CHAOS Holography And Other Science

# **Daniel Adam Roberts**

## CHAOS, Holography, And Other Science

by

Daniel Adam Roberts

Submitted to the Department of Physics in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

### MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

© Massachusetts Institute of Technology 2016. All rights reserved.

Certified by ...... Allan Wilfred Adams III Associate Professor of Physics Thesis Supervisor

Accepted by ...... Nergis Mavalvala Associate Department Head for Education

### CHAOS, Holography, And Other Science

by

Daniel Adam Roberts

Submitted to the Department of Physics on April 29, 2016, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics

#### Abstract

This thesis focuses on the relationship between black holes in holography, chaos in stronglycoupled quantum systems, and the computational complexity of holographic states. By directly considering the time evolution of local operators, I am led to a simple diagnostic of many-body chaos: a commutator of such operators separated in time and space. Using this diagnostic, I study the growth of operators—a manifestation of the butterfly effect—in a variety of quantum systems. By considering the butterfly effect in holography, I find evidence for a detailed correspondence between the tensor network (or quantum circuit) that builds the holographic state and the interior geometry (or Einstein-Rosen bridge) of the black hole. Ultimately, I try to understand these connections by considering entanglement across time: the entanglement between an output system following time evolution and a record or memory perfectly correlated with the initial system.

Thesis Supervisor: Allan Wilfred Adams III Title: Associate Professor of Physics

This thesis is based on the following four papers:

- $\bullet~[1]$  written with Douglas Stanford and Leonard Susskind
- $\bullet~[2]$  written with Adam Brown, Leonard Susskind, Brian Swingle, and Ying Zhao
- [3] written with Douglas Stanford
- $\bullet~[4]$  written with Pavan Hosur, Xiao-Liang Qi, and Beni Yoshida

## Acknowledgments

In many of my papers, acknowledgments have been a space for sincere thanks, literal acknowledgments (of funding sources), and hidden jokes. This is made possible by the fact that in all of my papers, the acknowledgments have come at the end. In this work, we are instructed to do things sdrawkcab.

It is my pleasure to thank many people:

My advisor, Allan Adams, for enthusiastically supporting all my varied interests and for always clearly smiling (frowning) when I would say something correct (incorrect) so that in discussions I always had "experimental" means of determining right from wrong.

My numerous collaborators ... for collaborating. Many of you literally helped me write this thesis. Special acknowledgment to Douglas Stanford, for essentially becoming my second advisor when he was only a graduate student. My entire approach to physics is derived from the things I learned during our collaborations. To Lenny Susskind, for adopting me at Stanford and always making me feel welcomed at the SITP. To Jesse Thaler, for doing the same for me at the CTP (and for working with me on the project that legitimately en*titles* me to add "Other Science"). To Yoni Kahn, for convincing me to come to MIT, for finding a way for us to collaborate even though we work on dissimilar things, (for introducing me to Jesse,) and for being the person who instantly alleviates all of my silly confusions. And to Beni Yoshida, Adam Brown, and Brian Swingle, for treating me like an adult physicist when I was probably more like an adolescent (if not a child). Finally, to Steve Shenker for his support, encouragement, and help.<sup>1</sup>

I'd also like to thank potential future collaborators (i.e. current and former CTP members): Tarek Anous (my Pufu), Michael Crossley, Ethan Dyer, Paolo Glorioso, my five-year officemate Dan Kolodrubetz (especially, for tolerating five years of distracting behavior and encouraging *all* my shenanigans), Mark Mezei, Ian Moult, Lina Necib, Nikhil Raghuram, Shu-Heng Shao, Zach Thomas, and Yifan Wang.

Next, I would like to more than merely acknowledge two of my funding sources: the Marshall Scholarship and the Fannie and John Hertz Foundation, both of which are responsible for providing all the opportunities that brought me where I am today (at any future date you happen to be reading this!). Also to the latter for actually standing by its mission

<sup>&</sup>lt;sup>1</sup>Though technically he is not (yet) a collaborator, he is acknowledged on nearly all of my papers and this thesis is indebted to his and Douglas's work.

statement and giving me the freedom to pursue all of my academic interests (i.e. chaos, holography, and other science), and for introducing me to John Frank, Kelly Moynihan, Josh Mueller, Vyas Ramanan, and Thomas Segall-Shapiro, whom I thank for science, *etc.* 

Although I feel (hope) that all my collaborators are close friends, the converse does not necessarily hold.<sup>2</sup> First, ack. to Kyle Mahowald, my longest running roommate (six years). That alone should indicate the level of thanks that he deserves. (It also occurs to me that I should probably thank him for teaching me the value of constraints and "constraint-based activities.") Thanks to my other roommate Robbie Kubala, who provides well-needed balance. Thanks to Max Kleiman-Weiner, for work on projects best classified as "other science" (i.e for cultivating my interest in machine learning and AI). And finally, an extra thanks to all three of them for all of the things.

Finally, I'd like to thank my family: my sister, Nicole Roberts, for agreeing to my strange request to design a cover page for my thesis (and actually designing it and making it absolutely amazing!); and my parents, David and Linda Roberts, for explaining everything to me before I learned how to learn.

This work was supported by the DOD through the NDSEG Program, by the Fannie and John Hertz Foundation, and by the U.S. Department of Energy under cooperative research agreement Contract Number DE-SC0012567.

This Ph.D. has been brought to you by the letters A, C, H, O, and S.

<sup>&</sup>lt;sup>2</sup>This means that I have friends who are not physicists.

"It's wanting to know that makes us matter." Tom Stoppard, Arcadia

"S =  $-\mathrm{tr} \rho \log \rho$ , dE = TdS,  $E = mc^2$ ." Tom Stoppard, Arcadia (translated to Greek)

"Graecum est, non legitur." William Shakespeare, *The Tragedy of Julius Caesar* (translated from English)

# Contents

1	$\mathbf{C}_{i}$	rculus	vitiosus: Introduction			
<b>2</b>	HIC SVNT DRACONES: Localized shocks					
	2.1	Introduction				
		2.1.1	Some terminology	41		
2.2 Qubit systems				41		
		2.2.1	Precursor growth	42		
		2.2.2	TN for multiple localized precursors	45		
	2.3 Holographic systems					
		2.3.1	Localized shock waves	49		
		2.3.2	Precursor growth	52		
		2.3.3	ERB dual to multiple localized precursors	54		
2.4 Discussion				60		
	2.5	ndix	61			
		2.5.1	Localized shocks at finite time	61		
		2.5.2	Localized shocks in Gauss-Bonnet	63		
		2.5.3	Maximal volume surface and decoupled surface	64		
3	$\mathbf{A}_{i}$	Agere sequitur credere: Complexity equals action 67				
4	$\mathbf{O}\mathbf{rdo}\ \mathbf{ab}\ \mathbf{chao}$ : Two-dimensional conformal field theory and the butterfly effect					
	4.1	Introd	luction	81		
4.2 CFT calculations						
		4.2.1	Conventions and review	83		
		4.2.2	Continuation to the second sheet	85		

