the construction of a reduced density matrix $\rho_R = \mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_L}} |\Psi\rangle\langle\Psi|$ on one boundary, and we want to compute its Rényi entropy[4]

$$\overline{\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n)}\,. \tag{4.3.1}$$

For concreteness we restrict to $|\Psi(\tilde\beta_L, \tilde\beta_R)\rangle = |e^{-\tilde\beta_L H_L/2}\, \mathcal{O}_\Psi\, e^{-\tilde\beta_R H_R/2}\rangle$ with $\mathcal{O}_\Psi$ a possibly non-local operator; examples include the thermofield double [79], partially entangled thermal states [33], and the two-sided shell states states. The conclusions extend to general superpositions by linearity.

Resolving the traces in (4.3.1) over the single-sided shell bases of Chapter 3 gives an explicit Hilbert-space prescription:

$$\overline{\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n)} \;=\; \lim_{k,p_1,\ldots,p_n \to -1} \overline{G^k_{R,cd}\, \langle d|_R \prod_{i=1}^{n} G^{p_i}_{L,a_i b_i}\, \langle b_i|_L \Psi\rangle\langle\Psi|a_i\rangle_L |c\rangle_R}\,, \tag{4.3.2}$$

a basis-resolved sum of products of overlaps with explicit shell insertions on the boundary (Figs. 4.5a, 4.5b).

[3] An analogous analysis should be possible for single-sided states with end-of-the-world branes behind the horizon [78], although these have not been shown to span the full Hilbert space.

[4] Strictly speaking the Rényi entropy is $\log \overline{\mathrm{Tr}(\rho^n)} - n \log \overline{\mathrm{Tr}(\rho)}$, but we work with (4.3.1) from which the entropy is easily obtained.

The holographic replica prescription instead computes $\mathrm{Tr}(\rho^n)$ as a single gravitational path integral on a periodic boundary obtained by sewing n copies of the boundary preparing $|\Psi\rangle\langle\Psi|$ along their $\mathcal{X}_L$ cuts (Fig. 4.3). The resulting boundary is a "torus" of n alternating $\mathcal{O}_\Psi, \mathcal{O}_\Psi^\dagger$ insertions, with consecutive insertions separated by $\tilde{\beta}_L$ on one side and $\tilde{\beta}_R$ on the other. We denote this path integral by $Z_\circ(\rho_\Psi^n)(\tilde{\beta}_L, \tilde{\beta}_R)$. We will show

$$\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n) \;=\; Z_\circ(\rho_\Psi^n)(\tilde{\beta}_L, \tilde{\beta}_R) \tag{4.3.3}$$

as a fine-grained equality, by two arguments paralleling those of Sec. 4.2: a general one based on factorisation, and an explicit one based on the shell basis. Concretely, two things must happen on the gravity-Hilbert-space side. The trace over $\mathcal{H}_{\mathcal{X}_L}$ in the construction of ρ_R must sew the bra and ket of $|\Psi\rangle\langle\Psi|$ into a closed boundary condition; and the outer trace over $\mathcal{H}_{\mathcal{X}_R}$ must glue n such cylinders into a torus. Neither is automatic from the path-integral construction of states.

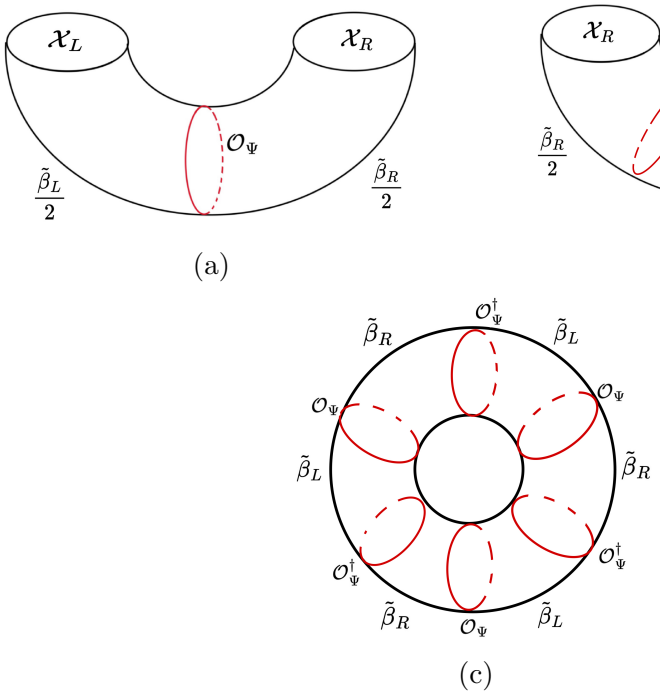


Figure 4.3: The QFT replica prescription. (**a**) The boundary condition for $|\Psi\rangle$, viewable as defining a map $M_\Psi : \mathcal{H}_{\mathcal{X}_L} \to \mathcal{H}_{\mathcal{X}_R}$. (**b**) The reduced density matrix ρ_R is obtained by sewing the $|\Psi\rangle$ ket to the bra $\langle\Psi|$ along $\mathcal{X}_L$, equivalently the operator $M_\Psi M_\Psi^\dagger$. (**c**) $\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n)$ is obtained by sewing n copies of ρ_R together along $\mathcal{X}_R$ into a periodic boundary.

4.3.1 The trace glues density matrices

The inner trace sews bra and ket. The reduced density matrix is, by definition,

$$\rho_R \;=\; \lim_{m\to -1} G^m_{L,ab}\,\langle b|_L\Psi\rangle\langle\Psi|a\rangle_L\;. \tag{4.3.4}$$

The state $|\Psi\rangle$ may be read as a map $M_\Psi : \mathcal{H}_{\mathcal{X}_L} \to \mathcal{H}_{\mathcal{X}_R}$ (Fig. 4.3a); its conjugate $M_\Psi^\dagger$ defines the reverse map. The sewn boundary condition of Fig. 4.3b corresponds to the operator $M_\Psi M_\Psi^\dagger$. Inserting an identity on $\mathcal{H}_{\mathcal{X}_L}$ resolved in the conjugate basis $\{|a^*\rangle\}$ of Sec. 1.1,

$$M_\Psi M_\Psi^\dagger \;=\; \lim_{m\to -1} G^m_{L,b^*a^*}\,M_\Psi|b^*\rangle_L\langle a^*|_L M_\Psi^\dagger\;. \tag{4.3.5}$$

For any $|n\rangle_R, |m\rangle_R \in \mathcal{H}_{\mathcal{X}_R}$, $\langle m|_R M_\Psi|b^*\rangle_L = \langle m|_R\langle b|_L|\Psi\rangle$ and $\langle a^*|_L M_\Psi^\dagger|n\rangle_R = \langle\Psi|a\rangle_L|n\rangle_R$, while $G_{L,b^*a^*} = G_{L,ab}$. The matrix elements of (4.3.5) and (4.3.4) coincide, so $\rho_R = M_\Psi M_\Psi^\dagger$.

Factorisation closes the chain. Compute $\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n) = \mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\prod_{i=1}^n M_{\Psi,i}M_{\Psi,i}^\dagger)$ by inserting an identity on $\mathcal{H}_{\mathcal{X}_L}$ between $M_{\Psi,k}^\dagger$ and the rest of the chain, and resolving both the outer trace and the inserted identity in conjugate bases. The two operator products on either side of the inserted identity define a bra $\langle\Sigma|$ and a ket $|\Phi\rangle$ in $\mathcal{H}_{\mathcal{X}_L\cup\mathcal{X}_R}$, with $\langle\Sigma|\Phi\rangle = Z_\circ(\rho_\Psi^n)(\tilde\beta_L,\tilde\beta_R)$ when the replica path integral is cut along $\mathcal{X}_L\cup\mathcal{X}_R$ (Fig. 4.4). The same dual-basis identifications as in Sec. 4.2.1 then yield

$$\overline{\mathrm{Tr}_{\mathcal{H}_{\mathcal{X}_R}}(\rho_R^n)} \;=\; \overline{\langle\Sigma|\,\mathbb{1}_{\mathcal{X}_L}\otimes\mathbb{1}_{\mathcal{X}_R}\,|\Phi\rangle} \;=\; \overline{\langle\Sigma|\Phi\rangle} \;=\; \overline{Z_\circ(\rho_\Psi^n)}\,, \tag{4.3.6}$$

the second equality using the factorisation $\mathbb{1}_{\mathcal{X}_L}\otimes\mathbb{1}_{\mathcal{X}_R} = \mathbb{1}_{\mathcal{X}_L\cup\mathcal{X}_R}$ of Chapter 3. The squaring trick of Sec. 2.1 lifts this to the fine-grained equality (4.3.3).

4.3.2 Explicit calculation in the shell basis

The shell-basis derivation, paralleling Sec. 4.2.2, again does not require factorisation. The boundary condition for (4.3.2) consists of n copies of the ρ_R boundary joined along their $\mathcal{X}_R$ cuts via G_R^k, with the outer trace closing the chain into a single R-shell index loop (Fig. 4.5b). In the limit $\kappa_L,\kappa_R\to\infty$ no shell index loop may be broken, so only fully-connected saddles contribute. These are constructed by gluing $k+\sum_{i=1}^n p_i$ shell strips into a single Euclidean disk along the corresponding shell worldvolumes (Fig. 4.6a), in direct analogy with the pinwheel construction of Sec. 1.2.1.7. The disk's asymptotic boundary carries n pairs of $\mathcal{O}_\Psi,\mathcal{O}_\Psi^\dagger$ insertions, with consecutive

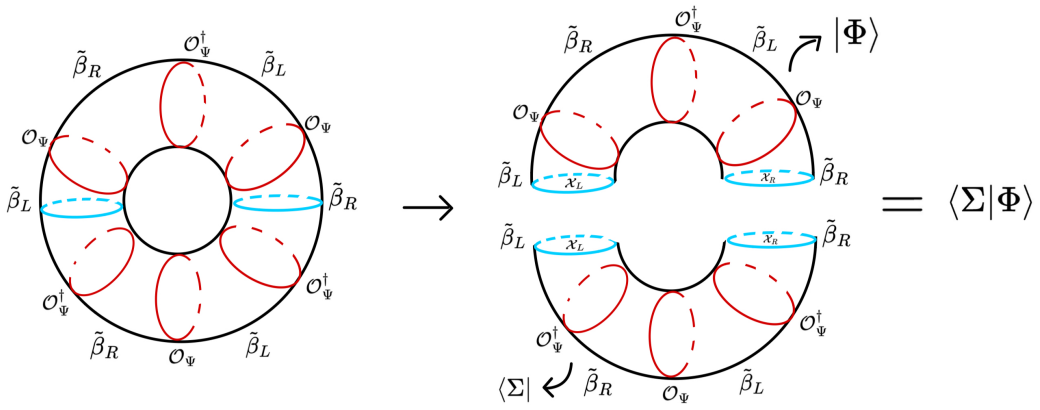


Figure 4.4: The path integral $\overline{Z_\circ(\rho_\Psi^n)(\tilde{\beta}_L, \tilde{\beta}_R)}$ rewritten as an overlap $\overline{\langle\Sigma|\Phi\rangle}$ by cutting it along $\mathcal{X}_L$ and $\mathcal{X}_R$.

pairs separated by $\tilde{\beta}_L + (p_i + 1)\beta_L$ on one side and $\tilde{\beta}_R + (k + 1)\beta_R$ on the other (Fig. 4.6b).

In the heavy mass limit the homology regions pinch off, the universal shell factors and strip path integrals cancel against the shell-state normalisations as in (1.2.21), and what remains is a Euclidean disk with the periodic operator structure described above. Taking $k, p_i \to -1$ collapses the operator separations to $\tilde{\beta}_{L,R}$ and the boundary becomes precisely that of $Z_\circ(\rho_\Psi^n)(\tilde{\beta}_L, \tilde{\beta}_R)$. The squaring trick lifts this to the fine-grained equality (4.3.3); cross-term saddles supported by additional $\mathcal{O}_\Psi$ insertions are handled in direct analogy with the constructions of Sec. 3.2.1, with full details in [3].

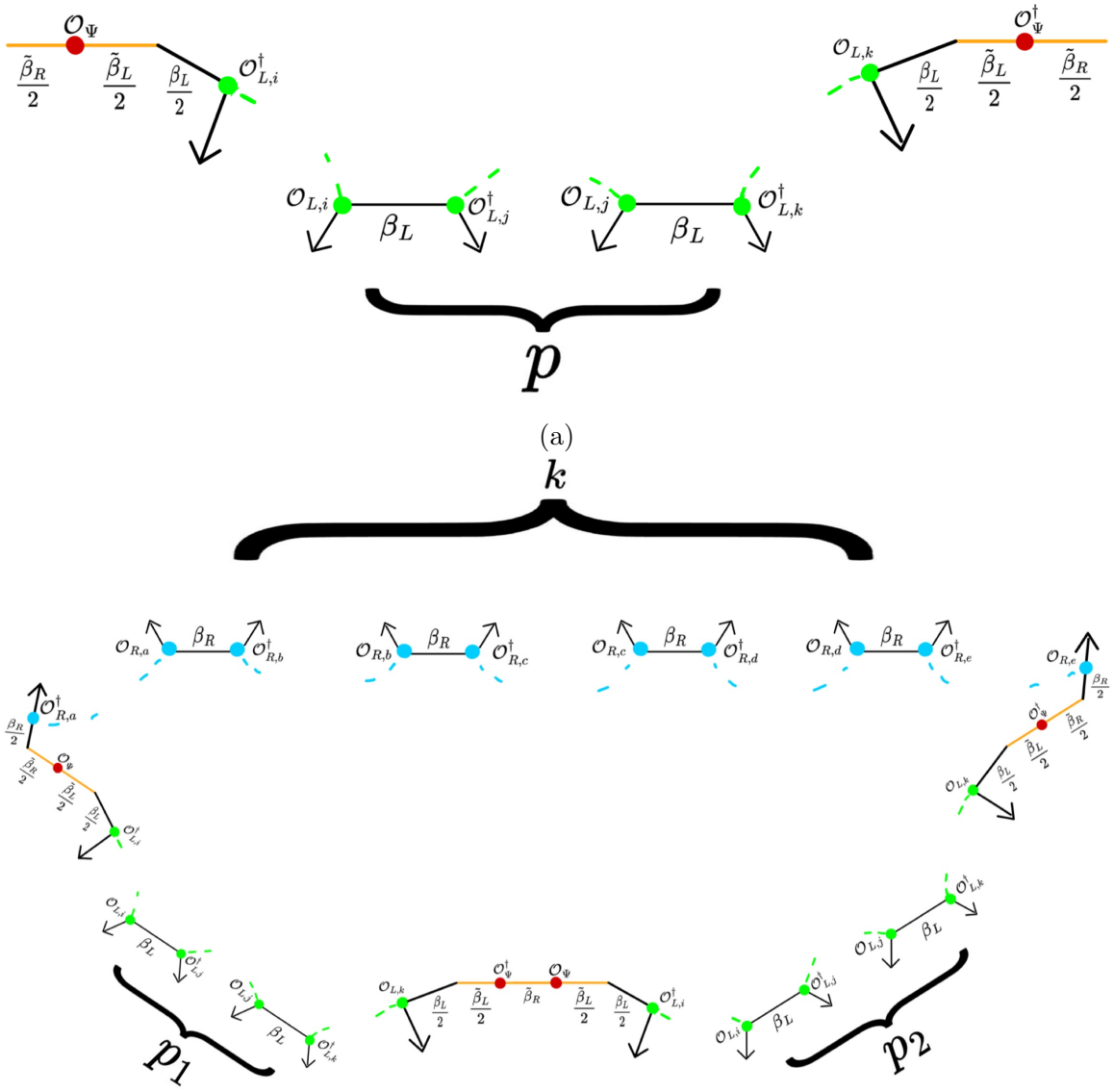
We restricted attention here to entanglement between disconnected boundary components. Extension to entanglement of subregions of a connected component is a natural direction for future work, with the operator-algebraic framework of [23] a natural starting point.

4.4 Discussion

The two derivations of $\mathrm{Tr} = \overline{Z}$ given in this chapter invoke logically distinct ingredients. The factorisation argument of Sec. 4.2.1 requires the mixed wormhole saddles connecting boundaries of single- and two-sided types that underlie the factorisation result of Chapter 3. The shell-basis argument of Sec. 4.2.2 requires only same-type wormholes connecting single-boundary states to one another. The latter is a strictly

weaker assumption: in principle, a theory of gravity could possess single-sided wormhole saddles — and hence Tr $= \overline{Z}$ — without admitting the mixed wormholes required for full Hilbert-space factorisation. The same dichotomy applies to the Rényi equality (4.3.3) and to the two-boundary trace factorisation (4.2.14). Whether the existence of one type of wormhole forces the existence of the other for consistency of the gravity path integral is an interesting question we leave open.

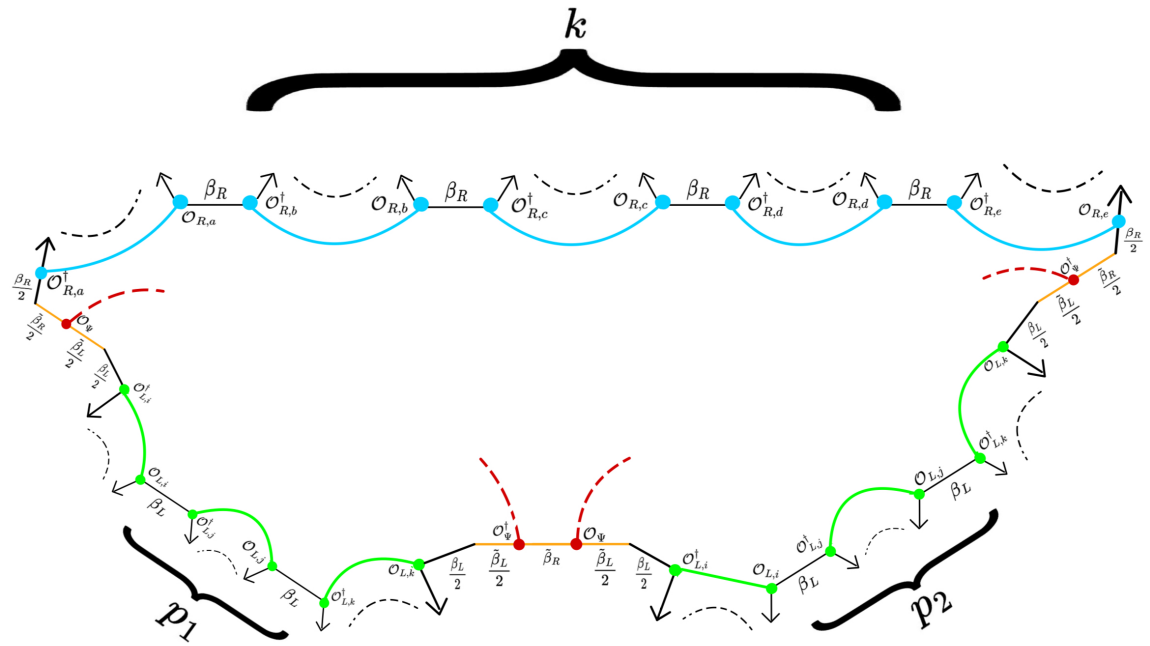
The arguments of this chapter make no holographic assumption and apply equally in asymptotically flat and AdS gravity. In holographic settings the equalities $\mathrm{Tr}_{\mathcal{H}_\mathfrak{X}}(e^{-\beta H}) = Z(\beta)$ and $\mathrm{Tr}_{\mathcal{H}_{\mathfrak{X}_R}}(\rho_R^n) = Z_\circ(\rho_\Psi^n)$ are tautological from the dual quantum mechanics; the non-trivial content here is that they can be derived directly from the gravitational path integral. The universality of the Bekenstein-Hawking area law shown in Sec. 4.2.3 is therefore equally a feature of asymptotically flat quantum gravity, which has no known holographic dual but admits the Euclidean path integral construction used here. The same applies to the gravitational interpretation of the holographic replica trick: the replica equality (4.3.3) does not require a CFT to make sense.



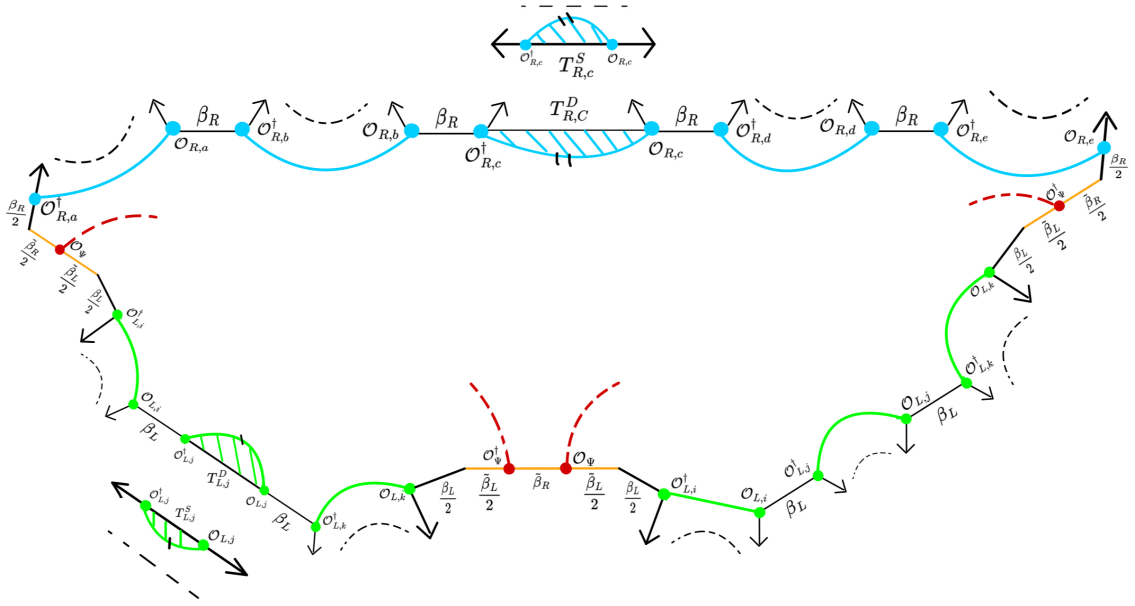
(a)

(b)

Figure 4.5: (**a**) Boundary condition for $\rho_R = \mathrm{Tr}_{\mathcal{H}_{\mathfrak{X}_L}} |\Psi\rangle\langle\Psi|$. (**b**) Boundary condition for $\mathrm{Tr}_{\mathcal{H}_{\mathfrak{X}_R}}(\rho_R^n)$ in the shell basis, depicted for $n = 2$.



(a)



(b)

Figure 4.6: (**a**) Maximally-connected saddle for (4.3.2). (**b**) Construction by gluing shell strips into a central disk along the shell worldvolumes (one L and one R shell shown for clarity).

"A man cannot inquire about that which he knows, because he knows it; and he cannot inquire about that which he does not know, for he does not know about what he is to inquire."
— *Plato,* Meno *80d–e*

Chapter 5

Decoding the Black Hole Interior

Chapter 4 established that the Bekenstein-Hawking formula counts black-hole microstates of the non-perturbative gravity Hilbert space. The construction of Chapters 2 and 3 provide an explicit basis of these microstates: the single- and two-sided shell states. A natural question follows: if such microstates are physical states of quantum gravity, can an asymptotic observer detect which microstate the universe is in? Classically, the horizon causally separates the interior from the exterior, and an asymptotic observer simply cannot. Quantum-mechanically the situation is more subtle, and several lines of argument suggest the microstates are detectable in principle, though only with sufficiently complex probes or sufficient patience: precise mass measurements work but take non-perturbatively long times [80]; resolving the data that distinguishes microstates from charges measurable at infinity requires Planck-scale precision [81]; and recent work from a quantum-information perspective shows that interior degrees of freedom are computationally complex to reconstruct from the boundary [56, 82–88]. A sharp version of this expectation was put forward in [89]: black-hole microstates are structurally complex combinations of the underlying degrees of freedom, so almost all probes give universal responses up to non-perturbatively small corrections, but *specially tuned* probes matched to the ac-

tual microstate yield a large response. The state is therefore expected to be hard to find but easy to verify — the structure of problems in the complexity class QMA, the quantum analogue of NP.

In this chapter we use the gravitational path integral to make this picture concrete for shell-state microstates. We construct asymptotic-boundary probes whose two-point functions $\langle \mathcal{O}_P \mathcal{O}_P^\dagger \rangle$ in a shell state $|i\rangle$ have the predicted structure: for a generic shell-operator probe $\mathcal{O}_P \neq \mathcal{O}_i$, the response is computed by a universal class of wormhole saddles whose action is independent of $\mathcal{O}_i$; for the fine-tuned choice $\mathcal{O}_P = \mathcal{O}_i$, an additional class of *detection* saddles contributes, increasing the response by an amount we compute. We carry this out for single-boundary shell states in Sec. 5.1, and for two-boundary shell states in Sec. 5.2, where two distinct probe strategies are available — coordinated insertions on both boundaries, or a single insertion localised to one boundary — and we work through both. All three settings exhibit the predicted universal-plus-detection structure. Sec. 5.3 then discusses what this means for the verification and search problems for shell states, and Sec. 5.4 concludes.

Three features of this analysis are worth flagging at the outset. First, the detection signal exists despite the shell worldvolume being causally disconnected from the asymptotic boundary in both type-A (black-hole-with-horizon) and type-B (compact closed-universe) shell states (Sec. 1.2.2). The signal arises from non-perturbative wormhole saddles that detect the fine quantum hair sourced by the shell, and is in this sense a non-perturbative effect of the kind that has driven all the Hilbert-space-level results in this thesis. Second, the construction extends to arbitrary superpositions of shell states; combined with the spanning result of Chapters 2 and 3, this means *any* non-perturbative gravity state can in principle be verified by a fine-tuned probe. Third, the two-boundary case has a particularly striking feature: a single-boundary probe localised on, say, the right asymptotic boundary can detect the state of a two-sided shell state, even though the bulk shell may not lie in the entanglement wedge of that boundary. The construction therefore probes regions of bulk geometry beyond the entanglement wedge of any single side — a feature with no perturbative analogue.

The results apply to any low-energy theory described by General Relativity, in any dimension, and to both asymptotically AdS and asymptotically flat gravity.

5.1 Observing a single-boundary universe

Suppose the universe is in the single-boundary shell state $|i\rangle$ of Sec. 1.2.2, prepared by inserting the shell operator $\mathcal{O}_i$ on a half-infinite Euclidean boundary and evolving for

time $\beta/2$ to the cut. We will show that there exist asymptotic-boundary probes that allow the operator $\mathcal{O}_i$ to be identified, despite the shell worldvolume being causally disconnected from the boundary — whether the dominant saddle for $|i\rangle$ is a black hole with the shell behind the horizon (type A) or thermal AdS with the shell on a disconnected closed universe (type B). The probe is a shell operator $\mathcal{O}_P$ of mass m_P inserted at $\tau = 0$, and the relevant observable is the magnitude squared of its expectation value,

$$\overline{|\langle i|\mathcal{O}_P|i\rangle|^2} \;=\; \overline{\langle i|\mathcal{O}_P|i\rangle\,\langle i|\mathcal{O}_P^\dagger|i\rangle}\;. \tag{5.1.1}$$

The single insertions $\overline{\langle i|\mathcal{O}_P|i\rangle}$ vanish in the saddlepoint approximation by the same orthogonality argument as in Sec. 1.2.1.1, but the magnitude squared captures the physical signal: the gravitational path integral computes coarse-grained averages, and (5.1.1) is the analogue of $\overline{|\langle i|j\rangle|^2} \neq 0$ encountered throughout this thesis.

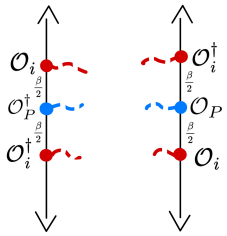
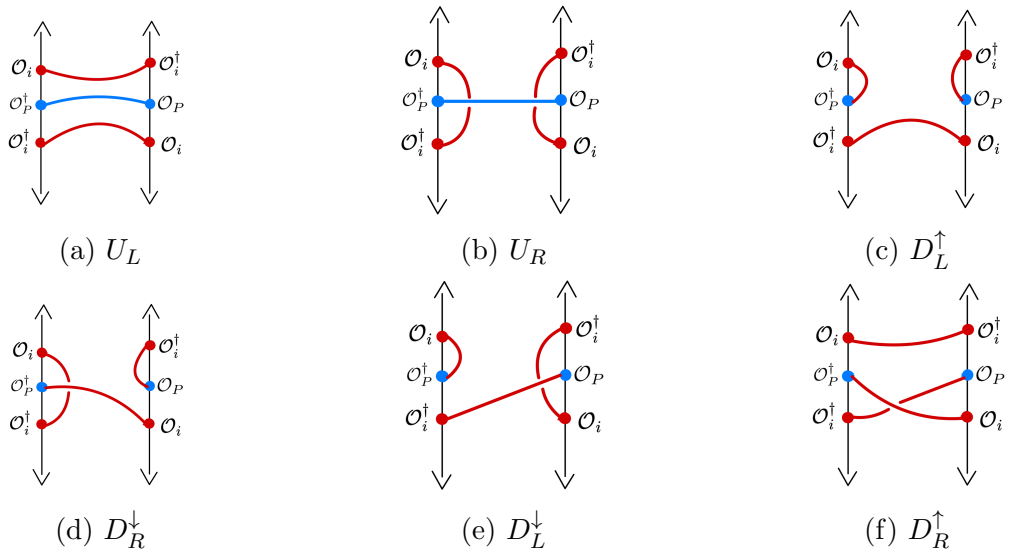


Figure 5.1: Asymptotic boundary condition for $\overline{\langle i|\mathcal{O}_P|i\rangle\langle i|\mathcal{O}_P^\dagger|i\rangle}$.

The boundary condition (Fig. 5.1) consists of two disconnected asymptotic pieces, each carrying one $\mathcal{O}_i$, one $\mathcal{O}_i^\dagger$, and one $\mathcal{O}_P$ (or $\mathcal{O}_P^\dagger$) insertion. The disconnected saddle vanishes by the orthogonality of the single-insertion overlaps, so all contributions come from wormhole saddles connecting the two boundaries. The structure of these saddles depends on whether the probe matches $\mathcal{O}_i$ or not. There are $3! = 6$ ways to pair up the three pairs of shell insertions across the two boundaries: two of these are compatible with any choice of $\mathcal{O}_P$ (the *universal* saddles), and four require $\mathcal{O}_P = \mathcal{O}_i$ (the *detection* saddles). The six classes are sketched in Fig. 5.2. We work in the heavy-mass limit for both $\mathcal{O}_i$ and $\mathcal{O}_P$ throughout, so each shell contributes a universal factor Z_m and the homology regions pinch off as in Sec. 1.2.1.4.

5.1.1 Universal saddles

When $\mathcal{O}_P \neq \mathcal{O}_i$, the probe shell on each boundary must annihilate against its conjugate on the other boundary, and the four $\mathcal{O}_i, \mathcal{O}_i^\dagger$ insertions must pair up among



(a) U_L (b) U_R (c) $D_L^{\uparrow}$

(d) $D_R^{\downarrow}$ (e) $D_L^{\downarrow}$ (f) $D_R^{\uparrow}$

Figure 5.2: The six classes of wormhole saddle contributing to $\overline{\langle i|\mathcal{O}_P|i\rangle\langle i|\mathcal{O}_P^\dagger|i\rangle}$. Top row: the two universal saddles $U_{L,R}$ that contribute for any $\mathcal{O}_P$. Lower two rows: the four detection saddles $D_{L,R}^{\uparrow,\downarrow}$ that contribute only when $\mathcal{O}_P = \mathcal{O}_i$.

themselves. Two pairings are compatible: each $\mathcal{O}_i$ on one boundary pairs with the $\mathcal{O}_i^\dagger$ on the same boundary (U_R, Fig. 5.2b), or with the $\mathcal{O}_i^\dagger$ on the other boundary (U_L, Fig. 5.2a).

U_L. This is the saddle in which the $\mathcal{O}_i$ shells, like the probe shell, pass between the two boundaries. The construction is precisely the wormhole discussed around (3.2.3). In the heavy-mass limit each shell contributes its universal factor and what remains is two disks of length β and two strips of length 0, see Fig. 5.3a:

$$U_L \;=\; \overline{Z}(\beta)^2 \times \overline{S}(0)^2 \times Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.2}$$

U_R. Here the $\mathcal{O}_i$ shells annihilate within their own boundary, and only the probe shell crosses between boundaries. The construction is a folded-wormhole-like geometry (Sec. 2.3.1): place i- and p-shell insertions along the boundary in alternating order on a disk of length 2β, identify the two $\mathcal{O}_P$ shell worldvolumes to fold the disk into a cylinder, and glue a shell strip into each of the $\mathcal{O}_i$ worldvolumes (Fig. 5.3b). In the heavy-mass limit,

$$U_R \;=\; \overline{Z}(2\beta) \times \overline{S}(0)^2 \times Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.3}$$

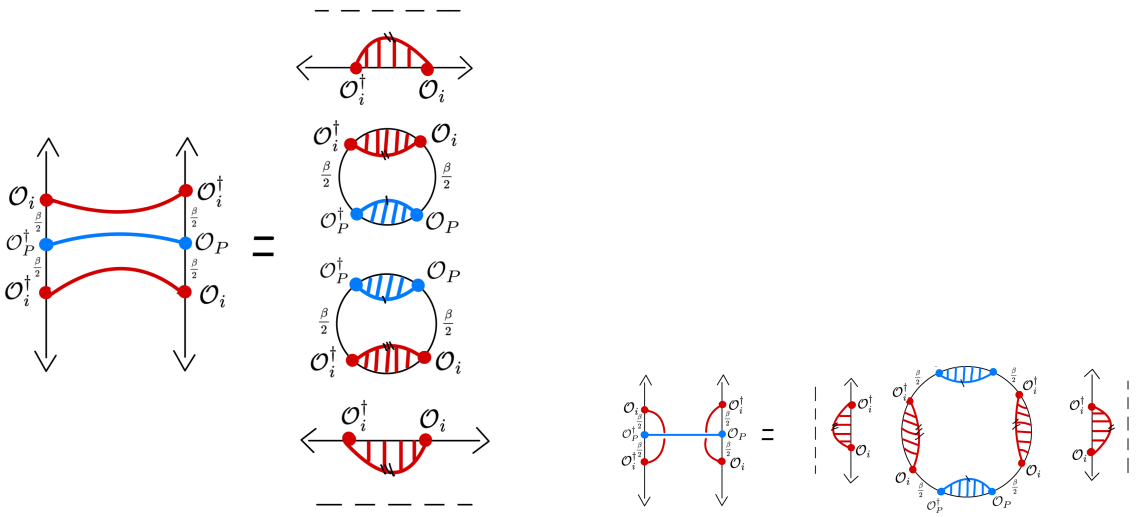
Total universal contribution. The universal piece is therefore

$$Z_U \;=\; U_L + U_R \;=\; \left(\overline{Z}(\beta)^2 + \overline{Z}(2\beta)\right) \overline{S}(0)^2\, Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.4}$$

After dividing by $\overline{\langle i|i\rangle^2} = (\overline{Z}(\beta)^2 + \overline{Z}(2\beta))\,\overline{S}(0)^2\,Z_{m_i}^2$ (computed by an analogous wormhole calculation, see Chapter 3), the normalised universal response is

$$\frac{Z_U}{\overline{\langle i|i\rangle^2}} \;=\; Z_{m_P}\;, \tag{5.1.5}$$

a function only of the probe shell mass and independent of the state $|i\rangle$. This is the universal-response prediction: a generic shell probe returns no information about the state.



(a) (b)

Figure 5.3: Construction of the universal saddles. (**a**) U_L by gluing two single-boundary norm saddles along the probe shell. (**b**) U_R by folding a four-shell disk along the probe worldvolumes and attaching shell strips at the $\mathcal{O}_i$ worldvolumes.

5.1.2 Detection saddles

When $\mathcal{O}_P = \mathcal{O}_i$, four further pairings open up. Label the boundary carrying $\mathcal{O}_P$ as B_1 and the boundary carrying $\mathcal{O}_P^\dagger$ as B_2, and write (X, Y) to mean the shell sourced by X is absorbed by Y. The four detection patterns are

$$D_L^{\uparrow} : (\mathcal{O}_P, \mathcal{O}_{i,B_1}^\dagger),\quad (\mathcal{O}_P^\dagger, \mathcal{O}_{i,B_2}),\quad (\mathcal{O}_{i,B_1}, \mathcal{O}_{i,B_2}^\dagger)\;, \tag{5.1.6}$$

$$D_R^{\uparrow} : (\mathcal{O}_P, \mathcal{O}_{i,B_2}^\dagger),\quad (\mathcal{O}_P^\dagger, \mathcal{O}_{i,B_1}),\quad (\mathcal{O}_{i,B_2}, \mathcal{O}_{i,B_1}^\dagger)\;, \tag{5.1.7}$$

$$D_L^{\downarrow} : (\mathcal{O}_P, \mathcal{O}_{i,B_2}^\dagger),\quad (\mathcal{O}_P^\dagger, \mathcal{O}_{i,B_2}),\quad (\mathcal{O}_{i,B_1}, \mathcal{O}_{i,B_1}^\dagger)\;, \tag{5.1.8}$$

$$D_R^{\downarrow} : (\mathcal{O}_P, \mathcal{O}_{i,B_1}^\dagger),\quad (\mathcal{O}_P^\dagger, \mathcal{O}_{i,B_1}),\quad (\mathcal{O}_{i,B_2}, \mathcal{O}_{i,B_2}^\dagger)\;. \tag{5.1.9}$$

Each pattern corresponds to a distinct wormhole topology, sketched in Fig. 5.2 (lower two rows) and constructed in Fig. 5.4.

$D_L^{\uparrow}$**.** Start with a strip carrying three shells separated by $\beta/2$, $\beta/2$ on the boundary, with propagation times T_{S1}, T_{S2}, T_{S3}. Glue a disk of boundary length $\beta/2 + T_{1D}$ into the first shell along its worldvolume, and similarly for the third shell with a disk of length $\beta/2 + T_{3D}$. Glue a shell strip into the second shell along its worldvolume. In the heavy-mass limit all propagation times vanish, the homology regions pinch off, and the saddle reduces to two disks of length $\beta/2$, one shell strip of length β, and one of length 0 (Fig. 5.4a):

$$D_L^{\uparrow} \;=\; \overline{Z}(\beta/2)^2\, \overline{S}(\beta)\, \overline{S}(0)\, Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.10}$$

$D_R^{\uparrow}$**.** Start with a central shell strip carrying five shells, labelled $1, \ldots, 5$, separated by boundary times $\beta/2$ and with propagation times $T_1, \ldots, T_5$. Glue a shell strip into the central strip along the worldvolume of shell 3; then glue shells 2 and 4 together by discarding their homology regions, and do the same for shells 1 and 5 (Fig. 5.4b). In the heavy-mass limit,

$$D_R^{\uparrow} \;=\; \overline{S}(2\beta)\, \overline{S}(0)\, Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.11}$$

$D_R^{\downarrow}$**.** Start with a central shell strip carrying four shells, labelled $1, \ldots, 4$ from top to bottom, separated by boundary times $\beta/2$ and with propagation times $T_1, \ldots, T_4$. Fold the central strip onto itself by identifying shells 1 and 3 along their worldvolumes, then glue a disk of boundary length $\beta/2 + T_D$ into the strip along the worldvolume of shell 4 and a shell strip of propagation time T_S along that of shell 2 (Fig. 5.4c). In the heavy-mass limit propagation times vanish and homology regions pinch off, leaving

$$D_R^{\downarrow} \;=\; \overline{S}(3\beta/2)\, \overline{Z}(\beta/2)\, \overline{S}(0)\, Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.12}$$

$D_L^{\downarrow}$**.** Constructed identically to $D_R^{\downarrow}$, with the roles of the two boundaries exchanged:

$$D_L^{\downarrow} \;=\; \overline{S}(3\beta/2)\, \overline{Z}(\beta/2)\, \overline{S}(0)\, Z_{m_i}^2\, Z_{m_P}\ . \tag{5.1.13}$$

Total detection contribution. Summing and normalizing,

$$Z_D = \frac{\left(2 \times \overline{Z}(\frac{\beta}{2}) \times \overline{S}(\frac{3\beta}{2}) \times \overline{S}(0) + \overline{S}(2\beta) \times \overline{S}(0) + \overline{Z}(\frac{\beta}{2})^2 \times \overline{S}(\beta) \times \overline{S}(0)\right)}{\left(\overline{Z}(\beta)^2 + \overline{Z}(2\beta)\right) \times \overline{S}(0)^2} \times Z_{m_P}. \tag{5.1.14}$$

5.1.3 Detection-to-universal ratio

The signal of interest is the ratio Z_D/Z_U, in which the shell-dependent factors $Z_{m_i}^2 Z_{m_P}$ and the $\overline{S}(0)^2$ from the shell-state normalisation cancel. Recall that the notation $\overline{S}(\alpha)$ is used to denote an infinite strip boundary condition of length $\lim_{\alpha\to\infty}[-\alpha,\beta+\alpha]$ [2]. The ratio $\frac{\overline{S}(\alpha)}{\overline{S}(0)} \equiv \hat{S}(\alpha)$ therefore corresponds to a "renormalized" strip action given by the volume of the Euclidean AdS strip of boundary length $[0,\alpha]$; i.e. dividing by $\overline{S}(0)$ subtracts off the infinite half line portions of the strip action. Furthermore, as the Euclidean strip has the same local geometry as thermal AdS but with uncompactified time, we have $\hat{S}(\alpha) = \overline{Z}_{TAdS}(\alpha)$ where $\overline{Z}_{TAdS}(\beta) \equiv e^{-I_{TAdS}(\beta)}$. Then

$$\frac{Z_D}{Z_U} \;=\; \frac{2\,\overline{Z}(\beta/2)\,\overline{Z}_{TAdS}(3\beta/2) + \overline{Z}_{TAdS}(2\beta) + \overline{Z}(\beta/2)^2\,\overline{Z}_{TAdS}(\beta)}{\overline{Z}(\beta)^2 + \overline{Z}(2\beta)}\,. \tag{5.1.15}$$

The size of (5.1.15) depends on which saddles dominate each $\overline{Z}$. For type-B states ($\beta/2 > \beta_{HP}$), thermal AdS dominates everywhere and $\overline{Z}_{TAdS}(nx) = \overline{Z}_{TAdS}(x)^n$ gives $Z_D/Z_U \approx 2$. For type-A states ($\beta < \beta_{HP}$), the dominant term is the third in the numerator, and $Z_D/Z_U \approx \overline{Z}_{BH}(\beta/2)^2\,\overline{Z}_{TAdS}(\beta)/\overline{Z}_{BH}(\beta)^2$ — a regime-dependent ratio which we now extract more cleanly by projecting onto microcanonical bands.