		4.2.3	The Virasoro identity block	88			
		4.2.4	Various comments	89			
	4.3	3 Discussion					
	4.4	1.4 Appendix					
		4.4.1	Bulk calculations	95			
		4.4.2	Ising model	99			
5	Scientia ipsa potentia est: Chaos in quantum channels						
	5.1	Introd	luction	103			
	5.2	Unitary quantum channels					
		5.2.1	Entanglement in time	109			
		5.2.2	Scrambling	110			
		5.2.3	Examples	113			
		5.2.4	Chaotic channels vs. integrable channels	118			
	5.3	Butter	rfly effect implies scrambling	119			
		5.3.1	Average over OTO correlators	119			
		5.3.2	Early-time behavior	123			
		5.3.3	Butterfly velocity vs. entanglement velocity	126			
	5.4	Numerics in qubit channels					
		5.4.1	Spin chains	130			
		5.4.2	Majorana fermion fast scrambler	136			
	5.5	Perfec	t tensor model	138			
		5.5.1	Ballistic growth of operators and linear growth of $I_3$	139			
		5.5.2	Recurrence time	143			
	5.6	Discus	ssion	146			
	5.7	Apper	ndix	147			
		5.7.1	Haar scrambling	147			
		5.7.2	Entanglement propagation in CFT	150			
		5.7.3	Proof of Eq. (5.32)	153			
		5.7.4	Tensor calculus	157			
٨	?	A gover	soquitur complexionem. Enilogue	161			
A	• 1	rgere s	bequitur complexionem: Ephogue	TOT			

## Chapter 1

## Circulus vitiosus

(Introduction)

The text in this introduction will appear in a future article elsewhere.

#### \*\*\*\*

If you run out of memory on your computer, you can purchase an external hard drive. If you fill up that as well, you can buy two more. (In 2016) each hard drive has a volume of roughly 50 cm<sup>3</sup>, weighs 50 grams, and can store  $1 \text{ TB} \approx 10^{13}$  bits. One can repeat this exercise with other physical media: written text, pictures, *etc.*, but the conclusion is always the same: information apparently takes of up space. And intuitively, it scales like the volume; every time you double the number of physical hard drives, you double the amount of information you can store.

How many hard drives can you really store? Maybe you have a really big walk-in closet where you put these hard drives, but eventually you'll even fill that up.<sup>1</sup> If you continue on this futile attempt at acquiring more hard drives, you're also going to run out of silicon and metal and subvert global industrial progress and grind the economy to a halt. But beyond all these mere "practical" concerns, eventually, if you put enough hard drives together, something very surprising is going to happen: they're going to collapse into a black hole [5-8]. It's a *vicious circle*. The gravitational attraction of the 50-gram hard drives to each

 $<sup>^{1}</sup>$ A typical walk-in closet has a volume of roughly 5 cubic meters, allowing for 100,000 hard drives storing 100 petabytes or approximately  $10^{19}$  bits.

other is outrageously tiny, but if you stick enough of them together they will eventually matter in a very serious way.<sup>2</sup>

Okay, but what if instead you just invent a better hard drive? As we have in the past, in the future we will probably develop technology that will let us increase storage capacity. But what's the best we can do? Should there even be a "best we can do?" Ignoring "practical" (i.e. engineering) challenges, could we in principle engineer the bit to volume ratio (i.e. the storage capacity per hard drive of fixed size) to be arbitrarily large?

Before we address that, more fundamentally how is it that "information" can take up *physical* space? How much does information weigh? Why does it weigh?

## Informative thermodynamics

Information takes up space because it must be represented. You need paper on which to inscribe your story. A picture—representing a visual scene—is a physical object that you frame. Hard drives store information digitally—encoded in binary—which means they must contain different physical states corresponding to different patterns of 1s and 0s. The exact physical state of the hard drive is known as a *microstate*. Therefore, the total number of possible microstates—the different possible sequences of 1s and 0s—characterizes the storage potential of the hard drive.

For such a binary sequence, the logarithm of the number of different sequences gives its length  $S = \log_2 W$ , where W is the number of microstates and S is the number of bits in the sequence. In physics, S goes by the name *entropy*. In www.apple.com, S goes by the name "hard drive storage capacity." The entropy characterizes the number of physical states in a system and therefore the amount of information that it can represent.

In more general physical systems, the different microstates may not be equally likely to occur. In that case, we can assign a probability  $p_i$  to each microstate i, and the entropy is

<sup>&</sup>lt;sup>2</sup>The Schwarzschild radius  $R_s = 2G_N M$  relates the mass of a black hole to its radius [5], where  $R_s$  is the Schwarzschild radius, M is the mass, and  $G_N = 6.67 \times 10^{-11} \,\mathrm{m}^3 \,\mathrm{kg}^{-1} \,\mathrm{s}^{-2}$  is Newton's constant. If you try to pack enough mass into a small enough region with linear dimensions R such that  $R/M = 2G_N$  you will form a black hole. For a collection of N objects of constant density, the mass of the collection scales like N but the radius of the collection scales like  $N^{1/3}$ . Thus,  $R/M \sim N^{-2/3}$  and you will always reach the black hole threshold by increasing the number of objects in your collection. In the case of the hard drives, you will need to pack together roughly  $10^{42}$  of them for them to collapse into a black hole. For comparison,  $10^{42}$  hard drives could store roughly  $10^{31}$  times the total information on the Internet (again, pegged the values that are reasonable when this thesis was published). Clearly nowhere near this many hard drives exist (on Earth).

given by

$$S = -\sum_{i} p_i \log p_i. \tag{1.1}$$

In information theory, this formula is known as the Shannon entropy [9]. The quantity  $-\log p_i$  is known as the *surprisal* characterizing how "unexpected" a given probabilistic outcome (i.e. observed microstate) is. Really rare outcomes with  $p_i \ll 1$  have really large surprisal. In this context, the entropy is a measure of the expected surprisal in the system. In physics, (1.1) is known as the Gibbs entropy [10] though a similar formula was used previously by Boltzmann. When the distribution is uniform  $p_i = 1/W$ , this formula reduces to the simpler formula  $S = \log W$  which was discovered by Boltzmann by 1875.<sup>3</sup>

Of course, everything in last few paragraphs is a bad approximation to the actual physical world. Such machinery of statistical physics and information theory is *classical*: information expressed using binary sequences. Our universe is *quantum*, and the physical states are (approximately described) by *qubits* rather than bits. The generalization of (1.1) is the von Neumann entropy [11]

$$S = -\mathrm{tr}\,\rho\log\rho,\tag{1.2}$$

where  $\rho$  is the density matrix (also introduced by von Neumann) and generalizes the probability distribution  $p_i$  to include both quantum and classical correlation.<sup>4</sup>

Now, to practically describe a system with a gigantic number of states, one needs to group similar states together to get a *macrostate*. For instance, if the system in question is a text (i.e. a book), you might ignore capitalization: the microstate "apple" would be considered to be the same macrostate as the microstate "Apple". Or you might choose to group microstates together by concepts so that "orange" and "apple" are treated the same. Or you might simply ignore the content and characterize the macrostate by the total number of each letter; this would just require only 26 numbers. This representation loses information about the microstate since it clearly cannot be inverted; there are many microstates with the same letter frequencies which would get mapped to the same macrostate.

<sup>&</sup>lt;sup>3</sup>The Gibbs formulation was published in 1878 [10], but Shannon's seminal work on communication theory (which also coined the term/unit "bit") wasn't published until 1948 [9]. The punchline is that despite being a single formula (1.1), it took 70 years for entropy as a measure of states in a physical system to be connected to an information-theoretic notion of surprisal and storage capacity.

<sup>&</sup>lt;sup>4</sup>The sorts of systems we will consider in this thesis are strongly-coupled quantum systems, and the von Neumann entropy and related quantities will play a central role in what follows. However, for the purpose of the first part of the introduction, the differences between quantum and classical information are not important.

If you encounter a system in the "wild," to learn about it you have to perform measurements. The sorts of things you can easily measure are coarse-grained properties, such as temperature or energy. However, there may be many states at a fixed temperature in a small energy range. Thus, entropy can also be thought of as the ignorance of incomplete measurement. In the same sense, if you read the label on a hard drive and it says 1 TB, this coarse grained property—the storage capacity—doesn't tell you what's actually stored on it. In fact, all 1 TB hard drives (of the same brand and model) look and feel identical. The 1 TB entropy captures the idea that there are approximately  $2^{10^{13}}$  possible configurations of the bits on the hard drive. On the other hand if you know that all the bits are 0, that the hard drive is precisely erased, then you know its exact microstate. From this perspective (entropy as ignorance), we would say that its entropy is 0.

Returning to our central point, the notion of entropy thus explained (whether in terms of physical states or information content) is statistical. However, there's one additional notion of "coarse-grained" entropy, which was originally introduced by Clausius in 1865 in an entirely different context: thermodynamics. For a physical process at constant volume, the first law of thermodynamics

$$dE = TdS, (1.3)$$

relates the change in energy dE in a system to the change in entropy dS when the system is at temperature T. The quantity TdS often goes by the name heat. Some comments are in order: (i) It was the genius of Boltzmann that offered the original statistical interpretation of entropy. (ii) Returning to our discussion of coarse-grained variables, it seems that the number of states accessible at a fixed temperature increases with energy.<sup>5</sup>

Thus, (1.3) almost provides an answer to our earlier question about how information can take up space and weigh. To store more information in a fixed volume, you need to add energy. To complete the picture, we need the most famous equation in physics [12]

$$E = mc^2, (1.4)$$

<sup>&</sup>lt;sup>5</sup>The reason we present the relationship between energy and temperature in differential form as (1.3)—rather than some function E(S,T)—is because E(S,T) in not universal. In general, the relationship will depend on the particular system being studied. For instance, a gas of photons has E = 3TS/4, but for more complicated systems E(S,T) will not necessarily have a simple functional form. Nevertheless, it's always expected that the density of states—the number of states available to the system at approximately the same energy—will increase monotonically with energy.

which states that mass and energy are equivalent.<sup>6</sup> This suggests that to engineer a hard drive (of fixed size) with a larger storage capacity, it will have to weigh more. And this makes intuitive sense: since we know information needs to be represented in the physical state of the system, if we want to represent more information we should increase the amount of "stuff" in the system.

In summary, "it's wanting to know that makes us matter," or, more specifically, the very act of acquiring:

$$S = -\operatorname{tr} \rho \log \rho, \qquad dE = TdS, \qquad E = mc^2. \tag{1.5}$$

And as we ourselves learn (1.5), we matter.

## Informative holography

The Schwarzschild radius

$$R_s = 2G_N M,\tag{1.6}$$

relates the mass M of a spherically-symmetric black hole to its radius  $R_s$  [5]. They are proportional, and (one half) the constant of proportionality is Newton's gravitational constant  $G_N = 6.67 \times 10^{-11} \,\mathrm{m^3 \, kg^{-1} \, s^{-2}}$ . This can be rearranged to be interpreted as bound on the amount of mass that can fit inside a (spherical) region of space of volume V

$$M \le \frac{1}{2G_N} \left(\frac{3V}{4\pi}\right)^{1/3}.$$
 (1.7)

Black holes saturate this bound. This means that if you didn't start with a black hole but you cram enough mass into a region to saturate (1.7), you will end up with a black hole. If you try to increase the mass of a black hole by a small amount  $\delta M$ , due to the relation (1.6) the volume must also increase by an amount proportional to  $M^2 \delta M$ .

So, returning to our question, can we always just invent a better hard drive? As already recounted, (in 2016) hard drives are roughly the size of a deck of cards, weight only half that of the same deck, and pack in enough space to store roughly 20 copies of the English language Wikipedia. This is an information density of  $10^{17}$  b/m<sup>3</sup>. We now know that

<sup>&</sup>lt;sup>6</sup>N.B. (1.4) is the only time in this thesis that the speed of light c will not be set to unity. (It seemed very underwhelming to announce "the most famous equation in physics" and then write E = m.)

increasing the capacity of a hard drive of fixed size will make it heavier. Engineers are very clever, but due to (1.7) this process of improvement can't go on forever. Once the deck-of-cards-sized hard drive weigh about  $10^{25}$  kg (roughly 2.5 times the mass of Earth), it will form a black hole.

Once it's a black hole, does it continue to be a hard drive? Certainly the technology of modern (2016 era) hard drives cannot be scaled up so that they weight more than Earth, but really the questions we are asking are: Can black holes can store information? Equivalently, do black holes have entropy?

In classical gravity (general relativity), every (non-rotating, neutral) black hole is described by exactly one number, its mass. That means that regardless of whether it was formed with a collection of blank hard drives, with a collection of hard drives containing the world's knowledge, or even by a collapsing star, the resulting black holes are identical. One might think that this means that there's only one black hole of a given mass—one microstate—and that therefore black holes can't represent any information or carry entropy.

In classical gravity, there's a problem with the idea that black holes don't carry entropy. Namely, it means that one can violate the second law of thermodynamics—the fact that the entropy of a closed system cannot decrease—by throwing highly entropic objects into the black hole to "disappear" their entropy. Bekenstein resolved this paradox by suggesting that black holes must have an entropy [13, 14]. Considering a theorem of Hawking's that the surface area of a black hole may never decrease [15], Bekenstein thought that the reason this sounded so similar to the second law of thermodynamics is because it *is* the second law of (black hole) thermodynamics. Via a series of thought experiments, he argued that black holes have an entropy proportional to their area and furthermore that they also have a temperature.

A black body with a temperature will emit thermal radiation. This is in tension with the idea that black holes cannot emit any radiation. (Obviously.) Hawking was disturbed by this and sought to falsify the idea that black holes could have a temperature via a semiclassical quantum field theory calculation in a black hole background. As the story goes, he found a very surprising result: black holes emit Hawking radiation [16, 17]! Furthermore, he was

able to derive a precise relationship between the black hole's area and it's entropy<sup>7</sup>

$$S_{BH} = \frac{A}{4G_N},\tag{1.8}$$

where A is the surface area of the black hole and  $S_{BH}$  is the Bekenstein-Hawking entropy (though if the physicists in question had different names it's possible the symbol for black hole entropy might still have ended up being the same).<sup>8</sup>

Various comments are in order.

1. From the perspective of building our black hole by gravitationally collapsing hard drives, we might expect that the information in the black hole must somehow contain a record of the information on the hard drives. Black holes made from hard drives storing different data must somehow be different, even if they look the same from the outside. That is, the black hole macrostate is specified by its mass, and the entropy represents the ignorance over the particular black hole—the microstate. Classical gravity is an effective field theory, but with a complete theory of quantum gravity one might expect to be able to count the black hole microstates of a given fixed energy to get the Bekenstein-Hawking entropy (1.8). Hawking derived (1.8) in 1974, but it wasn't until 1996 that Strominger and Vafa were able to count the black hole microstates is more general settings is still confusing and potentially may not be resolved until there is a complete theory of quantum gravity.