Microcanonical projection. We obtain a more illuminating expression for the ratio (5.1.15) by projecting the shell states into a microcanonical band. Expanding a canonical shell state in the energy basis,

$$|i,\beta\rangle \;\equiv\; e^{-\beta H/2}|\hat{\imath}\rangle \;\approx\; \int_0^\infty d\tilde{E}\;\rho(\tilde{E})\,e^{-\beta\tilde{E}/2}\,\langle\tilde{E}|\hat{\imath}\rangle\,|\tilde{E}\rangle\,, \tag{5.1.16}$$

the microcanonical projection onto the window $[E-\Delta E/2, E+\Delta E/2]$ is implemented by a Laplace transform,

$$|i,\beta,E\rangle \;=\; \int_{E-\Delta E/2}^{E+\Delta E/2} d\tilde{E}\,e^{-\beta\tilde{E}/2}\int d\tilde{\beta}\,e^{\tilde{E}\tilde{\beta}/2}\,|i,\tilde{\beta}\rangle\,. \tag{5.1.17}$$

The microcanonical probe correlator $\overline{|\langle i,\beta,E|\mathcal{O}_P|i,\beta,E\rangle|^2}$ is then a four-fold Laplace transform of the canonical correlator, one factor for each of the four preparation temperatures $\beta_1,\ldots,\beta_4$ associated with the four boundaries carrying $\mathcal{O}_i$ insertions,

$$\overline{|\langle i,\beta,E|\mathcal{O}_P|i,\beta,E\rangle|^2} \;= \\ \int_{E-\Delta E/2}^{E+\Delta E/2}\prod_{k=1}^4 dE_k\,e^{-\beta\sum_k E_k/2}\int\prod_{k=1}^4 d\beta_k\,e^{\sum_k E_k\beta_k/2}\;\times\;\overline{\langle\beta_1|\mathcal{O}_P|\beta_2\rangle\langle\beta_3|\mathcal{O}_P^\dagger|\beta_4\rangle}\,. \tag{5.1.18}$$

The detection and universal contributions to $\overline{\langle\beta_1|\mathcal{O}_P|\beta_2\rangle\langle\beta_3|\mathcal{O}_P^\dagger|\beta_4\rangle}$ are read off from (5.1.4) and (5.1.14) with the four preparation temperatures distinguished:

$$\frac{Z_{D,\text{micro}}}{\overline{S}(0)^2\, Z_{m_i}^2\, Z_{m_P}} = \Big(\overline{Z}(\beta_1/2)\,\overline{Z}(\beta_3/2)\,\overline{Z}_{TAdS}(\frac{\beta_2+\beta_4}{2}) + \overline{Z}_{TAdS}(\frac{\beta_1+\beta_2+\beta_3+\beta_4}{2}) + \overline{Z}(\beta_1/2)\,\overline{Z}_{TAdS}(\frac{\beta_2+\beta_3+\beta_4}{2}) + \overline{Z}(\beta_4/2)\,\overline{Z}_{TAdS}(\frac{\beta_1+\beta_2+\beta_3}{2})\Big)\,, \tag{5.1.19}$$

$$\frac{Z_{U,\text{micro}}}{\overline{S}(0)^2\, Z_{m_i}^2\, Z_{m_P}} = \Big(\overline{Z}(\frac{\beta_1+\beta_2+\beta_3+\beta_4}{2}) + \overline{Z}(\frac{\beta_1+\beta_4}{2})\,\overline{Z}(\frac{\beta_2+\beta_3}{2})\Big)\,. \tag{5.1.20}$$

Evaluating each Laplace transform in the saddlepoint approximation,

$$\int dx\, e^{Ex}\,\overline{Z}(x) \;\approx\; \sqrt{2\pi/h}\, e^{S_{BH}(E)}\,, \tag{5.1.21}$$

where $h = \partial_x^2 \ln \overline{Z}(x)|_{x=x^*}$ is the Hessian at the saddlepoint x^* and $S_{BH}(E) = (1 - \beta\partial_\beta)\ln\overline{Z}(\beta)|_{\beta=\beta(x^*)}$. Restricting to windows with $E \gg \Delta E$,

$$\int_{E-\Delta E/2}^{E+\Delta E/2} d\tilde{E}\sqrt{2\pi/h}\, e^{S_{BH}(\tilde{E})} \;\approx\; \Delta E\,\sqrt{2\pi/h}\, e^{S_{BH}(E)} \;\equiv\; e^{\mathbf{S}(E)}\,, \tag{5.1.22}$$

and

$$\int_{E-\Delta E/2}^{E+\Delta E/2} d\tilde{E}\sqrt{2\pi/h}\, e^{S_{BH}(\tilde{E})}\, e^{-\beta\tilde{E}} \;\approx\; e^{-\beta E}\, e^{\mathbf{S}(E)}\,. \tag{5.1.23}$$

Applying these identities to $Z_{D,\text{micro}}$ and $Z_{U,\text{micro}}$, the β-dependent prefactors collapse identically on numerator and denominator and the ratio reduces to

$$\left.\frac{Z_D}{Z_U}\right|_{\text{micro}} = \frac{e^{2\mathbf{S}(E)} + 2\, e^{\mathbf{S}(E)} + 1}{e^{2\mathbf{S}(E)} + e^{\mathbf{S}(E)}} \;=\; 1 + e^{-\mathbf{S}(E)}\,. \tag{5.1.24}$$

Above the black-hole threshold, $e^{\mathbf{S}(E)} \gg 1$ and $Z_D/Z_U \to 1$: detection saddles match universal saddles term for term, and the response doubles when $\mathcal{O}_P = \mathcal{O}_i$. Below threshold, $e^{\mathbf{S}(E)} = 1$ and $Z_D/Z_U = 2$, so the response triples. Either way, the $\mathcal{O}(1)$ jump is unambiguous.

This establishes the central claim of the section: although a generic shell probe gives a state-independent response, a probe matched to the operator that prepared the state gives a response larger by a calculable $\mathcal{O}(1)$ factor. The signal exists in both the type-A regime (where the shell is behind a horizon) and the type-B regime (where it is in a disconnected closed universe), and its strength is set by the microcanonical entropy.

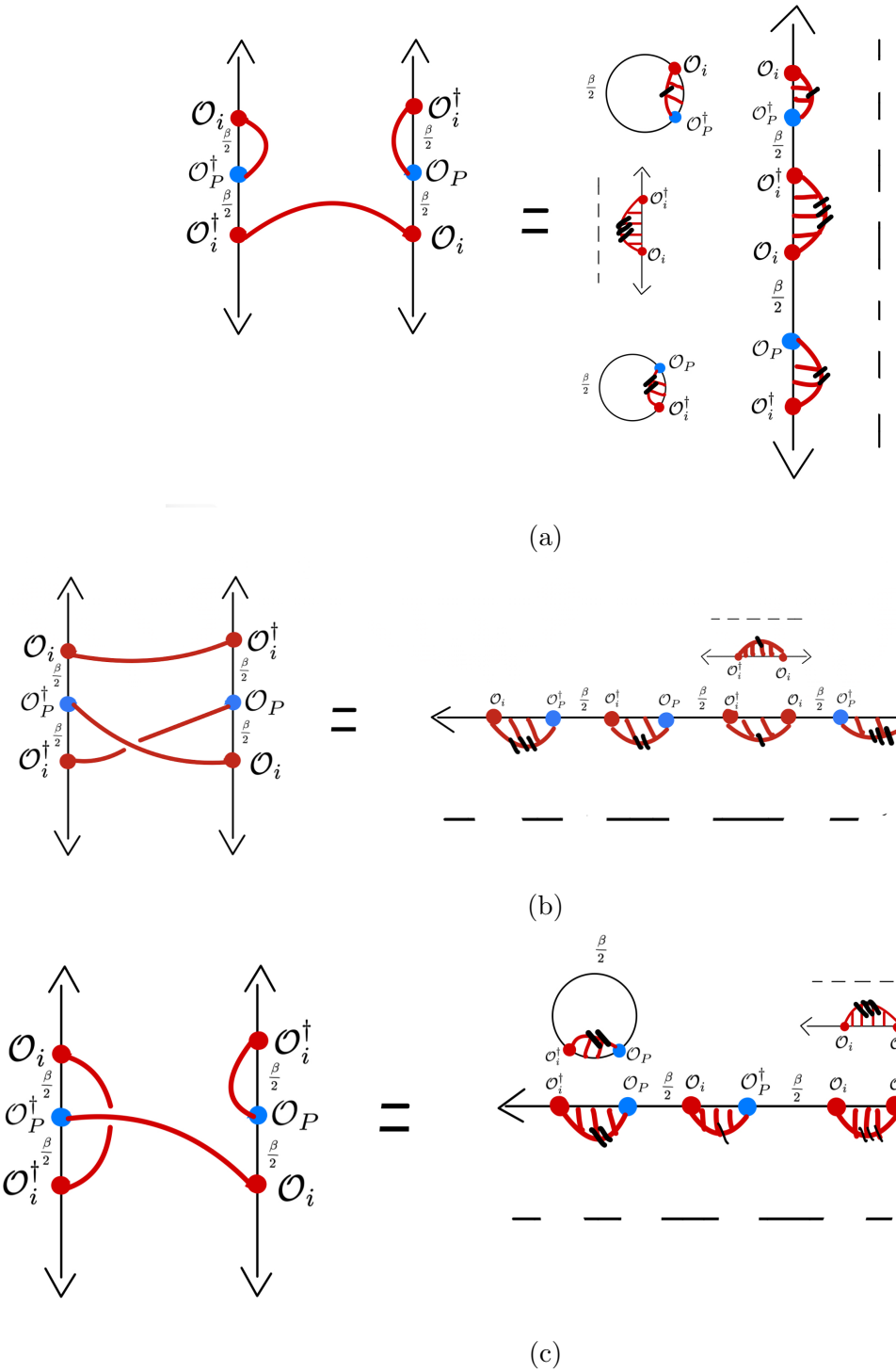


Figure 5.4: Construction of the detection saddles. (**a**) $D_L^{\uparrow}$. (**b**) $D_R^{\uparrow}$. (**c**) $D_R^{\downarrow}$ (and analogously $D_L^{\downarrow}$).

5.2 Observing a two-boundary universe

We now turn to two-boundary shell states $|\mathbf{i}\rangle_{2s}$ of Sec. 1.2.1, prepared by inserting a shell operator $\mathcal{O}_\mathbf{i}$ on a periodic Euclidean boundary at distances $\beta_L/2, \beta_R/2$ from the two cuts. There are two natural probe strategies for such a state. First, one can insert a probe operator and its conjugate on opposite asymptotic boundaries simultaneously, exploiting the fact that the bulk shell lies in the entanglement wedge of the union but not necessarily of either boundary alone (Sec. 5.2.1). Second, one can probe with a single operator localised on just one boundary, in direct analogy with the single-boundary construction of Sec. 5.1 (Sec. 5.2.2). Both strategies admit a universal-plus-detection structure, but the strengths and microcanonical scaling of the resulting signals are different.

5.2.1 Coordinated two-boundary probe

Consider inserting $\mathcal{O}_P^\dagger$ on the left boundary and $\mathcal{O}_P$ on the right boundary at $\tau = 0$, and computing

$$\overline{\langle \mathbf{i}|\mathcal{O}_{P,L}^\dagger\, \mathcal{O}_{P,R}|\mathbf{i}\rangle}\ , \tag{5.2.1}$$

in the boundary configuration of Fig. 5.5. The boundary condition has a single connected component, with four shell insertions in total: one $\mathcal{O}_\mathbf{i}$ pair from the state preparation and one mixed pair $\mathcal{O}_{P,L}^\dagger, \mathcal{O}_{P,R}$ from the probe. There are two kinds of saddle: *universal* ones that contribute for any probe operator, and *detection* ones that contribute only when $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$.

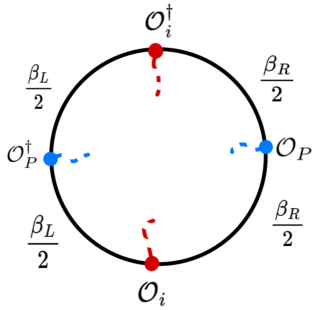


Figure 5.5: Asymptotic boundary condition for the coordinated two-boundary probe correlator (5.2.1).

Universal saddles. On a disk with four shell insertions $\mathcal{O}_\mathbf{i}, \mathcal{O}_P^\dagger, \mathcal{O}_\mathbf{i}^\dagger, \mathcal{O}_P$ in this cyclic order separated by boundary times $\beta_L/2, \beta_L/2, \beta_R/2, \beta_R/2$, identify the two $\mathcal{O}_\mathbf{i}$ shell worldvolumes and the two $\mathcal{O}_P$ shell worldvolumes. The resulting saddle has the

topology of a disk with one handle (a punctured torus): the two shells pass through the bulk without intersecting, with one taking the handle. In the large shell mass limit (Sec. 1.2.1.4), the saddle reduces to a single Z-saddle of length $\beta_L + \beta_R$:

$$Z_U \;=\; \overline{Z}(\beta_L + \beta_R)\, Z_{m_i}\, Z_{m_P} \;. \tag{5.2.2}$$

Detection saddles. When $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$, the $\mathcal{O}_P$ probe shell can also be absorbed by an $\mathcal{O}_\mathbf{i}^\dagger$ insertion. Construct a central disk with two shells separated by $\beta_L/2$ and $\beta_R/2$, glue an additional disk of length $\beta_L/2$ into one shell along its worldvolume, and another of length $\beta_R/2$ into the other (Fig. 5.6b). In the large shell mass limit,

$$Z_D \;=\; \overline{Z}(\beta_L/2)\, \overline{Z}(\beta_R/2)\, \overline{Z}((\beta_L + \beta_R)/2)\, Z_{m_i}\, Z_{m_P} \;. \tag{5.2.3}$$

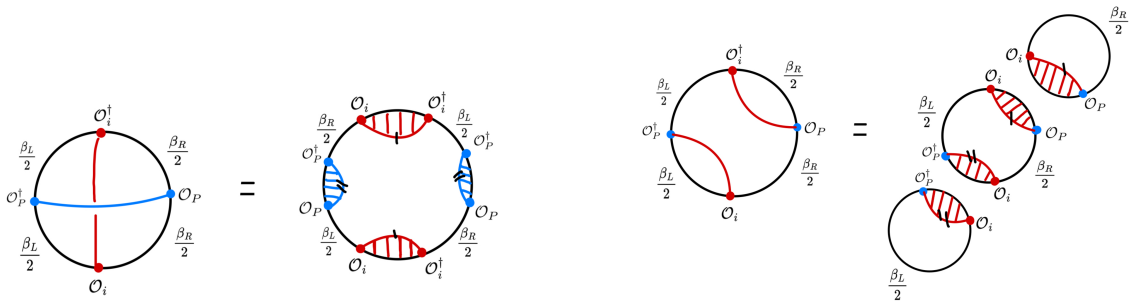


Figure 5.6: Saddle contributions to (5.2.1). (**a**) The universal saddle, of disk-with-handle topology. (**b**) The detection saddle, present only when $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$.

Detection-to-universal ratio. Taking the ratio,

$$\frac{Z_D}{Z_U} \;=\; \frac{\overline{Z}(\beta_L/2)\, \overline{Z}(\beta_R/2)\, \overline{Z}((\beta_L + \beta_R)/2)}{\overline{Z}(\beta_L + \beta_R)} \;. \tag{5.2.4}$$

For low temperatures $\beta_L + \beta_R > \beta_{HP}$ (thermal AdS dominant on the denominator while black-hole saddles can still dominate the numerator), the ratio is $\mathcal{O}(1)$; for high temperatures $\beta_L + \beta_R < \beta_{HP}$ (where black-hole saddles dominate everywhere), Z_D/Z_U is exponentially large.

As in the single-boundary case, the clean version of the result comes from projecting the shell states into microcanonical bands. The two-boundary projector is now a double Laplace transform, one for each boundary,

$$|i, \beta_{L,R}, E_{L,R}\rangle \;=\; \int_{E_L - \Delta E_L/2}^{E_L + \Delta E_L/2} \int_{E_R - \Delta E_R/2}^{E_R + \Delta E_R/2} d\tilde{E}^{L,R}\, e^{-\beta_{L,R}\tilde{E}^{L,R}/2} \int d\tilde{\beta}_{L,R}\, e^{\tilde{E}^{L,R}\tilde{\beta}_{L,R}/2}\, |i, \tilde{\beta}_{L,R}\rangle \;, \tag{5.2.5}$$

and the microcanonical correlator is a four-fold Laplace transform of the canonical one with E_1, E_3 integrated over the L window and E_2, E_4 over the R window:

$$\begin{aligned}\overline{\langle i,\beta_{L,R},E_{L,R}|\mathcal{O}^\dagger_{P,L}\mathcal{O}_{P,R}|i,\beta_{L,R},E_{L,R}\rangle}\\ = \int_{E_L-\Delta E_L/2}^{E_L+\Delta E_L/2} dE_1\,dE_3\,e^{-\beta_L(E_1+E_3)/2}\int_{E_R-\Delta E_R/2}^{E_R+\Delta E_R/2} dE_2\,dE_4\,e^{-\beta_R(E_2+E_4)/2}\\ \times\int\prod_{k=1}^{4}d\beta_k\,e^{\sum_k E_k\beta_k/2}\,\overline{\langle\beta_1,\beta_2|\mathcal{O}^\dagger_{P,L}\mathcal{O}_{P,R}|\beta_3,\beta_4\rangle}\ .\end{aligned} \tag{5.2.6}$$

The detection-to-universal ratio with the four preparation temperatures kept distinct, read off from (5.2.3) and (5.2.2), is

$$\frac{Z_D}{Z_U} \;=\; \frac{\overline{Z}(\beta_2/2)\,\overline{Z}(\beta_3/2)\,\overline{Z}((\beta_1+\beta_4)/2)}{\overline{Z}((\beta_1+\beta_2+\beta_3+\beta_4)/2)}\ . \tag{5.2.7}$$

The third factor in the numerator, $\overline{Z}((\beta_1+\beta_4)/2)$, mixes L and R preparation temperatures — the corresponding saddle is the one in which the detection wormhole connects the two boundaries through a disk of length $(\beta_L+\beta_R)/2$, with β_1 coming from the L side and β_4 from the R side. This is the source of the window-overlap requirement that emerges below.

Evaluating the Laplace transforms in the saddlepoint approximation using (5.1.21)–(5.1.23), the individual contributions are

$$Z_D \;=\; Z_{m_i}\,Z_{m_P}\,e^{-(\beta_L E_L/2+\beta_R E_R/2)}\,e^{\mathbf{S}(E_L)+\mathbf{S}(E_R)}\,\Phi(E_L,\Delta_L,E_R,\Delta_R,\beta_L,\beta_R) \tag{5.2.8}$$

and

$$Z_U \;=\; Z_{m_i}\,Z_{m_P}\,\Phi(E_L,\Delta_L,E_R,\Delta_R,2\beta_L,2\beta_R)\ , \tag{5.2.9}$$

where

$$\begin{aligned}&\Phi(E_L,\Delta_L,E_R,\Delta_R,\beta_L,\beta_R)\;=\\ &\int_{E_L-\Delta E_L/2}^{E_L+\Delta E_L/2} dE_1\,e^{-\beta_L E_1/2}\int_{E_R-\Delta E_R/2}^{E_R+\Delta E_R/2} dE_4\,e^{-\beta_R E_4/2}\,\delta(E_4-E_1)\,e^{\mathbf{S}((E_1+E_4)/2)}\ .\end{aligned} \tag{5.2.10}$$

As a result, Φ vanishes unless the L and R microcanonical windows overlap. When they do, the ratio is

$$\left.\frac{Z_D}{Z_U}\right|_{\text{micro}} \sim\; e^{\mathbf{S}(E_L)+\mathbf{S}(E_R)}\;\times\;\frac{\Phi(E_L,\Delta_L,E_R,\Delta_R,\beta_L,\beta_R)}{\Phi(E_L,\Delta_L,E_R,\Delta_R,2\beta_L,2\beta_R)}\ , \tag{5.2.11}$$

the exponential factor dominating once either $E_{L,R}$ is above the black-hole threshold. The detection signal is then exponentially larger than the universal response — in

stark contrast to the $\mathcal{O}(1)$ enhancement of the single-boundary case (5.1.24). Below threshold $e^{\mathbf{S}(E_{L,R})} = 1$ and the ratio is $\mathcal{O}(1)$.

The exponentially large enhancement reflects the fact that this probe does not require the wormhole-detection mechanism at all: the universal and detection saddles have different topologies (disk-with-handle vs. simply-connected) and the detection one is "classically larger" once the windows overlap. The corresponding picture in AdS/CFT is intuitive: the bulk shell lies in the entanglement wedge of the union of the two boundaries [90–98], and a probe supported on both sides accesses it directly. The interesting question, addressed next, is whether one can detect the same state with a probe localized on only one side.

5.2.2 Single-boundary probe of a two-boundary state

Now insert the probe $\mathcal{O}_P$ at $\tau = 0$ on just one boundary — say the right — and compute the magnitude squared

$$\overline{\langle \mathbf{i}|\mathcal{O}_R|\mathbf{i}\rangle \, \langle \mathbf{i}|\mathcal{O}_R^\dagger|\mathbf{i}\rangle} \;=\; \overline{|\langle \mathbf{i}|\mathcal{O}_R|\mathbf{i}\rangle|^2} \,. \tag{5.2.12}$$

The boundary condition (Fig. 5.7) consists of two disconnected closed boundaries, each carrying a $\mathcal{O}_\mathbf{i}$ pair on the left and right sides, and one $\mathcal{O}_R$ (or $\mathcal{O}_R^\dagger$) probe insertion. As in the single-boundary case (Sec. 5.1), the disconnected single-boundary saddles vanish and all contributions come from wormholes connecting the two boundaries. The wormholes split again into universal saddles compatible with any $\mathcal{O}_P$, and detection saddles that contribute only when $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$. The construction here requires a constraint not invoked in earlier sections: only saddles in which the orientation of Euclidean time-flow can be smoothly continued from the boundary into the bulk contribute (see App. B of [32]). We have not been explicit about this requirement in the preceding chapters as it does not lead to subtleties there. This excludes some otherwise-natural pairings; the surviving classes are detailed below.

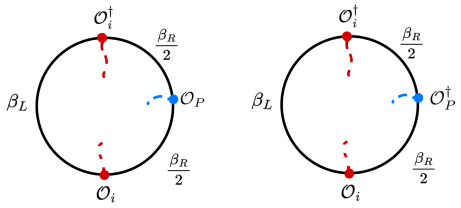


Figure 5.7: Asymptotic boundary condition for the single-boundary probe (5.2.12) of a two-boundary state.

Universal saddles. Two classes of universal wormhole saddle survive the orientation constraint. In the first (U_1, Fig. 5.8a), all three shell pairs propagate from one boundary to the other through a cylinder. Cutting the cylinder along the shell worldvolumes yields three sheets, each obtainable as a portion of a Z-saddle disk. The disk lengths are fixed by the boundary structure to be $2\beta_L, \beta_R, \beta_R$, giving

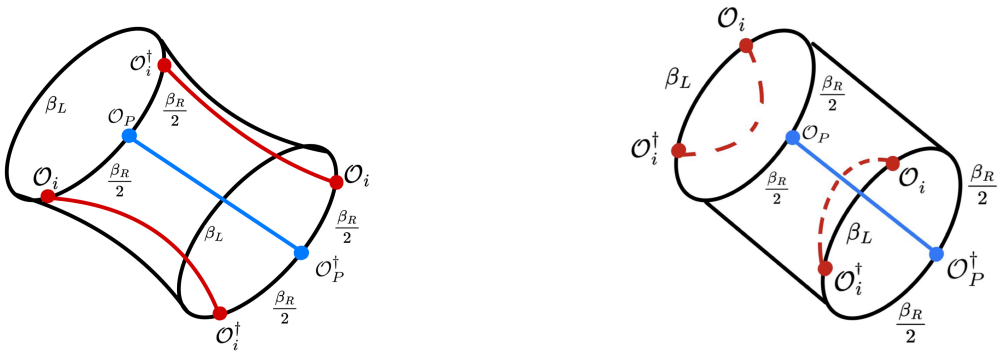
$$U_1 \;=\; \overline{Z}(2\beta_L)\,\overline{Z}(\beta_R)^2\,Z_{m_i}^2\,Z_{m_P}\;. \tag{5.2.13}$$

In the second (U_2, Fig. 5.8b), only the probe shell crosses between boundaries, while each $\mathcal{O}_\mathbf{i}$ pair annihilates within its own boundary. The construction is a folded-wormhole-style geometry analogous to the U_R saddle of Sec. 5.1.1. Start with a disk carrying two P-shell and two $\mathbf{i}$-shell insertions in alternating order $P, \mathbf{i}, P, \mathbf{i}$ on the boundary, and fold it into a cylinder by discarding the P-shell homology regions and identifying two P-shell worldvolumes. Then take two further disks, each of boundary length $\beta_L + \Delta T$ and carrying an $\mathbf{i}$-shell, and glue each into one end of the cylinder along the respective $\mathbf{i}$-shell worldvolume. In the heavy-mass limit all propagation times $\Delta T \to 0$ and the homology regions pinch off, leaving a central segment of length $2\beta_R$ (from the folded disk) joined to two disks of length β_L:

$$U_2 \;=\; \overline{Z}(2\beta_R)\,\overline{Z}(\beta_L)^2\,Z_{m_i}^2\,Z_{m_P}\;. \tag{5.2.14}$$

The total universal contribution is

$$Z_U \;=\; \left(\overline{Z}(2\beta_L)\,\overline{Z}(\beta_R)^2 + \overline{Z}(2\beta_R)\,\overline{Z}(\beta_L)^2\right) Z_{m_i}^2\,Z_{m_P}\;. \tag{5.2.15}$$



(a) (b)

Figure 5.8: The two universal saddles for the single-boundary probe of a two-boundary state. (**a**) U_1: all three shells cross between boundaries through a cylinder. (**b**) U_2: only the probe shell crosses; the $\mathcal{O}_\mathbf{i}$ shells annihilate within each boundary.

Detection saddles. Four classes of detection saddle contribute when $\mathcal{O}_P = \mathcal{O}_\mathbf{i}$. Three are direct adaptations of the universal saddles with the probe-to-shell pairings rerouted (Fig. 5.9); their actions are read off from the corresponding Z-disk lengths:

$$D_1 = \overline{Z}(\beta_R/2)^2\,\overline{Z}(\beta_R+2\beta_L)\,Z_{m_i}^2\,Z_{m_P}\,, \tag{5.2.16}$$

$$D_2 = D_3 = \overline{Z}(\beta_R/2)\,\overline{Z}(\beta_L)\,\overline{Z}(\beta_L+3\beta_R/2)\,Z_{m_i}^2\,Z_{m_P}\,. \tag{5.2.17}$$

The fourth class, D_4, is genuinely novel: a saddle of twice-punctured-torus topology in which both shells pass between the two boundaries through different handles, allowing the orientation of Euclidean time to remain consistent (Fig. 5.10). In the heavy-mass limit,

$$D_4 = \overline{Z}(2\beta_L+2\beta_R)\,Z_{m_i}^2\,Z_{m_P}\,. \tag{5.2.18}$$

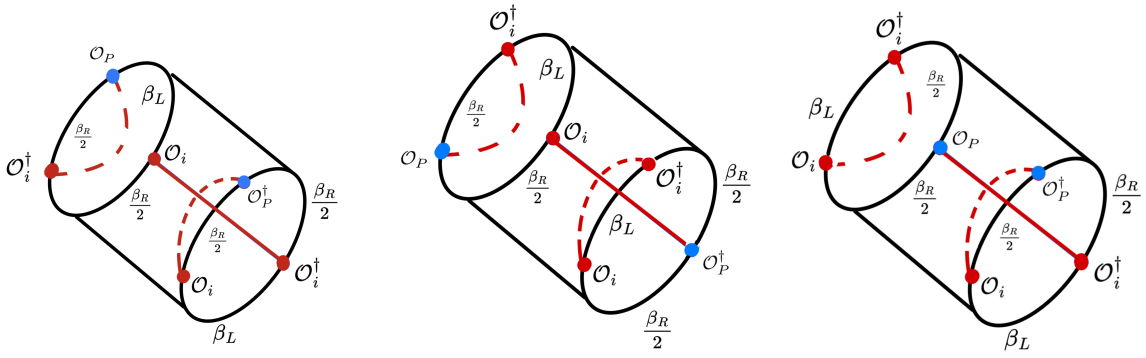


Figure 5.9: The first three classes of detection saddle for the single-boundary probe of a two-boundary state, $D_{1,2,3}$.

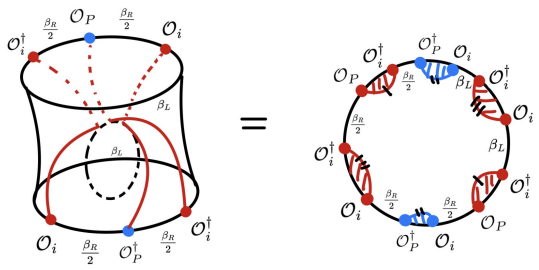


Figure 5.10: The fourth class of detection saddle, D_4, of twice-punctured-torus topology.

Detection-to-universal ratio. Summing $Z_D = D_1 + D_2 + D_3 + D_4$ and forming the ratio,

$$\frac{Z_D}{Z_U} = \frac{\overline{Z}(\beta_R/2)^2\,\overline{Z}(\beta_R+2\beta_L) + 2\,\overline{Z}(\beta_R/2)\,\overline{Z}(\beta_L)\,\overline{Z}(\beta_L+3\beta_R/2) + \overline{Z}(2\beta_L+2\beta_R)}{\overline{Z}(2\beta_R)\,\overline{Z}(\beta_L)^2 + \overline{Z}(2\beta_L)\,\overline{Z}(\beta_R)^2}\,, \tag{5.2.19}$$

which depends on the relative sizes of β_L and β_R in a complicated way. As in the previous cases, a cleaner expression emerges from projecting the shell states into microcanonical windows $[E_{L,R} - \Delta E_{L,R}/2, E_{L,R} + \Delta E_{L,R}/2]$ via the double Laplace transform (5.2.5). The Laplace transforms of the individual $\overline{Z}$ factors in (5.2.19) are evaluated using (5.1.21)–(5.1.23), exactly as in Sec. 5.2.1. We assume for ease of calculation that $E_L = E_R \equiv E$ and $\Delta_L = \Delta_R \equiv \Delta$; without this choice the L and R microcanonical windows would have to overlap to give a nonzero result, by the same L, R mixing mechanism as in Sec. 5.2.1.[1]

Under this assumption, every term in both numerator and denominator of (5.2.19) yields the same exponential prefactor $e^{-2E(\beta_L+\beta_R)+\mathbf{S}(E)}$ after the saddlepoint Laplace transforms: each contains exactly one factor of $\mathbf{S}(E)$ from the saddlepoint integration over the bounded E window, and the β-dependent exponential weights collapse against the canonical exponents of the $\overline{Z}$ factors. The exponential prefactor cancels between numerator and denominator, leaving only the multiplicities of terms in (5.2.19) — three in the numerator (counting $D_2 = D_3$ once each), two in the denominator — and

$$\left.\frac{Z_D}{Z_U}\right|_{\text{micro}} = \frac{3}{2}\,. \tag{5.2.20}$$

The detection signal therefore again sits at an $\mathcal{O}(1)$ enhancement over the universal response, as in the genuine single-boundary case (5.1.24). This is the central observation of this subsection: *the state of a two-sided shell state can be detected by a probe localised on just one boundary*, even though the bulk shell may be inaccessible to that boundary's entanglement wedge. Compared with the coordinated probe of Sec. 5.2.1, the signal is much smaller — $\mathcal{O}(1)$ rather than exponential — because detection here proceeds via wormhole saddles rather than entanglement-wedge access. As in Sec. 5.2.1, an overlap between the L and R microcanonical windows is required for the signal to be nonzero.

[1] The mixing arises here because the partition functions in the numerator and denominator of (5.2.19) carry arguments that combine β_L and β_R — e.g. $\overline{Z}(\beta_R + 2\beta_L)$ and $\overline{Z}(2\beta_L + 2\beta_R)$. Choosing $E_L = E_R$ trivially satisfies the resulting energy-window-overlap constraint.

5.3 Verification, superpositions, and the search problem

The previous two sections established a common pattern: probing a shell state $|i\rangle$ with a generic probe $\mathcal{O}_P \neq \mathcal{O}_i$ produces a universal response carrying no information about the state, while probing with the matched operator $\mathcal{O}_P = \mathcal{O}_i$ yields an additional detection contribution that increases the response by a calculable factor. This is the structure of a verification problem: a proposed state is easy to test even when finding the state from scratch is hard. We make this picture precise in three steps: first, we describe the experimental protocol an asymptotic observer would use to verify a proposed shell-state assignment; second, we extend the analysis from individual shell states to arbitrary superpositions, which combined with the spanning results in earlier chapters means *any* non-perturbative gravity state can be verified; third, we discuss why the corresponding *search* problem appears to be exponentially hard within the gravitational path integral, putting state identification in the complexity class QMA.

Verification protocol. The probe operator $\mathcal{O}_P$ is itself not Hermitian and not a physical observable. The Hermitian observables $A = (\mathcal{O}_P + \mathcal{O}_P^\dagger)/2$ and $B = (\mathcal{O}_P - \mathcal{O}_P^\dagger)/(2i)$ satisfy $\langle i|A|i\rangle^2 + \langle i|B|i\rangle^2 = |\langle i|\mathcal{O}_P|i\rangle|^2$, so an observer with access to many copies of $|i\rangle$ can measure them and reconstruct the probe correlator. Comparing the measured value to a reference baseline established with random probe choices reveals whether the matched-operator detection enhancement is present. Below the black-hole threshold this enhancement is a factor of three (5.1.24), well separated from the calibration baseline; above threshold it is a factor of two for single-boundary probes (5.1.24, 5.2.20) but exponential for the coordinated two-boundary probe (5.2.11). All are unambiguous.

The observer can also pre-select between type-A and type-B states by measuring the ambient temperature: the bulk EFT in the region accessible to the observer is, in the heavy shell mass limit, almost exactly thermal at the preparation temperature [6], with only fine quantum hair distinguishing the state from the genuine thermal state. Comparing the measured temperature to the Hawking-Page scale[2] fixes whether the proposal should describe a state with horizon (type A) or a horizonless thermal-plus-closed-universe geometry (type B). A proposal for $\mathcal{O}_i$ can then be checked against the matched probe correlator.

[2]In AdS, T_{HP} is fixed by the AdS radius ℓ, which the observer can determine by measuring the asymptotic curvature.

Detecting superpositions. Generic gravity states are not individual shell states but superpositions $|\psi\rangle = \sum_{i=1}^{N} \alpha_i |i\rangle$. The verification protocol extends to such states by considering linear-combination probes $\mathcal{O}_P = \sum_{i=1}^{N} \gamma_i \mathcal{O}_i$. The probe correlator $\overline{|\langle\psi|\mathcal{O}_P|\psi\rangle|^2}$ receives contributions from the same six classes of saddle as in Sec. 5.1, but with index structure inherited from the superposition coefficients. Tracking the index contractions,

$$\overline{\langle\psi|\mathcal{O}_P|\psi\rangle \langle\psi|\mathcal{O}_P^\dagger|\psi\rangle} = \langle\vec{\alpha}|\vec{\alpha}\rangle^2 \langle\vec{\gamma}|\vec{\gamma}\rangle (U_L + U_R) + \langle\vec{\alpha}|\vec{\alpha}\rangle |\langle\vec{\alpha}|\vec{\gamma}\rangle|^2 (D_L^{\uparrow} + D_R^{\uparrow} + D_L^{\downarrow} + D_R^{\downarrow}) , \tag{5.3.1}$$

where $(\vec{\alpha})_i = \alpha_i$, $(\vec{\gamma})_i = \gamma_i Z_{m_i}$, and the inner product is the standard one on $\mathbb{C}^N$. Normalising $\langle\vec{\alpha}|\vec{\alpha}\rangle = \langle\vec{\gamma}|\vec{\gamma}\rangle = 1$, the universal contribution becomes a $\vec{\alpha},\vec{\gamma}$-independent constant while the detection contribution scales with $|\langle\vec{\alpha}|\vec{\gamma}\rangle|^2 \in [0,1]$. For random independent $\vec{\alpha},\vec{\gamma}$, this overlap is generically of order $1/\sqrt{N}$, so the detection signal is suppressed and the correlator is universal. The detection signal is maximised at $\vec{\gamma} = \vec{\alpha}$, where the Cauchy-Schwarz inequality is saturated and the response recovers the full enhancement of Sec. 5.1.

A guess for the superposition coefficients $\vec{\alpha}$ can therefore be verified using the matched linear-combination probe $\mathcal{O}_P = \sum_i \alpha_i \mathcal{O}_i$. Combined with the spanning result of earlier chapters that sufficiently large sets of shell states form a basis of the non-perturbative gravity Hilbert space, this means any state in the Hilbert space can in principle be verified by an asymptotic observer. The verification of state proposals is thus a generic feature of the non-perturbative theory.

The search problem. While verification is straightforward, *finding* which state the universe is in — without prior knowledge — appears to be exponentially hard. Suppose the observer knows the universe is in one of N possible shell states with operators $\{\mathcal{O}_1, \ldots, \mathcal{O}_N\}$ and wants to identify which. Sequential probe testing needs $\mathcal{O}(N)$ trials; Grover search [99] reduces this to $\mathcal{O}(\sqrt{N})$. For Hilbert spaces of microcanonical dimension $N \sim e^{\mathbf{S}(E)}$, both scale exponentially in $\mathbf{S}(E)$ and become impractical above the black-hole threshold.

In principle, full quantum state tomography [100] would identify the state using polynomially many measurements in the Hilbert-space dimension. However, tomography requires access to the fine-grained inner products between basis states. The gravitational path integral at the saddle level has access only to coarse-grained data: it computes $\overline{\langle m|n\rangle} = \delta_{mn} Z_1$ but not the erratic phases whose averaging makes the

off-diagonal elements vanish (Sec. 1.1.1). The phase information required for tomography is, in this sense, beyond the reach of the saddlepoint description used here, and the search problem is correspondingly exponentially hard.

The search-vs-verification asymmetry is the defining feature of the complexity class QMA — the quantum analogue of NP, in which a proposed solution is verifiable in polynomial time but no efficient method to find one is known. Identifying the state of a black-hole or baby-universe gravity state therefore lies in QMA, in the sense made precise by the semicalssical gravitational path integral.

5.4 Discussion

We close with some observations on the implications of the construction.

5.4.1 Detection vs entanglement-wedge reconstruction

The most surprising result of the chapter is that a probe localised on a single asymptotic boundary can distinguish two-boundary shell-state microstates (Sec. 5.2.2), regardless of which boundary the probe sits on. In AdS/CFT, two-sided shell states are partially entangled thermal states (PETS) on $\mathcal{H}_{CFT,L}\otimes\mathcal{H}_{CFT,R}$ [33], and entanglement-wedge reconstruction [90–98] places the bulk shell in the entanglement wedge of the union, generically not of either boundary alone. The detection signal — carried by the wormhole saddles of Sec. 5.2.2, including the novel twice-punctured-torus saddle D_4 — accesses bulk information that no linear single-boundary operator can reconstruct. Verifying a microstate is therefore not the same operation as reconstructing the bulk effective field theory: the latter lies within the entanglement wedge, but the former does not. Making the relationship between these two notions of bulk access precise would be valuable.

5.4.2 Horizons and baby universes geometrize complexity

The universal response to simple operators alongside the non-universal response to a fine-tuned probe, both established in this chapter, are hallmarks of complex states [86,89]. For type-A states the shell sits behind a horizon, and to any generic probe the state looks indistinguishable from a thermal one — the classical picture of a hidden interior is recovered. But the microstate information is not lost: it is encoded in non-perturbative wormhole saddles, accessible only to the fine-tuned probe matched to the state. The horizon's causal disconnection from the boundary is therefore

not a fundamental limit on what can be known from infinity, but the saddlepoint shadow of the QMA-like complexity of the underlying microstate: the state is hidden because resolving it is exponentially hard, not because its information lies beyond the boundary's reach. In this sense the horizon itself — the very existence of a bulk region causally disconnected from the boundary — is a geometric manifestation of the complexity of the underlying quantum state.

Remarkably, the same logic applies to type-B states, where the shell lives on a baby universe disconnected from the asymptotic AdS factor altogether. The parallel suggests that the disconnection of the baby universe from the rest of spacetime is itself a semiclassical manifestation of the underlying complexity of the gravitational state. Notably, as discussed in the next chapter, a naive boundary counting renders the non-perturbative Hilbert space of closed universes one-dimensional [13, 21, 101, 102]. Within our framework, this can be understood as a consequence of the state's extreme complexity, which renders it nearly impossible to probe—effectively forming a perfect black box. Our findings show, however, that while probing the baby universe from the boundary is difficult, it is not impossible. This is consistent with the results of [31, 103], which recover a nontrivial baby universe Hilbert space by entangling it with a reference system—the "observer", which in our case is the disconnected AdS factor.

This apparent complexity of the baby universe is surprising, since its ADM energy lies below the black hole threshold at which the gravitational theory is expected to become highly chaotic [104–107]. Moreover, as noted in [6], the shell operator itself is also "simple", consisting merely of a gas of EFT particles. It thus remains unclear which feature of the time evolution gives rise to the complex characteristics exhibited by the state. We address these questions in the next chapter.