2. It is absolutely crazy that the entropy of the black hole is proportional to its surface area. Entropy, like energy but unlike temperature, is what is known as an *extensive* quantity. This means that it scales with the volume of the system. If you have one hard drive with an entropy of 1 TB and then you get a second identical hard drive, then the combined two-hard-drive system has an entropy 2 TB total.<sup>9</sup> Take two black holes and add them together (collapse them into the same region) and the information capacity as compared to the single black hole doesn't just double; it increases by a factor of 4! (This is because

<sup>&</sup>lt;sup>7</sup>N.B. below and in the rest of this thesis we have set various constants to unity:  $\hbar = k_B = c = 1$ . However, we will never set  $G_N = 1$  since it's not dimensionless! (Not to mention the fact that in AdS units we usually take  $G_N \ll 1$ .)

<sup>&</sup>lt;sup>8</sup>Resolving the earlier paradox, Bekenstein conjectured a "generalized" second law of thermodynamics which stated that the entropy of a black hole *and* everything else must never decrease [18]. Hawking then showed this is in fact the case when Hawking radiation is taken into account [19].

<sup>&</sup>lt;sup>9</sup>The same is not true for temperature, an *intensive* quantity. If your body temperature is 98.6 °F and your significant other's body temperature is also 98.6 °F, the combined system of you and your significant other (despite what you may have thought) is still only 98.6 °F.

 $S_{BH} \sim R_s^2 \sim M^2$ , and we're doubling the mass.) The point is that due to the fact that  $R_s = 2G_N M$ , doubling the mass (or energy) of the system doesn't just double the volume (as would be true for an extensive system), but increases the volume by a factor of 8 (since  $V \sim R_s^3 \sim M^3$ ).<sup>10</sup>

**3.** More generally, extensive quantities have the property that they scale with the number of degrees of freedom in the system. In systems without gravity, the number of degrees of freedom scales with the number of sites—the volume. (1.8) suggests that for gravitational systems the number of degrees of freedom scales with the surface area. This led 't Hooft [21] and then Susskind [22] to postulate a holographic bound stating that the entropy of any region of spacetime is bounded as

$$S \le \frac{A}{4G_N},\tag{1.9}$$

with saturation for black holes. Informally, black holes make the densest hard drives in nature.<sup>11</sup> The power of this idea is the *holographic principle*. It postulates that all d + 1-dimensional quantum-gravitational systems have a description in terms of their *d*dimensional non-gravitational boundaries (where it is suggested that the degrees of freedom "live").<sup>12</sup>

4. This notion of holography has actually been the focus of much of high-energy theoretical physics for the last 20 years. In 1997, Maldacena discovered a specific example of the holographic principle and made a conjecture that string theory on Anti-de Sitter space

$$S \le 2\pi ER,\tag{1.10}$$

<sup>&</sup>lt;sup>10</sup>There's a small caveat about this comparison. When you combine the two 1 TB hard drives, the temperature of the combined system (as compared to the original separated systems) doesn't change. When you combine the two black holes, the temperature for the combined system as compared to the temperature of the single black hole (for Schwarzschild black holes in flat space) decreases by a factor of 2 (since  $T \sim M^{-1}$ ). I thank Jesse Thaler for raising this point.

<sup>&</sup>lt;sup>11</sup>The bound (1.9) does not constrain the energy or mass of the system, only its area. A related informationtheoretic bound

known as the Bekenstein bound, holds for a region of linear dimension R and energy E [23] (see also [24, 25]). <sup>12</sup>Since the smallest finite entropy that makes sense is 1 bit, this bound also suggests the notion of a minimal area  $\sim G_N \approx 10^{-70}$  m<sup>2</sup>, often denoted the *Planck area*. In fact, there's an entire set of related minimal Planck units: length, mass, time, *etc.*, derived by Planck in 1899 by combining the fundamental constants of nature:  $c, \hbar, G_N$ , of relativity, quantum theory, and gravity, respectively [26]. (Of course, when  $c = \hbar = 1$ , these all become simply related by different powers of  $G_N$ .)

The history of this is all very confusing since Planck's 1899 paper on natural units [26] appears to have preceded his 1900 paper [27] where he postulates that the energy of light is quantized  $E = \hbar \omega$  and where he introduces Planck's constant  $h \equiv 2\pi\hbar$ . However, since both papers are in German and scanned in such a way that Google will not easily translate them, it's probably impossible to say at this point what actually happened.

(AdS) backgrounds has a dual description in terms of conformal field theory (CFT) living on the boundary of spacetime [28] (see also [29, 30] and, as of April 2016, over 11,000 others). The AdS string theory is a quantum gravitational theory that lives in the "bulk," and the CFT is a particular type of quantum field theory that lives on the boundary. This conjecture goes by the "AdS/CFT correspondence" or more generally "gauge/gravity duality," or just simply holography. The amazing thing about this duality is that it relates strongly-coupled theories to weakly-coupled theories. In the regime we understand reasonably well, the bulk is described by weak-coupled classical gravity and the boundary is a strongly-coupled quantum theory. In this thesis, we will make heavy use of holography both to learn about chaos in strongly-coupled quantum systems and to try to understand the computational power of black holes.

Okay, enough comments on the importance of (1.8).

Now that we know that black holes have entropy, we can address the question of how much information our technologically-enhanced black-hole hard drive can store. As a reminder, we engineered a deck-of-cards-sized hard drive to weigh about 2.5 times the mass of Earth so that it saturated the mass bound (1.7) and formed a black hole. Using the entropy relation (1.8), we see that our maximal hard drive can store  $10^{67}$  bits!<sup>13</sup> Originally it stored only 1 TB or  $10^{13}$  bits. This is just unfathomably large. The largest metric prefix only goes up to  $10^{24}$  (= Yotta), and the total information on the Internet (in 2016) is approximate 1 YB or (converting bytes to bits)  $10^{25}$  bits. Thus, this tiny black hole hard drive could store roughly  $10^{42}$  complete (2016) Internets!<sup>14</sup>

One final comparison: as discussed in footnotes 13–14, comparing the capacity of a

 $<sup>^{13}</sup>$ An astute reader armed with multiplication might be confused about how the capacity of this deck-ofcards-sized hard drive has an entropy  $10^{67}$  bits, but the  $10^{42}$  1 TB hard drives packed together discussed in footnote 2 would naively appear to have a capacity of only  $10^{55}$  bits. As a reminder, putting this many hard drives together will form a black hole of radius  $10^{12}$  m, which is 6 times the radius of the red supergiant star Betelgeuse.

The problem is that in the former case (the deck-of-cards-sized hard drive), we are using (1.8) to compute the storage capacity, but in the latter case (the  $10^{42}$  1 TB hard drives) we are simply multiplying ( $10^{42} \times 1$  TB =  $10^{55}$  bits) to compute the storage capacity of the hard drives qua hard drives. This encoding of information is very inefficient; most of the entropy in the system (in the material making up the hard drives) is not used for storage. In fact, using (1.8) we can compute the total entropy of the black hole hole formed by the  $10^{42}$  1 TB hard drives, giving  $10^{94}$  bits.

<sup>&</sup>lt;sup>14</sup>Note that in general we cannot calculate the storage capacity of a hard drive just by knowing its mass and volume because it depends on the details of the technology involved. It's only because the entropy of a black hole is related to its area that we can fix the capacity of our black-hole hard drive. That is, there can be no technological improvement to force more bits into the black hole without changing its surface area. Throw something into a black hole, and its area increase is dictated by the laws of thermodynamics. This is another way of understanding why black holes are the ultimate in information storage.

hard drive to the entropy of a black hole isn't exactly fair; the hard drive storage capacity doesn't include the configurational entropy of the constituent materials that make up the hard drive. Since a hard drive has many components, estimating the total thermodynamic entropy of the hard drive system would be quite difficult. As a substitute, it is interesting to compare the entropy of a black hole of mass M to the entropy of a gas of photons of the same mass and confined to the same volume.<sup>15</sup> Such a gas is known to have an energy given by  $M = \pi^2 V T^4/15$  and an entropy given by S = 4M/3T, where V is the volume, and T is the temperature. In this setup, the radius of the gas is given in terms of its mass by the Schwarzschild radius  $r = 2G_N M$ . Solving for the temperature in terms of the radius, we can write the entropy of the gas as

$$S_{gas} = \left(\frac{128\pi^3}{3645}\right)^{1/4} \frac{r^{3/2}}{G_N^{3/4}},\tag{1.11}$$

where the numerical constant in front has an O(1) value that is very O(1): 1.02. The important point is that the scaling with the radius of the entropy of the photon gas  $S_{gas} \sim r^{3/2}/G_N^{3/4}$  is much smaller than the scaling with the radius of the entropy of the black hole,  $S_{BH} \sim r^2/G_N$ . For comparison, a photon gas the size of a deck of cards has an entropy of  $10^{50}$  bits. This is much smaller than the entropy of the same sized black hole ( $S_{BH} = 10^{67}$ bits), but many orders of magnitude larger than the storage capacity of the 1 TB hard drive  $(10^{13} \text{ bits})$ .

In summary, information requires physical representation (i.e. needs to be stored somewhere), and these simple thought experiments show that black holes are the extreme limits of physical memory. Whether we try to store more information by considering more hard drives (i.e. by increasing the volume, holding the information density fixed), or by improving the technology (i.e. by increasing the information density, holding the volume fixed), we always eventually make a black hole.

### Informative processing

If a black hole is a hard drive, is it also a computer? If so, how fast of a computer? Even if it is only a hard drive, how do we access the "files"—i.e. recover the information?

<sup>&</sup>lt;sup>15</sup>This is similar to a calculation of 't Hooft's from his original paper on the holographic principle [21].

Computation is the processing of information. Since information needs to be encoded in a physical system, physical systems that change in time can be thought of as computers. As we will see, black holes are not just computers, they are nature's most extreme computers.

The first clue that black holes compute is in the fact that black holes are not actually black.<sup>16</sup> The semiclassical calculation of Hawking [16, 17] suggests that black holes can eventually completely evaporate. This in turn suggests that the information in the black hole (the specification of the particular microstate) is contained in the Hawking radiation. However, due to the semiclassical nature of his calculation, its regime of validity leaves much to be desired, and neither of these "suggestions" are at all well-understood. The question of whether the information in a black hole can be recovered is known as the "black hole information paradox" (unless you're Strominger in which case you call it the "black hole information puzzle").

Hawking himself suggested that the information is lost/destroyed [31]. His calculation predicts an exactly thermal spectrum—the radiation of a black body. This means that the radiation is essentially random and uncorrelated with the black hole, making it impossible to deduce the exact microstate.

The destruction of the information is in tension with a general (and very important) principle of quantum physics known as unitarity. This is the notion that the microscopic laws of physics (i.e. before any coarse-graining occurs) are time-reversal invariant; in other words, all physical processes are reversible. If the Hawking radiation contains no information, all black hole microstates lead to the same state of random Hawking radiation. Many states are mapped to one state, and this process is clearly not reversible.

On the other hand, when considering coarse-grained variables (i.e. thermodynamics) entropy can increase; information can be lost. This suggests that Hawking's calculation is incomplete, and the microstate can be recovered from the Hawking radiation. This turns out to be a lot like saying that the information in an encyclopedia (or, for the modern audience, the information in a hard drive containing Wikipedia) can be recovered from the ash resulting from burning it. While in principle such a burning process is reversible and correlations in the ash (shape, placement, *etc.*) contain the information to reconstruct the encyclopedia, it is impossible in practice.<sup>17</sup>

<sup>&</sup>lt;sup>16</sup>Let us reemphasize, Hawking's calculation [16, 17] incorporated quantum effects; black holes are actually black in classical gravity.

<sup>&</sup>lt;sup>17</sup>Other good examples of this phenomenon include *Humpty Dumpty* and the great fall [32].

Turning an encyclopedia into ash involves *processing* the information it contains, effectively *scrambling* it so that it's rendered unusable. That is, the information is no longer represented in a local way such that an individual consulting the resulting system can easily obtain it. Instead, you'd basically need to understand how to follow the ash back through time such that the encyclopedia can be reconstructed. Thus, for information to be processed it must interact and spread. In a large sense, this is what is meant by a (extremely not useful) computation.

If the information in a black hole is not destroyed, then black holes must be really good information scramblers so that the resulting Hawking radiation looks thermal. Let's try and understand what it means for black holes to process information. Let's say that you recently collapsed some hard drives together so they form a black hole. You knew exactly what is on the hard drives initially, and it's reasonable to expect that the black-hole microstate (initially) is closely related to the contents of the hard drives. Now, what happens when you try to store a few additional bits of information in the black hole?

In this framework, throwing a small particle into a black hole is equivalent to trying to store additional bits on the hard drive. The black hole is a very unusual hard drive. Rather than saving those bits in a well-defined memory address or physical location, the black hole will immediately spread and *scramble* them over all its many degrees of freedom.<sup>18</sup> It will scramble the additional bits so thoroughly that in a short time you will have to consult almost the entire hard drive just to recover them. Modern computer memory has the opposite property; the information is localized and can be recovered almost instantly, regardless of the overall storage capacity. Therefore, a black hole is really like a (randomizing) computer rather than a hard drive. This scrambling is essential so that when the black hole starts

<sup>&</sup>lt;sup>18</sup>Since the entropy is proportional to the area, many people expect that the degrees of freedom live on the surface area—the horizon. However, if you jump into the black hole, you are not supposed to experience anything as you cross the horizon. In classical gravity, the horizon is not a locally distinguishable place (i.e. it should be the vacuum), and you aren't supposed to be able to tell whether you've even crossed a horizon. The only way you know you're in a black hole is when you eventually get stretched apart at the singularity.

Susskind, Thorlacius, and Uglum attempted to resolve this paradox by postulating *black hole complementarity*: the experience of the infalling observer has a dual description on the surface of the black hole [33]. However, using a quantum-information argument based on entanglement Almheiri, Marolf, Polchinski, and Sully (AMPS) show that the idea of black hole complementarity is inconsistent with the black hole horizon being in a vacuum state [34]. An observer that crosses such a horizon in an excited state would experience what they call a "firewall." This has thrown into question whether black holes can even have interiors.