"In order to arrive at what you do not know
You must go by a way which is the way of ignorance"

— *T. S. Eliot,* East Coker

Chapter 6

On the Unexpected Virtue of Ignorance

Euclidean gravity appears to know more about the microscopic theory than a conventional low-energy effective field theory has a right to. As we have seen throughout this thesis, connected wormhole topologies do not compute amplitudes of a single quantum theory but *statistical* data: variances, higher moments, and other universal features of an underlying microstate ensemble. The factorisation argument of Chapter 3 and the partition-function-as-trace identity of Chapter 4 show that this ensemble interpretation is consistent with — indeed enforced by — the existence of a fine-grained Hilbert space. The natural next question is what exactly is being averaged, and on what physical principle the average is selected.

Existing ensemble interpretations are typically top-down: one postulates a microscopic ensemble and asks which gravitational saddles compute its moments. In gravity there is also an intrinsically bottom-up question. A realistic observer has access only to a restricted class of measurements, carried out over finite times and with finite precision, and cannot in principle recover the exact global state from the data available to them. The description they should adopt is therefore not the exact microstate but the least informative state compatible with what they can actually measure: a principle of *maximal ignorance*. This viewpoint becomes especially sharp in closed

universes, where naive wormhole contributions render the fine-grained baby-universe Hilbert space one-dimensional [13, 21, 101, 108–110], while a bulk observer should plainly experience non-trivial physics.[1] The recently proposed *observer rules* of [127] cure this by invoking the need to condition on the existence of a bulk observer's worldline, motivating a modified sum over topologies, recovering a non-trivial effective Hilbert space for closed universes. The resulting prescription is powerful but, as presented there, phenomenological: its microscopic CFT origin has remained unclear. This chapter proposes such an origin, and in doing so recasts the observer rules as unavoidable consequences of epistemic coarse-graining in chaotic holography.

Our starting point is the familiar fact that in a chaotic theory, simple operators obey the Eigenstate Thermalization Hypothesis (ETH) and therefore have exponentially suppressed off-diagonal matrix elements. Combined with finite temporal resolution, this implies that any pure state is operationally indistinguishable from a smoothed diagonal density matrix when probed by any sub-exponential product of simple observables. The maximum-entropy state compatible with a functionally complete set of simple measurements coincides with this smoothed diagonal state.

A central observation follows: the ensemble naturally computed by the gravitational path integral is the equivalence class of microstates that agree on all simple-operator data, together with the maximal-ignorance rule for predicting quantities outside that data. Operators with substantial off-diagonal weight in the energy basis — which we call *complex operators* — display $\mathcal{O}(1)$ fluctuations across microstates in this class and demand an ensemble description rather than a single effective density matrix.

Applied to heavy-shell states in AdS, whose dual CFT operators obey a generalised ETH ansatz [25, 41], this picture becomes quantitatively sharp. The ensemble-averaged density matrix

$$\overline{\rho_\Psi} = \frac{\sum_n e^{-\beta_{\rm obs} E_n - f(E_n)} \, |E_n\rangle\langle E_n|}{\sum_n e^{-\beta_{\rm obs} E_n - f(E_n)}} \,, \tag{6.0.1}$$

with smooth envelope $f(E)$ determined by the semiclassical shell data, is naturally identified with the coarse-grained state selected by simple-operator measurements in the dual CFT. Computing observables in $\overline{\rho_\Psi}$ is *not* equivalent to computing them in the top-down shell ensemble: for linear observables they agree, but for nonlinear quantities the two prescriptions differ in whether different state insertions are allowed

[1] See also [31, 34, 35, 56, 57, 103, 109–126] for related recent discussions of observers, baby universes, and quantum gravity in closed universes.

to cross-contract. The observer prescription based on a single effective density matrix retains operator contractions but removes state cross-contractions, which on the gravity side becomes a precise, microscopically motivated restriction on the topology sum.

The bulk content of this restriction is controlled by a single quantity: the shift δE in the dominant saddle energy induced by a probe operator. Three regimes appear. When the probe is light enough that $\delta E = 0$, the observer reconstructs only the leading saddle energy, and the maximal-ignorance state is thermal at an effective inverse temperature β_{eff}: the corresponding boundary geometry is the AdS–Schwarzschild exterior at β_{eff}, which we identify with the causal wedge [128,129]. We call the associated prescription *observer rule B1*. When one instead uses a functionally complete set of simple probes whose individual backreaction is small but non-zero, $\delta E \sim \mathcal{O}(G_N)$, comparing sufficiently many independent probes resolves higher derivatives of $f(E)$ around the saddle. The resulting effective state is the bulk dual of the Engelhardt–Wall simple wedge [130,131], and its entropy is the generalised entropy of the apparent horizon, reproducing the result of Chandra–Hartman [41]. The corresponding topology restriction is *observer rule A1* of [127], here derived from maximal ignorance rather than postulated. Finally, complex probes with $\delta E \sim \mathcal{O}(1)$ stabilise new wormhole saddles: these are the topologies excluded by the observer rules, and they encode the bulk region — the *Python's lunch* [83] — that lies beyond the outermost quantum extremal surface (QES). The three regimes thus map cleanly onto the three nested reconstruction wedges, causal $\subset$ simple $\subset$ entanglement wedge (of the union of all boundaries).

A second theme concerns the effective Hilbert space seen by the observer. Given a state ρ and an operator system $\mathcal{S}$ — not, in general, an algebra — the sesquilinear form $\langle O_i|O_j\rangle_\rho = \mathrm{Tr}(\rho\, O_i^\dagger O_j)$ defines an effective Hilbert space $\mathcal{H}_\rho$ via the KSGNS construction [132]. For a sufficiently large operator system and a mixed state of rank r whose probes reach a microcanonical target space of dimension d_R, the dimension of this Hilbert space is given by

$$d_{\text{ignorant}} \;=\; r \times d_R\,. \tag{6.0.2}$$

Replacing the microstate by its coarse-grained version therefore *enlarges* the operational Hilbert space beyond the fine-grained global one. We compute d_{ignorant} via the gravitational path integral and show that the result, away from the planar limit, is determined by rank theorems for generalised Wishart ensembles whose structural

content resolves analytic-continuation ambiguities that arise in naive replica calculations. The extra dimension counted by the modified topology sums acquires a sharp operational interpretation: it counts the operator directions that become non-null once inaccessible phase cancellations — correlated off-diagonal coherences of order $e^{-\mathbf{S}/2}$ — have been removed. Reconstructing the Python's lunch, in this language, corresponds to verifying the fine-tuned null relations that the observer's ignorance prevents them from verifying.

For nonlinear observables, however, maximal ignorance alone does not fix a unique answer: one must also specify a prior on the unobserved global structure of spacetime. A single-state prescription reproduces observer rules A1 and B1 and the enlarged dimension $r \times d_R$. Alternatively, correlations between the various state insertions motivate the use of a probability distribution over Hilbert space — a state *ensemble*. A pure-state ensemble over single-boundary microstates compatible with $\overline{\rho_\Psi}$ restores state cross-contractions, collapses the operational Hilbert space back to d_{global}, and recovers the full topology sum. A mixed-state ensemble, equivalently an ensemble of two-boundary purifications, interpolates between these and — in the relevant shell-state regime — reproduces the same modified topology sum as the single-state prescription. The fate of the full topology sum is therefore not fixed by the simple measurement record alone; it depends on what the observer assumes about the causally inaccessible parts of the universe.

The same logic has a striking consequence in closed universes. The fine-grained bilinear on an operator system is rank one — a direct consequence of the two-boundary Hilbert space factorisation of Chapter 3 together with the trace interpretation of periodic boundaries from Chapter 4 — and reproduces the standard triviality argument of [13, 21, 101, 108]. The coarse-grained bilinear, defined by replacing the fine-grained Ψ_{obs} projector by its ensemble average, defines a new positive bilinear form on the same operator system whose rank can be parametrically larger. A non-trivial effective closed-universe Hilbert space therefore emerges from coarse-grained boundary data alone, with no input from any ambient microscopic Hilbert space.

The chapter is organised as follows. Section 6.1 formulates epistemic coarse-graining in chaotic quantum systems. Section 6.2 applies this framework to heavy-shell states in AdS, derives observer rules A1 and B1 from epistemic coarse-graining, and exhibits the saddle-shift analysis that maps probe backreaction onto the three reconstruction wedges. Section 6.3 develops the effective Hilbert space associated to an operator system and a coarse-grained state and computes d_{ignorant} via the gravitational path integral. Section 6.4 constructs maximal-ignorance ensembles of pure

and mixed states and identifies the priors on global spacetime structure under which the modified and full topology sums are recovered. Section 6.5 applies the framework to closed universes. We conclude in Sec. 6.6.

6.1 Epistemic coarse-graining

In this section we develop the coarse-grained description of a chaotic CFT state available to an observer restricted to simple measurements, and derive rules that govern such an observer's predictions. Two ingredients drive the analysis. First, the Eigenstate Thermalization Hypothesis (ETH) implies that simple operators are effectively diagonal in the energy eigenbasis up to $e^{-\mathbf{S}/2}$ corrections, so any pure state is operationally indistinguishable from a diagonal density matrix ρ_{diag} under simple measurements (Sec. 6.1.1). Second, the maximum-entropy state compatible with a functionally complete set of simple-operator data agrees with a smooth approximation to ρ_{diag} (Sec. 6.1.3). For operators that violate ETH — which we call *complex operators* — expectation values exhibit $\mathcal{O}(1)$ fluctuations across microstates compatible with the diagonal data, requiring a state ensemble for statistical predictions (Sec. 6.1.2). The resulting prediction rules for nonlinear observables (Sec. 6.1.4) involve two logically independent sources of averaging — over states and over operators — whose treatment differs between top-down (state ensemble) and observer (density matrix) descriptions. The observer's rules retain operator contractions but suppress state cross-contractions; in later sections this becomes a restriction on the sum over topologies in the gravitational path integral.

6.1.1 Simple observables and dephasing

A realistic experimenter is limited in the probes they can prepare and in the precision with which they can measure. We call a set of operators $\{\mathcal{A}_\alpha\}$ *simple* if they satisfy the conventional ETH ansatz [133–135]

$$(\mathcal{A}_\alpha)_{ij} = \delta_{ij}\, m_\alpha(\overline{E}) + e^{-\mathbf{S}(\overline{E})/2}\, n_\alpha(\overline{E}, \omega)\, M_{ij}\,, \tag{6.1.1}$$

where $\overline{E} \equiv (E_i + E_j)/2$, $\omega \equiv E_i - E_j$, M_{ij} are independently drawn from a complex unit Gaussian, and m_α, n_α are order-one smooth envelope functions. In chaotic many-body systems, ETH is expected to hold for operators that probe only a finite amount of microscopic information in the thermodynamic or large-N limit. In lattice systems, this includes k-local operators with fixed k as system size $N \to \infty$, finite sums of local densities, and finite products thereof. Such operators do not resolve

fine-grained structure in individual energy eigenstates and exhibit matrix elements consistent with (6.1.1): smooth diagonal dependence on energy and exponentially suppressed, pseudo-random off-diagonal components. In holographic large-N CFTs this class consists of finite products of single-trace operators of $\mathcal{O}(1)$ scaling dimension; on the bulk side, it corresponds to semiclassical bulk fields and finite products thereof, evaluated in black hole microstates, reproducing thermal correlation functions up to $e^{-\mathbf{S}}$ corrections.

Dephasing channel. The expectation value of any simple operator Θ in a state $|\Psi_{\rm obs}\rangle = \sum_n c_n |n\rangle$ factorises as

$$\langle\Psi_{\rm obs}|\Theta|\Psi_{\rm obs}\rangle = \sum_n |c_n|^2\,\Theta_{nn} + \sum_{n\neq m} c_n^* c_m\,\Theta_{nm} = {\rm Tr}(\rho_{\rm diag}\,\Theta) + \mathcal{O}(e^{-\mathbf{S}/2})\,, \tag{6.1.2}$$

where

$$\rho_{\rm diag} = \sum_n |c_n|^2 |E_n\rangle\,\langle E_n|\ . \tag{6.1.3}$$

The off-diagonal term is an erratic sum of $\mathcal{O}(e^{2\mathbf{S}})$ contributions of typical magnitude $e^{-\mathbf{S}}$, with mean zero and variance $\mathcal{O}(e^{-\mathbf{S}})$; resolving it would require precision better than $e^{-\mathbf{S}/2}$. In the large entropy system of interest to us, the entropy scales as $\mathbf{S}\sim G_N^{-1}$, and hence any observation of a simple operator at realistic (polynomial) precision therefore accesses only the diagonal data ${\rm Tr}(\rho_{\rm diag}\,\Theta)$. Equivalently, the observer's measurement record is unchanged under the dephasing channel

$$\mathcal{D}(\Theta) = \sum_n |E_n\rangle\,\langle E_n|\ \Theta\ |E_n\rangle\,\langle E_n|\ . \tag{6.1.4}$$

The same channel is reproduced by time-averaging over a window long compared to the inverse energy gaps: $\langle\Theta\rangle_T = T^{-1}\int_0^T dt\,{\rm Tr}(\rho_\Psi(t)\Theta) \to {\rm Tr}(\rho_{\rm diag}\Theta)$ as $T\to\infty$. This dynamical dephasing applies to individual one-point functions; for nonlinear quantities involving multiple state insertions, temporal correlations between fluctuations at different times can produce additional contributions, discussed in Sec. 6.4.

6.1.2 Complex operators and state ensembles

The diagonal density matrix (6.1.3) captures all information accessible to sub-exponential products of simple operators. For operators with substantial off-diagonal weight in the energy eigenbasis — which we call *complex operators* — the off-diagonal term in

$$\langle\Psi|A|\Psi\rangle\ =\ \sum_{m,n}\sqrt{p_m p_n}\,A_{mn}\,e^{i(\theta_n-\theta_m)} \tag{6.1.5}$$

retains $\mathcal{O}(1)$ contributions, and the expectation value is sensitive to the specific microstate within the equivalence class

$$\mathcal{M}(\{p_n\}) \;=\; \Big\{ \sum_n \sqrt{p_n}\, e^{i\theta_n} |n\rangle \Big\} \tag{6.1.6}$$

of pure states sharing the same populations. The diagonal density matrix is then not predictive on a single microstate.

The state ensemble. Predictions for variance of complex observables across the different microstates are made by introducing a probability distribution μ over $\mathcal{M}(\{p_n\})$ satisfying

$$\mathbb{E}_\mu[|\psi\rangle\langle\psi|] \;=\; \rho_{\text{diag}}\,. \tag{6.1.7}$$

Several natural constructions satisfy (6.1.7): a random-phase ensemble (uniform θ_n, fixed amplitudes), a Haar/typicality ensemble on a microcanonical shell, a Gaussian (Ginibre) ensemble of expansion coefficients, and the maximum-entropy distribution at fixed first moment. In chaotic systems these constructions agree on all low-order moments up to $\mathcal{O}(1/d)$ corrections with $d \sim e^{\mathbf{S}}$; for a Gaussian ensemble $|\psi\rangle = \sum_n c_n R_n |n\rangle$ the Wick pairings yield

$$\mathbb{E}[X] \;=\; \sum_n p_n\, A_{nn}\,, \qquad \mathrm{Var}(X) \;=\; \sum_{m\neq n} p_m\, p_n\, |A_{mn}|^2 + \cdots, \tag{6.1.8}$$

for $X = \langle\psi|A|\psi\rangle$. The mean reproduces the diagonal-state prediction by (6.1.7); the variance quantifies fluctuations across microstates. See Appendix B.1 for the explicit Wick-contraction calculation establishing this equivalence in each of the four ensembles.

Self-averaging and its failure. For a simple operator satisfying (6.1.1), the normalised variance evaluates to $\mathrm{Var}(X)/\overline{X}^2 = \mathcal{O}(e^{-\mathbf{S}})$, far below realistic experimental resolution in the gravitational regime $\mathbf{S} \sim \mathcal{O}(G_N^{-1})$. Simple operators are therefore *self-averaging*: their expectation values are exponentially well-concentrated around the diagonal-state prediction. For a complex operator with $|A_{mn}| \sim \mathcal{O}(1)$ over an $e^{\mathbf{S}} \times e^{\mathbf{S}}$ block, one finds instead

$$\mathrm{Var}(X) \sim \sum_{m\neq n} e^{-\mathbf{S}} \cdot e^{-\mathbf{S}} \cdot \mathcal{O}(1) \sim \mathcal{O}(1)\,, \tag{6.1.9}$$

with a comparable $\mathcal{O}(1)$ mean: complex operators exhibit $\mathcal{O}(1)$ fluctuations across compatible microstates. Their expectation values genuinely depend on the phases,

and the diagonal state alone is not predictive. The state ensemble predicts the *distribution* of outcomes of complex probes across the microstates compatible with simple data: the variance (6.1.9) is a robust, ensemble-independent quantity.

A subtlety arises for products of complex operators. For $C = AA^\dagger$ the diagonal elements $C_{nn} = \sum_k |A_{nk}|^2$ are sums of positive terms with no phase cancellation, so $\langle\Psi|AA^\dagger|\Psi\rangle$ is self-averaging despite A being complex: positivity forces $AA^\dagger$ to satisfy an effective ETH. For a product AB of two independent complex operators this mechanism is absent, and the normalised variance can remain $\mathcal{O}(1)$. This distinction will reappear on the gravity side, where bilinears such as $\mathrm{Tr}(\rho\, O_i^\dagger O_j)$ inherit a smooth saddle-point description when $i = j$ even when individual matrix elements do not.

6.1.3 Capturing an observer's ignorance

We now discuss the operational relevance of the diagonal state to an experimenter probing a chaotic CFT with large central charge $c \sim \mathcal{O}(N^2)$. The experimenter's aim is to measure a set of observables to constrain the state. The state is unknown to them, and they are restricted to sub-exponential time measurements of sub-exponential products of simple operators. The experimenter thus probes the unknown CFT state by perturbing what is effectively a quantum black box — configured to produce identical copies of the state — and reading off the outcomes to a given precision. The set of simple operators does not form a closed algebra, since the product of sufficiently many simple operators can exceed any fixed complexity bound. The observable sector accessible to a realistic experimenter is therefore an *operator system* — a self-adjoint unital subspace of $\mathcal{B}(\mathcal{H})$ closed under $O \mapsto O^\dagger$ but not under multiplication. We will repeatedly use this structure in what follows.

Maximal ignorance from simple operators. Measuring a set of expectation values $\{\langle\Theta_\alpha\rangle\}$ within their operator system induces an equivalence class of microscopic states indistinguishable from this data alone. The maximum-entropy principle selects the unique, least-informative representative: maximising the von Neumann entropy of candidate states ρ subject to $\mathrm{Tr}(\rho\,\Theta_\alpha) = \langle\Theta_\alpha\rangle$ yields

$$\rho_\Theta^{\max} = \frac{e^{-\sum_\alpha \lambda_\alpha \Theta_\alpha}}{Z}\,. \tag{6.1.10}$$

By the ETH ansatz (6.1.1) the constraints depend only on diagonal matrix elements weighted by the smooth envelopes $m_\alpha(E)$, so $\rho_\Theta^{\max}$ is itself diagonal:

$$\rho_\Theta^{\max} = \frac{1}{Z}\sum_n e^{-g(E_n)}\,|n\rangle\langle n|\,, \qquad g(E) = \sum_\alpha \lambda_\alpha\, m_\alpha(E)\,, \tag{6.1.11}$$

a smooth function with $\mathcal{O}(1)$ features. The maximal-ignorance state cannot resolve the individual populations $|c_n|^2$ at each of the $\sim e^{\mathbf{S}}$ energy levels; it captures only the smooth thermodynamic profile.

The diagonal state and its smooth approximation. For concreteness we take the black box to be prepared by acting on the vacuum with a heavy operator Ψ of mass $m_\Psi \sim \mathcal{O}(N^2)$ followed by Euclidean evolution,

$$|\Psi_{\rm obs}\rangle \sim e^{-H\beta_{\rm obs}/2}\,\Psi\,|0\rangle\ , \tag{6.1.12}$$

matching the gravity setting of later sections. The dephasing channel projects this onto

$$\rho_{\rm diag} = \sum_n |c_n|^2 |n\rangle\langle n|\ , \qquad |c_n|^2 \equiv e^{-\beta_{\rm obs}E_n}|\Psi_{0n}|^2\,, \tag{6.1.13}$$

which is specified by $\sim e^{\mathbf{S}}$ erratic weights and is not an operationally reconstructed object. Since simple-operator measurements probe only smoothly-weighted averages of the $|c_n|^2$, we define the smooth approximation

$$\rho_{\rm smooth} = \frac{1}{\tilde{Z}}\sum_n e^{-f(E_n)}\,|n\rangle\langle n|\ , \tag{6.1.14}$$

where $e^{-f(E)}$ is the microcanonical average of $|c_n|^2$ over a window narrow enough that the ETH envelopes do not vary appreciably but wide enough to contain exponentially many eigenstates. Although Ψ is too heavy to satisfy the standard ETH ansatz, in a chaotic theory the matrix elements are expected to fluctuate erratically with a smooth microcanonical average $\overline{|\Psi_{0n}|^2}|_{\delta E} = e^{-\mathbf{S}(E_n)}F_\Psi(E_n)$, giving $f(E) = \beta_{\rm obs}E + \mathbf{S}(E) - \log F_\Psi(E)$. The states (6.1.13) and (6.1.14) agree on all simple observables to $\mathcal{O}(e^{-\mathbf{S}/2})$, since the difference $|c_n|^2 - e^{-f(E_n)}/\tilde{Z}$ is erratic and zero-mean within each microcanonical window and does not correlate with the smooth envelopes $m_\alpha(E_n)$.

Agreement of maximal ignorance and the smooth diagonal state. The maximal-ignorance state (6.1.11) is parametrised by a smooth function $g(E)$; the smooth diagonal state (6.1.14) by a smooth function $f(E)$. Both reproduce the measured data, and if the simple operators form a *functionally complete* set — meaning that their ETH envelopes $\{m_\alpha(E)\}$ span the space of smooth functions of energy on the relevant microcanonical window to within the observer's precision — the constraints uniquely fix $g(E) = f(E)$, so $\rho_\Theta^{\max} = \rho_{\rm smooth}$. This identification requires only $\mathcal{O}(1)$ independent measurements, since $f(E)$ is smooth with $\mathcal{O}(1)$ parameters in the thermodynamic limit. For more limited observables the maximal-ignorance state is

correspondingly coarser: if only the Hamiltonian is measured, $g(E) = \beta E$ and the resulting state is the Gibbs state $e^{-\beta H}/Z$.

6.1.4 Predictions

We now turn to predictions for observables involving a probe operator Ω distinct from the state-creating operator Ψ. Such predictions may be nonlinear in the state, and their structure depends on both the complexity of Ω and the level of description adopted.

Top-down (state ensemble). From the perspective of Sec. 6.1.2, the unknown microstate is modelled by a probability distribution μ over the equivalence class $\mathcal{M}(\{p_n\})$. For a linear observable, $\mathbb{E}_\mu[\mathrm{Tr}(\rho_\Psi\,\Omega)] = \mathrm{Tr}(\rho_{\mathrm{diag}}\,\Omega)$. For nonlinear quantities such as $Q \equiv (\mathrm{Tr}(\rho_\Psi\,\Omega))^k$, the ensemble average $\mathbb{E}_\mu[Q]$ involves *state cross-contractions* between different insertions of ρ_Ψ, generated by the higher moments of μ. When Ω itself admits a statistical description — via the ETH ansatz (6.1.1) or an independent operator ensemble $E_\Omega[\cdot]$ — a second source of contractions arises, and the full top-down prediction is

$$\mathbb{E}_\mu\Big[E_\Omega\Big[(\mathrm{Tr}(\rho_\Psi\,\Omega))^k\Big]\Big], \tag{6.1.15}$$

including both state cross-contractions (from μ) and operator cross-contractions (from E_Ω). For simple Ω, state cross-contractions are suppressed by $e^{-\mathbf{S}}$ via self-averaging; for complex Ω, both types of contraction are generically $\mathcal{O}(1)$.

Observer (density matrix prior). An observer who has adopted ρ_{diag} as their description via maximal ignorance computes predictions directly from this single density matrix:

$$Q_{\mathrm{obs}} = E_\Omega\Big[(\mathrm{Tr}(\rho_{\mathrm{diag}}\,\Omega))^k\Big]. \tag{6.1.16}$$

Compared with (6.1.15) the state ensemble average is absent: each insertion of ρ_Ψ is independently replaced by ρ_{diag}, generating no state cross-contractions, while operator contractions are retained. For simple Ω this agrees with the top-down result to $\mathcal{O}(e^{-\mathbf{S}})$; for complex Ω the two can differ at $\mathcal{O}(1)$, with the variance (6.1.9) quantifying the gap.

The distinction between (6.1.15) and (6.1.16) reflects a choice of epistemic prior, not a contradiction. The most conservative choice is to adopt ρ_{diag} as the complete description and compute all predictions from it, with no state cross-contractions.

Adopting the state ensemble μ instead amounts to augmenting the prior beyond what maximal ignorance strictly requires; this becomes natural when predicting not just the mean response of a complex probe but the spread of outcomes across compatible microstates. We develop the extended framework in Sec. 6.4. Looking ahead, we will apply these principles to an observer at the asymptotic boundary of a black hole spacetime prepared by the gravitational path integral. We will identify regimes in which implementing the observer's prediction rule modifies the sum over topologies, and others in which it modifies the boundary conditions of the path integral. We refer to any such modification as an *observer rule*.

6.2 Coarse-graining the black hole interior

We now apply the framework of Sec. 6.1 to the gravitational path integral, taking as our concrete setting the heavy-shell states introduced in Chapter 1. The CFT operators dual to such shells satisfy a generalised ETH ansatz [25,41], which we argue in Sec. 6.2.1 provides the gravitational realisation of the state ensemble compatible with simple-operator data. The ensemble-averaged density matrix $\overline{\rho_\Psi}$ then plays the role of ρ_{smooth}, and the prediction rules of Sec. 6.1.4 translate into a restriction on the sum over topologies. The central technical input (Sec. 6.2.2) is a saddle-point analysis showing that the energy shift δE induced by a probe operator on the smooth envelope $f(E)$ controls the regime of bulk reconstruction. Three regimes appear: $\delta E = 0$ (thermal effective state, causal wedge, observer rule B1, Sec. 6.2.3.2), $\delta E \sim \mathcal{O}(G_N)$ (small backreaction of multiple independent simple probes resolving higher derivatives of f, simple wedge, observer rule A1, Sec. 6.2.3), and $\delta E \sim \mathcal{O}(1)$ (complex probes stabilising new wormhole saddles excluded by the rules, encoding the Python's lunch beyond the outermost QES). Section 6.2.4 extends the discussion to two-boundary shell states.

6.2.1 Heavy shells and the gravitational state ensemble

We consider the single-boundary shell states introduced in Sec. 1.2.2. A heavy operator $\mathcal{O}_{\text{obs}}$ of mass $m \sim \mathcal{O}(G_N^{-1})$ acts on the vacuum and is followed by Euclidean evolution, producing

$$|\Psi_{\text{obs}}\rangle \equiv e^{-H\beta_{\text{obs}}/2}\,\mathcal{O}_{\text{obs}}\,|0\rangle\ . \tag{6.2.1}$$

For $\beta_{\text{obs}} < \beta_{\text{HP}}$ the dominant Lorentzian saddle is a single-boundary black hole with the shell behind the horizon; for $\beta_{\text{obs}} > \beta_{\text{HP}}$ it is thermal AdS entangled with a disconnected closed cosmology (Fig. 1.15a, 1.15b of Chapter 1).

The CFT operators dual to these heavy shells in the semiclassical limit are not well understood, and many distinct fine-grained CFT operators may share the same semiclassical description. Although a detailed construction of these operators has not been achieved,[2] certain universal features can be constrained by matching with semiclassical gravity. [25, 41] established that bulk saddles built from heavy shell gluing conditions match the generalised ETH ansatz

$$\mathcal{O}_{nm} = e^{-f_{\mathcal{O}}(E_n,E_m)/2}\, R_{nm}\,, \tag{6.2.2}$$

where $f_{\mathcal{O}}(E_n, E_m)$ is a smooth order-one function of the energies (and implicitly of the shell mass) determined by the gravitational and intrinsic shell action, and the R_{nm} are independently drawn from the unit complex Gaussian ensemble. Two features distinguish (6.2.2) from the standard ETH ansatz (6.1.1): the absence of a diagonal term, reflecting the vanishing one-point function of a semiclassical dust shell, and the absence of $e^{-\mathbf{S}/2}$ off-diagonal suppression, reflecting the significant backreaction of the heavy shell. The ansatz (6.2.2) is therefore a gravitational realisation of the statistical assumption introduced in Sec. 6.1.3: the matrix elements of $\mathcal{O}_{\text{obs}}$ are erratic with a smooth microcanonical envelope, but are not off-diagonally suppressed.

The envelope $f_{\mathcal{O}}(E_n, E_m)$ was determined in [25] by matching the saddle value of the GPI computing the two-boundary shell overlap to the semiclassical limit of (6.2.2); it is obtained as a function of the gravitational entropy and the intrinsic shell action. As the precise form is not relevant here, we leave it implicit and refer to [25,41] for details. We henceforth drop the subscript and write $f \equiv f_{\mathcal{O}}$.

Remarkably, the match between gravity and (6.2.2) extends to higher moments. In particular, the action of wormhole amplitudes resulting in the non-factorisation of GPI moments such as

$$\left[\langle\Psi_{\text{obs}}|\Psi_{\text{obs}}\rangle\,\langle\Psi_{\text{obs}}|\Psi_{\text{obs}}\rangle\right]_{\text{GPI}} \;\neq\; \left(\left[\langle\Psi_{\text{obs}}|\Psi_{\text{obs}}\rangle\right]_{\text{GPI}}\right)^2 \tag{6.2.3}$$

match the connected contributions of moments of the ensemble (6.2.2): each semiclassical saddle in the topology sum maps to a corresponding generalised-ETH contraction. Throughout this chapter we will use $[\,\cdot\,]_{\text{GPI}}$ for path-integral computed quantities and an overbar $\bar{\cdot}$ for ensemble averages.

[2] It is not yet established whether *any* CFT operators admit a semiclassical dual respecting the shell saddle constructions [34].

Identification with the simple-data ensemble. The GPI used in this manner therefore computes coarse-grained CFT information: the semiclassical shell operator $\mathcal{O}_{\rm obs}$ does not correspond to a precise microscopic CFT operator but to an ensemble of operators $\{\Psi\}$ satisfying (6.2.2). The duality is accordingly a mapping between the semiclassical bulk and a coarse-grained description of the CFT state, not between bulk geometries and exact microstates. A given black hole microstate is, of course, still expected to correspond to a single fine-grained pure state in the CFT. We now argue that the gravitational ensemble arises naturally as the effective description of such a microstate when probed by simple operators, in precise correspondence with the framework of Sec. 6.1.2.

Together (6.2.1) and (6.2.2) define an ensemble of pure states

$$\{|\Psi_{\rm obs}\rangle\} = \left\{ \sum_n c_n \, R_n \, |n\rangle \right\}, \qquad c_n \equiv \frac{e^{-E_n\beta_{\rm obs}/2 - f(E_n)/2}}{\sqrt{\hat{Z}(\beta_{\rm obs})}}, \tag{6.2.4}$$

with $f(E_n, 0) \equiv f(E_n)$ and $\hat{Z}(\beta_{\rm obs}) \equiv \sum_n e^{-E_n\beta_{\rm obs} - f(E_n)}$. The expansion coefficients are not strictly independent because of normalisation, but the norm is approximately self-averaging,

$$\frac{\mathrm{Var}(\overline{\langle\Psi_{\rm obs}|\Psi_{\rm obs}\rangle})}{(\overline{\langle\Psi_{\rm obs}|\Psi_{\rm obs}\rangle})^2} = \frac{3\,\hat{Z}(2\beta_{\rm obs})}{\hat{Z}(\beta_{\rm obs})^2} \sim e^{-\mathbf{S}}, \tag{6.2.5}$$

so we may treat the c_n as independent Gaussian and take the ratio of GPI quantities to obtain normalised observables. The ensemble-averaged density matrix is then the diagonal mixed state

$$\overline{\rho_\Psi} \approx \frac{\overline{|\Psi_{\rm obs}\rangle\langle\Psi_{\rm obs}|}}{\overline{\langle\Psi_{\rm obs}|\Psi_{\rm obs}\rangle}} = \frac{\sum_k e^{-f(E_k)-\beta_{\rm obs}E_k}\,|E_k\rangle\langle E_k|}{\sum_n e^{-f(E_n)-\beta_{\rm obs}E_n}}. \tag{6.2.6}$$

Comparing with Sec. 6.1.2, the gravitational ensemble (6.2.4) is precisely the Gaussian (Ginibre) construction with $g(E_n) = e^{-E_n\beta_{\rm obs}/2 - f(E_n)/2}$ centered around the diagonal state (6.2.6), which is the smoothed diagonal state $\rho_{\rm smooth}$ of Sec. 6.1.3 with $f(E) = \beta_{\rm obs}E - \log F_\Psi(E) + \mathbf{S}(E)$, i.e. the ensemble associated to the effective state compatible with simple operators. We show in Sec. 6.4 that this is also the maximum-entropy ensemble compatible with the diagonal data. Hence:

The ensemble computed by the gravitational path integral can be identified with the set of microstates compatible with the simple-operator data. Within this ensemble, simple observables are captured by a single coarse-grained density matrix, while complex operators exhibit $\mathcal{O}(1)$ fluctuations across microstates, or become self-averaging for suitable composites (e.g. $AA^\dagger$). Euclidean wormholes encode this coarse-grained information: they capture the statistical properties of complex observables across the ensemble, rather than their exact values in a fixed microstate.

This identification is one of the main conceptual results of this chapter. It is often noted that wormhole non-factorisation gives the GPI an ensemble interpretation; the sharper statement here is that this ensemble has a natural physical interpretation as the equivalence class of microstates compatible with the simple operator data of the dual CFT. The semiclassical GPI computes not exact microscopic data but the statistical information that survives coarse-graining to the simple sector. In particular, the non-factorisation of the semiclassical GPI amplitudes is not a failure of the underlying microscopic theory to factorise: it reflects the fact that the bulk computation is already coarse-grained. This coarse-grained information nevertheless retains enough structure for the genuinely gravitational questions studied later in the chapter, such as the dimension of the observer's effective Hilbert space (Sec. 6.3) and the consistency conditions on the underlying microscopic theory (Chapters 3 and 4). Hence, although the exact values of complex operators are not accessible to the semiclassical analysis, their typical fluctuations are, and these are sufficient to determine certain key features of the gravity theory.

Self-averaging on the gravity side. The universal results of Sec. 6.1.2 translate directly into wormhole estimates. The ensemble-average expectation value

$$\overline{\frac{\langle\Psi_{\rm obs}|\Theta|\Psi_{\rm obs}\rangle}{\langle\Psi_{\rm obs}|\Psi_{\rm obs}\rangle}} \approx \mathrm{Tr}(\overline{\rho_\Psi}\,\Theta) \tag{6.2.7}$$

is computed by the GPI ratio $[\langle\Psi_{\rm obs}|\Theta|\Psi_{\rm obs}\rangle]_{\rm GPI}/[\langle\Psi_{\rm obs}|\Psi_{\rm obs}\rangle]_{\rm GPI}$, with the variance computed by wormhole contributions to the corresponding squared expression. For a simple Θ, the leading wormhole contribution is built by gluing a disk to two strips, giving a normalised variance $\sim e^{\mathbf{S}}/e^{2\mathbf{S}} = e^{-\mathbf{S}}$, in agreement with self-averaging. For a complex probe (e.g. a shell operator of mass different from $\mathcal{O}_{\rm obs}$), $[\langle\Psi_{\rm obs}|\Theta|\Psi_{\rm obs}\rangle]_{\rm GPI} = 0$, but the variance receives a wormhole contribution from gluing two disks together, and a strip into each of these disks, which scales as $e^{2\mathbf{S}}$, yielding an $\mathcal{O}(1)$ normalised fluctuation. The wormhole stabilisation responsible for this $\mathcal{O}(1)$ contribution is the geometric realisation of the $\mathcal{O}(1)$ saddle shift for complex probes derived in Sec. 6.2.3.

6.2.2 Topological restriction and the simple wedge

Having identified the gravitational ensemble (6.2.4) as the gravity realisation of the state ensemble of Sec. 6.1.2, we now apply the prediction rules of Sec. 6.1.4 to obtain a topology restriction on the GPI.

Top-down vs. effective-state computations. A shell-state boundary condition for the GPI corresponds, in the dual CFT, to the insertion of a randomly drawn vector from (6.2.4); wormhole geometries compute connected contributions to moments of this ensemble. Probe operators that are themselves heavy enough to admit a generalised-ETH description introduce an independent operator ensemble $E_\Omega[\cdot]$ uncorrelated with the state ensemble, and the GPI computes the full top-down expression (6.1.15).

The effective state compatible with all simple measurement data is $\overline{\rho_\Psi}$. To compute predictions in this state directly while still treating heavy probes statistically, the rule of Sec. 6.1.4 is (6.1.16): retain operator cross-contractions but exclude state cross-contractions. While the pure shell density matrix $\rho_\Psi = |\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|$ is prepared by a GPI, the ensemble-averaged state $\overline{\rho_\Psi}$ is not. The prescription on the bulk side is therefore: starting from the GPI for the relevant quantity in the pure state, retain only those saddles in which a shell sourced by $\mathcal{O}_{\text{obs}}$ on a given boundary is absorbed by $\mathcal{O}^\dagger_{\text{obs}}$ on the *same* boundary (Fig. 6.2). Cross-contractions among Ω insertions are still allowed: heavy probes can propagate between different boundaries, generating their own wormhole contributions.

Next, we give a clear interpretation of $\overline{\rho_\Psi}$ as the dual of the simple wedge.

The simple wedge. The single-boundary shell state (6.2.1) is pure, so its von Neumann entropy is zero, and the Quantum Extremal Surface (QES) formula [74–76, 82, 136–138] recovers this: the minimal QES is empty. The same answer follows from the full sum over topologies for the gravitational replica trick computation of this entropy, since the boundary conditions for $[\mathrm{Tr}(\rho_\Psi^k)]_{\text{GPI}}$ and the normalisation $[(\mathrm{Tr}\,\rho_\Psi)^k]_{\text{GPI}}$ are identical, giving $S_{vN} = 0$. The black hole geometry, however, also contains a non-minimal QES: the apparent horizon. Within standard black hole thermodynamics, its generalised entropy should capture a coarse-grained entropy of the state. Indeed, given the match between the GPI and the state ensemble (6.2.4), Chandra and Hartman [41] showed that the replica partition functions $\mathrm{Tr}(\overline{\rho_\Psi}^k)$ are computed by $\mathbb{Z}_k$-symmetric multi-boundary wormholes branching around the apparent horizon, yielding

$$S(\overline{\rho_\Psi}) = \frac{A_{\text{apparent}}}{4G_N} + S^{\text{outer}}_{\text{bulk}}\,. \tag{6.2.8}$$

For the unperturbed shell states the apparent horizon and event horizon coincide, so the exposed region is empty. Once additional perturbative matter is added this degeneracy is lifted: probe backreaction moves the apparent horizon behind the event

horizon, and incorporating this backreaction into the $\mathbb{Z}_k$ -symmetric wormhole recovers the generalised entropy of the new apparent horizon [41].

Within our framework, $\overline{\rho_\Psi}$ is the effective state associated with simple measurements of $|\Psi_{\text{obs}}\rangle$: starting from the pure-state GPI for the k-th Rényi entropy and applying the topological restriction retains precisely the $\mathbb{Z}_k$-symmetric branching wormholes and reproduces the Chandra–Hartman result.

The identification $S(\overline{\rho_\Psi}) = A_{\text{apparent}}/(4G_N) + S^{\text{outer}}_{\text{bulk}}$ is precisely the *outer entropy* of [130, 131], which counts the degrees of freedom of the *simple wedge* — the bulk region between the asymptotic boundary and the outermost QES. This region was independently characterised in [83] as the set of bulk operators admitting polynomial-complexity CFT reconstruction: precisely the operators a simple-data observer can in fact probe. A coherent picture emerges. The effective state $\overline{\rho_\Psi}$, defined by what a simple observer can measure, is the holographic dual of the simple wedge, defined by what they can reconstruct; its von Neumann entropy reproduces the wedge's gravitational entropy, and the topological restriction promotes this from a statement about entropy to a prescription for any simple-probe observable.

The relationship of this region to the causal wedge, and the precise sense in which simple probes reach beyond the latter, are developed in Sec. 6.2.3.2.

The Python's lunch. The bulk region beyond the outermost non-minimal QES — the *Python's lunch*, bounded for a pure state by the empty RT surface — is not captured by the diagonal state. In the language of Sec. 6.1.2, the lunch information is encoded in the off-diagonal matrix elements $\langle m|A|n\rangle$ with $m \neq n$, precisely what the diagonal projection removes. Complex operators probe this region: their $\mathcal{O}(1)$ variance (6.1.9) across the gravitational ensemble is the information-theoretic signature of the lunch's inaccessibility. Reconstructing the lunch requires resources scaling as $e^{S_{\text{bulge}} - S_{\text{appetiser}}}$ [83], exactly the obstruction to resolving the off-diagonal structure that the topological restriction projects out. Conversely, the topologies removed by the restriction do encode statistical information about the typical complex-operator behaviour across compatible interior microstates — enough, as we will see in Sec. 6.3, to determine quantities such as the dimension of the effective Hilbert space.