In this thesis, we hope to relate the notion of the black hole interior to the notion of computational complexity. It has been suggested by Susskind that while "typical" states of high complexity might have firewalls, most accessible black hole states (meaning states that can be formed in a reasonable time from e.g. the collapse of a star) will be of low complexity and without firewalls [35].

evaporating via Hawking radiation, any small collection of the radiation will look thermal. It is only when all the radiation is taken together and decoded that the information can be recovered.<sup>19</sup>

How fast can black holes process and scramble (quantum) information? In a paper seemingly unrelated to black hole physics, Page noted that for generic states any bipartition that includes up to half the (quantum) bits will looks completely random [37]. The time it takes for a system to reach this "generic" condition from a finely-tuned state is known as the *scrambling time*. Applying Page's machinery to black holes, Hayden and Preskill argued that the state of the black hole scrambles in a time

$$t_* \sim \beta \log S,\tag{1.12}$$

where  $\beta$  is the inverse temperature 1/T, and  $t_*$  is a new timescale known as as the fast scrambling time [38].<sup>20</sup> This timescale was so novel that Sekino and Susskind conjectured that  $\sim \beta \log S$  is the fastest scrambling time in nature and that black holes saturate it [39] (though it remains an open question how exactly black holes manage this feat).

It's the log S scaling that makes it "fast." In order for scrambling to occur, distant parts of the system must (indirectly) interact. For most physical systems, interactions can only involve a few degrees of freedom. (This property is often referred to as k-locality, where k specifies the maximum number of degrees of freedom that can interact at a time.) If the system has a notion of spatial locality—if the interactions can only involve spatial close degrees of freedom—then we expect that scrambling takes a time proportional to the largest linear dimension of the system. Thus, for a system with d spatial dimensions containing Sdegrees of freedom, we'd expect a scrambling time  $t_* \sim \beta S^{1/d}$ . Taking  $d \to \infty$ , we'd find (1.12). Thus, black holes behaves as if they effectively live in an infinite number of spatial dimensions!<sup>21</sup>

In a seminal holographic computation (upon which much of the work in this thesis is indebted), Shenker and Stanford calculated the scrambling time of black holes in AdS

<sup>&</sup>lt;sup>19</sup>However, even if you have all of the radiation, it isn't expected to be easy to decode! This is like saying that if you have all of the ash you can recover the burnt encyclopedia. In [36], this Hawking radiation decoding question is precisely considered and shown to take a time that is exponential in the (already maximal given the spacetime region) entropy of the black hole.

<sup>&</sup>lt;sup>20</sup>The temperature is the natural scale here so that  $t_*$  carries the correct units. That is, we expect at infinite temperature that scrambling is very quick and at zero temperature (the ground state) that scrambling does not occur.

<sup>&</sup>lt;sup>21</sup>For additional work on scrambling, see [40-43].

[44]. They realized that throwing a particle into a black hole—processing a few extra bits through the computer—is a manifestation of the butterfly effect. They computed the time after which the infalling particle makes a large difference to the state of the black hole, finding

$$t_* = \frac{\beta}{2\pi} \log S,\tag{1.13}$$

precisely the scrambling time of [38, 39]. Note, this fixes the constant of proportionality  $\beta/2\pi$  in (1.12). This value is extremely special for black holes, and Kitaev has suggested that it's a necessary (and possibly sufficient) condition for a holographic system to have an interpretation as a black hole in Einstein gravity (i.e. classical general relativity) [45].

Scrambling requires a system to process information. This ought to be related to how strongly the system interacts, since if the information-carrying particles pass right through each other there can be no computation. Computers work the same way: a computer needs a *central* processing unit or CPU where information is combined and mixed together. If a computer doesn't have a CPU, then it is not a computer—it's only a hard drive.<sup>22</sup> In Chapter 3 (based on a collaboration with Brown, Susskind, Swingle, and Zhao [2]), we will introduce a conjecture in holography relating the computational complexity or complicatedness of the boundary state to a geometric property of the black hole interior. In our parlance, complexity is the minimal difficulty of taking the system from a simple reference state to a particular state of interest. It can also be related to the amount of time it takes for a computer do a computation—to convert the information stored on the hard drive from one state to another. Based on this holographic conjecture (and in the spirit of the fast scrambling conjecture [39]), we make and provide evidence for an additional conjecture that black holes are nature's fastest computers. In particular, we conjecture that the rate of complexity growth is bounded

$$\frac{d\text{Complexity}}{dt} \le \frac{2M}{\pi},\tag{1.14}$$

where M is the mass of the system. This bound was inspired by a conjectured complexity bound of Lloyd [47] based on bounds on the quantum orthogonality time by Margolus and

<sup>&</sup>lt;sup>22</sup>For a Turing machine [46], the CPU is the "head" that reads the "tape." In the frame of the head, the content of the tape (the information) is brought to the head where it is processed. For systems that are entirely decentralized, information exchange and processing still has to occur locally. I thank Jesse Thaler for this line of questioning and for emphasizing these points.

Levitin [48]. In the supplementary article [49], we provide evidence that black holes saturate this bound.

Just to put some numbers to (1.14), a (2016) computer can compute with  $10^{12}$  operations per second. A black hole of equal mass would compute at  $10^{50}$  operations per second. By comparison, the total (2016) human computational resources are estimated to be roughly  $10^{20}$  operations per second. So a black hole computer potentially has the power of  $10^{30}$  total human computational resources! Since chaotic systems, by definition, are almost impossible to simulate, one interpretation of (1.14) and its relation to scrambling is that only interacting systems are capable of computation; i.e. chaos is the capacity to do computation. We will return to this idea in Chapter 5.<sup>23</sup>

## Informative chaos

(The original title for this section was "Informative technicalities.")

A main goal of this thesis is to understand chaos in many-body quantum systems. This will be shown to be closely related to computation/complexification and be highly relevant for black hole physics because, as has hopefully been clearly argued, black holes are nature's most extreme computers. We will focus on two basic questions:

- Can a small perturbation have a macroscopic effect on a quantum system?
- How fast can information propagate in a quantum system?

We will attempt to answer both questions by studying the commutator of local operators separated in time and space. If such a commutator is zero, a measurement of one of the operators cannot affect the other. If such commutator is small, there can be no information transmission [51]. Thus, it is a natural observable to consider for answering the above questions. In fact, we will argue that the growth of such commutators is a manifestation of the butterfly effect in strongly-coupled quantum systems.

Consider two simple Hermitian local operators W(x) and V(0) located at x and at the origin so that they are separated by a positive distance x > 0. Since they have no

 $<sup>^{23}</sup>$ N.B. this does not necessarily imply that black holes are easy to program: due to the fact that they are maximally chaotic, setting up a particular computation is likely to take an exponential amount of time. For generic problems, one can think of the computation itself as being done by the encoding and decoding, making the black hole itself unnecessary [50]. From this perspective, the fact that a black hole saturates the complexity bound (1.14) can be interpreted as specifying the necessary computational resources required to simulate that black hole.

overlapping support, they will commute [W(x), V(0)] = 0. Now let's separate the operators in space and time by evolving W for time t using the Hamiltonian H. To compare to V at time 0, we consider the Heisenberg operator  $W(x,t) = e^{iHt}W(x)e^{-iHt}$  in the Schrödinger picture. This is known as a *precursor* because W(x,t) precurses the action of applying W(x) at time t. In general, W(x,t) will be a complicated operator comprised of sums of products of local operators and will not necessarily commute with V(0).