6.2.3 The simple wedge observer and observer rule A1

The need to compute in $\overline{\rho_\Psi}$ admits a direct observer-theoretic motivation. The coarse CFT experimenter of Sec. 6.1.3 corresponds to an operationally restricted observer

at the asymptotic boundary.[3] Such an observer measures simple probes to constrain the global state. Provided they have access to a functionally complete set of simple operators over the relevant microcanonical window, the most informative description they can obtain is the smoothed diagonal state $\rho_{\rm smooth}$, which by Sec. 6.2.1 coincides with $\overline{\rho_\Psi}$. For an observer with fewer probes, the maximum-entropy state $\rho_\Theta^{\max}$ is coarser still. Adopting the prior that ignorance is encoded in a single density matrix, $\overline{\rho_\Psi}$ defines the observer's prior state of knowledge, and predictions are made by applying the topology restriction of Sec. 6.2.2. We promote this restriction to a principle:

> **Observer rule A1**: *To capture the most informative description of a gravitational shell state obtainable by a boundary observer restricted to coarsely measuring simple operators, restrict the sum over topologies to geometries in which each shell propagates only between the bra and ket belonging to the same ρ_Ψ insertion.*

This is the rule originally proposed in [127], although the physical motivation we present is conceptually distinct. There, an operator insertion like $\mathcal{O}_{\rm obs}$ was viewed as sourcing a *bulk* gravitational observer, and the rule was motivated by demanding that wormhole geometries not produce non-zero overlaps between states with and without an observer. Here we do not view any operator used in preparing the state as sourcing a bulk observer. Instead, an "observer" is defined as a list of simple observables measured to a given accuracy. The observer rule is the manifestation, in the topology sum, of the resulting epistemic ignorance. The principle generalises straightforwardly to any semiclassical gravity state in which some fine-grained details are inaccessible to finite-time, finite-precision measurements; in the shell-state case, the inaccessibility arises because the shell is hidden behind a horizon or in a disconnected baby universe.

6.2.3.1 Operational requirements and probe backreaction

Having identified observer rule A1, we now clarify the operational requirements for the observer to actually reconstruct the state $\overline{\rho_\Psi}$. For an operator Θ with smooth diagonal envelope $m_\Theta(E)$,

$$\mathrm{Tr}(\overline{\rho_\Psi}\,\Theta) = \frac{\int dE\, e^{\mathbf{S}(E)-\beta_{\rm obs}E-f(E)}\, m_\Theta(E)}{\int dE\, e^{\mathbf{S}(E)-\beta_{\rm obs}E-f(E)}}\,. \tag{6.2.9}$$

Absorbing $\log m_\Theta$ into the exponent, the numerator and denominator have distinct saddle points E'_* and E_* satisfying

$$\mathbf{S}'(E_*) = \beta_{\rm obs} + f'(E_*)\,, \qquad \mathbf{S}'(E'_*) = \beta_{\rm obs} + f'(E'_*) - (\log m_\Theta)'(E'_*)\,. \tag{6.2.10}$$

[3] A more refined description would involve a quasi-local bulk observer, but since semiclassical bulk geometry itself emerges only after coarse-graining, any such observer must also be emergent. We leave this to future work.

Linearising around E_* gives the saddle shift

$$\delta E \equiv E'_* - E_* \;\approx\; \frac{(\log m_\Theta)'(E_*)}{\mathbf{S}''(E_*) - f''(E_*)}\,. \tag{6.2.11}$$

This is precisely the bulk statement that the probe backreacts on the geometry: $-(\log m_\Theta)'$ is the rate at which Θ shifts the action, and $\mathbf{S}'' - f''$ is the thermodynamic stiffness of the unprobed state. Three regimes follow.

Zero backreaction ($\delta E = 0$). For a simple operator the ETH envelope is order-one, so $(\log m_\Theta)' \sim \mathcal{O}(1)$, while $\mathbf{S}'' \sim \mathcal{O}(G_N^{-1})$. The saddle shift $\delta E \sim \mathcal{O}(G_N) \to 0$ in the semiclassical limit, and at leading order

$$\mathrm{Tr}(\overline{\rho_\Psi}\,\Theta) \;\approx\; m_\Theta(E_*)\,, \tag{6.2.12}$$

with E_* fixed by the unprobed saddle condition. The envelope $f(E)$ enters only through the location E_*, via the saddle condition (6.2.10). A linear f is therefore indistinguishable from a shift of the preparation temperature: $f'(E) = \delta\beta$ has the same effect on simple-operator expectation values as $\beta_{\mathrm{obs}} \to \beta_{\mathrm{obs}} + \delta\beta$. A single simple-operator measurement, evaluated at the saddle of the unprobed geometry, therefore constrains f only at the level of $f'(E_*)$, which is operationally indistinguishable from a thermal state at an effective temperature. This is the regime of observer rule B1 (Sec. 6.2.3.2).

Small backreaction ($\delta E \sim \mathcal{O}(G_N)$). Including the saddle shift (6.2.11) at first subleading order,

$$\mathrm{Tr}(\overline{\rho_\Psi}\,\Theta) \;\approx\; m_\Theta(E_*) + \frac{m'_\Theta(E_*)\,(\log m_\Theta)'(E_*)}{\mathbf{S}''(E_*) - f''(E_*)} + \mathcal{O}(G_N^2)\,. \tag{6.2.13}$$

The correction is suppressed by one power of G_N but depends explicitly on $f''(E_*)$. Probe backreaction is therefore the channel through which simple operators become sensitive to higher derivatives of f beyond $f'(E_*)$. This clarifies what "functional completeness" means in the gravitational context: with multiple independent simple operators whose envelopes have different functional dependence on energy, the observer effectively samples $f(E)$ across a neighbourhood of E_*, and a functionally complete set reconstructs the full smooth envelope and hence $\overline{\rho_\Psi}$.

Order-one backreaction ($\delta E \sim \mathcal{O}(1)$). For a complex operator Ω with $\log m_\Omega \sim \mathcal{O}(G_N^{-1})$ the saddle-point formula breaks down in two senses. The perturbative expansion around E_* becomes uncontrolled, since the saddle shift is no longer parametrically small; and the formula (6.2.12) implicitly assumes a smooth diagonal envelope $m_\Omega(E)$ to evaluate at the shifted saddle, which the generalised ETH ansatz (6.2.2) does not provide: the diagonal matrix elements R_{nn} are erratic $\mathcal{O}(1)$ numbers with no smooth profile in E. Decomposing the expectation value into diagonal and off-diagonal pieces, the diagonal contribution $X_{\text{diag}} = \sum_n p_n \Omega_{nn}$ is an incoherent random walk of magnitude $|X_{\text{diag}}| \sim e^{-\mathbf{S}/2}$, while the off-diagonal piece evaluates to $|X_{\text{off}}| \sim \mathcal{O}(1)$ by the random-walk estimate of Sec. 6.1.2, and dominates. The diagonal state $\overline{\rho_\Psi}$ does not predict the outcome in any single microstate; it predicts only the ensemble mean

$$\mathbb{E}_\mu[\langle\Psi|\Omega|\Psi\rangle] = \mathrm{Tr}(\overline{\rho_\Psi}\,\Omega) \sim \mathcal{O}(e^{-\mathbf{S}/2})\,, \tag{6.2.14}$$

with $\mathcal{O}(1)$ variance, in precise agreement with the $[\langle\Psi_{\text{obs}}|\Omega|\Psi_{\text{obs}}\rangle]_{\text{GPI}} \approx 0$ and wormhole-stabilised variance found in Sec. 6.2.1.

The operationally relevant state therefore depends on what the observer is asking. If they want to predict their outcome before measuring, the ensemble of compatible microstates is what they have access to: $\overline{\rho_\Psi}$ gives the mean and the wormhole variance gives the spread, both at $\mathcal{O}(1)$. If they want to use the measurement to pin down which microstate they are in, the relevant state is the underlying microstate itself — as complex probes can resolve this — but whether they can identify it depends on what they know about the probe.

What a complex measurement reveals. In the natural semiclassical setting, the observer prepares a heavy shell of known mass — controlling the smooth envelope f_Ω — but does not know the specific CFT operator implementing it. The outcome $\langle\Psi|\Omega|\Psi\rangle = \mathcal{O}(1)$ is then a projection of the $e^{\mathbf{S}}$-dimensional state vector onto an unknown direction in operator space, and no microstate information can be extracted: sub-exponentially many such measurements remain statistically consistent with $\overline{\rho_\Psi}$, and the ensemble description is the most informative the observer can construct.

The situation changes if the observer has full microscopic knowledge of Ω_{mn}. Each measurement is then a known linear functional of the unknown coefficients $c_m^* c_n$, and because complex operators have $|\Omega_{mn}| \sim \mathcal{O}(1)$ each measurement couples to the off-diagonal phases at leading order. $\mathcal{O}(e^{\mathbf{S}})$ such measurements suffice for full state tomography, matching the Python's lunch complexity bound $\sim e^{S_{\text{bulge}} - S_{\text{appetiser}}}$ [83].[4]

[4] See [4] for closely related considerations.

Probe class	$\boldsymbol{\delta E}$	**Data resolved**	**Effective description**
Low-backreaction	0	Leading saddle energy only: $f'(E_*)$.	Thermal at β_{eff}; causal wedge; observer rule B1.
Functionally complete simple	$\mathcal{O}(G_N)$	Local Taylor data $f'', f''', \ldots$ of f around E_*.	$\overline{\rho_\Psi}$; simple wedge; observer rule A1.
Complex	$\mathcal{O}(1)$	Saddle-point expansion breaks down. With unknown phases: ensemble statistics only. With known phases: $\mathcal{O}(e^{\mathbf{S}})$ measurements give state tomography.	$\overline{\rho_\Psi}$ unfalsifiable; Python's lunch inaccessible without microscopic control.

Table 6.1: The three backreaction regimes and their correspondence to the nested reconstruction wedges. The first two admit a well-defined saddle-point description and a single effective density matrix $\overline{\rho_\Psi}$. In the third regime the saddle-point expansion breaks down for individual complex expectation values; bilinear quantities such as Gram matrix elements retain a saddle-point description via the $AA^\dagger$ mechanism (Sec. 6.1.2). The wedge identifications in the last column are made precise in Sec. 6.2.3.2.

In this regime the operationally relevant state collapses from the ensemble back to the underlying microstate, and the information unlocked is exactly the off-diagonal data that the diagonal projection threw away — the Python's lunch. Simple operators cannot drive this transition: $|\Theta_{mn}| \sim e^{-\mathbf{S}/2}$ means each measurement couples to the off-diagonal phases only exponentially weakly, so even with full microscopic knowledge of Θ, exponential per-measurement precision would be required.

6.2.3.2 The causal wedge observer and observer rule B1

For an observer with fewer probes than functional completeness, the maximum-entropy state $\rho_\Theta^{\max}$ is coarser than $\overline{\rho_\Psi}$ and is generally not prepared by any natural set of GPI boundary conditions. The thermal limit is a notable exception. Consider an observer who has measured only the ADM Hamiltonian and a small number of low-backreaction probes; by Sec. 6.2.3.1 they constrain only $f'(E_*)$. The shell-state exterior is, by the gluing construction of Sec. 1.2.2, identical to that of the AdS–Schwarzschild black hole at the inverse temperature

$$\beta_{\text{eff}} = \beta_{\text{obs}} + \Delta T_D \,, \tag{6.2.15}$$

where ΔT_D is the disk propagation time of the shell determined by the junction conditions, dual to the thermofield-double state [79]. The reduced density matrix on either boundary is thermal, so the expectation value of the ADM Hamiltonian and of any low-backreaction probe outside the horizon coincides with that of a thermal state at β_{eff}. The maximum-entropy state available to the observer is therefore

$$|\Psi_{\text{obs}}\rangle \langle\Psi_{\text{obs}}| \to \frac{e^{-\beta_{\text{eff}} H}}{Z(\beta_{\text{eff}})}\,, \tag{6.2.16}$$

which admits a clear GPI interpretation (Fig. 6.1). This can be implemented as an observer rule:

> **Observer rule B1**: *To capture the epistemic ignorance of an observer who has measured a small set of simple, low-backreaction operators in a shell state $|\beta_{\text{obs}}, m_i\rangle \langle\beta_{\text{obs}}, m_i|$ using a single effective state, replace the GPI boundary condition preparing this density matrix with that of the thermofield-double state at inverse temperature $\beta_{\text{obs}} + \Delta T(m_i)$, and compute the quantities of interest using the full sum over topologies for the resulting GPI.*

While the predictions of B1 are broadly similar to those of A1, this example shows that observer rules need not all take the form of topology restrictions on a fixed boundary condition: some instead modify the boundary condition itself.

Identification with the causal wedge. Observer rule B1 is the natural gravitational realisation of *causal-wedge reconstruction*. A single low-backreaction probe has $\delta E = 0$ at leading order, so its expectation value is fixed by $f'(E_*)$ alone — a one-point function in the unperturbed exterior background, the data HKLL [36, 139] reconstructs in the causal wedge. The maximally ignorant CFT state compatible with such single-probe measurements is the thermofield double at β_{eff}, whose dual is the AdS–Schwarzschild exterior bounded by the bifurcation surface. This realises, in the shell-state setting, the proposal of [129] that the maximum-entropy state fixing all simple one-point functions on a domain of dependence (the *one-point entropy*) is dual to one quarter the area of the causal surface (the *causal holographic information*); since the shell-state exterior is stationary, the causal surface, event horizon, and apparent horizon coincide, and the entropy of (6.2.16) equals their common area. Observer rule A1 sharpens this: a *functionally complete* set of simple probes has individual saddle shifts $\delta E \sim \mathcal{O}(G_N)$ (6.2.11), and the collective dependence of their expectation values on probe choice resolves $f''(E_*), f'''(E_*), \ldots$, extending the reconstruction from the causal wedge to the simple wedge [84, 130, 131]. Each probe

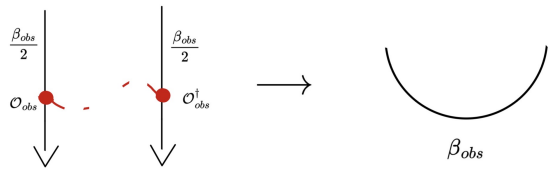


Figure 6.1: Observer rule B1, corresponding to the coarse-graining (6.2.16). The right-hand side is the thermofield-double at the effective temperature β_{eff}.

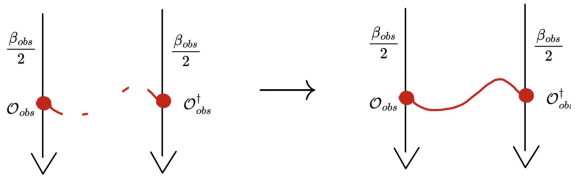


Figure 6.2: Observer rule A1, corresponding to the coarse-graining (6.2.6).

individually is HKLL-reconstructible in the causal wedge; the family parametrised by probe choice carries strictly more information than any single member.

Heavy shell mass limit. In the limit $m_{\mathcal{O}_{\text{obs}}} \to \infty$ the two coarse-grainings coincide. As shown in [6, 19], this limit reduces $f(E_n, E_m)$ to a function of the shell mass alone, so $f''(E)$ and all higher derivatives vanish, and even a functionally complete observer recovers no information beyond the leading saddle energy. Substituting into (6.2.6), $\overline{\rho_\Psi} \to e^{-\beta_{\text{obs}} H}/Z(\beta_{\text{obs}})$ becomes exactly thermal.

6.2.4 Two-boundary gravity states

The above considerations carry over naturally to two-boundary shell states. Consider the state on $\mathcal{H}_{\text{AdS},L} \otimes \mathcal{H}_{\text{AdS},R}$ prepared by inserting the heavy shell operator $\mathcal{O}_{\text{obs}}$ between Euclidean evolutions on the two boundaries,

$$|\boldsymbol{\Psi_{obs}}\rangle = |e^{-\frac{\beta_L H_L}{2}}\, \mathcal{O}_{\text{obs}}\, e^{-\frac{\beta_R H_R}{2}}\rangle \,. \tag{6.2.17}$$

The boundary dual of $\mathcal{O}_{\text{obs}}$ satisfies (6.2.2), giving the ensemble-averaged state

$$\overline{\rho_\Psi} = \frac{\sum_{n,m} e^{-f(E_n,E_m)-\beta_L E_n-\beta_R E_m}\, |E_n, E_m\rangle \langle E_n, E_m|}{\sum_{n,m} e^{-f(E_n,E_m)-\beta_L E_n-\beta_R E_m}} \,. \tag{6.2.18}$$

The new feature is that the smooth envelope does not factorise:

$$f(E_L, E_R) \neq f_L(E_L) + f_R(E_R) \,. \tag{6.2.19}$$

The non-factorising part of f encodes correlations between the two boundaries that are absent in a product of single-sided ensembles. These correlations are not visible to single-boundary measurements, but are accessible to suitably correlated two-sided probes, as we make precise below.

Saddle shifts and the role of probe correlations. The single-boundary saddle-shift analysis (6.2.11) generalises directly. For a probe Θ with smooth two-sided envelope $m_\Theta(E_L, E_R)$, the saddle displacement vector is

$$\delta E_a = \sum_{b=L,R} (K^{-1})_{ab}\, \partial_b \log m_\Theta(E_L^*, E_R^*)\,, \qquad K_{ab} = \delta_{ab}\mathbf{S}_a''(E_a^*) - \partial_a\partial_b f(E_L^*, E_R^*)\,, \tag{6.2.20}$$

with K the thermodynamic stiffness matrix. Its off-diagonal entry $-\partial_L\partial_R f$ measures the coupling between the two energy directions: a perturbation on one side can induce a response on the other through the inverse stiffness, but only if f does not factorise.

This separates one-sided from two-sided probes. For a one-sided probe $\Theta = \Theta_L \otimes \mathbb{1}_R$, the leading-order expectation value is $m_L(E_L^*)$, with E_L^* fixed by the unprobed saddle condition, and the right boundary enters only indirectly through $\partial_L f(E_L^*, E_R^*)$. One-sided low-backreaction data therefore cannot distinguish (6.2.18) from a product of thermal states with effective inverse temperatures

$$\beta_{\text{eff},L} = \beta_L + \partial_L f(E_L^*, E_R^*)\,, \qquad \beta_{\text{eff},R} = \beta_R + \partial_R f(E_L^*, E_R^*)\,. \tag{6.2.21}$$

The corresponding two-boundary version of observer rule B1 is

$$|\boldsymbol{\Psi}_{\boldsymbol{obs}}\rangle\langle\boldsymbol{\Psi}_{\boldsymbol{obs}}| \longrightarrow \rho_{\beta_{\text{eff},L}} \otimes \rho_{\beta_{\text{eff},R}}\,, \tag{6.2.22}$$

which agrees with the direct gravitational picture: the gluing construction makes the geometry exterior to the L and R horizons identical to AdS–Schwarzschild exteriors at the corresponding effective temperatures, and the observer reconstructs the two single-sided causal wedges.

At first subleading order in G_N, the one-sided source $\partial_L \log m_\Theta$ shifts both saddle energies via $\delta E_L = (K^{-1})_{LL}\partial_L \log m_\Theta$ and $\delta E_R = (K^{-1})_{RL}\partial_L \log m_\Theta$, but only δE_L feeds back into the probe's expectation value because Θ_L is blind to E_R. The response therefore depends on $\partial_L\partial_R f$ only through the single combination $(K^{-1})_{LL} = K_{RR}/\det K$. Equivalently, it sees the curvature of the left marginal distribution

$$p_L(E_L) \propto e^{\mathbf{S}_L(E_L)-\beta_L E_L} \int dE_R\, e^{\mathbf{S}_R(E_R)-\beta_R E_R - f(E_L,E_R)}\,, \tag{6.2.23}$$

in which the joint envelope appears only after E_R has been integrated out. A functionally complete set of one-sided simple probes on the left reconstructs p_L in a neighbourhood of E_L^* — and analogously on the right — but no one-sided data can disentangle the contribution of $\partial_L \partial_R f$ from that of $\partial_L^2 f$, so non-factorisation of the joint envelope cannot be detected. Single-sided simple wedges are the corresponding bulk regions: the left observer reconstructs the left simple wedge, the right observer reconstructs the right simple wedge, and these regions are disjoint in the unperturbed shell-state geometry.

Genuinely two-sided simple probes, with $\partial_L \log m_{LR}$ and $\partial_R \log m_{LR}$ both nonzero, sample the joint envelope $f(E_L, E_R)$ along non-trivial directions in the (E_L, E_R) plane and access the mixed derivatives directly. A functionally complete set of two-sided simple probes therefore reconstructs the full smooth envelope and the full effective state (6.2.18), detecting the thermodynamic imprint of the Einstein–Rosen bridge. Reconstructing the geometric interior of the bridge would require $\mathcal{O}(1)$ backreaction, which are sensitive to the underlying microstate at leading order, or exponentially precise measurements; that information lies in the Python's lunch.

Complex probes. The single-boundary analysis carries over: for $\log m_\Omega \sim \mathcal{O}(G_N^{-1})$ the saddle displacement is $\mathcal{O}(1)$ in both components, the saddle-point expansion breaks down, and the diagonal state $\overline{\rho_\Psi}$ predicts only the ensemble mean, with $\mathcal{O}(1)$ variance across compatible microstates. Without microscopic control of the probe, the ensemble is the most informative description available. With full microscopic knowledge of Ω_{mn}, $\mathcal{O}(e^{\mathbf{S}_L + \mathbf{S}_R})$ measurements suffice to pin down the underlying microstate.

The new feature is what one-sided versus two-sided complex probes can actually reach. A one-sided complex probe $\Omega_L \otimes \mathbb{1}_R$ couples only to the reduced density matrix $\rho_L = \mathrm{Tr}_R |\Psi\rangle\langle\Psi|$ on its boundary; with microscopic control, $\mathcal{O}(e^{\mathbf{S}_L})$ measurements reconstruct ρ_L in full, including its off-diagonal phases, unlocking the one-sided Python's lunch. Two pure microstates with identical ρ_L but distinct two-sided correlations are indistinguishable to all one-sided probes, however complex. Resolving the genuinely two-sided correlations — that records the Einstein–Rosen bridge — requires genuinely two-sided complex probes with microscopic control on both boundaries.

6.3 The observer's effective Hilbert space

An observer who cannot determine the global state of a black hole replaces it by a coarse-grained density matrix ρ. The coarse-graining discards microscopic information — such as the off-diagonal phases responsible for the purity of the state — and the resulting mixed state ρ necessarily describes a larger space of possibilities than the single pure state it replaces. This is a familiar phenomenon: an observer on one boundary of the thermofield double sees a thermal state $\rho_R = e^{-\beta H}/Z$ of rank $e^{\mathbf{S}}$, even though the global state is pure. The observer's effective description involves $e^{\mathbf{S}}$ degrees of freedom not because new degrees of freedom have appeared, but because the observer lacks the information that would constrain them away.

The observer rules of Sec. 6.2 implement this kind of epistemic ignorance in the gravitational path integral, replacing $|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|$ by $\overline{\rho_\Psi}$. The effective Hilbert space associated to this coarse-grained description is generically *larger* than the global Hilbert space $\mathcal{H}_{\text{AdS}}$. The central question of this section is: how large, and what does its dimension count physically?

To answer this we must specify two ingredients: the coarse-grained state ρ, and the operator system $\mathcal{S}$ used to probe it. We computed $\overline{\rho_\Psi}$ in Sec. 6.2 from the observer's simple measurements; the natural operator system would then be the simple operators $\mathcal{S}_{\text{simple}}$ accessible to the observer. However, the resulting d_{ignorant} for the pair $(\mathcal{S}_{\text{simple}}, \rho)$ conflates two logically distinct limitations: the observer's ignorance of the microstate (encoded in ρ) and their restricted operational access (encoded in $\mathcal{S}_{\text{simple}}$). To isolate the first — and thereby the intrinsic dimensionality of the coarse-grained state space — we will compute d_{ignorant} with $\mathcal{S}$ taken to be a spanning subset of $\mathcal{B}(\mathcal{H}_{\text{AdS}})$,[5] saturating the algebraic bound. This requires complex operators beyond the observer's simple set — specifically, the heavy shell operators of Sec. 6.2. No logical circularity arises: the state $\overline{\rho_\Psi}$ is fixed by simple measurements and is not revisited; the complex operators are used only to extract the dimension of the Hilbert space that $\overline{\rho_\Psi}$ defines. The physical question being answered is: given the coarse-grained state $\overline{\rho_\Psi}$, what is the maximal effective Hilbert space it can support under the algebraic bound (6.3.5)?

6.3.1 The effective Hilbert space from operator overlaps

Given a state ρ and an operator system $\mathcal{S}$ — a self-adjoint subspace of $\mathcal{B}(\mathcal{H}_{\text{AdS}})$ containing the identity, closed under $O \mapsto O^\dagger$ but not necessarily under products —

[5] $\mathcal{B}(\mathcal{H})$ denotes the algebra of bounded operators on $\mathcal{H}$.

the sesquilinear form

$$\langle O_i|O_j\rangle_\rho \equiv \mathrm{Tr}(\rho\, O_i^\dagger O_j) \tag{6.3.1}$$

defines a pre-inner product on $\mathcal{S}$. Quotienting by the null subspace $\mathcal{N}_\rho = \{O \in \mathcal{S} : \mathrm{Tr}(\rho\, O^\dagger O) = 0\}$ and completing gives the effective Hilbert space

$$\mathcal{H}_\rho \cong \mathcal{S}/\mathcal{N}_\rho\,, \tag{6.3.2}$$

whose dimension we denote d_{ignorant}; concretely, $d_{\text{ignorant}} = \mathrm{rank}\, G$, where $G_{ij} = \mathrm{Tr}(\rho\, O_i^\dagger O_j)$ is the Gram matrix of any basis $\{O_i\}$ of $\mathcal{S}$. This construction requires only the linear and †-structure of $\mathcal{S}$ together with positivity of ρ on $\mathrm{span}(\mathcal{S}^\dagger\mathcal{S})$; it does not require $\mathcal{S}$ to be closed under products. The products $O_i^\dagger O_j$ are of course defined in the ambient algebra $\mathcal{B}(\mathcal{H}_{\text{AdS}})$ — they simply need not lie in $\mathcal{S}$. This is physically essential: as discussed in Sec. 6.1.3, the simple operators accessible to a realistic observer do not form an algebra.

When $\mathcal{S}$ is a full $*$-algebra, $\mathcal{H}_\rho$ is the standard GNS Hilbert space of ρ, and (6.3.1) is the usual GNS inner product. For a general operator system it is instead a subspace of that space: writing $(\pi_\rho, \mathcal{H}_{\text{GNS}}, \Omega_\rho)$ for the GNS data of ρ on $\mathcal{B}(\mathcal{H}_{\text{AdS}})$, the map $O \mapsto \pi_\rho(O)\Omega_\rho$ is linear with kernel exactly $\mathcal{N}_\rho$, whence

$$\mathcal{H}_\rho \;\cong\; \overline{\pi_\rho(\mathcal{S})\,\Omega_\rho} \;\subseteq\; \mathcal{H}_{\text{GNS}}\,, \tag{6.3.3}$$

the closed linear span of $\{\pi_\rho(O)\Omega_\rho : O \in \mathcal{S}\}$. Because $\mathcal{S}$ is not closed under products, this subspace is *not* invariant under $\pi_\rho(\mathcal{S})$: acting with a second accessible operator generically leaves $\mathcal{H}_\rho$. The induced compression $P_\rho\, \pi_\rho(\cdot)\, P_\rho|_{\mathcal{S}}$, with P_ρ the projector onto $\mathcal{H}_\rho$, is accordingly unital completely positive rather than a $*$-homomorphism — the general form of such maps being the content of Stinespring's theorem [132]. Only when $\mathcal{S}$ is an algebra is $\pi_\rho(\mathcal{S})\Omega_\rho$ dense and the compression multiplicative, recovering the standard GNS construction.

Dimension bounds and Hilbert space interpretation. Let $\mathcal{H}_{\text{AdS}}$ have dimension d and ρ have rank r with support projector P. An element $O \in \mathcal{S}$ is null iff $OP = 0$. The map $\Phi : O \mapsto OP$ has $\mathcal{N}_\rho = \ker(\Phi)$, and by the rank-nullity theorem

$$\dim \mathcal{H}_\rho = \dim \mathcal{S} - \dim\{O \in \mathcal{S} : OP = 0\}\,. \tag{6.3.4}$$

Since $\mathrm{Im}(\Phi) \subseteq \mathrm{Hom}(\mathrm{supp}(\rho), \mathcal{H}_{\text{AdS}})$,

$$d_{\text{ignorant}} \leq r \times d\,, \tag{6.3.5}$$

saturated when $\mathcal{S} = \mathcal{B}(\mathcal{H}_{\text{AdS}})$. When the operators map into a microcanonical window $\mathcal{H}_R$ of dimension d_R, the bound sharpens to

$$d_{\text{ignorant}} = r \times d_R \tag{6.3.6}$$

for a sufficiently large operator system. The dimension is equivalently the rank of the Gram matrix $G_{ij} = \text{Tr}(\rho\, O_i^\dagger O_j)$.

For a pure state $\rho = |\Psi\rangle\langle\Psi|$, $\mathcal{H}_\rho$ is the cyclic subspace $\mathcal{S}\,|\Psi\rangle \subset \mathcal{H}_{\text{AdS}}$ — a genuine subspace of the global Hilbert space, with $d_{\text{ignorant}} \leq d_{\text{AdS}}$. For a mixed state, $\mathcal{H}_\rho$ is *not* a subspace of $\mathcal{H}_{\text{AdS}}$: it is a quotient of an operator space, and its dimension can exceed d_{AdS}. Nevertheless d_{ignorant} is a genuine Hilbert space dimension. Any mixed state $\rho = \sum_a p_a\,|a\rangle\langle a|$ admits the canonical purification $|\sqrt{\rho}\rangle = \sum_a \sqrt{p_a}\,|a\rangle_L\,|a\rangle_R \in \mathcal{H}_L \otimes \mathcal{H}_R$, and

$$\hat{G}_{ij} = \text{Tr}(\rho\, O_i^\dagger O_j) = \langle\sqrt{\rho}|(O_i^\dagger O_j \otimes \mathbb{1})|\sqrt{\rho}\rangle\ , \tag{6.3.7}$$

so $\mathcal{H}_\rho$ is isomorphic to the subspace $(\mathcal{S}\otimes\mathbb{1})\,|\sqrt{\rho}\rangle \subset \mathcal{H}_L \otimes \mathcal{H}_R$ of dimension $r \times d_R$. For the shell state in the heavy mass limit, $\overline{\rho_\Psi}$ is thermal at β_{eff}, and its canonical purification is the thermofield double — the two-boundary eternal black hole. The factor $r = e^{S_{\text{apparent}}}$ counts the interior microstates entangled with the exterior but individually inaccessible. The gravitational computation of Sec. 6.3.4 confirms that the statistics of the ignorant Gram matrix coincide with those of the two-boundary Gram matrix.

Why the rank increases under coarse-graining. The mechanism is explicit in the dephasing framework of Sec. 6.1. Write $O_v = \sum_i v_i O_i$ for the operator associated with a vector v in the kernel. The global Gram matrix decomposes as

$$G_{ij} = \underbrace{\sum_a |c_a|^2 \langle a|O_i^\dagger O_j|a\rangle}_{\text{diagonal}} + \underbrace{\sum_{a\neq b} c_a^* c_b\,\langle a|O_i^\dagger O_j|b\rangle}_{\text{off-diagonal}}\,, \tag{6.3.8}$$

the first term being the ignorant Gram matrix

$$\hat{G}_{ij} = \text{Tr}(\rho\, O_i^\dagger O_j) = \sum_a |c_a|^2 \langle a|O_i^\dagger O_j|a\rangle\ . \tag{6.3.9}$$

Both G and $\hat{G}$ are positive semi-definite, being Gram matrices of the vectors $O_i\,|\Psi\rangle$ and of the same vectors in the GNS inner product induced by ρ respectively. The distinction lies in how their kernels are realised. A null direction of G requires $O_v\,|\Psi\rangle = \sum_a c_a O_v\,|a\rangle = 0$, which is a *coherent* cancellation: it can occur even when no

individual $O_v\,|a\rangle$ vanishes, provided the erratic phases $c_a^* c_b$ conspire with the pseudo-random ETH matrix elements $\langle a|O_i^\dagger O_j|b\rangle \sim e^{-\mathbf{S}/2} R_{ab}$ appearing in the off-diagonal term of (6.3.8). For $\hat{G}$ no such cancellation is available. Each $G_{ij}^{(a)} = \langle a|O_i^\dagger O_j|a\rangle$ is separately positive semi-definite, so

$$v^\dagger \hat{G} v = \sum_a |c_a|^2 \| O_v\,|a\rangle \,\|^2, \tag{6.3.10}$$

a sum of non-negative terms, which vanishes only if $O_v\,|a\rangle = 0$ for every microstate in the support of ρ. Hence $\ker \hat{G} = \bigcap_a \ker G^{(a)} \subseteq \ker G$, and

$$\operatorname{rank} \hat{G} \;\geq\; \operatorname{rank} G \tag{6.3.11}$$

with the gap controlled by the number of coherent null relations destroyed by dephasing. The rank of $\hat{G}$ counts operator directions that act non-trivially on *at least* one microstate — the union across the mixture — whereas $\operatorname{rank} G$ is sensitive to the particular superposition realised. The observer does not access more physics than exists in $\mathcal{H}_{\text{AdS}}$; the enlargement reflects their inability to distinguish the realised microstate from the others in the mixture, not the appearance of new degrees of freedom.

Simple vs. complex operator systems. Two physically distinct questions correspond to two choices of operator system. With $\mathcal{S} = \mathcal{S}_{\text{simple}}$, $d_{\text{ignorant}}^{\text{simple}} = r \times d_R^{\text{simple}}$ counts what the observer can operationally probe; since simple operators in a microcanonical window are at most polynomial in $\mathbf{S}$, this dimension is $r \times \text{poly}(\mathbf{S})$. With $\mathcal{S}$ enlarged to a spanning subset of $\mathcal{B}(\mathcal{H}_{\text{AdS}})$, the resulting $d_{\text{ignorant}} = r \times d_R$ counts not what the observer can probe but what the coarse-grained state can *support*. In a microcanonical window $d_R = e^{\mathbf{S}(E_R)}$, exponentially larger. Both are well-defined and physically meaningful; the former is bounded by complexity, the latter by entropy.

A finer breakdown distinguishes three quantities. The dimension d_{ignorant} defined by (6.3.6) is the theoretical maximum set by epistemic ignorance alone, computed below by saturating $\mathcal{S}$ with heavy shell operators. The dimension $d_{\text{ignorant}}^{\text{simple}}$ is the result for the observer's actual simple operator system; it is generically smaller because $d_R^{\text{simple}} \ll d_R$. The dimension d_{measured} accessible to an observer with finite probe count κ, finite precision ε, and finite observer entropy S_{obs} is smaller still. The hierarchy is

$$d_{\text{measured}} \;\leq\; d_{\text{ignorant}}^{\text{simple}} \;\leq\; d_{\text{ignorant}} \;\leq\; r \times d_R\,, \tag{6.3.12}$$

with d_{ignorant} at the top and d_{measured} at the bottom. The gravitational computations that follow determine d_{ignorant} by saturating the algebraic bound; the operationally accessible d_{measured} would require non-planar Gram matrix eigenvalue statistics at $\kappa \ll e^{\mathbf{S}}$, which we do not develop here. We return to this distinction in the discussion.

6.3.2 Setup and path integral methods

We take $\mathcal{S}$ to include a set of time-evolved heavy shell operators $\{e^{-\beta_R H/2}\,\mathcal{O}_i\}$ — complex operators in the sense of Sec. 6.2.3. As discussed above, these are used not as the observer's measurement probes but as a theoretical basis for saturating (6.3.5); the coarse-grained state $\overline{\rho_\Psi}$ was already fixed by the observer's simple measurements and is not revisited. Acting on the reference state produces

$$|i\rangle = e^{-H\beta_R/2}\,\mathcal{O}_i\,e^{-H\beta_{\rm obs}/2}\,\mathcal{O}_{\rm obs}\,|0\rangle\ . \tag{6.3.13}$$

Each state contains two heavy shells: $\mathcal{O}_i$ labelling the state and $\mathcal{O}_{\rm obs}$ shared among all states. We obtain κ such states by varying the mass of $\mathcal{O}_i$, with $|m_i - m_j|$ sufficiently large that $[\langle i|j\rangle]_{\rm GPI} \sim \delta_{ij}$. The Gram matrix of overlaps is

$$G_{ij} = \langle\Psi_{\rm obs}|\,\mathcal{O}_i^\dagger\,e^{-H\beta_R}\,\mathcal{O}_j\,|\Psi_{\rm obs}\rangle\ , \tag{6.3.14}$$

and the observer's epistemic uncertainty replaces $|\Psi_{\rm obs}\rangle\langle\Psi_{\rm obs}| \to \rho_\Theta^{\rm max}$, giving the ignorant Gram matrix

$$\hat{G}_{ij} \equiv {\rm Tr}(\rho_\Theta^{\rm max}\,\mathcal{O}_i^\dagger\,e^{-H\beta_R}\,\mathcal{O}_j)\,. \tag{6.3.15}$$

We compute $\hat{G}$ for both the ensemble-averaged state (6.2.6) and the thermal state (6.2.16), using observer rules A1 and B1 respectively.

Path integral methods. The rank of a Gram matrix $G_{ij} \equiv \langle i|j\rangle$ equals the number of nonzero eigenvalues, and is given by

$${\rm rank}(G) = \lim_{n\to 0}{\rm Tr}(G^n)\,. \tag{6.3.16}$$

In our setting ${\rm Tr}(G^n)$ is a gravitational path integral $[{\rm Tr}(G^n)]_{\rm GPI}$ with n asymptotic boundaries, and the heavy shells stabilise many connected wormhole topologies. We use three complementary techniques to evaluate the resulting sum, two of which were introduced in Chapter 2 for similar Gram matrices.

In the planar limit $\kappa\,e^{-1/G_N} \sim \mathcal{O}(1)$ the topology sum admits a Schwinger–Dyson resummation [6, 13] via the resolvent matrix $R_{ij}(\lambda) = \frac{\delta_{ij}}{\lambda} + \sum_n G_{ij}^n/\lambda^{n+1}$ as in Sec. 2.3.2. Defining the fully connected normalised moments

$$Z_n \equiv \frac{[\langle i_1|i_2\rangle\,\langle i_2|i_3\rangle\cdots\langle i_n|i_1\rangle]_{\rm GPI,\,connected}}{[\langle i_1|i_1\rangle\,\langle i_2|i_2\rangle\cdots\langle i_n|i_n\rangle]_{\rm GPI}}\,, \tag{6.3.17}$$

which simplify in the heavy shell limit because the universal shell factors cancel, the resummation reduces to $\lambda R(\lambda) = \kappa + \sum_n Z_n R(\lambda)^n$. In all cases considered here this

yields $Z_n|_{\text{micro}} = e^{(1-n)\mathbf{S}(E)}$ after a microcanonical projection for some function $\mathbf{S}$ that is related to the Bekenstein–Hawking entropy. The resulting geometric series sums, and the resulting eigenvalue density gives

$$\text{rank}(G)|_{\text{micro}} = \min(\kappa,\, e^{\mathbf{S}(E)})\,. \tag{6.3.18}$$

In the canonical ensemble Z_n does not simplify; we instead take the strict large-κ limit $\kappa\, e^{-1/G_N} \to \infty$, in which (as discussed in Sec. 2.3.4) only the maximally connected topologies survive and $[\text{Tr}(G^n)]_{\text{GPI}} = \kappa^n Z_n$ directly.

In some cases we encounter neither limit renders the moments analytic in n. To proceed there we note that the full non-planar finite-κ sum reproduces the moments of generalised Wishart ensembles (Appendix B.2), and the rank can be extracted without analytic continuation by appealing to rank theorems for these ensembles.

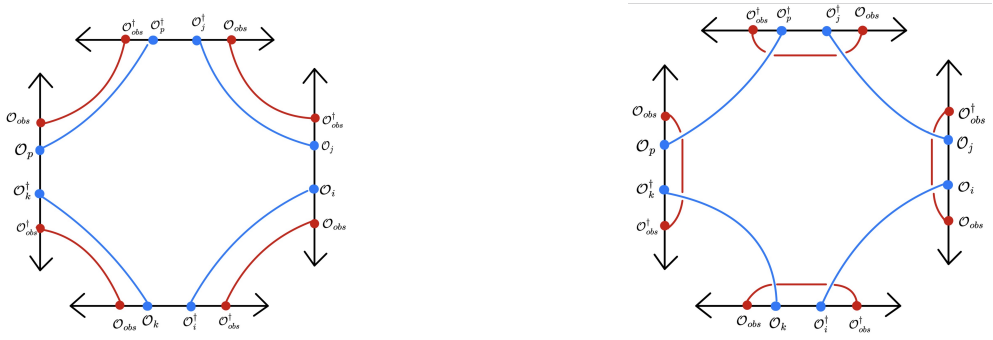
6.3.3 Global case

We compute the rank of (6.3.14) in the heavy shell mass limit. The rank of a similar Gram matrix — without the common operator $\mathcal{O}_{\text{obs}}$ — was computed in [2] and gave $\min(\kappa, e^{\mathbf{S}(E)})$, with the shell states spanning the microcanonical Hilbert space once $\kappa \geq e^{\mathbf{S}(E)}$ and the canonical Hilbert space as $\kappa \to \infty$. We now show that adding the common operator $\mathcal{O}_{\text{obs}}$ does not change these conclusions, although the topology sum is more involved.

Topology of the sum. A shell sourced by $\mathcal{O}_i$ can only be absorbed by $\mathcal{O}_j^\dagger$ if $i = j$, so the index loops in $\text{Tr}(G^n)$ require maximally connected topologies for the $\mathcal{O}_i$ shells. The shared shell $\mathcal{O}_{\text{obs}}$, however, can be absorbed by any $\mathcal{O}_{\text{obs}}^\dagger$, giving $n!$ connectivity patterns labelled by a permutation $\sigma \in S_n$. Cutting the diagrams along all shell worldvolumes decomposes them into closed loops (filled by Z-saddle disks of the corresponding length) and strips (filled by Euclidean vacuum AdS). In the heavy mass limit the junction conditions confine the shells to the asymptotic region, each shell of mass m_i contributes a universal factor Z_{m_i}, and a permutation σ with cycle lengths $\{k_l(\sigma)\}$, contributes $\prod_{l=1}^{q(\sigma)} \overline{Z}(k_l(\sigma)\beta_{\text{obs}})$, where $q(\sigma)$ is the number of cycles of the permutation.

For the normalisation $[\langle i_1|i_1\rangle \cdots \langle i_n|i_n\rangle]_{\text{GPI}}$, each $\mathcal{O}_i$ self-contracts producing a loop of length β_R, so

$$[\langle i_1|i_1\rangle \cdots \langle i_n|i_n\rangle]_{\text{GPI}} = \overline{Z}(\beta_R)^n \sum_{\sigma\in S_n} \prod_{j=1}^{q(\sigma)} \overline{Z}(k_j(\sigma)\beta_{\text{obs}}) \times \mathcal{P}_n\,, \tag{6.3.19}$$



(a) (b)

Figure 6.3: Two classes of replica-symmetric saddles contributing to (6.3.20), depicted for $n = 4$. (**a**) n subsidiary disks of length $\beta_{\rm obs}$, each carrying one strip, glued into a central disk of length $n\beta_R$. (**b**) A single disk of length $n\beta_{\rm obs}$ with n strips, glued to the central disk.

where $\mathcal{P}_n = \overline{S}(0)^n \prod_i Z_{m_i}(Z_{m_{\rm obs}})^n$ collects the universal shell prefactors. For the numerator $[\langle i_1|i_2\rangle \cdots \langle i_n|i_1\rangle]_{\rm GPI}$ the unbroken index loop fills with a single loop of length $n\beta_R$, while the $\mathcal{O}_{\rm obs}$ shells now form cycles of $\tau^{-1} \circ \sigma$ where $\tau(k) = k + 1$ mod n:

$$[\langle i_1|i_2\rangle \cdots \langle i_n|i_1\rangle]_{\rm GPI} = \overline{Z}(n\beta_R) \sum_{\sigma \in S_n} \prod_{l=1}^{q(\tau^{-1}\circ\sigma)} \overline{Z}(k_l(\tau^{-1} \circ \sigma)\beta_{\rm obs}) \times \mathcal{P}_n \,. \tag{6.3.20}$$

Since τ is invertible, $\sigma \mapsto \tau^{-1} \circ \sigma$ is a bijection on S_n and the two sums in (6.3.19) and (6.3.20) are equal. The remarkably simple result is

$$Z_n = \frac{\overline{Z}(n\beta_R)}{\overline{Z}(\beta_R)^n} \,, \tag{6.3.21}$$

with the dependence on the intermediate temperature $\beta_{\rm obs}$ dropping out entirely.[6]

Global dimension. Microcanonical projection of (6.3.21) (achieved by Laplace-transforming with respect to β_R) yields $Z_n|_{\rm micro} = e^{(1-n)\mathbf{S}(E_R)}$, giving $\text{rank}(G) = \min(\kappa, e^{\mathbf{S}(E_R)})$ from the planar resolvent. The addition of $\mathcal{O}_{\rm obs}$ therefore does not change the fact that the shell states span the microcanonical Hilbert space once $\kappa \geq$

[6] The argument is robust under restriction to replica-symmetric saddles: if a non-replica-symmetric saddle for $\gamma \in S_n$ is included in (6.3.19), the cancellation continues to hold provided the saddle for $\tau^{-1} \circ \gamma$ is included in (6.3.20). As discussed in [31], including non-replica-symmetric saddles raises difficulties for analytic continuation away from positive integer $n > 1$ when cycle lengths are intermediate, so summing over all saddles or just the replica-symmetric ones gives the same result. In Sec. 6.3.5.2 below we consider the two-boundary version of this set-up and find that no such cancellation occurs.

$e^{\mathbf{S}(E_R)}$. In the canonical ensemble the large-κ limit gives $[\mathrm{Tr}(G^n)]_{\mathrm{GPI}} = \kappa^n \overline{Z}(n\beta_R)/\overline{Z}(\beta_R)^n$, from which $\mathrm{rank}(G) \to \overline{Z}(0) \equiv d_{\mathrm{canon}}$. The gravity partition function $\overline{Z}(\beta) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{AdS}}}(e^{-\beta H})$ diverges as $\beta \to 0$, reflecting the infinite dimensionality of the canonical ensemble; we use $d_{\mathrm{canon}} \approx \int_0^{E_{\mathrm{UV}}} dE\, e^{\mathbf{S}(E)}$ as a regulated bookkeeping device.

These countings recover the algebraic result (6.3.6) with $r = 1$ (since $|\Psi_{\mathrm{obs}}\rangle$ is pure) and $d_R = e^{\mathbf{S}(E_R)}$ or $d_R = d_{\mathrm{canon}}$ in the micro- and canonical cases respectively.

Finite κ. Beyond the planar limit, the moments $[\mathrm{Tr}(G^n)]_{\mathrm{GPI}}$ receive contributions from all topologies, labelled by an additional permutation $\eta \in S_n$ controlling the connectivity of the $\{\mathcal{O}_i\}$ shells. Introducing $\mathrm{Tr}_\pi(D) \equiv \prod_{c\in\mathrm{Cycles}(\pi)} \mathrm{Tr}(D^{|c|})$, the microcanonical moments become

$$\frac{[\langle i_1|i_2\rangle \cdots \langle i_n|i_1\rangle]_{\mathrm{GPI,\,micro}}}{\mathcal{P}_n} = \sum_{\sigma,\eta\in S_n} \kappa^{q(\tau^{-1}\circ\eta)} \mathrm{Tr}_\eta(D_R)\mathrm{Tr}_{\eta^{-1}\circ\sigma}(D_{\mathrm{obs}})\,, \tag{6.3.22}$$

with the analogous denominator. The β_{obs}-dependent factors cancel by the same τ-bijection argument, and after dividing one obtains

$$[\mathrm{Tr}(G^n)]_{\mathrm{GPI}} = \frac{\sum_{\eta\in S_n} \kappa^{q(\tau^{-1}\circ\eta)} \mathrm{Tr}_\eta(D_R)}{\sum_{\theta\in S_n} \kappa^{q(\theta)} \mathrm{Tr}_\theta(D_R)}\,, \tag{6.3.23}$$

where $D_R = e^{-\beta_R H}$. Under the identification $\overline{Z}(\beta) = \mathrm{Tr}(e^{-\beta H})$, this is the n-th moment of a scale-mixture Wishart matrix (Appendix B.2.1) with population covariance D_R. The non-planar corrections are precisely the non-planar Wick contractions suppressed at large κ, and they generically spoil analyticity in n. Since the moments of G match those of W at all positive integer n, the two matrices share the same eigenvalue distribution, and the rank of G is therefore determined by the Wishart rank theorem:

$$\mathrm{rank}(G) = \min(\kappa,\, r_R) \qquad \text{almost surely (a.s.)}, \tag{6.3.24}$$

with $r_R = e^{\mathbf{S}(E_R)}$ in a microcanonical window, reproducing the planar result. The non-planar corrections affect the eigenvalue distribution but not the rank.

6.3.4 The ignorant Gram matrix

We now apply observer rules A1 and B1 to the global Gram matrix above.

Observer rule A1. A1 restricts the $\mathcal{O}_{\rm obs}$ shells to self-contract, fixing $\sigma = e$ in the finite-κ expressions and giving $\mathrm{Tr}_{\eta\circ e}(D_{\rm obs}) = \mathrm{Tr}_\eta(D_{\rm obs})$. The A1 moments become

$$[\mathrm{Tr}(\hat{G}^n)]_{\mathrm{GPI,\,A1}} = \frac{\sum_{\eta\in S_n} \kappa^{q(\tau^{-1}\circ\eta)} \mathrm{Tr}_\eta(D_R)\mathrm{Tr}_{\eta^{-1}}(D_{\rm obs})}{\sum_{\theta\in S_n} \kappa^{q(\theta)} \mathrm{Tr}_\theta(D_R)\mathrm{Tr}_{\theta^{-1}}(D_{\rm obs})}\,. \tag{6.3.25}$$

Compared with (6.3.23), the $\beta_{\rm obs}$-dependent factors no longer cancel: the traces $\mathrm{Tr}_\eta(D_{\rm obs})$ are now coupled to the same permutation η controlling the $\{\mathcal{O}_i\}$ connectivity. Noting $\mathrm{Tr}_\eta(D_{\rm obs}) = \mathrm{Tr}_{\eta^{-1}}(D_{\rm obs})$ and recognizing $\mathrm{Tr}_\eta(D_R)\mathrm{Tr}_\eta(D_{\rm obs}) = \mathrm{Tr}_\eta(D_R \otimes D_{\rm obs})$, this is the n-th moment of a scale-mixture Wishart with population covariance $D_R \otimes D_{\rm obs}$ in the product space. This is the Wishart counterpart of the purification identity (6.3.7): the ignorant Gram matrix is the Gram matrix of the canonical purification, with $D_R \otimes D_{\rm obs}$ the covariance of the operators $\mathcal{O}_i \otimes \mathbb{1}$ acting on the purified state. The rank theorem applies as in Sec. 6.3.3 giving,

$$\mathrm{rank}(\hat{G}) = \min(\kappa,\, r_R \times r_{\rm obs}) \qquad \text{a.s.}, \tag{6.3.26}$$

giving $\min(\kappa, e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\rm obs})})$ in a microcanonical window.

Observer rule B1. For the thermal coarse-graining (6.2.16),

$$\hat{G}_{ij} = \frac{1}{Z(\beta_{\rm eff})}\,\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta_{\rm eff}H}\,\mathcal{O}_i^\dagger\, e^{-H\beta_R}\,\mathcal{O}_j)\,, \tag{6.3.27}$$

which is, up to normalisation, the Gram matrix of two-boundary shell states $|\mathbf{i}\rangle = e^{-\beta_{\rm eff}H_L/2}\,\mathcal{O}_i\, e^{-\beta_R H_R/2}$. At finite κ this is a trace-construction Wishart (Appendix B.2.2) with deterministic state matrix $C_A = e^{-\beta_{\rm eff}H}$ and operator covariance $D_2 = e^{-\beta_R H}$, and the rank theorem gives

$$\mathrm{rank}(\hat{G}) = \min(\kappa,\, e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\rm obs})})\,, \tag{6.3.28}$$

in agreement with (6.3.26).

Planar limit and observer rules. The planar limit corresponds to keeping only the dominant permutations: $\eta = \tau$ in the numerator (maximising $q(\tau\circ\eta) = n$) and $\theta = e$ in the denominator. Retaining these and using $\mathrm{Tr}_\tau(D) = \overline{Z}(n\beta)$ and $\mathrm{Tr}_e(D) = \overline{Z}(\beta)^n$, the fully connected moment is

$$\hat{Z}_n = \frac{\overline{Z}(n\beta_R)}{\overline{Z}(\beta_R)^n} \times \frac{\overline{Z}(n\beta_{\rm obs})}{\overline{Z}(\beta_{\rm obs})^n}\,. \tag{6.3.29}$$

Both observer rules give the same expression: under B1, the leading contribution to the product of n two-boundary shell overlaps is the n-boundary pinwheel of Sec. 1.2.1.7, giving the same result. We use a hat to distinguish the coarse-grained moments ($\hat{Z}_n$) from the global case (Z_n).

The planar moments are analytic in n, so the rank can be extracted via $\text{rank}(\hat{G}) = \lim_{n\to 0} \text{Tr}(\hat{G}^n)$. Comparing with the global case (6.3.21), $\hat{Z}_n = Z_n \times \overline{Z}(n\beta_{\text{obs}})/\overline{Z}(\beta_{\text{obs}})^n$. If $\beta_{\text{obs}} > \beta_{\text{HP}}$, the leading saddle for both $\overline{Z}(n\beta_{\text{obs}})$ and $\overline{Z}(\beta_{\text{obs}})$ is thermal AdS at positive integer n, so the additional factor is unity and $\hat{Z}_n = Z_n$: we recover the global result. In this regime $|\Psi_{\text{obs}}\rangle$ is a Type 2 shell state (empty AdS entangled with a closed cosmology), and the coarse-graining does not affect Gram matrix statistics.

In the black hole phase $\beta_{\text{obs}} < \beta_{\text{HP}}$, projecting to a microcanonical window in E_{obs} (in addition to E_R) gives $\hat{Z}_n = e^{(1-n)(\mathbf{S}(E_R)+\mathbf{S}(E_{\text{obs}}))}$ and so $\text{rank}(\hat{G})|_{\text{micro}} = e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\text{obs}})}$, in agreement with (6.3.26). In the large-κ limit $\text{Tr}(\hat{G}^n) = \kappa^n \hat{Z}_n$ and the analytic continuation $n \to 0$ gives the canonical result $\text{rank}(\hat{G}) = d^2_{\text{canon}}$, with mixed micro/canonical projections giving $e^{\mathbf{S}(E_R)} \times d_{\text{canon}}$ or $d_{\text{canon}} \times e^{\mathbf{S}(E_{\text{obs}})}$ in the obvious way. In all cases in the black hole regime, the rank exceeds the global Gram matrix rank, in accordance with the algebraic bound (6.3.6).

Agreement of A1 and B1. The agreement is unsurprising: in the heavy shell mass limit used throughout, the ensemble-averaged state becomes exactly thermal and the two prescriptions coincide. Finite-mass corrections [32] would distinguish them in the eigenvalue distribution, but not in the rank, which is a topological quantity. In both cases the statistics coincide with those of the two-boundary Gram matrix, supporting the interpretation that the maximal-ignorance description of the single-boundary black hole admits a canonical purification by introducing a second boundary.

Physical interpretation. The product structure $d_{\text{ignorant}} = r \times d_R$ has a transparent meaning. The factor $r = \text{rank}(\overline{\rho_\Psi}) \sim e^{S_{\text{apparent}}}$ counts the microstates compatible with the exterior geometry, while $d_R \sim e^{\mathbf{S}(E_R)}$ counts the distinct excitations the probes can create within each microstate. The total $d_{\text{ignorant}} \sim e^{S_{\text{apparent}}+\mathbf{S}(E_R)}$ exceeds the global value $d_{\text{global}} \sim e^{\mathbf{S}(E_R)}$ by exactly the "gauge" factor r.

The relationship to the Python's lunch is direct. In the global computation, cross-contractions of the $\mathcal{O}_{\text{obs}}$ shells across boundaries stabilise the wormhole topologies that distinguish microstates, reducing the rank from $r \times d_R$ to d_R; observer rule A1 forbids precisely these contractions. The lunch therefore does not manifest as missing directions in $\hat{G}$ but as the *excess*: $d_{\text{ignorant}} > d_{\text{global}}$ because the observer cannot verify

the fine-tuned null relations between energy eigenstates that would collapse the effective space. Restoring those relations would require both per-measurement precision $\varepsilon \ll e^{-\mathbf{S}/2}$ (to see the off-diagonal coherences $c_a^* c_b$) and $\mathcal{O}(e^{S_{\text{apparent}}})$ measurements (matching the Python's lunch complexity bound, since the appetiser is the empty RT surface and the bulge the apparent horizon). The information content of the lunch therefore lives in the gap $d_{\text{ignorant}}/d_{\text{global}} = r$: it is the difference between the constrained Hilbert space available to a microscopic observer and the unconstrained one available to the coarse-grained observer.

6.3.5 Two-boundary states

The above considerations carry over to two-boundary shell states. Starting from

$$|\boldsymbol{\Psi}_{\boldsymbol{obs}}\rangle = |e^{-\frac{\beta m_L H_L}{2}}\, \mathcal{O}_{\text{obs}}\, e^{-\frac{\beta m_R H_R}{2}}\rangle \,, \tag{6.3.30}$$

and acting with additional heavy shell operators, two cases naturally arise: an observer with access to both boundaries, and one with access only to the right. We treat them in turn, focusing on the most physically informative new features.

6.3.5.1 Two-boundary probes

Consider the states

$$|\mathbf{i},\mathbf{j}\rangle = e^{-\beta_L H_L/2}\, \mathcal{O}_i\, e^{-\beta_{m_L} H_L/2}\, \mathcal{O}_{\text{obs}}\, e^{-\beta_{m_R} H_R/2}\, \mathcal{O}_j\, e^{-\beta_R H_R/2} \,, \tag{6.3.31}$$

with $\{\mathcal{O}_i\}_{i=1}^{\kappa_L}$ and $\{\mathcal{O}_j\}_{j=1}^{\kappa_R}$ drawn from independent left and right operator sets of mutually non-degenerate masses. The non-degeneracy ensures that a shell sourced by a right operator $\mathcal{O}_{j,R}$ can only be absorbed by $\mathcal{O}_{n,R}^\dagger$ with $n = j$, and similarly for left operators, so that there are no left-to-right cross-contractions. Elements of the Gram matrix are

$$G_{i,j,m,n} \equiv \langle ij|mn\rangle = \langle \Psi_{\text{obs}}|\, \mathcal{O}_i^\dagger e^{-H\beta_L} \mathcal{O}_m \otimes \mathcal{O}_j^\dagger e^{-H\beta_R} \mathcal{O}_n\, |\Psi_{\text{obs}}\rangle \,. \tag{6.3.32}$$

We compute the resulting global dimension first before moving to the ignorant case governed by the observer rules.

Planar limit. In the planar (or large-κ) limit, the sum over topologies for Z_n involves a single permutation $\sigma \in S_n$ controlling the connectivity of the $\mathcal{O}_{\text{obs}}$ shells;

the state-counting shells on each boundary contribute a single unbroken index loop. The result is

$$Z_n = \frac{\overline{Z}(n\beta_L)\,\overline{Z}(n\beta_R)}{\overline{Z}(\beta_L)^n\,\overline{Z}(\beta_R)^n} \times \frac{\sum_{\sigma\in S_n}\prod_{j=1}^{q(\tau^{-1}\circ\sigma)}\overline{Z}(\tilde{k}_j(\tau^{-1}\circ\sigma)\beta_{m_L})\,\overline{Z}(\tilde{k}_j(\tau^{-1}\circ\sigma)\beta_{m_R})}{\sum_{\omega\in S_n}\prod_{w=1}^{q(\omega)}\overline{Z}(k_w(\omega)\beta_{m_L})\,\overline{Z}(k_w(\omega)\beta_{m_R})}\,. \tag{6.3.33}$$

As in the single-boundary global case (6.3.21), the map $\sigma \mapsto \tau^{-1}\circ\sigma$ is a bijection on S_n. Both sums therefore run over the same set, and the ratio of the second factor is identically unity. The dependence on the intermediate parameters β_{m_L} and β_{m_R} — encoding the preparation temperatures of the observer state — drops out entirely:

$$Z_n = \frac{\overline{Z}(n\beta_L)\,\overline{Z}(n\beta_R)}{\overline{Z}(\beta_L)^n\,\overline{Z}(\beta_R)^n}\,. \tag{6.3.34}$$

The rank is determined by the probe temperatures alone: in a microcanonical window $\mathrm{rank}(G) = \min(\kappa_R\kappa_L, e^{\mathbf{S}(E_R)+\mathbf{S}(E_L)})$, and in the large-$\kappa$ limit $\mathrm{rank}(G) \to d^2_{\mathrm{canon}}$. This should be contrasted with the single-boundary-probe case of Sec. 6.3.5.2, where the intermediate parameters do not cancel.

Finite κ. At finite κ the cancellation of intermediate parameters breaks down at the level of moments. Including all topologies, the moments are labelled by three permutations $\sigma, \eta, \phi \in S_n$ controlling the connectivity of the $\mathcal{O}_{\mathrm{obs}}$, right $\mathcal{O}_{j,R}$, and left $\mathcal{O}_{i,L}$ shells respectively. Introducing $\mathrm{Tr}_\pi(D) \equiv \prod_{c\in\mathrm{Cycles}(\pi)}\mathrm{Tr}(D^{|c|})$ and the diagonal matrices $D_R = e^{-\beta_R H}$, $D_L = e^{-\beta_L H}$, $D_{m_R} = e^{-\beta_{m_R} H}$, $D_{m_L} = e^{-\beta_{m_L} H}$, the numerator reads

$$\frac{[\langle \mathbf{i_1}|\mathbf{i_2}\rangle\cdots\langle \mathbf{i_n}|\mathbf{i_1}\rangle]_{\mathrm{GPI}}}{\mathcal{P}_n} = \sum_{\sigma,\eta,\phi\in S_n} \kappa_R^{q(\tau^{-1}\circ\eta)}\kappa_L^{q(\tau^{-1}\circ\phi)}\,\mathrm{Tr}_\eta(D_R)\,\mathrm{Tr}_{\eta^{-1}\circ\sigma}(D_{m_R})\,\mathrm{Tr}_{\phi^{-1}\circ\sigma}(D_{m_L})\,\mathrm{Tr}_\phi(D_L)\,, \tag{6.3.35}$$

and the analogous denominator

$$\frac{[\langle \mathbf{i_1}|\mathbf{i_1}\rangle\cdots\langle \mathbf{i_n}|\mathbf{i_n}\rangle]_{\mathrm{GPI}}}{\mathcal{P}_n} = \sum_{\xi,\gamma,\theta\in S_n} \kappa_R^{q(\theta)}\kappa_L^{q(\gamma)}\,\mathrm{Tr}_\theta(D_R)\,\mathrm{Tr}_{\theta^{-1}\circ\xi}(D_{m_R})\,\mathrm{Tr}_{\gamma^{-1}\circ\xi}(D_{m_L})\,\mathrm{Tr}_\gamma(D_L)\,. \tag{6.3.36}$$

The additional permutations η, ϕ (or θ, γ) control the sub-leading connectivity of the right and left state-counting shells. In the planar limit they are fixed to maximise the number of index loops, reducing the triple sum to a single sum over σ and producing the exact cancellation (6.3.34). At finite κ they are summed over freely, and the κ-dependent weights $\kappa_R^{m(\tau^{-1}\circ\eta)}, \kappa_L^{m(\tau^{-1}\circ\phi)}$ in the numerator differ from $\kappa_R^{m(\theta)}, \kappa_L^{m(\gamma)}$ in the denominator. The substitution $\sigma \mapsto \tau^{-1}\circ\sigma$ no longer suffices to equate numerator and denominator, and the normalised moments depend explicitly on β_{m_L} and β_{m_R}.

Bi-ensemble Wishart structure. Nevertheless, the rank remains independent of the intermediate parameters. Under the identification $\overline{Z}(\beta) = \mathrm{Tr}(e^{-\beta H})$, the finite-$\kappa$ moments (6.3.35)–(6.3.36) match those of a bi-ensemble Wishart matrix (Appendix B.2.3) with $D_2 = D_R$ and $D_4 = D_L$ controlling the right and left operator insertions, and $\mathcal{O}_{\mathrm{obs}}$ playing the role of the bridge operator $B = D_3^{1/2} A D_1^{1/2}$ (where $D_1 = D_{m_L}$ and $D_3 = D_{m_R}$). The mass non-degeneracy across the two operator sets is essential: it ensures that the right and left ensembles contribute independent Gaussian draws with no cross-contractions, matching the mutual independence of $\{\mathcal{O}_L\}$ and $\{\mathcal{O}_R\}$ in Appendix B.2.3.

As in the single-boundary case, the non-planar corrections render the moments non-analytic in n, but the rank follows from structural properties of the ensemble. By the rank theorem for the bi-ensemble construction,

$$\mathrm{rank}(G) = \min(\kappa_R \kappa_L,\, r_R\, r_L) \quad \text{a.s.}, \tag{6.3.37}$$

where $r_R = \mathrm{rank}(D_R)$ and $r_L = \mathrm{rank}(D_L)$. In a microcanonical window this gives $\min(\kappa_R\kappa_L, e^{\mathbf{S}(E_R)+\mathbf{S}(E_L)})$, reproducing the planar result. The bridge operator B enters only through a non-degeneracy condition ($B \neq 0$, guaranteed by the non-vanishing of the shell overlap). Thus, while the intermediate parameters genuinely affect the eigenvalue distribution at finite κ — unlike the planar limit, where they cancel from the moments entirely — they do not affect the rank. The Wishart rank theorem provides a structural explanation: the independence of the rank from the bridge parameters is a consequence of the bi-ensemble structure itself, not an artefact of the planar resummation.

Ignorant Gram matrix. We now apply the observer rules. *Observer rule A1* restricts the $\mathcal{O}_{\mathrm{obs}}$ shells to self-contract, fixing $\sigma = e$ in (6.3.35)–(6.3.36). Setting $\sigma = e$ gives $\mathrm{Tr}_{\eta\circ e}(D_{m_R}) = \mathrm{Tr}_\eta(D_{m_R})$ and $\mathrm{Tr}_{\phi\circ e}(D_{m_L}) = \mathrm{Tr}_\phi(D_{m_L})$, so the sums over η and ϕ decouple. The A1 moments factorise:

$$[\mathrm{Tr}(\hat{G}^n)]_{\mathrm{GPI,\,A1}} = \frac{\sum_\eta \kappa_R^{q(\tau^{-1}\circ\eta)} \mathrm{Tr}_\eta(D_R)\mathrm{Tr}_{\eta^{-1}}(D_{m_R})}{\sum_\theta \kappa_R^{q(\theta)} \mathrm{Tr}_\theta(D_R)\mathrm{Tr}_{\theta^{-1}}(D_{m_R})} \times \frac{\sum_\phi \kappa_L^{q(\tau^{-1}\circ\phi)} \mathrm{Tr}_\phi(D_L)\mathrm{Tr}_{\phi^{-1}}(D_{m_L})}{\sum_\gamma \kappa_L^{q(\gamma)} \mathrm{Tr}_\gamma(D_L)\mathrm{Tr}_{\gamma^{-1}}(D_{m_L})}\,. \tag{6.3.38}$$

Each factor is precisely the single-boundary A1 ignorant moment (6.3.25): one for the R boundary (with covariance $D_R \otimes D_{m_R}$) and one for the L boundary (with covariance $D_L \otimes D_{m_L}$). The factorisation arises because A1 fixes $\sigma = e$, removing the coupling between the L and R sectors that was mediated by the compositions $\eta^{-1}\circ\sigma$ and $\phi^{-1}\circ\sigma$ in the global moments. In the global case this coupling was harmless

— it cancelled by the τ-bijection argument — but under A1 it is absent altogether. *Observer rule B1* replaces the two-boundary state by $\rho_{\beta_{\text{eff},L}} \otimes \rho_{\beta_{\text{eff},R}}$ via (6.2.22), and the ignorant Gram matrix factorises directly into single-boundary B1 ignorant Gram matrices.

In both cases the effective dimension factorises:

$$d_{\text{ignorant}} = d_{\text{ignorant},L} \times d_{\text{ignorant},R}\,. \tag{6.3.39}$$

By the Wishart rank theorems, the product rank is the product of the individual ranks:

$$\text{rank}(\hat{G}) = \min(\kappa_L,\, e^{\mathbf{S}(E_L)+\mathbf{S}(E_{\text{obs},L})}) \times \min(\kappa_R,\, e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\text{obs},R})}) \tag{6.3.40}$$

in the microcanonical case, and $d^2_{\text{canon}} \times d^2_{\text{canon}} = d^4_{\text{canon}}$ in the fully canonical large-κ limit. This is enlarged relative to the global $e^{\mathbf{S}(E_L)+\mathbf{S}(E_R)}$ by the factor $e^{\mathbf{S}(E_{\text{obs},L})+\mathbf{S}(E_{\text{obs},R})}$ counting the unobserved interior microstates on each side.

6.3.5.2 Single-boundary probes

The qualitatively new feature appears when probes are restricted to one boundary. Take the right boundary, giving the states

$$|\tilde{\mathbf{i}}\rangle = e^{-\frac{\beta_{m_L} H_L}{2}}\, \mathcal{O}_{\text{obs}}\, e^{-\frac{\beta_{m_R} H_R}{2}}\, \mathcal{O}_i\, e^{-\beta_R H_R/2}\,. \tag{6.3.41}$$

In the planar limit, a single permutation $\sigma \in S_n$ again controls the $\mathcal{O}_{\text{obs}}$ shells, but now the L-side contribution to the numerator involves the cycles of σ alone (no τ shift, since there are no L state-counting shells imposing one) while the R-side intermediate factor involves the cycles of $\tau^{-1} \circ \sigma$. The result is

$$Z_n = \frac{\overline{Z}(n\beta_R)}{\overline{Z}(\beta_R)^n} \times \frac{\sum_{\sigma\in S_n} \prod_{j=1}^{q(\sigma)} \overline{Z}(k_j(\sigma)\beta_{m_L}) \prod_l \overline{Z}(\tilde{k}_l(\tau^{-1}\circ\sigma)\beta_{m_R})}{\sum_{\gamma\in S_n} \prod_{w=1}^{q(\gamma)} \overline{Z}(k_w(\gamma)\beta_{m_L})\overline{Z}(k_w(\gamma)\beta_{m_R})}\,. \tag{6.3.42}$$

Crucially, unlike (6.3.21), the permutation sums do *not* cancel: the cycles of σ and $\tau^{-1}\circ\sigma$ are generically of different lengths. The single-boundary result is recovered in the limit $\beta_{m_L} \to \infty$, where for $\beta_{m_L} > \beta_{\text{HP}}$ the partition function satisfies $\overline{Z}(\beta_{m_L})^n = \overline{Z}(n\beta_{m_L})$ and the L-dependent factor becomes the same constant in numerator and denominator.

Min-vs-max analytic continuation. The second factor in (6.3.42) does not admit a straightforward analytic continuation to $n \to 0$, because intermediate cycle lengths give $\overline{Z}(k_j\beta)$ contributions with negative local temperatures [31]. Restricting to replica-symmetric saddles — the stabilisers $\sigma = \tau^k$ of τ — and projecting to microcanonical windows in E_{m_L}, E_{m_R}, E_R gives

$$Z_n|_{E_{m_L},E_{m_R},E_R} = e^{(1-n)\mathbf{S}(E_R)} \left(\frac{e^{(1-n)\mathbf{S}(E_{m_L})} + e^{(1-n)\mathbf{S}(E_{m_R})}}{\varphi(n)\, e^{(1-n)(\mathbf{S}(E_{m_L})+\mathbf{S}(E_{m_R}))} + 1} \right), \tag{6.3.43}$$

where $\varphi(n)$ is the Euler totient. This expression is non-analytic in n, and the rank depends on the order of two operations that do not commute. *Approximating first*: for $n > 1$ the φ term is subdominant and one finds

$$\log d \approx \mathbf{S}(E_R) + \min[\mathbf{S}(E_{m_L}), \mathbf{S}(E_{m_R})]\,. \tag{6.3.44}$$

Continuing first: in the large-κ limit, naive substitution with $\varphi(0) = 0$ gives

$$\log d \approx \mathbf{S}(E_R) + \max[\mathbf{S}(E_{m_L}), \mathbf{S}(E_{m_R})]\,. \tag{6.3.45}$$

The two routes give min versus max. The max result is problematic — it can exceed d^2_{canon}, violating the algebraic bound on the two-boundary Hilbert space. From linear algebra one expects $d \geq \text{rank}(\mathcal{O}_{\text{obs}}) \times e^{\mathbf{S}(E_R)}$, and for a Gaussian random $\mathcal{O}_{\text{obs}}$ this rank is $\min(r_{m_L}, r_{m_R})$ almost surely, favouring (6.3.44).

The Wishart resolution. The tension is resolved by working at finite κ. Including all topologies, the moments are labelled by two permutations $\sigma, \eta \in S_n$:

$$\frac{[\langle\tilde{\mathbf{i}}_1|\tilde{\mathbf{i}}_2\rangle \cdots \langle\tilde{\mathbf{i}}_n|\tilde{\mathbf{i}}_1\rangle]_{\text{GPI}}}{\mathcal{P}_n} = \sum_{\sigma,\eta} \kappa^{q(\tau^{-1}\circ\eta)} \text{Tr}_\eta(D_R)\text{Tr}_{\sigma^{-1}\circ\eta}(D_{m_R})\text{Tr}_\sigma(D_{m_L})\,, \tag{6.3.46}$$

and analogously for the denominator. Unlike the two-boundary-probes case, the permutation sums over σ do not cancel between numerator and denominator. Under the identification $\overline{Z}(\beta) = \text{Tr}(e^{-\beta H})$, these moments match those of the trace construction (Appendix B.2.2) with $D_1 = D_{m_L}$, $D_2 = D_R$, $D_3 = D_{m_R}$, where the state-dependent random matrix $C_A = D_1^{1/2} A^\dagger D_3 A D_1^{1/2}$ encodes the $\mathcal{O}_{\text{obs}}$ bridge. The non-planar corrections render the moments non-analytic in n, but the rank theorem gives

$$\text{rank}(G) = \min(\kappa,\, r_2 \cdot \min(r_1, r_3)) \qquad \text{a.s.}, \tag{6.3.47}$$

which in a microcanonical window with both E_{m_L}, E_{m_R} above threshold reduces to

$$\text{rank}(G) = \min(\kappa,\, e^{\mathbf{S}(E_R)+\min[\mathbf{S}(E_{m_L}),\mathbf{S}(E_{m_R})]})\,. \tag{6.3.48}$$

This is (6.3.44), not (6.3.45): the Wishart rank theorem settles the min–max tension in favour of min. Physically, the rank is controlled by the bottleneck $\min(r_1, r_3)$ — the smaller of the two internal Hilbert space dimensions through which the bridge C_A maps.

Ignorant Gram matrix. Under observer rule A1, fixing $\sigma = e$ in the finite-κ moments gives $\mathrm{Tr}_e(D_{m_L}) = \overline{Z}(\beta_{m_L})^n$ in both numerator and denominator. The dependence on β_{m_L} *drops out entirely*: with the obs shell forced to self-contract on each boundary, the closed loop of length β_{m_L} on the L side cancels in the ratio. This is physically expected: under A1 the L sector is completely traced out, since the probes act only on R. The remaining structure

$$[\mathrm{Tr}(\hat{G}^n)]_{\mathrm{GPI,\,A1}} = \frac{\sum_\eta \kappa^{q(\tau^{-1}\circ\eta)}\mathrm{Tr}_\eta(D_R)\mathrm{Tr}_\eta(D_{m_R})}{\sum_\theta \kappa^{q(\theta)}\mathrm{Tr}_\theta(D_R)\mathrm{Tr}_\theta(D_{m_R})} \tag{6.3.49}$$

is precisely the single-boundary A1 ignorant moment with $D_{\mathrm{obs}} \to D_{m_R}$. By the Wishart rank theorem,

$$\mathrm{rank}(\hat{G})_{\mathrm{A1}} = \min(\kappa,\, e^{\mathbf{S}(E_R)+\mathbf{S}(E_{m_R})})\,, \tag{6.3.50}$$

which exceeds the global result (6.3.48) whenever $\mathbf{S}(E_{m_L}) < \mathbf{S}(E_{m_R})$: the observer's ignorance has eliminated the bottleneck constraint from the boundary they cannot probe. Observer rule B1 reduces directly to the single-boundary B1 ignorant Gram matrix and gives the same result in the heavy mass limit. The resulting ignorant dimension is d^2_{canon}, $e^{\mathbf{S}(E_R)} \times d_{\mathrm{canon}}$, $d_{\mathrm{canon}} \times e^{\mathbf{S}(E_{\mathrm{obs},R})}$, or $e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\mathrm{obs},R})}$ depending on the projections.

6.4 Maximal ignorance ensembles

The observer rules of Secs. 6.2 and 6.3 replace the unknown pure state by a coarse-grained density matrix — $\overline{\rho_\Psi}$ for A1, $\rho_{\beta_{\mathrm{eff}}}$ for B1 — and compute all quantities in this effective state. For linear observables this is sufficient. The rank of the ignorant Gram matrix, however, is not linear: it is $\lim_{n\to 0}\mathrm{Tr}(\hat{G}^n)$ and involves n insertions of the state. Since all insertions refer to the same underlying microstate they are in principle correlated, and a single density matrix does not encode such correlations: each insertion is replaced independently, with no cross-terms. To capture the correlations one must promote the single effective state to a probability distribution — an ensemble — whose higher moments encode the inter-insertion correlations. The same ensemble can also be used to quantify the variance of complex observables across compatible microstates. The form of the ensemble depends on a further choice of prior.

The origin of the mixedness. This prior decides on a question that the observer's simple measurements cannot resolve: *why* is the effective state mixed? The single-boundary observer sees $\overline{\rho_\Psi}$ of rank $r \sim e^{\mathbf{S}}$. The mixedness is a fact of their description, but its origin admits several physically distinct interpretations:

(i) **Globally pure, single boundary.** The mixedness is entirely an artefact of coarse-graining: dephasing projects the pure state onto $\overline{\rho_\Psi}$, but the "true" state is pure. Different microstates in the support of $\overline{\rho_\Psi}$ share the same single-boundary Hilbert space, and the correlations between insertions of ρ are those of a single pure state.

(ii) **Globally pure, multiple components.** The universe has additional asymptotic boundaries or closed-universe partners. The single-boundary state is genuinely mixed because it is entangled with inaccessible components. Inter-insertion correlations are diluted across the purifying system.

(iii) **No global assumption.** The observer simply adopts $\overline{\rho_\Psi}$ as their complete description and computes all quantities — linear and nonlinear — from it directly, treating each insertion independently.

For linear observables all three give identical predictions. The choice between them is a prior on global topology, not a measurement outcome, and becomes consequential only for nonlinear and/or complex quantities.

Three priors, three predictions. We will show that each interpretation leads to a distinct prescription for $\mathrm{rank}(\hat{G})$. *Scenario (iii)* is the single-state approach of Secs. 6.2–6.3: each insertion of ρ in $\mathrm{Tr}(\hat{G}^n)$ is independently replaced by $\overline{\rho_\Psi}$, generating no cross-contractions; on the gravity side this is the topology restriction of observer rule A1, with $d_{\text{ignorant}} = r \times d_R$. *Scenario (i)* is a pure-state ensemble: the observer models their ignorance as a distribution over pure states $|\Psi\rangle$ compatible with $\overline{\rho_\Psi}$; since all insertions are drawn from the same pure state, the ensemble average generates cross-contractions — precisely the topologies that A1 projects out. These cross-contractions restore the fine-tuned null relations (6.3.8) that reduce the rank, and the effective dimension collapses to d_{global}. *Scenario (ii)* is a mixed-state ensemble (a distribution over single-boundary mixed states, equivalently over two-boundary purifications): cross-contractions are present but act on the purifying system $\mathcal{H}_L$, not the observer's boundary $\mathcal{H}_R$, and do not restore the null relations on the observer's Hilbert space. The result is again $d_{\text{ignorant}} = r \times d_R$.

A single-boundary observer who sees a mixed state has no operational basis for assuming the universe is globally pure with a single boundary — scenario (i). The observed mixedness could equally well arise from tracing out a second boundary, a closed universe, or any other inaccessible component. Scenarios (ii) and (iii) make no such assumption, and their agreement on $d_{\text{ignorant}} = r \times d_R$ suggests that this is the robust answer; the single-state observer rules of Sec. 6.3 are reproduced by the physically appropriate ensemble construction.

The ensembles themselves are not ad hoc. They arise from two complementary routes: a top-down route in which the generalised ETH ansatz (6.2.2) defines a Gaussian (Ginibre) ensemble, and constraining a maximal-entropy distribution to have mean $\overline{\rho_\Psi}$ reproduces this ensemble independently from the max-ent principle. The second is bottom-up: an observer who has measured a set of simple-operator expectation values constructs a maximal-entropy ensemble consistent with this data, arriving at the same result when the measurements are functionally complete. The one exception is the type B setting, where the observer's measurements are too limited to fix the full diagonal weights.

In the remainder of this section we develop the maximal-entropy framework (Sec. 6.4.1), construct the pure-state ensembles corresponding to scenario (i) (Sec. 6.4.2), the mixed-state ensembles corresponding to scenario (ii) (Sec. 6.4.3), and compare the resulting observer rules and effective dimensions (Sec. 6.4.4).

6.4.1 The maximal ignorance framework

Rather than selecting a single representative density matrix, one may describe the equivalence class of states compatible with the observer's data by a probability measure $\mu_\Theta(\rho)$ on the space of states, constrained by

$$\int d\rho\, \mu_\Theta(\rho)\, \mathrm{Tr}(\rho\,\Theta) \stackrel{!}{=} \langle\Theta\rangle \qquad \forall\,\Theta \in \mathcal{S}_{\text{obs}}\,. \tag{6.4.1}$$

We adopt the measure that maximises the classical Shannon entropy $\mathcal{S}[\mu] = -\int d\rho\, \mu(\rho) \log \mu(\rho)$ subject to (6.4.1), following [26].[7] For ensembles of pure states $|\Psi\rangle = \sum_i A_i\,|i\rangle$ in a D-dimensional Hilbert space, the prior is the $U(D)$-invariant Haar measure $d\,|\psi\rangle = dA\,\delta(1 - AA^*)$, and extremising the Shannon entropy subject to constraints on $\langle O_\alpha\rangle = C_\alpha$ or $\overline{|\psi\rangle\,\langle\psi|} = \hat{\rho}$ yields

$$\mu(A) = \frac{1}{N}\,\exp(-\langle\Psi|\Lambda|\Psi\rangle) = \frac{1}{N}\,\exp(-\Lambda_{ij}A_i^*A_j)\,, \tag{6.4.2}$$

[7] Other choices of entropy functional or prior would generically produce inequivalent ensembles. Shannon entropy with Haar prior is the natural analogue of the von Neumann entropy used to derive the single effective state.

where Λ is Hermitian: $\Lambda = \sum_\alpha \lambda_\alpha O_\alpha$ in the first case, or the matrix of Lagrange multipliers in the second. The fixed-norm condition $\delta(1-AA^*)$ introduces correlations between the A_i.