In this thesis we will study such commutators [W(x,t), V(0)] in a variety of systems to understand how they grow. In particular, we will often study the quantity

$$C(x,t) \equiv -\langle [W(x,t), V(0)]^2 \rangle_\beta, \qquad (1.15)$$

where  $\langle \cdot \rangle_{\beta} = Z^{-1} \operatorname{tr} \{ \cdot e^{-\beta H} \}$  indicates an average over the thermal ensemble, and  $Z = \operatorname{tr} e^{-\beta H}$ . The minus sign ensures  $C(x,t) \geq 0$ . In principle, we could consider this commutator in any state, but for the most part we will be interested in chaos as a property of thermal systems.

A commutator (1.15) is a natural diagnostic of quantum chaos. It gives the effect of a perturbation V(0) on later measurements of W(x) at time t; i.e. a sensitivity to initial conditions. A system is said to be chaotic if commutators C(x,t) will grow large (and stay large for a time that scale double-exponentially with the entropy  $O(e^{e^S})$ ) for almost all choices of simple operators W, V [3, 45]. It is by this notion that a small perturbation have a macroscopic effect on a quantum system.

This measure of chaos was first pointed out by Larkin and Ovchinnikov in 1969 in the context of semiclassical chaos [52] and then rediscovered recently in the context of the firewall paradox by Almheiri, Marolf, Polchinski, Stanford, and Sully [53]. Shenker and Stanford studied this and related quantities in holography, showing that AdS black holes behave chaotically [44, 54]. By considering spherically symmetric perturbations, they were able to derive the log S black hole fast scrambling time (1.13) by considering the time it takes for the commutator C(x,t) to become O(1).<sup>24</sup> Their paper launched the study of chaos in strongly-coupled quantum systems and inspired a large body of follow-up work [1, 3, 4, 55–73], including much of this thesis.

Now, let us try to understand how C(x, t) behaves in different systems. We are interested

 $<sup>^{24}</sup>$ In Chapter 5 we will explain the connection between this commutator definition of the scrambling time and the information-theoretic definition presented in the previous subsection.

in evolution by Hamiltonians with a notion of locality. If a Hamiltonian can be expressed in a basis of simple operators such that the number of operators multiplied in each term is never greater than k, then it is k-local

$$H_k = \sum_{i_1\dots i_k} J_{i_1\dots i_k} \mathcal{O}_{i_1}\dots \mathcal{O}_{i_k}.$$
(1.16)

The subscript on  $\mathcal{O}$  serves as a label for the operator. The number of operators in a given term is sometimes referred to as the *weight*.

For systems with a notion of spatial locality, the basis where  $H_k$  is k-local should be the position basis. In that case, the subscript on  $\mathcal{O}$  labels both type and position, and  $J_{i_1...i_k} = 0$  for terms where the operators are not spatially neighboring.<sup>25</sup> For 0-dimensional systems, k-locality provides a notion of locality despite the absence of a notion of space.

For Lorentz invariant relativistic systems, spacelike operators must commute

$$[W(x,t), V(0)] = 0, \qquad t < x, \qquad \text{(relativistic systems)}. \tag{1.17}$$

This is the usual notion of causality. In non-relativistic systems, the commutator [W(x, t), V(0)]can be nonzero for any finite time t > 0. We can gain intuition for this by using the Baker-Campbell-Hausdorff formula to expand W(x, t)

$$W(x,t) = W(x) + it[H, W(x)] - \frac{t^2}{2!}[H, [H, W(x)]] - \frac{it^3}{3!}[H, [H, [H, W(x)]]] + \dots$$
(1.18)

The *j*th-order nested commutator (H, ..., [H, W(x)] ...] can lead term terms in W(x, t) comprised of as many as (k - 1)j + 1 local operators. The overall weight of W(x, t) will generically grow under time evolution. For systems with spatial locality, W(x, t) will generically grow outwards in space as a "ball." As long as the subsystems containing W(x) and V(0) are connected through the interaction graph of the Hamiltonian, there will be terms in W(x,t) that have overlapping support with V(0). Thus, in non-relativistic systems W(x,t) and V(0) will generically cease to exactly commute for any t > 0, irrespective of their separation.

However, this does not mean the commutator can grow arbitrarily quickly! For instance,

 $<sup>^{25}\</sup>mathrm{See}$  §2.2.1 for a specific example in a concrete system.

for large-N gauge theories with  $O(N^2)$  degrees of freedom per thermal cell, the initial growth of C(x, t) is suppressed, taking the form

$$C(x,t) = \frac{C_0}{N^2} e^{\lambda_L (t - x/v_B)} + O(N^{-4}), \qquad (1.19)$$

with constants  $C_0$ ,  $\lambda_L$ , and  $v_B$ . Note that these constants can depend on the choices of operators W, V.<sup>26</sup>

 $C_0$  is uninteresting since it does not pertain to growth in space or time. In general, it is related to the norm of the operators, though in certain cases it can vanish and then the early-time expansion (1.19) will begin with the  $O(N^{-4})$  term. In this thesis, we will usually normalize the commutator so that C(x,t) = 1 at  $t = x/v_B + \lambda_L^{-1} \log N^2 = x/v_B + t_*$  and so that it has a late-time limit C(x,t) = 2.

 $\lambda_L$  has the interpretation of a "Lyapunov" exponent [45] since it is an exponent characterizing the rate of early-time exponential growth in a chaotic system.<sup>27</sup> In [58], Maldacena, Shenker, and Stanford showed that quantum mechanics imposes a bound

$$\lambda_L \le \frac{2\pi}{\beta}.\tag{1.20}$$

Holographic systems with near-horizon geometries described by Einstein gravity saturate this bound; they are the most chaotic systems allowed by nature.<sup>28</sup> Kitaev has proposed that the saturation of (1.20) is necessary (and possibly sufficient) for a system to have a holographic description with a bulk described by Einstein gravity [45] and then found such an example quantum system [73, 77] (see also [78]). However, Maldacena and Stanford suggest that such a condition is not sufficient [72]: although the proposed SYK model [73, 77, 78] saturates the bound on chaos (1.20), its bulk dual is not local in the usual sense.

$$C(x,t) = C_0 e^{\lambda_L (t - t_* - x/v_B)} + O(N^{-4}).$$
(1.21)

<sup>&</sup>lt;sup>26</sup>N.B. as an operator, the growth of [W(x,t), V(0)] can be controlled by the Lieb-Robinson bound [74–76]. However, since such an operator bound must be valid in all states, the resulting coefficients controlling the growth are UV-sensitive and therefore the bound is not very tight. In [51], we argue that by considering matrix elements in a state—such as C(x,t)—one can obtain a tighter state-dependent bound.

<sup>&</sup>lt;sup>27</sup>However, in [61] it is argued that the classical limit of  $\lambda_L$  is not the classical Lyapunov exponent. Hence the use of scare quotes.

<sup>&</sup>lt;sup>28</sup>The inverse of  $\lambda_L$  appears as the coefficient in (1.13). This can be seen by taking  $N^{-2}$  up into the exponential and noting  $S \sim N^2$ . Then, we can rewrite (1.19) as

This means that black holes are the fastest scramblers in two different senses: not only are they fast because scrambling scales with entropy as  $\sim \log S$ , they are also fast because they have the largest possible coefficient.