The Gaussian approximation. Working in the eigenbasis of Λ (eigenvalues a_i) and defining $p_i \equiv |A_i|^2$, the partition function reduces to a simplex integral

$$Z(a) = \int_{\Delta_{D-1}} dp\, e^{-\sum_i a_i p_i} = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} ds\, \frac{e^s}{\prod_{i=1}^D (s+a_i)}\,, \tag{6.4.3}$$

giving the exact finite-D result $Z(a) = \sum_i e^{-a_i}/\prod_{j\neq i}(a_j - a_i)$. Moments are obtained from $\overline{p_i} = -\partial_{a_i} \log Z$. In the thermodynamic limit $D \to \infty$ the Laplace integral is dominated by a saddle at $s = s_*$ satisfying

$$1 = \sum_{i=1}^{D} \frac{1}{s_* + a_i}\,, \tag{6.4.4}$$

and the moments become those of independent exponentials $\mu(p_i) \propto e^{-(s_*+a_i)p_i}$, factorising at large D. The fixed-norm ensemble is therefore well approximated by an unconstrained shifted Gaussian. The unconstrained Gaussian $\mu(p_i) \propto e^{-a_i p_i}$ (i.e. $s_* = 0$) is a valid approximation when (a) the norm is sharply concentrated, $\mathrm{Tr}(\rho^2)/(\mathrm{Tr}\rho)^2 \to 0$ as $D \to \infty$ (requiring ρ sufficiently mixed); and (b) the identity lies in the span of the constraints, $\sum_\alpha b_\alpha O_\alpha = \mathbb{1}$ with $\sum_\alpha b_\alpha C_\alpha = 1$. Details are given in Appendix B.3.

Two key examples. *Fixing the mean density matrix.* If the ensemble is constrained to fix $\overline{\rho} = \hat{\rho}$ with eigenvalues χ_i, then $a_i = \chi_i^{-1}$ and $s_* = 0$. The algebraic condition is automatic, and self-averaging holds whenever $\hat{\rho}$ is sufficiently mixed. The resulting ensemble is

$$|\Psi\rangle = \sum_i \sqrt{\chi_i}\, R_i\, |i\rangle\,, \tag{6.4.5}$$

with R_i independently drawn from a complex unit Gaussian. *Fixing a single observable.* If only $\langle O\rangle = C$ is constrained, the identity is generically not in the span of $\{O\}$, the algebraic condition fails, $s_* \neq 0$, and the mean state of the resulting ensemble can deviate significantly from the single-state max-entropy result $\propto e^{-\lambda O}$. This failure will matter for the type B ensembles below.

6.4.2 Ensemble of pure states

We construct max-entropy ensembles of single-boundary pure states for the two settings of Sec. 6.2: type A (functionally complete simple measurements, leading to A1) and type B (Hamiltonian and low-backreaction probes, leading to B1). For Type A we present three logically independent derivations — top-down, observer, and Euclidean time — and check that they all yield the same Gaussian ensemble.

6.4.2.1 Type A: functionally complete simple measurements

Top-down construction. The effective state compatible with all simple measurement data was identified in Sec. 6.2.1 as $\overline{\rho_\Psi}$ of (6.2.6). By the consistency condition (6.1.7) of Sec. 6.1.2, any ensemble whose mean is $\overline{\rho_\Psi}$ reproduces the correct predictions for all simple operators. The maximal-entropy ensemble with this mean is obtained from (6.4.5) with $\hat{\rho} = \overline{\rho_\Psi}$:

$$|\Psi\rangle = \sum_i \frac{e^{-f(E_i)/2-\beta_{\rm obs}E_i/2}}{\sqrt{\hat{Z}(\beta_{\rm obs})}} R_i \, |i\rangle \ , \tag{6.4.6}$$

where the R_i are independently drawn from a complex unit Gaussian and $\hat{Z}(\beta_{\rm obs}) = \sum_n e^{-f(E_n)-\beta_{\rm obs}E_n}$. This is precisely the gravitational state ensemble (6.2.4) derived in Sec. 6.2.1 from the generalised ETH ansatz. The self-averaging argument of Sec. 6.2.1 confirms that the Gaussian approximation is valid: $\mathrm{Tr}(\overline{\rho_\Psi}^2)/(\mathrm{Tr}\overline{\rho_\Psi})^2 \sim e^{-\mathbf{S}}$.

Observer construction. An observer who has measured a functionally complete set of simple operators $\{\mathcal{A}_\alpha\} \subset \mathcal{S}_{\rm obs}$ imposes the constraints

$$\langle \mathcal{A}_\alpha \rangle = [\langle \Psi_{\rm obs}|\mathcal{A}_\alpha|\Psi_{\rm obs}\rangle]_{\rm GPI} \, . \tag{6.4.7}$$

Since these operators satisfy the ETH ansatz (6.1.1), they are approximately diagonal in the energy eigenbasis. The max-ent distribution takes the Gaussian form (6.4.2) with

$$\mu(A) = \frac{1}{N} \exp\left(- \sum_i \frac{|A_i|^2}{\sigma(E_i)} \right) , \qquad \sigma(E_i)^{-1} = \sum_\alpha \lambda_\alpha \, m_\alpha(E_i) \, , \tag{6.4.8}$$

where $m_\alpha(E)$ are the smooth ETH envelopes. As argued in Sec. 6.1.3, functional completeness means that the $\{m_\alpha(E)\}$ span the space of smooth functions of energy over the relevant microcanonical window, so the variance $\sigma(E_i)$ is uniquely determined by the constraints. Since the gravitational amplitudes on the right-hand side of (6.4.7) are themselves computed from $\overline{\rho_\Psi}$, matching requires $\sigma(E_i)$ to be the eigenvalues of $\overline{\rho_\Psi}$, reproducing (6.4.6).

Euclidean time evolution constraint. The same ensemble can be derived from the continuous family of constraints

$$\langle e^{-TH}\rangle = [\langle\Psi_{\rm obs}|e^{-TH}|\Psi_{\rm obs}\rangle]_{\rm GPI} \qquad \forall T \geq 0\,. \tag{6.4.9}$$

This is the single-boundary version of the construction analysed in [26]. Extremising the Shannon entropy subject to (6.4.9) with Lagrange multiplier $\lambda(T)$ yields the Gaussian measure (6.4.2) with $\sigma(E_i)^{-1} = \int dT\,\lambda(T)\,e^{-TE_i}$ — the Laplace transform of $\lambda(T)$. Solving the constraint using the GPI amplitude (6.2.7) for the right-hand side gives

$$\sigma(E_i) = \frac{e^{-f(E_i)-\beta_{\rm obs}E_i}}{\sum_n e^{-f(E_n)-\beta_{\rm obs}E_n}}\,, \tag{6.4.10}$$

again reproducing (6.4.6). The Euclidean time evolution constraint is not a directly measurable quantity in any straightforward operational sense; its utility is that it packages all diagonal information into a single continuous family parametrised by T. The expectation value $\langle e^{-TH}\rangle = \sum_i |A_i|^2 e^{-TE_i}$ is the Laplace transform of the energy distribution $\{|A_i|^2\}$, and constraining it for all T is equivalent to fixing the full smooth energy profile $\sigma(E)$. Setting $T = 0$ enforces normalisation $\langle\mathbb{1}\rangle = 1$, automatically satisfying the algebraic condition. This is precisely the information content of a functionally complete set of simple operators, packaged in generating-function form.

Observer rule A2 and the effective dimension. The Gaussian ensemble (6.4.6) is precisely the gravitational state ensemble (6.2.4). Since the gravitational path integral computes ensemble averages over this distribution, adopting it as the observer's prior reproduces the full sum over topologies, including both self- and cross-contractions of the $\mathcal{O}_{\rm obs}$ shells. The corresponding observer rule is:

> **Observer rule A2**: *To capture the epistemic ignorance of an observer who has measured a functionally complete set of simple operators in a shell state $|\Psi_{\rm obs}\rangle$ using an ensemble of single-boundary pure states, perform the full sum over topologies.*

The rank of the ignorant Gram matrix collapses to the global Gram matrix computation of Sec. 6.3.3:

$$d^{\rm A2}_{\rm ignorant} = d_{\rm canon} \quad \text{(canonical)} \quad \text{or} \quad e^{\mathbf{S}(E_R)} \quad \text{(microcanonical)}\,, \tag{6.4.11}$$

strictly smaller than the result $r \times d_R$ of A1: the state cross-contractions restore the fine-tuned null relations (6.3.8) that A1 projected out, collapsing the effective space

back to the global Hilbert space. The discrepancy reflects the choice of prior, not an error in either rule. A1 adopts the single density matrix $\overline{\rho_\Psi}$ as the complete description, with each insertion of ρ in the nonlinear quantity $\mathrm{Tr}(\hat{G}^n)$ independently replaced by $\overline{\rho_\Psi}$. A2 supplements this with the Gaussian distribution over compatible microstates, encoding correlations between different state insertions that arise because all refer to the same underlying microstate. The two rules agree on all linear observables (by (6.1.7)) but differ for nonlinear quantities. From the observer's perspective, adopting A2 amounts to a strong prior on the global structure: that the universe is globally a single-boundary pure state, with the observed mixedness an artefact of coarse-graining rather than of tracing out additional degrees of freedom.

6.4.2.2 Type B: Hamiltonian and low-backreaction probes

We now construct the maximal-entropy pure-state ensemble compatible with the data that led to observer rule B1 in Sec. 6.2.3.2: the expectation value of the ADM Hamiltonian and a small number of low-backreaction probes, which together determine only the effective inverse temperature β_{eff}.

The Hamiltonian constraint. The most natural starting point is to constrain only the expectation value of the Hamiltonian: $\langle H\rangle = [\langle\Psi_{\text{obs}}|H|\Psi_{\text{obs}}\rangle]_{\text{GPI}}$. However, this does *not* yield a physically meaningful ensemble. The Hamiltonian is a single non-trivial observable, so the identity is not in its span and the algebraic condition (6.4.4) fails. The saddle condition then requires a nonzero shift s_*, and in the thermodynamic limit the mean density matrix $\overline{\rho}$ deviates significantly from the Gibbs state $e^{-\beta H}/Z(\beta)$. The physical origin of this failure is structural: a single constraint in the D-dimensional space of pure states leaves $D-1$ directions unconstrained, and the fixed-norm condition forces the distribution to be nearly uniform across these directions, producing a mean state much closer to the maximally mixed state $D^{-1}\mathbb{1}$ than to the Gibbs state. As shown in Appendix B.4, while the mean probability density of the total energy does concentrate around the saddle-point energy E_* in a Gaussian window of thermodynamic width σ, the state-level distribution (the full density matrix) is far from thermal.

Constraining the mean density matrix. A more physical implementation of "thermal" coarse-grained data is to constrain the mean density matrix directly:

$$\overline{\rho} \stackrel{!}{=} \rho_{\beta_{\text{eff}}} \equiv \frac{e^{-\beta_{\text{eff}}H}}{Z(\beta_{\text{eff}})}\,. \tag{6.4.12}$$

This is precisely the situation analysed around (6.4.5). Since $\mathrm{Tr}(\rho_{\beta_{\rm eff}}) = 1$ the algebraic condition is automatically satisfied; since $\mathrm{Tr}(\rho^2_{\beta_{\rm eff}}) \sim e^{-\mathbf{S}} \to 0$ the self-averaging condition holds. Applying (6.4.5) yields the ensemble

$$|\Psi\rangle = \sum_i \frac{e^{-\beta_{\rm eff} E_i/2}}{\sqrt{Z(\beta_{\rm eff})}} R_i \, |i\rangle \ , \tag{6.4.13}$$

which coincides with the type A ensemble (6.4.6) in the heavy shell mass limit, where $f(E) \to 0$ and $\beta_{\rm eff} \to \beta_{\rm obs}$.

Since the thermal ensemble (6.4.13) coincides with the type A ensemble in the heavy shell mass limit, it has a clear gravitational interpretation:

> **Observer rule B2**: *To capture the epistemic ignorance of an observer who has measured a small set of simple, low-backreaction operators in a shell state $|\beta_{\rm obs}, m_i\rangle$ using an ensemble of single-boundary pure states, replace the shell state with the infinite-mass shell state $|\beta_{\rm eff}, m_i = \infty\rangle$ and perform the full sum over topologies.*

Since the rank is computed in the heavy mass limit, $d^{\rm B2}_{\rm ignorant} = d^{\rm A2}_{\rm ignorant} = d_{\rm global}$. As we show next, relaxing the pure-state prior to a mixed-state ensemble reproduces the single-state result.

6.4.3 Ensemble of mixed states

A single-boundary observer who sees a mixed state $\overline{\rho_\Psi}$ need not assume the universe is globally pure. The observed mixedness could equally arise from tracing out additional boundaries, closed universes, or other inaccessible components. A less restrictive prior models the observer's ignorance as a distribution over single-boundary mixed states rather than pure states. Any mixed state ρ_R on $\mathcal{H}_R$ admits a canonical purification as a two-boundary pure state: $|\Psi\rangle_{LR} = \sum_{i,j} A_{ij} \, |j\rangle_L \, |i\rangle_R$ with $\rho_R = AA^\dagger$. Constructing the mixed-state ensemble via a distribution over two-boundary purifications automatically enforces positivity and the trace condition. The natural prior is the $U(D)_L \times U(D)_R$-invariant measure on the space of two-boundary pure states with fixed-trace condition $\delta(1 - \mathrm{Tr}(AA^\dagger))$.

The analysis closely parallels that of Sec. 6.4.2. We impose constraints on the single-boundary data — expectation values of operators in the R-boundary Hilbert space — derived from the same gravitational amplitudes that led to the observer rules of Sec. 6.2. The crucial new feature is that the variance of each coefficient A_{ij} is suppressed by the dimension D of the purifying Hilbert space $\mathcal{H}_L$, qualitatively changing the effective dimension.

6.4.3.1 Type A: functionally complete simple measurements

Top-down construction. We constrain the mean reduced state on $\mathcal{H}_R$ to be $\overline{\rho_\Psi}$ of (6.2.6). Writing $dA = \prod_{i,j} dA_{ij}\, dA_{ij}^*$, the max-ent distribution with Lagrange multiplier Λ_i acting on the R-index gives

$$\mu(A) = \frac{1}{N} \exp\left(- \sum_{i,j} \frac{|A_{ij}|^2}{\sigma(E_i)} \right) , \tag{6.4.14}$$

where the variance depends only on the R-index energy E_i. Note that the L-index j runs freely over the purifying Hilbert space. Solving the constraint $(\rho_R)_{ii} = \sum_j |A_{ij}|^2 = D\,\sigma(E_i) \overset{!}{=} (\overline{\rho_\Psi})_{ii}$ gives

$$\sigma(E_i) = \frac{1}{D} \frac{e^{-f(E_i)-\beta_{\rm obs}E_i}}{\sum_n e^{-f(E_n)-\beta_{\rm obs}E_n}} . \tag{6.4.15}$$

The crucial difference from the pure-state case (6.4.10) is the factor of D^{-1}: the variance of each coefficient is suppressed by the dimension of the purifying Hilbert space. For a microcanonical shell state $D \sim e^{\mathbf{S}(E)}$; for a canonical state $D \sim d_{\rm canon}$. In either case the variance is at most $\mathcal{O}(e^{-1/G_N})$. The resulting two-boundary ensemble is

$$|\Psi\rangle_{LR} = \sum_{i,j} \frac{e^{(-f(E_i)-\beta_{\rm obs}E_i)/2}}{(D\hat{Z}(\beta_{\rm obs}))^{1/2}} R_{ji}\, |j\rangle_L\, |i\rangle_R , \tag{6.4.16}$$

where the R_{ji} are independently drawn from a complex unit Gaussian. For microcanonical gravitational input the sum over i is restricted to the relevant window, while the sum over j runs over the entire spectrum.

The mean reduced state on $\mathcal{H}_R$ is

$$\overline{\rho}_R = \overline{\sum_k A_{ik} A_{jk}^*} = \delta_{ij} \frac{e^{-f(E_i)-\beta_{\rm obs}E_i}}{\sum_n e^{-f(E_n)-\beta_{\rm obs}E_n}} = (\overline{\rho_\Psi})_{ij} , \tag{6.4.17}$$

which is precisely the ensemble-averaged state (6.2.6), as required. The fluctuations take the form

$$\overline{\delta\rho_{ij}\delta\rho_{kl}} = \delta_{il}\delta_{jk}\, \sigma(E_i)\sigma(E_j) \times D , \tag{6.4.18}$$

which has the same structure as the pure-state fluctuations but is suppressed by $D^{-1} \sim e^{-\mathbf{S}}$ relative to them (since $\sigma \sim D^{-1}$, the product $\sigma^2 \times D \sim D^{-1}$ compared to $\sigma^2 \sim 1$ in the pure case). This suppression reflects the dilution of fluctuations across the D purifying degrees of freedom.

Observer construction. An observer who has measured a functionally complete set of simple operators on $\mathcal{H}_R$ imposes the same constraints (6.4.7) as in the pure-state case, but now on the two-boundary ensemble. Since the constraints involve only R-boundary operators, the max-ent distribution factorises along the L-index, producing (6.4.14) with the same variance (6.4.15). The observer construction therefore reproduces the top-down result.

Euclidean time evolution constraint. The constraint (6.4.9) applied to the two-boundary ensemble reads

$$\langle e^{-TH_R}\rangle = D\sum_i e^{-TE_i}\,\sigma(E_i) \stackrel{!}{=} [\langle\Psi_{\rm obs}|e^{-TH}|\Psi_{\rm obs}\rangle]_{\rm GPI}\,, \tag{6.4.19}$$

and is solved by the same $\sigma(E_i)$ as (6.4.15). Setting $T=0$ again enforces normalisation. The Euclidean time evolution constraint provides a unified derivation for both the pure- and mixed-state ensembles; the only difference is whether the unconstrained index sum runs over one (j absent) or D ($j=1,\ldots,D$) purifying degrees of freedom.

Gravitational realisation. To determine whether the mixed-state ensemble leads to a tractable observer rule, we must identify its gravitational preparation. Define the operator

$$\Gamma \equiv e^{-f(E_i)/2}\,R_{ij}\,|i\rangle\,\langle j|\,, \tag{6.4.20}$$

where $f(E_i)\equiv f(E_i,0)$ depends only on one of the two energies. This differs from the heavy shell operator $\Psi = e^{-f(E_i,E_j)/2}\,R_{ij}\,|i\rangle\,\langle j|$ of (6.2.2), whose envelope depends on both. Instances of the single-boundary ensemble (6.4.6) are prepared by acting with Ψ on the vacuum, while instances of the two-boundary ensemble (6.4.16) are prepared by inserting Γ between the two boundaries.

At finite shell mass $\Gamma\neq\Psi$, and the mixed-state ensemble does not correspond to any known gravitational preparation. In the heavy shell mass limit $m_{\mathcal{O}_{\rm obs}}\to\infty$, however, $f(E_n,E_m)$ becomes independent of the energies [6, 13], so $\Gamma\to\Psi$. In this limit both ensembles can be obtained as limits of a common reference state: the reduced density matrix $\mathrm{Tr}_{\mathcal{H}_L}(|\beta_L,\beta_{\rm obs}\rangle\,\langle\beta_L,\beta_{\rm obs}|)$ of a two-boundary shell state with preparation temperatures β_L and $\beta_{\rm obs}$. The single-boundary ensemble is recovered as $\beta_L\to\infty$ (the L-side geometry collapses to the vacuum); the mixed-state ensemble is recovered as $\beta_L\to 0$ (the L boundary opens up to the canonical dimension). The resulting observer rule is:

Observer rule A3: *To capture the epistemic ignorance of an observer who has measured a functionally complete set of simple operators in an infinite-mass shell state* $|\beta_{\text{obs}}, \infty\rangle$ *using an ensemble of mixed states, replace the single-boundary state with a two-boundary infinite-mass shell state* $|\beta_L, \beta_{\text{eff}}, m_i = \infty\rangle_{LR}$*, compute the quantity of interest using the full sum over topologies, and take* $\beta_L \to 0$ *at the end. The finite-mass case does not lead to a clear observer rule.*

Observer rule A2 is recovered from the same reference state by taking $\beta_L \to \infty$ instead. The two ensemble priors — single-boundary pure states (A2) and two-boundary purifications (A3) — are therefore continuously connected, with the preparation temperature β_L interpolating between them.

6.4.3.2 Type B: thermal measurements

As in the pure-state case, constraining only the expectation value of the Hamiltonian in the two-boundary ensemble encounters the algebraic obstruction: the identity on $\mathcal{H}_L \otimes \mathcal{H}_R$ is not in the span of $\{H_R\}$. The physically motivated implementation is to constrain the mean reduced state directly: $\overline{\rho}_R = \rho_{\beta_{\text{eff}}}$. This gives the two-boundary ensemble

$$|\Psi\rangle_{LR} = \sum_{i,j} \frac{e^{-\beta_{\text{eff}} E_i/2}}{\sqrt{D\, Z(\beta_{\text{eff}})}} R_{ij}\, |i\rangle_R\, |j\rangle_L \,, \tag{6.4.21}$$

which coincides with (6.4.16) in the heavy shell mass limit where $f(E) \to 0$ and $\beta_{\text{obs}} \to \beta_{\text{eff}}$. The average two-boundary state factorises as $\overline{|\Psi\rangle\langle\Psi|}_{LR} = \mathbb{1}_L/D \otimes \rho_{\beta_{\text{eff}},R}$. The thermal vs. Gibbs subtlety of Sec. 6.4.2.2 applies equally here: the mean reduced state must be determined by the single-state max-ent procedure of Sec. 6.2.3.2 before the ensemble is constructed, and the ensemble max-ent procedure then builds the least-informative distribution of two-boundary states centred on this thermal target. The corresponding observer rule is A3 with $\beta_{\text{obs}} \to \beta_{\text{eff}}$.

6.4.3.3 Effective dimension from the mixed-state ensemble

The rank of the ignorant Gram matrix under observer rule A3 is computed by applying the full topology sum to a two-boundary shell state with preparation temperatures β_L on L and β_{eff} on R, with probes acting only on R, in the heavy shell mass limit, followed by $\beta_L \to 0$. This is precisely the single-boundary-probe set-up of Sec. 6.3.5.2, with the identifications $\beta_{m_L} \to \beta_L$, $\beta_{m_R} \to \beta_{\text{eff}}$, $\beta_R \to \beta_R$. The finite-κ rank from (6.3.47) is

$$\text{rank}(G) = \min(\kappa,\, r_2 \cdot \min(r_1, r_3)) \,, \tag{6.4.22}$$

where $r_1 = \mathrm{rank}(e^{-\beta_L H})$, $r_2 = \mathrm{rank}(e^{-\beta_R H})$, and $r_3 = \mathrm{rank}(e^{-\beta_{\mathrm{eff}} H})$.

The limit $\beta_L \to 0$ sends $r_1 = \mathrm{rank}(e^{-\beta_L H}) \to d_{\mathrm{canon}}$, since all states in the Hilbert space contribute as $\beta_L \to 0$. Since $r_3 \leq d_{\mathrm{canon}}$ for any finite β_{eff}, the bottleneck factor becomes $\min(r_1, r_3) = \min(d_{\mathrm{canon}}, r_3) = r_3$. The rank is therefore

$$\mathrm{rank}(G)|_{\beta_L \to 0} = \min(\kappa,\, r_2 \times r_3) = \min(\kappa,\, e^{\mathbf{S}(E_R)+\mathbf{S}(E_{\mathrm{obs}})}) \tag{6.4.23}$$

in a microcanonical window. In the fully canonical case, $r_2 = r_3 = d_{\mathrm{canon}}$ and $\mathrm{rank}(G) = \min(\kappa, d_{\mathrm{canon}}^2)$. Physically, the $\beta_L \to 0$ limit removes the bottleneck that the bridge operator $\mathcal{O}_{\mathrm{obs}}$ imposed on the global rank in Sec. 6.3.5.2. In the global computation the rank was limited by $\min(r_1, r_3)$ — the smaller of the two internal Hilbert space dimensions through which the bridge maps. Taking $r_1 \to d_{\mathrm{canon}}$ ensures that the L factor is never the bottleneck, leaving only $r_3 = e^{\mathbf{S}(E_{\mathrm{obs}})}$ as the internal constraint. This is the same mechanism by which the observer's ignorance enlarges the effective Hilbert space: the purifying system $\mathcal{H}_L$ is unconstrained, so it cannot restrict the dimension of the observer's boundary.

In all cases, the mixed-state ensemble reproduces the effective dimension of observer rules A1 and B1:

$$d_{\mathrm{ignorant}}^{\mathrm{A3}} = r \times d_R\,, \tag{6.4.24}$$

where $r = \mathrm{rank}(\overline{\rho_\Psi})$ and d_R is the probe image dimension.

Connection to the purification picture. The agreement between A1 and A3 is not accidental. Section 6.3.1 showed that the ignorant Gram matrix is the global Gram matrix of the canonical purification:

$$\hat{G}_{ij} = \langle \sqrt{\overline{\rho_\Psi}} | (O_i^\dagger O_j \otimes \mathbb{1}) | \sqrt{\overline{\rho_\Psi}} \rangle\ . \tag{6.4.25}$$

Observer rule A3 computes the Gram matrix of a two-boundary state in the $\beta_L \to 0$ limit. In the heavy shell mass limit, this two-boundary state is precisely the canonical purification of the thermal state $\rho_{\beta_{\mathrm{eff}}}$, which in turn coincides with $\overline{\rho_\Psi}$. The two computations therefore agree by construction: A3 provides the gravitational realisation of the algebraic identity (6.4.25).

6.4.4 Comparison of observer rules

The results of this section and Sec. 6.3 are summarised in Table 6.2.

The pattern is clear: $d_{\mathrm{ignorant}} = r \times d_R$ whenever the prior is compatible with the observer seeing a genuinely mixed state — whether because the observer adopts

Rule	**Prior**	**GPI prescription**	$d_{ignorant}$
A1	Single mixed state ($\overline{\rho_\Psi}$)	Topology restriction (self-contraction)	$r \times d_R$
B1	Single thermal state ($\rho_{\beta_{\text{eff}}}$)	Modified boundary condition (TFD)	$r \times d_R$
A2	Pure-state ensemble (Gaussian)	Full topology sum	d_R
B2	Thermal pure ensemble	Full sum, $m_i \to \infty$	d_R
A3	Mixed-state ensemble (two-boundary)	Two-boundary, $\beta_L \to 0$	$r \times d_R$

Table 6.2: Observer rules, their epistemic priors, gravitational path integral prescriptions, and resulting effective dimensions. Here $r = \text{rank}(\overline{\rho_\Psi})$ and d_R is the probe image dimension. Rules A1, B1, A3 agree on d_{ignorant}; rules A2 and B2 give the smaller global dimension.

a single density matrix (A1, B1) or because the prior allows for a purifying system the observer cannot access (A3). The dimension collapses to d_R when the prior assumes the universe is a single-boundary pure state (A2, B2), because the state cross-contractions restore the fine-tuned null relations between energy eigenstates.

The agreement between A1 and A3 provides a non-trivial consistency check. A1 was derived in Sec. 6.2.3 from epistemic coarse-graining — restrict the topology sum to self-contracting geometries; A3 was derived here from a maximal-entropy ensemble with the physically appropriate prior. The two constructions are logically independent — the first modifies the sum over topologies, the second averages over a distribution of two-boundary states — yet they agree. The additional topologies included by A3 relative to A1 are cross-contractions within the purifying system $\mathcal{H}_L$, which do not involve the observer's boundary and therefore do not restore the null relations on $\mathcal{H}_R$ that would reduce the dimension. By contrast, the cross-contractions included by A2 are between different insertions of the same single-boundary state, which do restore the null relations and collapse the effective space.

6.5 A non-trivial closed universe emerges through coarse-graining

We now turn to closed universes. The conventional argument that the fine-grained baby-universe Hilbert space is one-dimensional [21, 108] will be shown below to be a direct consequence of the factorisation of the two-boundary AdS Hilbert space

(Chapter 3) together with the trace interpretation of periodic boundary conditions (Chapter 4). At face value this is in tension with the observation that semiclassical physics on a cosmological slice looks conventional: an observer in an AdS big-crunch universe should be able to measure expectation values, assign probabilities to histories, and construct a non-trivial Hilbert space of excitations. The resolution proposed in this section is that the observer's effective description is, once again, an artefact of coarse-graining — specifically, of replacing the fine-grained one-point function of the state-preparing operator by its coarse-grained magnitude.

The logic parallels Secs. 6.2–6.3, with one important conceptual shift. In the black hole case, the coarse-grained description was derived from a microscopic Hilbert space that already existed: the observer was taken to reside on an asymptotic boundary of $\mathcal{H}_{\rm AdS}$. In the closed-universe case, no ambient microscopic Hilbert space is available to the bulk observer: the fine-grained two-boundary construction is rank one, and naively leaves no room for the observer at all. The operator system framework of Sec. 6.3.1 is therefore not merely convenient but essential: the closed-universe Hilbert space must be reconstructed from the data the observer can operationally access, using only the positive sesquilinear structure of the coarse-grained path integral.

6.5.1 The coarse-grained bilinear form

Consider the magnitude squared of the thermal one-point function of the heavy shell operator $\Psi_{\rm obs}$:

$$|\langle\Psi_{\rm obs}\rangle_\beta|^2 = \mathrm{Tr}_{\mathcal{H}_{\rm CFT}}(e^{-\beta H}\,\Psi_{\rm obs})\,\mathrm{Tr}_{\mathcal{H}_{\rm CFT}}(e^{-\beta H}\,\Psi_{\rm obs}^\dagger)\,. \tag{6.5.1}$$

As emphasised in Sec. 6.1, thermal one-point functions of heavy operators in a chaotic CFT depend on erratic fine-grained phase data that a realistic observer cannot resolve. The operationally accessible quantity is therefore the coarse-grained value, computed by the ensemble average via the generalised ETH ansatz (6.2.2) — equivalently, the natural output of the gravitational path integral:

$$\overline{\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta H}\Psi_{\rm obs})\,\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta H}\Psi_{\rm obs}^\dagger)} = [\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta H}\Psi_{\rm obs})\,\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta H}\Psi_{\rm obs}^\dagger)\,]_{\rm GPI}. \tag{6.5.2}$$

The leading saddle to (6.5.2) has cylinder topology, obtained by gluing two opposite sides of a disk along the shell worldvolume using the Israel junction conditions; cutting along its bottleneck and continuing to Lorentzian signature yields an AdS big-crunch cosmology — the closed universe in question.

The bilinear form ζ. The physics of this closed universe can be probed by inserting operators on the $t = 0$ slice of the Lorentzian geometry. Let $\mathcal{K}$ denote the operator system of such bulk insertions. By passing to Euclidean signature and propagating each bulk operator to the asymptotic boundaries of the cylinder, $\mathcal{K}$ is represented by a corresponding boundary operator system $\mathcal{S} \subset \mathcal{B}(\mathcal{H}_{\mathrm{AdS}})$. A simple example is a geodesic matter operator $\phi(\tau, X)$ acting on the Euclidean boundary at time τ and angular position X: the state prepared by cutting open

$$\overline{\mathrm{Tr}_{\mathcal{H}_{\mathrm{AdS}}}(e^{-(\frac{\beta}{2}+\tau)H}\phi(X)e^{-(\frac{\beta}{2}-\tau)H}\Psi_{\mathrm{obs}})\,\mathrm{Tr}_{\mathcal{H}_{\mathrm{AdS}}}(e^{-(\frac{\beta}{2}+\tau)H}\phi^\dagger(X)e^{-(\frac{\beta}{2}-\tau)H}\Psi_{\mathrm{obs}}^\dagger)} \tag{6.5.3}$$

describes a closed universe carrying an additional ϕ-excitation whose location on the $t = 0$ slice is controlled by (τ, X). More generally, low-backreaction boundary insertions prepare non-trivial QFT states of the bulk quantum fields on the $t = 0$ slice, although for generic insertions the resulting excitations may be delocalised across that slice.

Using operators from $\mathcal{S}$ the Euclidean path integral defines a sesquilinear form

$$\zeta(A, B) = \overline{\mathrm{Tr}_{\mathcal{H}_{\mathrm{AdS}}}(e^{-\frac{\beta}{2}H}\,A^\dagger\,e^{-\frac{\beta}{2}H}\,\Psi_{\mathrm{obs}}^\dagger)\,\mathrm{Tr}_{\mathcal{H}_{\mathrm{AdS}}}(e^{-\frac{\beta}{2}H}\,B\,e^{-\frac{\beta}{2}H}\,\Psi_{\mathrm{obs}})}\,, \tag{6.5.4}$$

which is Hermitian and positive semi-definite by construction, and therefore defines a state $\omega(A^\dagger B) \equiv \zeta(A, B)$ on $\mathcal{S}$. Following the operator-system construction of Sec. 6.3.1, this defines an effective Hilbert space $\mathcal{H}_{\mathrm{CU}} = \mathcal{S}/\mathcal{N}_\omega$ via KSGNS, with dimension equal to the rank of the Gram matrix $G_{ij} = \zeta(O_i, O_j)$ for any spanning probe set.[8]

6.5.2 The closed-universe Gram matrix and observer rule A1

Evaluating $\zeta(O_i, O_j)$ via the gravitational path integral and the generalised ETH ansatz (6.2.2) (assuming $O_i, O_j \neq \Psi_{\mathrm{obs}}$):

$$\zeta(O_i, O_j) = \sum_{n,m} O^\dagger_{i,nm} O_{j,mn}\, e^{-f(E_n, E_m)}\,. \tag{6.5.5}$$

Hierarchy of averages. The rank of G is inferred from $\mathrm{Tr}(G^n) = \zeta(O_{i_1}, O_{i_2}) \cdots \zeta(O_{i_n}, O_{i_1})$. The shell average defines ζ (and hence the inner product on $\mathcal{H}_{\mathrm{CU}}$); a probe ensemble E_Ω, if present, acts on *products* of such inner products. The shell average is internal to each ζ factor, while the probe ensemble may correlate different factors. On the

[8] The associated von Neumann algebra depends on the C^*-envelope of $\mathcal{S}$ and not on $(\mathcal{S}, \omega)$ alone, so the present framework determines the effective Hilbert space and its dimension but does not, by itself, fix the algebraic type of the closed-universe algebra.

gravity side, the shell sourced by $\Psi_{\rm obs}$ on one boundary of a given ζ insertion must be absorbed by the $\Psi_{\rm obs}^\dagger$ on the other boundary of the *same* insertion; cross-contractions between $\Psi_{\rm obs}$ shells in different factors of ζ are forbidden. This is the closed-universe analogue of observer rule A1, reproducing the prescription of [127]. Here it is not imposed by hand: it follows from the definition of ζ together with the observer's hierarchy of coarse-grainings.

Heavy shell mass limit. In the limit $m_\Psi \to \infty$ the envelope $f(E_n, E_m) \to I_{\rm shell}(m_\Psi)$ becomes a constant [6, 13], and (6.5.5) simplifies to

$$\zeta(A,B) = e^{-I_{\rm shell}(m_\Psi)}\,\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\beta H}A^\dagger e^{-\beta H}B)\,. \tag{6.5.6}$$

Up to the constant prefactor, this is the Gram matrix of two-boundary states $|\tilde{A}\rangle \equiv |e^{-\beta H/2}Ae^{-\beta H/2}\rangle \in \mathcal{H}_L \otimes \mathcal{H}_R$, so

$$\zeta(A,B) = e^{-I_{\rm shell}(m_\Psi)}\ \langle \tilde{A}|\tilde{B}\rangle_{\mathcal{H}_L\otimes\mathcal{H}_R}\ . \tag{6.5.7}$$

The closed-universe Gram matrix therefore reduces to the two-boundary Gram matrix analysed in Sec. 6.3.5: $\mathrm{rank}(G) = \dim(\mathrm{span}\{|\tilde{O}_i\rangle\}) \in [1, d_{\rm AdS}^2]$. Choosing $\{O_i\}$ to be a set of κ heavy shell operators of masses $m_i \neq m_\Psi$, the analysis of Sec. 6.3.5 gives

$$\mathrm{rank}(G) = \min(\kappa,\, d_{\rm AdS}^2)\,, \tag{6.5.8}$$

exponentially larger than one in the large-κ limit. The closed-universe effective Hilbert space is manifestly non-trivial. Equation (6.5.7) also realises $\mathcal{H}_{\rm CU}$ as (isomorphic to) a subspace of $\mathcal{H}_L \otimes \mathcal{H}_R$.

6.5.3 Triviality of the fine-grained Hilbert space and its enlargement

Fine-grained bilinear. It is illuminating to ask what would have been obtained had we instead used the fine-grained version of (6.5.4). Consider the same bilinear (6.5.4) *without* the shell average:

$$\chi(A,B) = \mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\frac{\beta}{2}H}A^\dagger e^{-\frac{\beta}{2}H}\Psi_{\rm obs}^\dagger)\,\mathrm{Tr}_{\mathcal{H}_{\rm AdS}}(e^{-\frac{\beta}{2}H}Be^{-\frac{\beta}{2}H}\Psi_{\rm obs})\,. \tag{6.5.9}$$

Using the factorisation of the two-boundary AdS Hilbert space established in Chapter 3,

$$\chi(O_i,O_j) = \langle\tilde{O}_i|\Psi_{\rm obs}\rangle\,\langle\Psi_{\rm obs}|\tilde{O}_j\rangle_{\mathcal{H}_L\otimes\mathcal{H}_R} = (VV^\dagger)_{ij}\,, \qquad V_i := \langle\Psi_{\rm obs}|\tilde{O}_i\rangle\ , \tag{6.5.10}$$

so χ is a rank-one Hermitian matrix — the projector onto the one-dimensional subspace spanned by $|\Psi_{\text{obs}}\rangle$, viewed through the lens of the probes. The same conclusion follows from the gravitational replica trick: the boundary conditions for $\mathrm{Tr}(G^n)$ enjoy a permutation symmetry between the n replica insertions, implying $\mathrm{Tr}(G^n) = (\mathrm{Tr}\, G)^n$ and rank one. This is the standard argument for the triviality of the baby-universe Hilbert space [21], here exhibited as a direct consequence of the two structural results of this thesis: periodic boundaries compute traces over $\mathcal{H}_{\text{AdS}}$ (Chapter 4), and the two-boundary AdS Hilbert space factorises (Chapter 3). In the fine-grained calculation, the full sum over topologies is essential: partial sums fail to reproduce the projector $|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|_{\mathcal{H}_L\otimes\mathcal{H}_R}$ that is the source of the rank-one collapse.

Origin of the non-triviality. The coarse-grained bilinear ζ is obtained from χ by the replacement

$$|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|_{\mathcal{H}_L\otimes\mathcal{H}_R} \longrightarrow \overline{|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|}_{\mathcal{H}_L\otimes\mathcal{H}_R}\,, \tag{6.5.11}$$

where the right-hand side is the mixed two-boundary state (6.2.18). Algebraically, the rank-one fine-grained state corresponds to a maximal null ideal $\mathcal{N}_\omega$: there exist coefficients v_i such that $G_{ij} = v_i^* v_j$, and every operator difference $O_i - v_i\mathbb{1}$ is null — all probes act as the same one-dimensional direction up to a scalar. The replacement (6.5.11) shrinks this null ideal,

$$\mathcal{N}_{\text{CU}} \subsetneq \mathcal{N}_\omega\,, \tag{6.5.12}$$

because the exact phase cancellations responsible for the fine-grained null relations are no longer accessible. Operator directions that were null become genuinely distinguishable, and the effective Hilbert space is enlarged. Importantly, $\mathcal{H}_{\text{CU}}$ is not a subspace of any fine-grained Hilbert space: it is built from a different positive form, and inherits a different geometry. The mechanism is the same as in Sec. 6.3 — coarse-graining shrinks the null ideal — but here it produces a non-trivial Hilbert space where the fine-grained one had no content at all, using only the coarse-grained path integral and its positivity, with no input from any ambient microscopic Hilbert space.

6.5.4 Predictability in the closed universe

A natural objection to this picture is that, unlike the black hole case, the shell responsible for the closed-universe state is not hidden behind a horizon: in Lorentzian signature it sits at finite distance from a bulk observer on the $t = 0$ slice, who could

in principle approach it and perform refined measurements to identify the underlying microstate. One may then wonder whether the effective Hilbert space collapses back to rank one. We argue here that it does not.

The key point is that the notion of a bulk observer following a semiclassical trajectory already presupposes a coarse-grained description of the state. The fine-grained description, given by the rank-one projector (6.5.10), has no a priori smooth-spacetime interpretation with well-defined locality, causal structure, or worldlines. These features arise only after restricting attention to an appropriate class of observables and averaging over fine-grained data. The semiclassical geometry of the closed universe is therefore an effective description associated with the coarse-grained bilinear ζ, rather than a property of the underlying rank-one state. This is analogous to, but more extreme than, the standard emergence of the black hole interior in AdS/CFT, where smooth spacetime descriptions rely on suppressing fine-grained phase information at the level of $e^{-\mathbf{S}}$ corrections [82].

This distinction can be sharpened by introducing an auxiliary "super-observer" with access to the full two-boundary description. Such an observer would describe the state as a pure vector $|\Psi_{\text{obs}}\rangle \in \mathcal{H}_L \otimes \mathcal{H}_R$ and assign the rank-one bilinear χ, but they are not localised within the closed-universe geometry: they have access to both asymptotic boundaries simultaneously. By contrast, an observer confined to the closed-universe region is naturally described in terms of bulk operators supported on the $t = 0$ slice, which map to the boundary operator system $\mathcal{S}$ and whose overlaps are governed by ζ. Within this restricted class of probes, attempts to extract detailed microstate information from the shell correspond to inserting increasingly complex operators of $\mathcal{O}(G_N^{-1})$ backreaction; by the analysis of Sec. 6.2.3 these at most provide statistical information about the equivalence class of compatible states, not resolution of a specific microstate. The non-trivial closed-universe Hilbert space obtained from ζ is therefore stable within the regime of semiclassical observables, even in the absence of a horizon: the limitation is not strictly geometric inaccessibility but the restricted sensitivity of the observer's operator system to fine-grained phase information.

A related concern is whether replacing the rank-one projector by the coarse-grained bilinear ζ might introduce sample dependence, with different microstates leading to inequivalent effective Hilbert spaces. As in Sec. 6.2.1, the variance of $\zeta(O^\dagger O)$ across the equivalence class is suppressed by $\mathcal{O}(e^{-\mathbf{S}})$ relative to its mean. Within the precision relevant for semiclassical observables, the effective Hilbert space can therefore be regarded as a property of the equivalence class rather than of individual microstates.

Summary. The picture that emerges is the following. The fine-grained two-boundary inner product cleanly encodes the rank-one projector $|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|$, and its action on any operator system is correspondingly one-dimensional. Coarse-graining replaces this pure projector with the mixed state $\overline{|\Psi_{\text{obs}}\rangle\langle\Psi_{\text{obs}}|}$ of rank $\sim e^{\mathbf{S}}$, and the induced sesquilinear form on the operator system has an exponentially enlarged image. An operationally limited CFT experimenter cannot detect the underlying rank-one structure, since doing so would require resolving the erratic $e^{-\mathbf{S}/2}$ phase data in the thermal one-point function of Ψ_{obs}. Their effective description instead matches that of a quantum field theory on a fixed AdS big-crunch background, and this emergent closed universe carries a non-trivial operational Hilbert space. Unlike the constructions of Secs. 6.3–6.4, this construction assumes no microscopic ambient Hilbert space for the closed universe: only the coarse-grained path integral and its positivity are used, and the effective Hilbert space is reconstructed from operational data alone. The closed universe is in this sense the most conservative realisation of the observer framework.

"... the moment there was something, there being nothing else, that something was the universe."

— *Italo Calvino,* Cosmicomics

6.6 Discussion

We have formulated a principle of maximal ignorance for gravitational observers restricted to simple measurements in chaotic holographic systems and traced its consequences through the gravitational path integral. The central conceptual claim is that the gravitational ensemble computed by the Euclidean path integral is naturally associated to the equivalence class of microstates compatible with the simple-operator data of an asymptotic observer; the observer rules [127] that govern its sum over topologies are unavoidable consequences of this identification rather than phenomenological inputs from a postulated bulk worldline.

Concretely, in Sec. 6.1 we used ETH to show that any pure state in a chaotic theory is operationally indistinguishable, under sub-exponential products of simple operators, from a smoothed diagonal density matrix ρ_{smooth}. Operators that violate ETH — complex operators — exhibit $\mathcal{O}(1)$ fluctuations across compatible microstates and require an ensemble description. In Sec. 6.2 we identified the gravitational ensemble defined by the generalised ETH ansatz on heavy shell operators with the simple-data ensemble of Sec. 6.1, and showed that the saddle-shift induced by a probe controls a

clean three-fold correspondence: zero backreaction (causal wedge, observer rule B1), $\mathcal{O}(G_N)$ backreaction of multiple independent simple probes (simple wedge, observer rule A1, reproducing the Chandra–Hartman result $S(\overline{\rho_\Psi}) = A_{\text{apparent}}/4G_N + S^{\text{outer}}_{\text{bulk}}$), and $\mathcal{O}(1)$ backreaction stabilising new wormhole saddles that encode coarse features of the Python's lunch.

In Sec. 6.3 the pair $(\mathcal{S}, \overline{\rho_\Psi})$ of an operator system and a coarse-grained state was shown to define an effective Hilbert space $\mathcal{H}_\rho$ via the sesquilinear form $\langle O_i | O_j \rangle_\rho = \mathrm{Tr}(\overline{\rho_\Psi}\, O_i^\dagger O_j)$, with dimension $d_{\text{ignorant}} = r \times d_R$ exceeding the global value. The increase has a precise origin: coarse-graining removes the fine-tuned cancellations between diagonal and off-diagonal Gram matrix contributions that would otherwise enforce null relations and collapse the rank to d_{global}. We computed d_{ignorant} via the gravitational path integral using both planar resolvent methods and finite-κ Wishart rank theorems, the latter resolving analytic-continuation ambiguities that arise in the replica trick. Section 6.4 showed that for nonlinear observables the answer depends on a prior on why the effective state is mixed: a single-state prior reproduces A1/B1 with $d_{\text{ignorant}} = r \times d_R$; a pure-state ensemble (assuming a globally pure single-boundary universe) collapses the dimension back to d_{global}; a mixed-state ensemble (allowing for a purifying system) reproduces $d_{\text{ignorant}} = r \times d_R$. The agreement between the single-state prescription and the mixed-state ensemble is a non-trivial consistency check.

In Sec. 6.5 the same logic applied to AdS closed universes produced a non-trivial effective Hilbert space from coarse-grained boundary data alone, with no input from any ambient microscopic Hilbert space. The fine-grained bilinear $\chi(A, B)$ is rank one — a direct consequence of factorisation together with the trace interpretation of periodic boundaries — but the coarse-grained bilinear $\zeta(A, B)$, obtained by replacing $|\Psi_{\text{obs}}\rangle \langle\Psi_{\text{obs}}|$ by its ensemble average, has rank as large as d^2_{AdS}. The closed universe, its Hilbert space, and the observer within it are all simultaneously emergent features of the coarse-grained description.

6.6.1 Future directions

The framework developed here suggests several directions for further investigation, of which we mention a few.

Algebraic type and crossed products. The KSGNS construction used in Sec. 6.5 delivers a Hilbert space without fixing the algebraic type (I, II, or III) of any associated von Neumann algebra; determining the type requires specifying the C^*-envelope,

which depends on the algebraic closure of the bulk operator system. A connection to the crossed-product constructions of [35, 60–64, 112, 113, 140–153] seems natural and would relate the present framework to the type-II description of de Sitter space and the black hole exterior.

Dynamical spacetimes. Our derivation of observer rules A1/B1 relied on the stationarity of the shell-state exterior. Extending the framework to non-stationary settings — gravitational collapse or evaporation — would test whether epistemic coarse-graining picks out the causal wedge, the simple wedge, or something in between. A natural Lorentzian counterpart is the colliding-particle states of [154], which prepare exact CFT states dual to the formation and (sub-threshold) evaporation of small AdS black holes; the survey of coarse-graining prescriptions there overlaps with two of those used here max-entropy density matrix, energy-basis dephasing), and a unified account of the remaining ones — Hamiltonian averaging, OPE-coefficient averaging, eigenvector/eigenvalue averaging and time-window averaging — within the operator-system framework would be a natural extension.

dS observers. The closed-universe construction relied on AdS asymptotics, but the logic of maximal ignorance applies equally to de Sitter or inflating backgrounds, where observers cannot resolve fine-grained structure beyond their cosmological horizon. The shell-state construction was extended to the static patch of de Sitter space in [155]: the full sum over topologies again gives a trivial Hilbert space, while imposing observer rule A1 on a sufficiently large collection of shell states reproduces the de Sitter entropy. Identifying a generalised-ETH ansatz for operators in the putative dS dual theory on $\mathcal{I}^{\pm}$ that reproduces the wormhole moments computed there would establish a precise parallel between the de Sitter static patch and the AdS closed universe case discussed here.

Emergence of the closed universe. The argument that the closed universe is inherently emergent has been established here only for the shell-state construction of Sec. 6.5, where both the rank-one character of $\chi(A, B)$ and the rank $\sim e^{\mathbf{S}}$ of its coarse-grained replacement $\zeta(A, B)$ rely on properties of heavy shells in asymptotically AdS gravity. Whether this simultaneous emergence of geometry, Hilbert space, and observer is a generic feature of closed-universe physics, or special to the construction we analyse, is a question we leave for future work. Suggestive evidence for the former can be drawn from [125], which reaches a closely related conclusion in a

setting with no shells, no asymptotic boundary, and no ETH-driven dephasing. In a one-dimensional topological model of de Sitter space, the gravitational path integral defines a commutative baby-universe Hilbert space $\mathcal{H}_{\rm BU} = L^2(\Omega,\, p_\alpha\, d\alpha)$ packaging the classical statistics of α-sectors, while bulk quantum mechanics is realised on a separate Hilbert space $\mathcal{H}_U \cong \mathbb{C}^{Z_L}$ of dimension equal to the sphere partition function. The two are connected by "patch operators", which in each α-sector reduce to c-numbers specifying an unnormalised Haar-random pure state $|\psi_{\rm HH}\rangle_\alpha \in \mathcal{H}_U$. Within each fixed α-sector the Hartle–Hawking wave function on the two spatial points is merely a product state, and the maximally entangled structure responsible for the semiclassical geometric interpretation of the closed universe arises strictly from averaging over α: no single fixed α supports a bulk closed-universe description at all. This is precisely the qualitative claim we make for the shell-state construction, reached by mechanistically distinct means — a topological sum over patch insertions in one dimension versus ETH-driven dephasing of heavy shells in asymptotically AdS. The two frameworks differ on a finer point: [125] takes $\mathcal{H}_U$ to be present already per-α, with only the geometric and dynamical content emergent from averaging, whereas in our construction the Hilbert space itself is built bottom-up via the KSGNS quotient of ζ and has no per-microstate counterpart. The point of convergence, however, is the physically substantive one: in both settings a non-trivial closed-universe description cannot be recovered without ensemble averaging. That two independent routes converge on this conclusion suggests that the emergence established here for shell states is one instance of a more general principle, whose precise formulation we leave for future investigation.

Causally restricted observers. A complementary perspective has been developed by Nomura and Ugajin [115, 116] in the canonical Lorentzian framework. Starting from the same rank-one α-sector state as our χ, they factor the kinematical Hilbert space as $\mathcal{H}_0 = \mathcal{H}_{\rm acc} \otimes \mathcal{H}_{\rm env}$ and trace out $\mathcal{H}_{\rm env}$, obtaining a reduced density matrix of rank much greater than one with variance suppressed by $e^{-S_{\rm env}}$. The mechanism is the one that turns χ into ζ in Sec. 6.5: tracing in the kinematical Hilbert space severs the replica-wormhole connectivity that would otherwise enforce rank one. The distinction lies in the operational limitation. Theirs is causal — no access outside the observer's causal past. Ours is complexity-theoretic — nominal access to the whole universe, but only through the sub-exponential operator system $\mathcal{S}$, which cannot resolve the erratic $e^{-S/2}$ phase information that would expose the fine-grained rank-one structure. We expect other operational limitations — thermodynamic, computational, or otherwise

— to give similar results, with the rank, suppression scale, and algebraic content of the effective Hilbert space tracking the class of measurements the observer can perform.

Decoherent observers and baby-universe field theory. A conceptually distinct observer rule was proposed in [109] in which the degrees of freedom defining a gravitating observer are cloned out of the system and act as a non-gravitational reference, suppressing certain diagrams in the topology sum by factors scaling with the observer's entropy S_{Ob}. This proposal has recently been extended to the third-quantised setting of baby-universe field theory (BUFT) in [156], where the same decohering channel is reinterpreted as an average over CFT boundary sources J and shown to recover BUFT from a holographic theory with one-dimensional fine-grained Hilbert space, up to errors of order $e^{-S_{Ob}/2}$. The mechanism shares the spirit of Sec. 6.5 — non-trivial closed-universe physics emerging from coarse-grained boundary data. Determining whether the decoherence channel resulting in J-averaging identified in [156] is morally equivalent to the operator-system coarse-graining used here, or a logically independent prescription with a coincident effective description, and computing the measured dimension $d_{\text{measured}} \leq e^{S_{Ob}}$ accessible at realistic κ and finite observer entropy, would clarify the relationship between the two frameworks.

Appendix A

Gram matrix of overlaps

In this appendix we detail the Gram matrix formalism for obtaining an orthonormal basis from an overcomplete set of states. We consider a d-dimensional Hilbert space $\mathcal{H}$ spanned by a set of $\kappa \geq d$, generally non-orthogonal, vectors $\mathcal{U} = \{|u_i\rangle\}$. The Gram matrix of overlaps

$$G_{ij} = \langle u_i | u_j \rangle \tag{A.0.1}$$

is Hermitian and can therefore be diagonalized by a unitary matrix U,

$$G = UDU^\dagger, \tag{A.0.2}$$

where D is a diagonal matrix with real, non-negative eigenvalues λ_i. The rank of G is given by the number of nonzero eigenvalues of D, which must equal the dimension d of $\mathcal{H}$, and the $(\kappa - d)$-dimensional kernel of G is the space of null states. Using the decomposition (A.0.2), powers of the Gram matrix can be defined for any real n as

$$G^n \equiv UD^nU^\dagger, \tag{A.0.3}$$

where

$$D^n \equiv \mathrm{diag}(\lambda_1^n, \lambda_2^n, \ldots, \lambda_d^n, 0, 0, \ldots, 0).$$

For non-integer powers k, this definition is understood as an analytic continuation of the positive integer n result:

$$G^k = \lim_{n \to k} G^n.$$

For example, the generalized inverse

$$G^{-1} = UD^{-1}U^\dagger$$

is defined by inverting all non-zero eigenvalues and leaving the rest untouched, satisfying

$$GG^{-1}G \equiv \lim_{n\to -1} G\,G^n\,G = \lim_{n\to -1} UD^{n+2}U^\dagger = G. \tag{A.0.4}$$

As it is intractable to explicitly diagonalize the Gram matrix in the cases of interest in this thesis, we write expressions like G^{-1} directly by their analytic continuation

$$G^{-1} \equiv \lim_{n\to -1} G^n$$

from G^n at positive integer n, without reference to D.

Orthonormal basis. This formalism can also be used to obtain an orthonormal basis for $\mathcal{H}$ in terms of the overcomplete set $\mathcal{U}$. Consider the states $\{|v_a\rangle\}$ defined as

$$|v_a\rangle = \sum_{ij} G^{-1/2}_{ji} U_{ia} |u_j\rangle = \sum_{jb} U_{jb} D^{-1/2}_{ba} |u_j\rangle. \tag{A.0.5}$$

The first d states in this set are orthonormal and span $\mathcal{H}$, while the remaining $\kappa - d$ states are null:

$$\begin{aligned}\langle v_a|v_b\rangle &= \sum_{jkcd} D^{-1/2}_{ac} U^\dagger_{cj} \langle u_j|u_k\rangle U_{kd} D^{-1/2}_{db} \\ &= \sum_{cd} D^{-1/2}_{ac} D_{cd} D^{-1/2}_{db} = (D^0)_{ab},\end{aligned} \tag{A.0.6}$$

where by the above definition

$$D^0 = \mathrm{diag}(1, 1, \ldots, 1, 0, 0, \ldots, 0),$$

with d entries equal to 1.

As $\mathcal{H}$ is a Hilbert space, the null states must all be the zero vector, hence the states $\{|v_a\rangle\}$ with $a = 1, \ldots, d$ provide an orthonormal basis for $\mathcal{H}$.

Identity, inner products and traces. The resolution of the identity in the orthonormal basis is given by

$$\sum_{a=1}^{d} |v_a\rangle\langle v_a|$$

and can be expressed in terms of the overcomplete basis using (A.0.5):

$$\begin{aligned}\mathbb{1}_{\mathcal{H}} &= \sum_a |v_a\rangle\langle v_a| \\ &= \sum_{abcjk} |u_j\rangle U_{jb} D^{-1/2}_{ba} D^{-1/2}_{ac} U^\dagger_{ck} \langle u_k| \\ &= \sum_{ij} G^{-1}_{ij} |u_i\rangle\langle u_j| \\ &\equiv \lim_{n\to -1} \sum_{ij} G^n_{ij} |u_i\rangle\langle u_j|. \end{aligned} \tag{A.0.7}$$

Here the sum over a effectively runs over $1, \ldots, d$ since contributions from $a > d$ vanish. Inserting this identity into the inner product, we obtain

$$\langle A|B\rangle = \langle A|\mathbb{1}_{\mathcal{H}}|B\rangle = \sum_{jk} \langle A|u_j\rangle G^{-1}_{jk} \langle u_k|B\rangle. \tag{A.0.8}$$

We can similarly derive a formula for the trace of an operator $\mathcal{O}$:

$$\begin{aligned}\mathrm{Tr}_{\mathcal{H}}(\mathcal{O}) &= \sum_a \langle v_a|\mathcal{O}|v_a\rangle \\ &= \sum_{abcjk} D^{-1/2}_{ab} U^\dagger_{bj} \langle u_j|\mathcal{O}|u_k\rangle U_{kc} D^{-1/2}_{ca} \\ &= \sum_{jk} G^{-1}_{kj} \langle u_j|\mathcal{O}|u_k\rangle. \end{aligned} \tag{A.0.9}$$

Hence this trace is formally defined by the analytic continuation

$$\mathrm{Tr}_{\mathcal{H}}(\mathcal{O}) = \lim_{n\to -1} \sum_{jk} G^n_{kj} \langle u_j|\mathcal{O}|u_k\rangle. \tag{A.0.10}$$

The trace can be used to extract the dimension of $\mathcal{H}$:

$$\begin{aligned} d &= \mathrm{Tr}_{\mathcal{H}}(\mathbb{1}_{\mathcal{H}}) \\ &= \sum_{jk} G^{-1}_{kj} \langle u_j|\mathbb{1}_{\mathcal{H}}|u_k\rangle \\ &= \sum_{jk,lm} G^{-1}_{kj} G^{-1}_{lm} \langle u_j|u_l\rangle\langle u_m|u_k\rangle \\ &= \sum_{jk,lm} G^{-1}_{kj} G^{-1}_{lm} G_{jl} G_{mk} \\ &= \mathrm{Tr}(G^{-1}GG^{-1}G) = \mathrm{Tr}(G^{-1}G), \end{aligned} \tag{A.0.11}$$

where the final equality follows from the definition of the generalized inverse,

$$GG^{-1}G = G.$$

Hence the dimension can be extracted directly from the overcomplete basis by the analytic continuation

$$d = \lim_{n \to -1} \sum_{i,j} G_{ij}^n G_{ji}. \tag{A.0.12}$$

The fact that the identity, trace and inner product (although not individual matrix elements) can be expressed in terms of the overcomplete set $\mathcal{U}$ while being independent of U will be very useful in the gravity constructions considered in this thesis, as finding U explicitly is generally intractable in this setting.

Appendix B

Ensemble statistics and random matrix calculations

B.1 Universality of ensemble statistics

In Sec. 6.1.2 we claimed that the mean and variance of $X = \langle\psi|A|\psi\rangle$ agree across the random-phase, Haar, Ginibre, and maximum-entropy ensembles at leading order in large $d = |\mathrm{supp}(\overline{\rho_\Psi})| \sim e^S$. In this appendix we verify this claim by explicit computation in each ensemble.

Throughout, we write $|\psi\rangle = \sum_n \alpha_n |n\rangle$ where $\{|n\rangle\}$ are energy eigenstates, and denote $p_n = \mathbb{E}[|\alpha_n|^2]$. The target expressions are

$$\mathbb{E}[X] = \sum_n p_n \, A_{nn} \,, \qquad \mathrm{Var}(X) = \sum_{m \neq n} p_m \, p_n \, |A_{mn}|^2 \,. \tag{B.1.1}$$

B.1.1 Random-phase ensemble

Set $\alpha_n = \sqrt{p_n}\, e^{i\theta_n}$ with the θ_n independently and uniformly distributed on $[0, 2\pi)$. Then

$$X = \sum_{m,n} \sqrt{p_m p_n}\, A_{mn}\, e^{i(\theta_n - \theta_m)} \,. \tag{B.1.2}$$

The second- and fourth-order phase averages are

$$\mathbb{E}[e^{i(\theta_n - \theta_m)}] = \delta_{nm} \,, \qquad \mathbb{E}[e^{i(\theta_n - \theta_m + \theta_p - \theta_q)}] = \delta_{nm}\delta_{pq} + \delta_{nq}\delta_{mp} - \delta_{nm}\delta_{pq}\delta_{nq}\delta_{mp} \,. \tag{B.1.3}$$

The last (subtraction) term contributes only when $m = n = p = q$ and removes the double-counting.

Mean.

$$\mathbb{E}[X] = \sum_{m,n} \sqrt{p_m p_n}\, A_{mn}\, \delta_{mn} = \sum_n p_n\, A_{nn}\,. \tag{B.1.4}$$

Variance. We compute $\mathbb{E}[|X|^2] = \sum_{m,n,p,q} \sqrt{p_m p_n p_p p_q}\, A_{mn}\, \overline{A}_{pq}\, \mathbb{E}[e^{i(\theta_n - \theta_m + \theta_p - \theta_q)}]$. Using (B.1.3), the two non-trivial pairings give

$$\mathbb{E}[|X|^2] = \Big(\sum_n p_n\, A_{nn}\Big)\Big(\sum_p p_p\, \overline{A}_{pp}\Big) + \sum_{m,n} p_m\, p_n\, |A_{mn}|^2 - \sum_n p_n^2\, |A_{nn}|^2\,, \tag{B.1.5}$$

where the first term is $|\mathbb{E}[X]|^2$ and the subtraction removes the diagonal overlap. Therefore

$$\mathrm{Var}_{\mathrm{RP}}(X) = \sum_{m \neq n} p_m\, p_n\, |A_{mn}|^2\,. \tag{B.1.6}$$

This is *exact*—no subleading corrections arise, because the amplitudes $\sqrt{p_n}$ are held fixed and only the phases are averaged.

B.1.2 Haar ensemble on the microcanonical shell

Let $|\psi\rangle$ be drawn Haar-randomly from $\mathbb{C}^d$, so that $p_n = 1/d$ for all n in the microcanonical window. The moments of the coefficients follow from the Weingarten calculus for $U(d)$:

$$\mathbb{E}[\alpha_m^*\, \alpha_n] = \frac{\delta_{mn}}{d}\,, \qquad \mathbb{E}[\alpha_m^*\, \alpha_n\, \alpha_p^*\, \alpha_q] = \frac{\delta_{mn}\, \delta_{pq} + \delta_{mq}\, \delta_{np}}{d(d+1)}\,. \tag{B.1.7}$$

Mean.

$$\mathbb{E}[X] = \sum_{m,n} A_{mn}\, \frac{\delta_{mn}}{d} = \frac{\mathrm{Tr}(A)}{d} = \sum_n p_n\, A_{nn}\,, \tag{B.1.8}$$

which agrees with (B.1.1).

Variance. Using the fourth moment in (B.1.7),

$$\mathbb{E}[|X|^2] = \sum_{m,n,p,q} A_{mn}\, \overline{A}_{pq}\, \frac{\delta_{mn}\, \delta_{pq} + \delta_{mq}\, \delta_{np}}{d(d+1)} = \frac{|\mathrm{Tr}(A)|^2 + \|A\|_F^2}{d(d+1)}\,, \tag{B.1.9}$$

where $\|A\|_F^2 = \mathrm{Tr}(A^\dagger A) = \sum_{m,n} |A_{mn}|^2$ is the Frobenius norm. Therefore

$$\mathrm{Var}_{\mathrm{Haar}}(X) = \frac{|\mathrm{Tr}(A)|^2 + \|A\|_F^2}{d(d+1)} - \frac{|\mathrm{Tr}(A)|^2}{d^2}\,. \tag{B.1.10}$$

Combining fractions,

$$\mathrm{Var}_{\mathrm{Haar}}(X) = \frac{d\,\|A\|_F^2 - |\mathrm{Tr}(A)|^2}{d^2(d+1)}\,. \tag{B.1.11}$$

Writing $\|A\|_F^2 = \sum_{m\neq n}|A_{mn}|^2 + \sum_n |A_{nn}|^2$ and $|\mathrm{Tr}(A)|^2 \leq d\sum_n |A_{nn}|^2$ (by Cauchy–Schwarz), the leading term at large d is

$$\mathrm{Var}_{\mathrm{Haar}}(X) = \frac{1}{d^2}\sum_{m\neq n}|A_{mn}|^2 + O\!\left(\frac{\|A\|_F^2}{d^3}\right). \tag{B.1.12}$$

Since $p_n = 1/d$, the leading term is precisely $\sum_{m\neq n} p_m\, p_n\, |A_{mn}|^2$, in agreement with the random-phase result (B.1.6). The relative correction is $O(1/d)$.

B.1.3 Ginibre ensemble

Draw $z_n \sim \mathcal{CN}(0, p_n)$ independently (i.e. $\mathbb{E}[z_m^* z_n] = p_n\,\delta_{mn}$) and set $\alpha_n = z_n/\mathcal{N}$ with $\mathcal{N} = \|z\| = (\sum_k |z_k|^2)^{1/2}$. We first compute the unnormalised moments, then account for the normalisation.

Unnormalised moments. The complex Gaussian Wick theorem (Isserlis' theorem) gives the fourth moment

$$\mathbb{E}[z_m^*\, z_n\, z_p^*\, z_q] = \mathbb{E}[z_m^* z_n]\,\mathbb{E}[z_p^* z_q] + \mathbb{E}[z_m^* z_q]\,\mathbb{E}[z_p^* z_n] = p_n\, p_q\, \delta_{mn}\,\delta_{pq} + p_q\, p_n\, \delta_{mq}\,\delta_{pn}\,. \tag{B.1.13}$$

Defining $\widetilde{X} = \sum_{m,n} z_m^*\, A_{mn}\, z_n$, the unnormalised mean and variance are

$$\mathbb{E}[\widetilde{X}] = \sum_n p_n\, A_{nn}\,, \qquad \widetilde{\mathrm{Var}}(\widetilde{X}) = \sum_{m,n} p_m\, p_n\, |A_{mn}|^2\,, \tag{B.1.14}$$

where the variance includes the $m = n$ diagonal. Subtracting,

$$\widetilde{\mathrm{Var}}(\widetilde{X}) = \sum_{m\neq n} p_m\, p_n\, |A_{mn}|^2 + \sum_n p_n^2\, |A_{nn}|^2\,. \tag{B.1.15}$$

Effect of normalisation. The actual random variable is $X = \widetilde{X}/\mathcal{N}^2$. The norm-squared $\mathcal{N}^2 = \sum_k |z_k|^2$ is a sum of d independent terms with $\mathbb{E}[\mathcal{N}^2] = \sum_k p_k = 1$ (assuming the populations are normalised) and

$$\mathrm{Var}(\mathcal{N}^2) = \sum_k p_k^2\,. \tag{B.1.16}$$

For microcanonical populations $p_k = 1/d$, this gives $\mathrm{Var}(\mathcal{N}^2) = 1/d$. More generally, by the Cauchy–Schwarz inequality $\sum_k p_k^2 \geq 1/d$, so $\mathcal{N}^2$ concentrates around 1 with fluctuations of order $d^{-1/2}$ or smaller.

The normalisation correction to the variance is

$$\mathrm{Var}(X) = \mathrm{Var}\left(\frac{\widetilde{X}}{\mathcal{N}^2}\right) = \frac{\widetilde{\mathrm{Var}}(\widetilde{X})}{(\mathbb{E}[\mathcal{N}^2])^2} + O\left(\frac{\mathrm{Var}(\mathcal{N}^2)}{(\mathbb{E}[\mathcal{N}^2])^4}\right) = \widetilde{\mathrm{Var}}(\widetilde{X}) + O\left(\sum_k p_k^2\right), \tag{B.1.17}$$

where the cross-terms between $\widetilde{X}$ and $\mathcal{N}^2$ contribute at the same subleading order. The diagonal contribution $\sum_n p_n^2 |A_{nn}|^2$ in (B.1.15) is itself of order $\sum_n p_n^2 \leq \max_n p_n \leq 1/d$ (for microcanonical populations with $\|A\|_\infty$ bounded), so

$$\mathrm{Var}_{\mathrm{Gin}}(X) = \sum_{m \neq n} p_m\, p_n\, |A_{mn}|^2 + O\left(\frac{1}{d}\right), \tag{B.1.18}$$

in agreement with (B.1.6).

B.1.4 Maximum-entropy ensemble

The maximum-entropy distribution over pure states subject to the constraint $\mathbb{E}[|\psi\rangle\langle\psi|] = \overline{\rho_\Psi}$ is the measure of maximal entropy (in the sense of Jaynes) on the space of rank-one projectors compatible with the given first moment.

For the microcanonical case $p_n = 1/d$, the constraint is $U(d)$-invariant, and the unique maximum-entropy measure is the Haar measure. The result then coincides with Sec. B.1.2.

For general (non-uniform) populations $\{p_n\}$, the maximum-entropy ensemble can be constructed as follows. Let $|e\rangle$ be a fixed reference state and U a Haar-random unitary. Define

$$|\psi\rangle = \frac{\overline{\rho_\Psi}^{1/2}\, U\, |e\rangle}{\|\overline{\rho_\Psi}^{1/2}\, U\, |e\rangle\|}. \tag{B.1.19}$$

In the energy eigenbasis, $(\overline{\rho_\Psi}^{1/2})_{mn} = \sqrt{p_n}\, \delta_{mn}$, so the unnormalised vector has components $\tilde{\alpha}_n = \sqrt{p_n}\, U_{n1}$. Since U_{n1} for Haar-distributed U are (up to normalisation) i.i.d. complex Gaussians, the $\tilde{\alpha}_n$ are independent $\mathcal{CN}(0, p_n/d)$ variables before normalisation.[1] This is precisely the Ginibre ensemble of Sec. B.1.3 (with p_n rescaled so that $\sum p_n = 1$), and the same calculation applies. In particular,

$$\mathrm{Var}_{\mathrm{ME}}(X) = \sum_{m \neq n} p_m\, p_n\, |A_{mn}|^2 + O\left(\frac{1}{d}\right). \tag{B.1.20}$$

[1] More precisely, the first column of a Haar unitary has the distribution of a normalised isotropic Gaussian vector in $\mathbb{C}^d$, so $U_{n1} \sim \mathcal{CN}(0, 1/d)$ marginally with correlations of order $1/d^2$. These correlations contribute only at $O(1/d^2)$ to the moments we need.

B.1.5 Summary

Collecting results, the four ensembles give

Ensemble	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	
Random-phase	$\sum_n p_n A_{nn}$	$\sum_{m\neq n} p_m p_n \lvert A_{mn}\rvert^2$	(exact)
Haar	$\sum_n p_n A_{nn}$	$\sum_{m\neq n} p_m p_n \lvert A_{mn}\rvert^2 + O(1/d)$	
Ginibre	$\sum_n p_n A_{nn}$	$\sum_{m\neq n} p_m p_n \lvert A_{mn}\rvert^2 + O(1/d)$	
Max-entropy	$\sum_n p_n A_{nn}$	$\sum_{m\neq n} p_m p_n \lvert A_{mn}\rvert^2 + O(1/d)$	

(B.1.21)

All four ensembles agree on the mean (exactly) and the variance (up to $O(1/d) \sim O(e^{-S})$ corrections). The random-phase ensemble is distinguished by giving the exact answer at all orders, since the amplitudes are not randomised; the other three ensembles introduce $O(1/d)$ corrections through the Weingarten denominators (Haar) or normalisation fluctuations (Ginibre, max-entropy). For the quantities appearing in the main text—in particular the self-averaging estimate $\mathrm{Var}(X) \sim e^{-S}$ for simple operators and $\mathrm{Var}(X) \sim O(1)$ for complex operators—these corrections are negligible, and the choice of ensemble is immaterial.

B.2 Generalised Wishart ensembles

In this appendix we collect the random matrix results needed for the Gram matrix computations of Secs. 6.3 and 6.4. All Gram matrices encountered in the main text are built from Gaussian random matrices with structured deterministic insertions, and fall into three classes depending on how the observer state and operator ensembles enter: via a vector contraction (rank-one case), a trace (matrix-valued case), or a pair of independent operator ensembles (bi-ensemble case). We treat each in turn. Throughout, matrices are $d\times d$ unless stated otherwise, and indices $i, j = 1, \ldots, \kappa$ label κ independent draws O_i of complex Gaussian matrices with i.i.d. standard entries.

B.2.1 Rank-one construction: scale-mixture Wishart

Consider the Gram matrix

$$G_{ij} = \langle 0|\, A\, D_1^{1/2}\, O_i\, D_2\, O_j^\dagger\, D_1^{1/2}\, A^\dagger\, |0\rangle \ , \tag{B.2.1}$$

where $|0\rangle$ is a fixed reference vector, A is complex Gaussian, and $D_1, D_2 \succeq 0$ are positive semi-definite diagonal matrices encoding the spectral weights of the deterministic insertions. This structure arises in the global case (Sec. 6.3.3), where state enters the Gram matrix through a single ket, so that all randomness associated with the state is absorbed into a single vector. Defining

$$w := D_1^{1/2} A^\dagger |0\rangle \,, \tag{B.2.2}$$

the Gram matrix factorises as

$$G_{ij} = w^\dagger O_i D_2 O_j^\dagger w = x_i^\dagger x_j \,, \qquad x_i := D_2^{1/2} O_i^\dagger w \,, \tag{B.2.3}$$

so that G is the Gram matrix of the vectors $\{x_i\}_{i=1}^{\kappa}$. Conditioned on w, the x_i are i.i.d. complex Gaussian with covariance $\mathbb{E}[x_i x_i^\dagger \mid w] = |w|^2 D_2$, and so $G = X^\dagger X$ with $X = |w| D_2^{1/2} Z$ and Z having i.i.d. standard complex Gaussian entries. Conditioned on w, this is a generalised Wishart matrix with population covariance $|w|^2 D_2$. Marginalising over $|w|^2$ gives a scale-mixture (compound) Wishart: the overall scale of the covariance fluctuates, but its shape is fixed by D_2.

Moments.
Since $|w|^2$ is a scalar independent of the O_i ensemble, the moments factorise into a state contribution and an operator contribution:

$$\mathbb{E}[\mathrm{Tr}(G^m)] = \mathbb{E}[|w|^{2m}] \cdot \mathbb{E}[\mathrm{Tr}(W^m)] \,, \qquad W = Z^\dagger D_2 Z \,. \tag{B.2.4}$$

The state factor evaluates to a sum over S_m weighted by traces of powers of D_1:

$$\mathbb{E}[|w|^{2m}] = |o|^{2m} \sum_{\pi \in S_m} \prod_{c \in \mathrm{Cycles}(\pi)} \mathrm{Tr}(D_1^{|c|}) \,, \tag{B.2.5}$$

with $|o|^2 = |\langle 0|0\rangle|^2$ and each cycle of length $|c|$ contributing $\mathrm{Tr}(D_1^{|c|})$ from the Wick contractions of A. The operator factor is the standard Wishart moment:

$$\mathbb{E}[\mathrm{Tr}(W^m)] = \sum_{\sigma \in S_m} \kappa^{\#(\sigma)} \prod_{c \in \mathrm{Cycles}(\gamma\sigma)} \mathrm{Tr}(D_2^{|c|}) \,, \qquad \gamma = (1\,2\,\cdots\,m) \,, \tag{B.2.6}$$

where γ is the cyclic permutation induced by the trace over replica indices and $\#(\sigma)$ denotes the number of cycles of σ. Combining,

$$\mathbb{E}[\mathrm{Tr}(G^m)] = |o|^{2m} \left(\sum_{\pi \in S_m} \prod_{c \in \mathrm{Cycles}(\pi)} \mathrm{Tr}(D_1^{|c|}) \right) \left(\sum_{\sigma \in S_m} \kappa^{\#(\sigma)} \prod_{c \in \mathrm{Cycles}(\gamma\sigma)} \mathrm{Tr}(D_2^{|c|}) \right). \tag{B.2.7}$$

The factorised structure of (B.2.7) is a direct consequence of the rank-one nature of the state insertion: the state contributes only an overall scalar $|w|^2$ to the covariance, decoupling from the matrix structure controlled by D_2.

Rank.
The rank of G equals the dimension of the span of $\{x_i\}$. These vectors lie in the column space of $D_2^{1/2}$, of dimension $r_2 = \mathrm{rank}(D_2)$, and for generic Gaussian O_i are almost surely linearly independent up to that dimension, so

$$\mathrm{rank}(G) = \min(\kappa,\, r_2) \qquad \text{a.s.} \tag{B.2.8}$$

In particular, if D_2 is full rank then $\mathrm{rank}(G) = \min(\kappa, d)$ almost surely. The result is independent of D_1 and of the distribution of w: the state randomness affects the eigenvalue distribution of G but not its rank.

B.2.2 Trace construction: matrix-valued Wishart mixture

We now consider the structure that arises when the observer state enters through a trace rather than a vector contraction:

$$G_{ij} = \mathrm{Tr}(D_3\, A\, D_1^{1/2}\, O_i\, D_2\, O_j^\dagger\, D_1^{1/2}\, A^\dagger)\,, \tag{B.2.9}$$

with $D_1, D_2, D_3 \succeq 0$ diagonal. This structure appears in Sec. 6.3.5.2.

Using cyclicity of the trace, rewrite (B.2.9) as

$$G_{ij} = \mathrm{Tr}(C_A\, O_i\, D_2\, O_j^\dagger)\,, \qquad C_A := D_1^{1/2}\, A^\dagger\, D_3\, A\, D_1^{1/2}\,, \tag{B.2.10}$$

with C_A a positive semi-definite random matrix encoding all dependence on the state through A, D_1, and D_3. Vectorising the O_i puts this in the form

$$G_{ij} = \mathrm{vec}(O_i)^\dagger\, (D_2 \otimes C_A)\, \mathrm{vec}(O_j)\,, \tag{B.2.11}$$

so that, conditioned on A, G is a generalised Wishart matrix with separable covariance $\Sigma_A = D_2 \otimes C_A$. Marginalising over A produces a compound Wishart ensemble with random separable covariance. Unlike the rank-one case, the state randomness now enters the full matrix structure of the covariance rather than merely rescaling it.

Moments.
Conditioned on A, the Wishart moment formula gives

$$\mathbb{E}_O[\mathrm{Tr}(G^m) \mid A] = \sum_{\sigma \in S_m} \kappa^{\#(\sigma)} \prod_{c \in \mathrm{Cycles}(\gamma\sigma)} \mathrm{Tr}(D_2^{|c|})\, \mathrm{Tr}(C_A^{|c|})\,, \tag{B.2.12}$$

the products over cycles now containing paired traces from the two tensor factors. Averaging over A is carried out by the Gaussian Wick theorem applied to powers of C_A;

each such trace generates an additional sum over the symmetric group. Introducing the shorthand

$$\mathrm{Tr}_\pi(M) \equiv \prod_{c\in \mathrm{Cycles}(\pi)} \mathrm{Tr}(M^{|c|})\,, \tag{B.2.13}$$

the fully averaged moment is

$$\mathbb{E}[\mathrm{Tr}(G^m)] = \sum_{\sigma,\tau\in S_m} \kappa^{\#(\sigma)}\ \mathrm{Tr}_{\gamma\sigma}(D_2)\ \mathrm{Tr}_\tau(D_1)\ \mathrm{Tr}_{\gamma\sigma\tau}(D_3)\,. \tag{B.2.14}$$

The structural difference from the rank-one case (B.2.7) is that (B.2.14) involves a *coupled* double sum over $(\sigma,\tau)\in S_m\times S_m$, reflecting the interleaving of the Wick contractions of A with the replica permutation σ.

Rank.
Write $r_k = \mathrm{rank}(D_k)$ for $k=1,2,3$. The random matrix $C_A = D_1^{1/2}\,A^\dagger\,D_3\,A\,D_1^{1/2}$ has its image confined to the intersection of the column spaces of D_1 and the image of D_3 under A; since A is generically full rank,

$$\mathrm{rank}(C_A) = \min(r_1,r_3) \qquad \text{a.s.} \tag{B.2.15}$$

The covariance $\Sigma_A = D_2\otimes C_A$ therefore has rank $r_2\cdot\min(r_1,r_3)$, and G, being a Wishart matrix with this covariance from κ samples, satisfies

$$\mathrm{rank}(G) = \min(\kappa,\ r_2\cdot\min(r_1,r_3)) \qquad \text{a.s.} \tag{B.2.16}$$

When all three weights are full rank this reduces to $\min(\kappa,d^2)$. Comparing to (B.2.10), the trace construction yields the larger rank $r_2\cdot\min(r_1,r_3)$ instead of r_2 alone. This enhancement reflects the additional degrees of freedom available when the observer state enters through a trace: the internal Hilbert space factor over which the trace is taken provides $\min(r_1,r_3)$ independent "channels", each of which contributes r_2 dimensions to the effective covariance. In the gravitational setting, this distinction manifests as the difference between the single-boundary and two-boundary Gram matrices, where the trace over the second boundary factor enlarges the rank by a factor controlled by the entropy of that factor.

B.2.3 Bi-ensemble construction: two-sided Gram–Wishart

A richer structure arises when the Gram matrix is built from two independent Gaussian ensembles $\{O_i\}_{i=1}^{\kappa_R}$ and $\{Q_j\}_{j=1}^{\kappa_L}$ acting on opposite sides of the state insertion as in Sec. 6.3.5.1. We consider

$$G_{(i,j),(m,n)} = \mathrm{Tr}\Big(D_3^{1/2}\,A\,D_1^{1/2}\,O_i\,D_2\,O_m^\dagger\,D_1^{1/2}\,A^\dagger\,D_3^{1/2}\,Q_n^\dagger\,D_4\,Q_j\Big)\,, \tag{B.2.17}$$

with $D_1, D_2, D_3, D_4 \succeq 0$ diagonal, and the two ensembles $\{O_i\}$, $\{Q_j\}$ mutually independent and independent of A. This structure arises when operator insertions are placed on two distinct boundary factors: the O_i act on one factor with weight D_2, and the Q_j on the other with weight D_4. The Gram matrix is indexed by pairs and is of size $\kappa_R\kappa_L \times \kappa_R\kappa_L$.

Ensemble interpretation.
The natural variable for the bi-ensemble is the *bridge*

$$B := D_3^{1/2}\, A\, D_1^{1/2}\,, \tag{B.2.18}$$

in terms of which (B.2.17) reads

$$G_{(i,j),(m,n)} = \mathrm{Tr}(B\, O_i\, D_2\, O_m^\dagger\, B^\dagger\, Q_n^\dagger\, D_4\, Q_j)\,. \tag{B.2.19}$$

Defining the rectangular operator

$$Y_{ij} := D_4^{1/2}\, Q_j\, B\, O_i\, D_2^{1/2}\,, \tag{B.2.20}$$

we have $G_{(i,j),(m,n)} = \mathrm{Tr}(Y_{ij}\, Y_{mn}^\dagger) = \langle Y_{mn} \,|\, Y_{ij}\rangle_{\mathrm{HS}}$, so that G is the Gram matrix of the $\kappa_R\kappa_L$ Hilbert–Schmidt vectors $\{Y_{ij}\}$. Each Y_{ij} is bilinear in two independent Gaussian ensembles, so even conditioned on A the matrix G is *not* of generalised Wishart form; it is the Gram matrix of products of two independent Gaussians, a strict generalisation of the Wishart structure of the trace case. Marginalising over A adds a third independent Gaussian source.

Moments.
The moment computation proceeds as in the single-ensemble case, except that each of the three independent Gaussian ensembles $\{O_i\}$, $\{Q_j\}$, and A generates its own replica permutation. Carrying out the Gaussian integrals over $\{O_i\}$ and $\{Q_j\}$ separately and averaging over A via Wick contractions,

$$\mathbb{E}[\mathrm{Tr}(G^r)] = \sum_{\sigma_R,\sigma_L,\tau\in S_r} \kappa_R^{\#(\sigma_R)}\, \kappa_L^{\#(\sigma_L)}\, \mathrm{Tr}_{\gamma\sigma_R}(D_2)\, \mathrm{Tr}_{\gamma\sigma_L}(D_4)\, \mathrm{Tr}_{\tau}(D_1)\, \mathrm{Tr}_{\sigma_L^{-1}\sigma_R\tau}(D_3)\,. \tag{B.2.21}$$

The cycle structure on D_3 reflects the split insertion $D_3^{1/2}\cdots D_3^{1/2}$ in (B.2.17): the joint permutation $\gamma\sigma_R\sigma_L\tau$ interleaves the contractions of all three Gaussian sources. The trace-construction formula (B.2.14) is recovered by removing the Q ensemble entirely (dropping the σ_L sum) and setting $D_4 = \mathbb{1}$; note that merely taking $\kappa_L = 1$ does *not* suffice, because the Wick contractions of the single Gaussian sample of Q would still generate a sum over $\sigma_L \in S_r$ with weights $d^{\#(\gamma\sigma_L)}$.

Rank.
Both O_i and Q_j act on the same d-dimensional space, so the rank analysis requires more care than in the single-ensemble case. From (B.2.20), the Gram matrix realises G as the Gram matrix of $\kappa_R\kappa_L$ vectors in the Hilbert–Schmidt space $\mathcal{B}(\mathbb{C}^d) \cong \mathbb{C}^{d^2}$, so $\operatorname{rank}(G) \le d^2$.

Setting $r_k = \operatorname{rank}(D_k)$, the bridge $B = D_3^{1/2} A D_1^{1/2}$ is almost surely nonzero whenever $r_1, r_3 > 0$, since A has i.i.d. Gaussian entries. The map $(O_i, Q_j) \mapsto Y_{ij}$ then generates, almost surely, all matrices with row support in $\operatorname{Im}(D_4)$ and column support in $\operatorname{Im}(D_2)$, a space of dimension $r_2 r_4$. Therefore

$$\operatorname{rank}(G) = \min(\kappa_R\kappa_L,\ r_2\, r_4) \qquad \text{a.s.}, \tag{B.2.22}$$

provided $r_1 > 0$ and $r_3 > 0$; if any of the four weight matrices has vanishing support then $G = 0$ identically. When all four are full rank, $\operatorname{rank}(G) = \min(\kappa_R\kappa_L,\ d^2)$.

The crucial point is that the ambient space is $\mathcal{B}(\mathbb{C}^d)$ of dimension d^2: the second Gaussian ensemble increases the number of samples but does not enlarge the underlying vector space. The effective rank is therefore controlled by the outer supports D_4 and D_2, which determine the row and column supports of Y_{ij}, while the inner factors D_1 and D_3 enter only as a non-degeneracy condition. This differs from the trace construction (B.2.16), in which the inner factors actively contribute to the rank through $\min(r_1, r_3)$.

Origin of the rank structure.
The difference between the trace and bi-ensemble rank formulas has a structural origin in the Wick contraction patterns of the two ensembles. The $\{O_i\}$ are i.i.d. Gaussian random matrices, so O_i and O_j are independent for $i \neq j$ and identical for $i = j$. The $\{Q_j\}$ are an independent family of Gaussians; in particular O_i and Q_j are independent for *all* pairs (i, j), including $i = j$.

In the trace construction, $G_{ij} = \operatorname{Tr}(O_j^\dagger C_A O_i D_2)$ is bilinear in the $\{O_i\}$, with the diagonal $i = j$ identifications $O_j^\dagger = O_i^\dagger$ producing Wick contractions that couple the left and right factors through C_A. The resulting averages collapse G into generalised Wishart form with covariance $C_A \otimes D_2$, constraining the rank to $\min(r_1, r_3) \cdot r_2$.

In the bi-ensemble construction, no choice of indices makes O_i and Q_j the same random variable, so (O, Q) Wick contractions vanish identically at every order. The kernel B between them remains a fixed deterministic bridge rather than a covariance. The moment formula (B.2.21) makes this manifest: σ_R and σ_L range over independent symmetric groups and never share cycles, so the O- and Q-Wick graphs are entirely

disjoint. The inner factors D_1, D_3, which enter only through B, accordingly reduce to a non-degeneracy condition ($B \neq 0$ a.s.), and the rank is set by the outer supports D_2, D_4 together with the number of samples $\kappa_R \kappa_L$.

B.3 Ensemble normalisation and the Gaussian approximation

This appendix contains the technical details underlying the Gaussian approximation to the fixed-norm pure-state ensemble used in Sec. 6.4.2. The body of the paper stated the exact finite-D result, the saddle-point conditions, and the two sufficient conditions for the Gaussian approximation. Here we supply the contour-integral derivation, carry out the saddle-point evaluation in full, and work the two illustrative examples.

B.3.1 Contour-integral derivation of the partition function

The maximum-entropy ensemble of pure states $|\Psi\rangle = \sum_i A_i |i\rangle$ in a D-dimensional Hilbert space, subject to the normalisation constraint $\sum_i |A_i|^2 = 1$ and a set of diagonal constraints, is specified by the distribution (6.4.2). In the eigenbasis of the Lagrange multiplier matrix Λ (eigenvalues a_i), and setting $p_i \equiv |A_i|^2$, the partition function reduces to an integral over the probability simplex:

$$\mathcal{Z}(\mathbf{a}) = \int_{\Delta_{D-1}} dp \, \exp\left(- \sum_{i=1}^{D} a_i \, p_i \right), \tag{B.3.1}$$

with $\Delta_{D-1} = \{(p_1, \ldots, p_D) : p_i \geq 0, \, \sum_i p_i = 1\}$ and the flat measure $dp = \prod_i dp_i$.

The simplex constraint is enforced via the Laplace representation $\delta(\sum_i p_i - 1) = \frac{1}{2\pi i} \int ds \, e^{s(\sum_i p_i - 1)}$, with the contour running parallel to the imaginary axis to the right of all poles. Extending each p_i integral to $[0, \infty)$ and performing the resulting Gaussian integrals,

$$\mathcal{Z}(\mathbf{a}) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} ds \, \frac{e^s}{\prod_{i=1}^{D} (s + a_i)}, \tag{B.3.2}$$

with convergence requiring $\mathrm{Re}(s) > -\min_i(a_i)$, i.e. $\mathrm{Re}(a_i) > 0$ for all i.

B.3.2 Exact finite-D result and moments

The integrand in (B.3.2) is meromorphic in s with simple poles at $s = -a_i$, all of which lie in the left half-plane when $a_i > 0$. Closing the contour to the left and

applying the residue theorem yields

$$\mathcal{Z}(\mathbf{a}) = \sum_{i=1}^{D} \frac{e^{-a_i}}{\prod_{j\neq i}(a_j - a_i)}\,, \tag{B.3.3}$$

valid for distinct a_i; coincident eigenvalues are accommodated by the usual limiting procedure. Moments of the simplex distribution follow by differentiation,

$$\overline{p_i} = -\frac{\partial}{\partial a_i}\log\mathcal{Z}(\mathbf{a})\,, \qquad \overline{p_i\,p_j} = \frac{\partial^2}{\partial a_i\,\partial a_j}\log\mathcal{Z}(\mathbf{a})\,, \tag{B.3.4}$$

and similarly for higher moments.

B.3.3 Saddle-point evaluation

In the thermodynamic limit $D \to \infty$, the Laplace integral (B.3.2) is evaluated by saddle-point. Defining

$$\Sigma(s) \equiv s - \sum_{i=1}^{D}\log(s + a_i)\,, \tag{B.3.5}$$

one has $\mathcal{Z}(\mathbf{a}) \approx e^{\Sigma(s_*)}$ with s_* determined by $\Sigma'(s_*) = 0$, i.e.

$$1 = \sum_{i=1}^{D}\frac{1}{s_* + a_i}\,. \tag{B.3.6}$$

This is the normalisation condition $\sum_i \overline{p_i} = 1$ expressed in terms of the saddle. In the saddle-point approximation the moments factorise:

$$\overline{p_i} \approx \frac{1}{s_* + a_i}\,, \qquad \overline{p_i^n} \approx \frac{n!}{(s_* + a_i)^n}\,, \qquad \overline{p_i^k\,p_j^l}\Big|_{i\neq j} \approx \overline{p_i^k} \times \overline{p_j^l}\,. \tag{B.3.7}$$

These are the moments of independent exponential distributions $\mu(p_i) \propto e^{-(s_*+a_i)p_i}$, confirming that the fixed-norm ensemble is well approximated by a shifted Gaussian in the thermodynamic limit.

B.3.4 Conditions for the Gaussian approximation

The unconstrained Gaussian ensemble $\mu(p_i) \propto e^{-a_i p_i}$ (i.e. with $s_* = 0$) is a valid approximation to the fixed-norm ensemble precisely when two conditions are met:

Self-averaging condition.
The norm $N = \sum_i p_i$ must be sharply concentrated around its mean in the Gaussian ensemble:

$$\frac{\mathrm{Var}(N)}{(\overline{N})^2} = \frac{\sum_i a_i^{-2}}{(\sum_i a_i^{-1})^2} = \frac{\mathrm{Tr}(\overline{\rho}^2)}{(\mathrm{Tr}\,\overline{\rho})^2} \to 0 \qquad \text{as } D \to \infty\,. \tag{B.3.8}$$

This requires $\overline{\rho}$ to be sufficiently mixed; if $\overline{\rho}$ is pure, its purity equals unity and the condition fails.

Algebraic condition.
The saddle shift must vanish: $s_* = 0$. By (B.3.6), this requires the average norm in the unconstrained Gaussian to already be unity, which in terms of the constrained observables $\{\mathcal{O}_\alpha\}$ becomes

$$\exists\,\{b_\alpha\} \quad \text{s.t.} \quad \sum_\alpha b_\alpha\,\mathcal{O}_\alpha = \mathbb{1} \quad \text{and} \quad \sum_\alpha b_\alpha\,C_\alpha = 1\,, \tag{B.3.9}$$

where C_α is the measured value of $\mathcal{O}_\alpha$. The first condition states that the identity must lie in the span of the constrained operators; the second requires that the constraints automatically fix the norm.

When both conditions hold, the moments (B.3.7) with $s_* = 0$ reproduce the Gaussian moments $\overline{p_i} = a_i^{-1}$ and the factorisation of higher moments, up to $\mathcal{O}(D^{-1})$ corrections.

B.3.5 Two illustrative examples

Fixing the mean density matrix.
If the ensemble is constrained to fix $\overline{\rho} = \hat{\rho}$ with eigenvalues χ_i, the constraint forces $\overline{p_i} = \chi_i$ and $\sum_i \chi_i = 1$, solved by $a_i = \chi_i^{-1}$ and $s_* = 0$. The algebraic condition is automatically satisfied since $\mathrm{Tr}(\hat{\rho}) = 1$; the self-averaging condition reduces to $\mathrm{Tr}(\hat{\rho}^2) \to 0$, which holds whenever $\hat{\rho}$ is sufficiently mixed. The resulting ensemble is $|\Psi\rangle = \sum_i \sqrt{\chi_i}\,R_i\,|i\rangle$ with R_i independent complex unit Gaussians.

Fixing a single observable.
If only $\langle\mathcal{O}\rangle = C$ is constrained (with $\mathcal{O}$ having eigenvalues ξ_i), the identity is not in the span of $\{\mathcal{O}\}$ unless $\mathcal{O}$ is trivial, so the algebraic condition fails. The saddle condition requires a non-zero shift s_* determined by

$$\sum_i \frac{1}{s_* + \lambda\,\xi_i} = 1\,, \tag{B.3.10}$$

with λ fixed by $\sum_i \xi_i/(s_* + \lambda\,\xi_i) = C$. The naive Gaussian ($s_* = 0$) gives $\lambda = D/C$, which does not satisfy the saddle condition for a generic spectrum. The unconstrained Gaussian is therefore not a valid proxy for the normalised ensemble in this case, a fact with important consequences for the type B ensemble construction of Sec. 6.4.2.2.

B.4 Thermal saddle-point analysis

This appendix analyses the simplex ensemble constrained by the expectation value of the Hamiltonian, which is the starting point of the type B pure-state ensemble construction in Sec. 6.4.2.2. We show that while the mean energy distribution concentrates in a thermodynamic Gaussian around the saddle energy E_*, the full state-level distribution deviates significantly from the Gibbs state, explaining why the thermal character of the type B effective state must be imposed at the level of the density matrix rather than derived from the pure-state ensemble.

B.4.1 Continuum limit and saddle-point evaluation

In the continuum limit the partition function (B.3.3) becomes a functional integral. Introducing the density of states $\rho(E) = e^{\mathbf{S}(E)}$, setting $a(E) = \omega E$ with ω the Lagrange multiplier for the Hamiltonian constraint, and defining

$$\Phi(E) = -a(E) + \mathbf{S}(E) - \int dE'\,\rho(E')\,\ln(a(E') - a(E))\,, \tag{B.4.1}$$

the partition function becomes $\mathcal{Z} = \int dE\, e^{\Phi(E)}$. Assuming $\Phi(E)$ has an isolated maximum at $E = E_*$, expansion to quadratic order gives

$$\Phi(E) = \Phi(E_*) - \frac{(E - E_*)^2}{2\sigma^2} + \cdots\,, \qquad \sigma^2 \equiv [-\Phi''(E_*)]^{-1}\,, \tag{B.4.2}$$

so $\mathcal{Z} \simeq e^{\Phi(E_*)}\sqrt{2\pi}\,\sigma$.

The mean probability density is obtained by functional differentiation:

$$\overline{p(E)} = -\frac{\delta}{\delta a(E)}\ln\mathcal{Z} = -\frac{1}{\mathcal{Z}}\int d\hat{E}\,\frac{\delta\Phi(\hat{E})}{\delta a(E)}\,e^{\Phi(\hat{E})}\,. \tag{B.4.3}$$

B.4.2 Evaluation of the mean probability

The functional derivative of Φ produces three terms,

$$\frac{\delta\Phi(\hat{E})}{\delta a(E)} = -\delta(\hat{E} - E) - \frac{\rho(E)}{a(E) - a(\hat{E})} + \delta(\hat{E} - E)\int d\tilde{E}\,\frac{\rho(\tilde{E})}{a(\tilde{E}) - a(\hat{E})}\,, \tag{B.4.4}$$

each of which can be evaluated in saddle-point. The first and third are straightforward and give, respectively, $I_1(E) = e^{\Phi(E)}$ and $I_3(E) = e^{\Phi(E)}\,K(E)$, with $K(E) \equiv \int d\tilde{E}\,\rho(\tilde{E})/(a(\tilde{E}) - a(E))$. For the second, introducing $\hat{x} = \hat{E} - E_*$ and linearising $a(\hat{E}) \approx a(E_*) + \kappa_{\text{slope}}\hat{x}$ around the saddle (using κ_{slope} to distinguish the local slope of a from the number of states κ used in the main text), the integral reduces to a principal-value Gaussian

$$I_2(E) \simeq \rho(E)\,e^{\Phi(E_*)}\,\frac{\pi}{\kappa_{\text{slope}}}\,e^{-c^2}\,\text{erfi}(c)\,, \qquad c \equiv (E - E_*)/(\sqrt{2}\,\sigma)\,, \tag{B.4.5}$$

where $\text{erfi}(c) = (2/\sqrt{\pi})\int_0^c e^{t^2}\,dt$ is the imaginary error function. Collecting terms and dividing by $\mathcal{Z} \simeq e^{\Phi(E_*)}\sqrt{2\pi}\,\sigma$,

$$\overline{p(E)} \simeq \frac{e^{-c^2}}{\sqrt{2\pi}\,\sigma}\left[1 + \rho(E)\,\frac{\pi}{\kappa_{\text{slope}}}\,\text{erfi}(c) - K(E)\right]. \tag{B.4.6}$$

B.4.3 Asymptotic behaviour

Far from the saddle ($|E - E_*| \gg \sigma$).
For $c \gg 1$, $\text{erfi}(c) \sim e^{c^2}/(\sqrt{\pi}\,c)$. Substituting into (B.4.6), the Gaussian prefactor e^{-c^2} combines with the e^{c^2} growth of the erfi function to yield a polynomial rather than exponential falloff:

$$\overline{p(E)} \propto \frac{\rho(E)}{\kappa_{\text{slope}}(E - E_*)}\,. \tag{B.4.7}$$

This power-law tail is very different from the exponential suppression characteristic of the Gibbs state.

Near the saddle ($|E - E_*| \lesssim \sigma$).
Expanding $\text{erfi}(c)$ and $K(E)$ to leading order in c and using the saddle condition, the bracket in (B.4.6) reduces to a smooth function finite at $E = E_*$, and the leading dependence is the Gaussian prefactor:

$$\overline{p(E)} \simeq \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-(E-E_*)^2/(2\sigma^2)} \times (\text{smooth prefactor})\,. \tag{B.4.8}$$

The mean probability distribution of the total energy is therefore approximately Gaussian in a window of width σ around E_*, qualitatively similar to the canonical ensemble.

B.4.4 Comparison with the Gibbs state

The Gibbs state $\rho_{\beta_{\text{eff}}} = e^{-\beta_{\text{eff}}H}/Z(\beta_{\text{eff}})$ has eigenvalues $p_i^{\text{Gibbs}} = e^{-\beta_{\text{eff}}E_i}/Z(\beta_{\text{eff}})$, exponentially decaying in the energy. By contrast, the mean eigenvalues of the constrained

ensemble exhibit power-law tails, and the state-level distribution is close to uniform over the Hilbert space, with concentration only in the total energy.

This discrepancy has a simple origin. Constraining a single scalar quantity (the energy) in a D-dimensional state space leaves $D-1$ directions unconstrained, and the maximum Shannon entropy measure over normalised pure states subject to this single constraint is dominated by nearly uniform states, with only a mild energetic bias. The maximum von Neumann entropy *state* subject to the same constraint is the Gibbs state, but this is a property of the single-state optimisation over density matrices, not of the ensemble optimisation over pure states. The two max-ent procedures—von Neumann over density matrices (Sec. 6.1.3) and Shannon over pure-state distributions (Sec. 6.4.2)—optimise different functionals over different spaces, and their results are generically distinct when the number of constraints is small relative to D.

This is why the type B ensemble construction of Sec. 6.4.2.2 constrains $\overline{\rho} = \rho_{\beta_{\text{eff}}}$ directly, rather than constraining only $\langle H \rangle$: the thermal character of the observer's effective state must be imposed by the single-state max-ent procedure *before* the ensemble is constructed.

Bibliography

[1] V. Balasubramanian and T. Yildirim, *A Nonperturbative Toolkit for Quantum Gravity*, 2504.16986.

[2] V. Balasubramanian and T. Yildirim, *The nonperturbative Hilbert space of quantum gravity with one boundary*, *JHEP* **03** (2026) 040 [2506.04319].

[3] V. Balasubramanian and T. Yildirim, *How to Count States in Gravity*, 2506.15767.

[4] V. Balasubramanian and T. Yildirim, *Observing Spacetime*, 2509.09763.

[5] J. Unwin and T. Yildirim, *A QCD R-axion*, *Phys. Rev. D* **111** (2025) 015048 [2407.17557].

[6] V. Balasubramanian, A. Lawrence, J.M. Magan and M. Sasieta, *Microscopic Origin of the Entropy of Black Holes in General Relativity*, *Phys. Rev. X* **14** (2024) 011024 [2212.02447].

[7] J.M. Bardeen, B. Carter and S.W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161.

[8] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333.

[9] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199.

[10] G.W. Gibbons and S.W. Hawking, *Action Integrals and Partition Functions in Quantum Gravity*, *Phys. Rev. D* **15** (1977) 2752.

[11] S.W. Hawking and D.N. Page, *Thermodynamics of Black Holes in anti-De Sitter Space*, *Commun. Math. Phys.* **87** (1983) 577.

[12] G.W. Gibbons, S.W. Hawking and M.J. Perry, *Path Integrals and the Indefiniteness of the Gravitational Action*, *Nucl. Phys. B* **138** (1978) 141.

[13] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, *JHEP* **03** (2022) 205 [1911.11977].

[14] 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].

[15] S.W. Hawking, *Quantum Coherence Down the Wormhole*, *Phys. Lett. B* **195** (1987) 337.

[16] S.B. Giddings and A. Strominger, *Axion Induced Topology Change in Quantum Gravity and String Theory*, *Nucl. Phys. B* **306** (1988) 890.

[17] S.R. Coleman, *Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence*, *Nucl. Phys. B* **307** (1988) 867.

[18] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, 1903.11115.

[19] V. Balasubramanian, A. Lawrence, J.M. Magan and M. Sasieta, *Microscopic Origin of the Entropy of Astrophysical Black Holes*, *Phys. Rev. Lett.* **132** (2024) 141501 [2212.08623].

[20] L.V. Iliesiu, A. Levine, H.W. Lin, H. Maxfield and M. Mezei, *On the non-perturbative bulk Hilbert space of JT gravity*, *JHEP* **10** (2024) 220 [2403.08696].

[21] 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].

[22] D. Marolf, O. Parrikar, C. Rabideau, A. Izadi Rad and M. Van Raamsdonk, *From Euclidean Sources to Lorentzian Spacetimes in Holographic Conformal Field Theories*, *JHEP* **06** (2018) 077 [1709.10101].

[23] E. Colafranceschi, X. Dong, D. Marolf and Z. Wang, *Algebras and Hilbert spaces from gravitational path integrals. Understanding Ryu-Takayanagi/HRT as entropy without AdS/CFT*, *JHEP* **10** (2024) 063 [2310.02189].

[24] P. Saad, *Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity*, 1910.10311.

[25] M. Sasieta, *Wormholes from heavy operator statistics in AdS/CFT*, *JHEP* **03** (2023) 158 [2211.11794].

[26] J. de Boer, D. Liska, B. Post and M. Sasieta, *A principle of maximum ignorance for semiclassical gravity*, *JHEP* **02** (2024) 003 [2311.08132].

[27] J. de Boer, D. Liska and B. Post, *Multiboundary wormholes and OPE statistics*, *JHEP* **10** (2024) 207 [2405.13111].

[28] A. Climent, R. Emparan, J.M. Magan, M. Sasieta and A. Vilar López, *Universal construction of black hole microstates*, *Phys. Rev. D* **109** (2024) 086024 [2401.08775].

[29] V. Balasubramanian, B. Craps, J. Hernandez, M. Khramtsov and M. Knysh, *Factorization of the Hilbert space of eternal black holes in general relativity*, *JHEP* **01** (2025) 046 [2410.00091].

[30] D. Marolf, *On the nature of ensembles from gravitational path integrals*, *J. Phys. A* **58** (2025) 035204 [2407.04625].

[31] S. Antonini, M. Sasieta and B. Swingle, *Cosmology from random entanglement*, *JHEP* **11** (2023) 188 [2307.14416].

[32] V. Balasubramanian, B. Craps, J. Hernandez, M. Khramtsov and M. Knysh, *Counting microstates of out-of-equilibrium black hole fluctuations*, *JHEP* **06** (2025) 083 [2412.06884].

[33] A. Goel, H.T. Lam, G.J. Turiaci and H. Verlinde, *Expanding the Black Hole Interior: Partially Entangled Thermal States in SYK*, *JHEP* **02** (2019) 156 [1807.03916].

[34] J. Kudler-Flam and E. Witten, *Emergent Mixed States for Baby Universes and Black Holes*, 2510.06376.

[35] H. Liu, *Towards a holographic description of closed universes*, 2509.14327.

[36] A. Hamilton, D.N. Kabat, G. Lifschytz and D.A. Lowe, *Holographic representation of local bulk operators*, *Phys. Rev. D* **74** (2006) 066009 [hep-th/0606141].

[37] W. Israel, *Singular hypersurfaces and thin shells in general relativity*, *Nuovo Cim. B* **44S10** (1966) 1.

[38] V. Balasubramanian and P. Kraus, *A Stress tensor for Anti-de Sitter gravity*, *Commun. Math. Phys.* **208** (1999) 413 [hep-th/9902121].

[39] K. Skenderis, *Lecture notes on holographic renormalization*, *Class. Quant. Grav.* **19** (2002) 5849 [hep-th/0209067].

[40] D. He, J. Hernandez and M. Knysh, *Quantum corrected black hole microstates and entropy*, 2510.02997.

[41] J. Chandra and T. Hartman, *Coarse graining pure states in AdS/CFT*, *JHEP* **10** (2023) 030 [2206.03414].

[42] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [1005.3035].

[43] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781 [1306.0533].

[44] J. Boruch, L.V. Iliesiu, G. Lin and C. Yan, *How the Hilbert space of two-sided black holes factorises*, *JHEP* **06** (2025) 092 [2406.04396].

[45] J. Cotler and K. Jensen, *Gravitational Constrained Instantons*, *Phys. Rev. D* **104** (2021) 081501 [2010.02241].

[46] J. Cotler and K. Jensen, *A precision test of averaging in AdS/CFT*, *JHEP* **11** (2022) 070 [2205.12968].

[47] J. Cotler and K. Jensen, *Wormholes and black hole microstates in AdS/CFT*, *JHEP* **09** (2021) 001 [2104.00601].

[48] J.L.F. Barbon and E. Velasco-Aja, *A note on black hole entropy and wormhole instabilities*, *JHEP* **08** (2025) 103 [2502.00769].

[49] P.-S. Hsin, L.V. Iliesiu and Z. Yang, *A violation of global symmetries from replica wormholes and the fate of black hole remnants*, *Class. Quant. Grav.* **38** (2021) 194004 [2011.09444].

[50] J. Boruch, L.V. Iliesiu and C. Yan, *Constructing all BPS black hole microstates from the gravitational path integral*, *JHEP* **09** (2024) 058 [2307.13051].

[51] T. Banks, M.R. Douglas, G.T. Horowitz and E.J. Martinec, *AdS dynamics from conformal field theory*, hep-th/9808016.

[52] L. Fidkowski, V. Hubeny, M. Kleban and S. Shenker, *The Black hole singularity in AdS / CFT*, *JHEP* **02** (2004) 014 [hep-th/0306170].

[53] V. Balasubramanian, B. Czech, K. Larjo, D. Marolf and J. Simon, *Quantum geometry and gravitational entropy*, *JHEP* **12** (2007) 067 [0705.4431].

[54] K. Papadodimas and S. Raju, *Remarks on the necessity and implications of state-dependence in the black hole interior*, *Phys. Rev. D* **93** (2016) 084049 [1503.08825].

[55] K. Papadodimas and S. Raju, *Local Operators in the Eternal Black Hole*, *Phys. Rev. Lett.* **115** (2015) 211601 [1502.06692].

[56] C. Akers, N. Engelhardt, D. Harlow, G. Penington and S. Vardhan, *The black hole interior from non-isometric codes and complexity*, *JHEP* **06** (2024) 155 [2207.06536].

[57] S. Antonini, V. Balasubramanian, N. Bao, C. Cao and W. Chemissany, *Non-isometry, state dependence and holography*, *JHEP* **02** (2025) 150 [2411.07296].

[58] D.L. Jafferis, *Bulk reconstruction and the Hartle-Hawking wavefunction*, 1703.01519.

[59] B. Concepcion, Y. Nomura, K. Ritchie and S. Weiss, *Complementarity for a dynamical black hole*, *Phys. Rev. D* **110** (2024) 086002 [2405.15849].

[60] S. Leutheusser and H. Liu, *Causal connectability between quantum systems and the black hole interior in holographic duality*, *Phys. Rev. D* **108** (2023) 086019 [2110.05497].

[61] S. Leutheusser and H. Liu, *Subregion-subalgebra duality: Emergence of space and time in holography*, *Phys. Rev. D* **111** (2025) 066021 [2212.13266].

[62] V. Chandrasekaran, G. Penington and E. Witten, *Large N algebras and generalized entropy*, *JHEP* **04** (2023) 009 [2209.10454].

[63] V. Chandrasekaran, R. Longo, G. Penington and E. Witten, *An algebra of observables for de Sitter space*, *JHEP* **02** (2023) 082 [2206.10780].

[64] E. Witten, *Gravity and the crossed product*, *JHEP* **10** (2022) 008 [2112.12828].

[65] S. Banerjee, J. Erdmenger and J. Karl, *Nonlocality induces isometry and factorisation in holography*, *Phys. Rev. D* **112** (2025) L021902 [2411.09616].

[66] J.M. Magan and Q. Wu, *Two types of quantum chaos: testing the limits of the Bohigas-Giannoni-Schmit conjecture*, 2411.08186.

[67] H. Verlinde, *Wormholes in Quantum Mechanics*, 2105.02129.

[68] D. Harlow, *Wormholes, Emergent Gauge Fields, and the Weak Gravity Conjecture*, *JHEP* **01** (2016) 122 [1510.07911].

[69] D. Harlow and D. Jafferis, *The Factorization Problem in Jackiw-Teitelboim Gravity*, *JHEP* **02** (2020) 177 [1804.01081].

[70] D. Stanford and E. Witten, *JT Gravity and the Ensembles of Random Matrix Theory*, *Adv. Theor. Math. Phys.* **24** (2020) 1475 [1907.03363].

[71] A. Belin, J. de Boer, D.L. Jafferis, P. Nayak and J. Sonner, *Approximate CFTs and Random Tensor Models*, *JHEP* **09** (2024) 163 [2308.03829].

[72] D.L. Jafferis, L. Rozenberg and G. Wong, *3d Gravity as a random ensemble*, *JHEP* **02** (2025) 208 [2407.02649].

[73] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090 [1304.4926].

[74] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [hep-th/0603001].

[75] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045 [hep-th/0605073].

[76] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A Covariant Holographic Entanglement Entropy Proposal*, *JHEP* **07** (2007) 062 [0705.0016].

[77] B. Banihashemi and T. Jacobson, *The enigmatic gravitational partition function*, *Gen. Rel. Grav.* **57** (2025) 43 [2411.00267].

[78] H. Geng and Y. Jiang, *Microscopic origin of the entropy of single-sided black holes*, *JHEP* **04** (2025) 133 [2409.12219].

[79] J.M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021 [hep-th/0106112].

[80] V. Balasubramanian, D. Marolf and M. Rozali, *Information Recovery From Black Holes*, *Gen. Rel. Grav.* **38** (2006) 1529 [hep-th/0604045].

[81] V. Balasubramanian, B. Czech, K. Larjo and J. Simon, *Integrability versus information loss: A Simple example*, *JHEP* **11** (2006) 001 [hep-th/0602263].

[82] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, *JHEP* **09** (2020) 002 [1905.08255].

[83] A.R. Brown, H. Gharibyan, G. Penington and L. Susskind, *The Python's Lunch: geometric obstructions to decoding Hawking radiation*, *JHEP* **08** (2020) 121 [1912.00228].

[84] N. Engelhardt, G. Penington and A. Shahbazi-Moghaddam, *A world without pythons would be so simple*, *Class. Quant. Grav.* **38** (2021) 234001 [2102.07774].

[85] N. Engelhardt, G. Penington and A. Shahbazi-Moghaddam, *Finding pythons in unexpected places*, *Class. Quant. Grav.* **39** (2022) 094002 [2105.09316].

[86] N. Engelhardt, Å. Folkestad, A. Levine, E. Verheijden and L. Yang, *Cryptographic Censorship*, *JHEP* **01** (2025) 122 [2402.03425].

[87] V. Balasubramanian, A. Kar, C. Li and O. Parrikar, *Quantum error correction in the black hole interior*, *JHEP* **07** (2023) 189 [2203.01961].

[88] V. Balasubramanian, A. Kar, C. Li, O. Parrikar and H. Rajgadia, *Quantum error correction from complexity in Brownian SYK*, *JHEP* **08** (2023) 071 [2301.07108].

[89] V. Balasubramanian, J. de Boer, V. Jejjala and J. Simon, *The Library of Babel: On the origin of gravitational thermodynamics*, *JHEP* **12** (2005) 006 [hep-th/0508023].

[90] M. Headrick, V.E. Hubeny, A. Lawrence and M. Rangamani, *Causality & holographic entanglement entropy*, *JHEP* **12** (2014) 162 [1408.6300].

[91] D.L. Jafferis, A. Lewkowycz, J. Maldacena and S.J. Suh, *Relative entropy equals bulk relative entropy*, *JHEP* **06** (2016) 004 [1512.06431].

[92] 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].

[93] N. Bao and I.H. Kim, *Precursor problem and holographic mutual information*, 1601.07616.

[94] T. Faulkner and A. Lewkowycz, *Bulk locality from modular flow*, *JHEP* **07** (2017) 151 [1704.05464].

[95] 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].

[96] C.-F. Chen, G. Penington and G. Salton, *Entanglement Wedge Reconstruction using the Petz Map*, *JHEP* **01** (2020) 168 [1902.02844].

[97] M.J. Kang and D.K. Kolchmeyer, *Holographic Relative Entropy in Infinite-Dimensional Hilbert Spaces*, *Commun. Math. Phys.* **400** (2023) 1665 [1811.05482].

[98] E. Gesteau and M.J. Kang, *Thermal states are vital: Entanglement Wedge Reconstruction from Operator-Pushing*, 2005.07189.

[99] L.K. Grover, *A fast quantum mechanical algorithm for database search*, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219, 1996.

[100] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (6, 2012).

[101] M. Usatyuk, Z.-Y. Wang and Y. Zhao, *Closed universes in two dimensional gravity*, *SciPost Phys.* **17** (2024) 051 [2402.00098].

[102] L. Wang and R. Li, *Entanglement islands and the Page curve of Hawking radiation for rotating Kerr black holes*, *Phys. Rev. D* **110** (2024) 066012 [2406.13949].

[103] S. Antonini, P. Rath, M. Sasieta, B. Swingle and A. Vilar López, *The baby universe is fine and the CFT knows it: on holography for closed universes*, *JHEP* **12** (2025) 159 [2507.10649].

[104] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [1306.0622].

[105] J.S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S.H. Shenker et al., *Black Holes and Random Matrices*, *JHEP* **05** (2017) 118 [1611.04650].

[106] P. Saad, S.H. Shenker and D. Stanford, *A semiclassical ramp in SYK and in gravity*, 1806.06840.

[107] J.-M. Schlenker and E. Witten, *No ensemble averaging below the black hole threshold*, *JHEP* **07** (2022) 143 [2202.01372].

[108] M. Usatyuk and Y. Zhao, *Closed universes, factorization, and ensemble averaging*, *JHEP* **02** (2025) 052 [2403.13047].

[109] D. Harlow, M. Usatyuk and Y. Zhao, *Quantum mechanics and observers for gravity in a closed universe*, *JHEP* **02** (2026) 108 [2501.02359].

[110] N. Engelhardt and E. Gesteau, *Further Evidence Against a Semiclassical Baby Universe in AdS/CFT*, 2504.14586.

[111] J. Maldacena, *Real observers solving imaginary problems*, 2412.14014.

[112] J. De Vuyst, S. Eccles, P.A. Höhn and J. Kirklin, *Gravitational entropy is observer-dependent*, 2405.00114.

[113] J. De Vuyst, S. Eccles, P.A. Höhn and J. Kirklin, *Crossed products and quantum reference frames: on the observer-dependence of gravitational entropy*, *JHEP* **07** (2025) 063 [2412.15502].

[114] K. Jensen, S. Raju and A.J. Speranza, *Holographic observers for time-band algebras*, *JHEP* **06** (2025) 242 [2412.21185].

[115] Y. Nomura and T. Ugajin, *Nonperturbative quantum gravity in a closed Lorentzian universe*, *JHEP* **10** (2025) 166 [2505.20390].

[116] Y. Nomura and T. Ugajin, *Physical Predictions in Closed Quantum Gravity*, 2602.13387.

[117] A. Sahu and M. Van Raamsdonk, *Holographic black hole cosmologies*, *JHEP* **05** (2025) 233 [2411.14673].

[118] C. Akers, G. Bueller, O. DeWolfe, K. Higginbotham, J. Reinking and R. Rodriguez, *On observers in holographic maps*, *JHEP* **05** (2025) 201 [2503.09681].

[119] H.Z. Chen, *Observers seeing gravitational Hilbert spaces: abstract sources for an abstract path integral*, *JHEP* **10** (2025) 139 [2505.15892].

[120] Z. Wei, *Observers and Timekeepers: From the Page-Wootters Mechanism to the Gravitational Path Integral*, 2506.21489.

[121] A. Blommaert, J. Kudler-Flam and E.Y. Urbach, *Absolute entropy and the observer's no-boundary state*, *JHEP* **11** (2025) 113 [2505.14771].

[122] N. Engelhardt, E. Gesteau and D. Harlow, *Observer complementarity for black holes and holography*, 2507.06046.

[123] K. Higginbotham, *On tests for baby universes in AdS/CFT*, *JHEP* **09** (2025) 038 [2507.05337].

[124] T. Mori and B. Yoshida, *Baby universe as logical qubits: information recovery in random encoding*, 2511.20747.

[125] Y. Zhao, *"It from Bit": The Hartle-Hawking state and quantum mechanics for de Sitter observers*, 2602.05939.

[126] S. Antonini and P. Rath, *Do holographic CFT states have unique semiclassical bulk duals?*, *Int. J. Mod. Phys. D* **34** (2025) 2544025 [2408.02720].

[127] A.I. Abdalla, S. Antonini, L.V. Iliesiu and A. Levine, *The gravitational path integral from an observer's point of view*, *JHEP* **05** (2025) 059 [2501.02632].

[128] V.E. Hubeny and M. Rangamani, *Causal Holographic Information*, *JHEP* **06** (2012) 114 [1204.1698].

[129] W.R. Kelly and A.C. Wall, *Coarse-grained entropy and causal holographic information in AdS/CFT*, *JHEP* **03** (2014) 118 [1309.3610].

[130] N. Engelhardt and A.C. Wall, *Decoding the apparent horizon: coarse-grained holographic entropy*, *Phys. Rev. Lett.* **121** (2018) 211301 [1706.02038].

[131] N. Engelhardt and A.C. Wall, *Coarse graining holographic black holes*, *JHEP* **05** (2019) 160 [1806.01281].

[132] V.I. Paulsen, *Completely Bounded Maps and Operator Algebras*, vol. 78 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press (2002), 10.1017/CBO9780511546631.

[133] M. Srednicki, *Chaos and quantum thermalization*, *Phys. Rev. E* **50** (1994) 888.

[134] J.M. Deutsch, *Quantum statistical mechanics in a closed system*, *Phys. Rev. A* **43** (1991) 2046.

[135] N. Lashkari, A. Dymarsky and H. Liu, *Eigenstate Thermalization Hypothesis in Conformal Field Theory*, *J. Stat. Mech.* **1803** (2018) 033101 [1610.00302].

[136] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [1307.2892].

[137] N. Engelhardt and A.C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [1408.3203].

[138] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [1908.10996].

[139] A. Hamilton, D.N. Kabat, G. Lifschytz and D.A. Lowe, *Local bulk operators in AdS/CFT: A Boundary view of horizons and locality*, *Phys. Rev. D* **73** (2006) 086003 [hep-th/0506118].

[140] S. Leutheusser and H. Liu, *Emergent times in holographic duality*, *Phys. Rev. D* **108** (2023) 086020 [2112.12156].

[141] E. Witten, *Why Does Quantum Field Theory In Curved Spacetime Make Sense? And What Happens To The Algebra of Observables In The Thermodynamic Limit?*, in *Dialogues Between Physics and Mathematics: C. N. Yang at 100*, M.-L. Ge, A.J. Niemi, K.K. Phua and L.A. Takhtajan, eds., pp. 241–284, Springer (2022), DOI [2112.11614].

[142] E. Witten, *Algebras, regions, and observers*, *Proc. Symp. Pure Math.* **107** (2024) 247 [2303.02837].

[143] E. Witten, *A background-independent algebra in quantum gravity*, *JHEP* **03** (2024) 077 [2308.03663].

[144] G. Penington and E. Witten, *Algebras and States in JT Gravity*, 2301.07257.

[145] D.K. Kolchmeyer, *von neumann algebras in jt gravity*, *JHEP* **06** (2023) 067 [2303.04701].

[146] K. Jensen, J. Sorce and A.J. Speranza, *Generalized entropy for general subregions in quantum gravity*, *JHEP* **12** (2023) 020 [2306.01837].

[147] S. Ali Ahmad and R. Jefferson, *Crossed product algebras and generalized entropy for subregions*, *SciPost Phys. Core* **7** (2024) 020 [2306.07323].

[148] M.S. Klinger and R.G. Leigh, *Crossed products, extended phase spaces and the resolution of entanglement singularities*, *Nucl. Phys. B* **999** (2024) 116453 [2306.09314].

[149] T. Faulkner and A.J. Speranza, *Gravitational algebras and the generalized second law*, *JHEP* **11** (2024) 099 [2405.00847].

[150] J. Kudler-Flam, S. Leutheusser and G. Satishchandran, *Generalized black hole entropy is von neumann entropy*, *Phys. Rev. D* **111** (2025) 025013 [2309.15897].

[151] M. Ali and V. Suneeta, *Generalized entropy in higher curvature gravity and entropy of algebra of observables*, *Phys. Rev. D* **108** (2023) 066017 [2307.00241].

[152] M. Ali and V. Suneeta, *Local generalized second law in crossed product constructions*, *Phys. Rev. D* **111** (2025) 024015 [2404.00718].

[153] J. Kudler-Flam, S. Leutheusser and G. Satishchandran, *Algebraic observational cosmology*, 2406.01669.

[154] J. de Boer, J. Hollander and A. Rolph, *A Reverse Black Hole Information Problem*, 2601.22077.

[155] J. de Boer, D. Liska and K. Salehi Vaziri, *Microstate counting from defects in de Sitter*, 2511.09624.

[156] D. Harlow, *Observers, α-parameters, and the Hartle-Hawking state*, 2602.03835.