Figure 1-1: Signal propagation cones in a relativistic system. The causal light cone is defined by t = x, and for t < x the commutator C(x,t) must vanish by causality. For  $t < x/v_B + t_*$ , C(x,t) will be approximately zero. The butterfly cone is defined by  $t = x/v_B + t_*$ . C(x,t) = 1 on the butterfly cone, but the commutator will quickly obtain the value C(x,t) = 2 on the inside.

 $v_B$  is a velocity. It is the speed of the butterfly effect; the speed at which the operator W(x,t) grows as measured by V(0). This defines an effective "light" cone or "butterfly" cone for non-relativistic systems; even in such systems there's still a limit on how fast signals can propagate. For relativistic systems we must have  $v_B \leq 1$  by causality, and there are many examples of systems that have a nontrivial  $v_B < 1$ .

In a collaboration with Swingle that follows the work in this thesis, we show that if the commutators between operators at spatially separated regions have a magnitude no greater than  $\epsilon$ , then the classical channel capacity between these regions must be  $O(\epsilon)$  [51]. This answers our second question about how fast information can propagate in a quantum system:  $v_B$  can be thought of as a signaling speed because it sets the speed that small perturbations can affect distant regions, see Fig. 1-1.

## Informative outline

This thesis comprises four previously-published papers [1–4].

In §1, we provide an introduction. You are (mostly) already aware of what it contains.

In §2 based on [1], we consider the growth of operators in strongly-coupled systems. Following [44, 54] (which papers did not focus on localized operators), we use spin-chain numerics and holography to investigate the speed at which local operators grow under chaotic dynamics. A central object of study is the speed of operator growth, the "butterfly" velocity  $v_B$ . Additionally, we show that for a large class of states that the geometry of the black hole interior matches the geometry of the minimal tensor network (or quantum circuit) that builds the state. This suggests a strong connection between tensor-networks/complexity and the geometry of the black hole interior.

In §3 based on [2], we present a conjecture that the computational complexity of a holographic state is dual to the spacetime action of the black hole interior. This improves on a previous conjecture relating complexity and black hole geometry [79]. Evidence for this conjecture can be found in the previous chapter §2 and in the supplementary article [49]. As a corollary, we conjecture that black holes are the fastest computers in nature (in the sense of the complexity growth of the state).

In §4 based on [3], we study chaos in two-dimensional conformal field theory. We emphasize a particular diagnostic of many-body chaos simply related to C(x,t), the outof-time-order (OTO) four-point correlation function  $\langle W(x,t) V(0) W(x,t) V(0) \rangle_{\beta}$ . With this, we provide model calculations in both chaotic and integrable theories: a large-*c* CFT and the two-dimensional Ising model. The large-*c* computation provides a purely CFT understanding of the holographic butterfly effect studied in [44].

Finally, in §5 based on [4], we attempt to study chaos in information-theoretic terms. We discuss the idea of *scrambling* [38, 39] and show that chaos (as measured by the diagnostic presented in Chapter 4) implies scrambling. To do so, we study entanglement in "time" and show that unitary time evolution that leads to the creation of multipartite entanglement will scramble localized information. The mechanism for this is many-body chaos. Finally, we conclude as we hypothesize a relationship between chaos and complexification in terms of the ideas studied in this thesis.

## Chapter 2

# HIC SVNT DRACONES

## (Localized shocks)

The initial investigations of chaos in holography by Shenker and Stanford [44, 54] mostly focused on holographic perturbations with spherical symmetric. These were effectively studies of chaos within a single thermal cell in the boundary field theory. A cell has  $O(N^2)$  "matrix" degrees of freedom that are all coupled together, though each term in the Hamiltonian only involves O(1) degrees of freedom. The system has no notion of spatial locality, and the time it takes for the entanglement in the system to be disrupted—the fast scrambling time  $t_*$ —was found to be

$$t_* = \frac{\beta}{2\pi} \log N^2, \tag{2.1}$$

making precise and providing evidence for an earlier conjecture about black holes of Sekino and Susskind [39].

Here, we will focus on holographic systems whose boundary consists of many thermal cells. Each thermal cell still has  $O(N^2)$  of freedom, but perturbations will be localized to be in a particular thermal cell. This lets us study the spatial growth of operators: the speed of the butterfly effect. The butterfly "velocity"  $v_B$  is a central object of study in its own right [51, 57, 69, 71]. In a modern framing [51], we can think of the butterfly velocity as a state-dependent *effective* Lieb-Robinson velocity [74–76], which is a mathematically precise bound on the rate of growth of commutators of local operators separated in time and space.

The initial impetus of this chapter was the question of whether  $v_B$  is universal. In [44],

the butterfly velocity was discovered and computed for Einstein gravity to be

$$v_B = \sqrt{\frac{d}{2(d-1)}},\tag{2.2}$$

where d is the spacetime dimension of the boundary field theory. In an Appendix of this chapter §2.5.2, we show that it is modified from this value in Gauss-Bonnet gravity. In [51, 71], it was later shown that for hyperscaling violating geometries  $v_B$  can depend nontrivially on thermodynamic quantities such as the temperature and critical exponents.

A main focus of this chapter is the study of how localized operators grow in time. The growth is ballistic, with velocity  $v_B$ , and in chaotic theories the operator grows outward like a "ball." (By contrast, in integrable theories operators will expand like "shells.") This pattern of growth is understood to arise from the locality of the Hamiltonian. In [51], it is further shown that the growth of commutators also controls the rate of signaling;  $v_B$  is the speed at which information travels in strongly-interacting systems.

The second focus of this chapter is on the relationship between chaotic growth of operators on the boundary and the geometry of the Einstein-Rosen (ER) bridge or wormhole in the bulk. This is related to ideas that suggest quantum entanglement—EPR [80]—and the connectivity of asymptotic spacetime regions—ER [81]—are holographically dual [82] (see also [83])<sup>1</sup>

$$ER = EPR. \tag{2.3}$$

Building on ideas of Swingle [84] and Hartman and Maldacena [85, 86], Stanford and Susskind conjectured that the volume of a spatial slice through the Einstein-Rosen bridge is holographically dual to the computational complexity of the field theory state [79]. Here we provide evidence that, more than just the scalar volume, the geometry of the black hole interior is in detailed correspondence with the minimal tensor network geometry that would construct the boundary state.

The understanding of the black hole interior is a major open problem in quantum gravity, e.g. see Fig. 2-1. The puzzle of the interior was recently made sharp in terms of an argument involving entanglement, leading to the conclusion that the interior does not exist [34, 53]. However, as emphasized by Susskind, "Entanglement is not enough" [87].

<sup>&</sup>lt;sup>1</sup>It used to have to be pointed out that one shouldn't confuse these very different phenomenon. It's not clear what one should point out now.


Figure 2-1: Here be dragons.<sup>2</sup>

**\*\*\*** 

This chapter was first presented as [1] under the title "Localized shocks," and is a collaboration with Douglas Stanford and Leonard Susskind,

JHEP 03 (2015) 051, arXiv:1409.8180 [hep-th].

## Abstract

We study products of precursors of spatially local operators,  $W_{x_n}(t_n) \dots W_{x_1}(t_1)$ , where  $W_x(t) = e^{-iHt} W_x e^{iHt}$ . Using chaotic spin-chain numerics and gauge/gravity duality, we

<sup>&</sup>lt;sup>2</sup>Adapted from the *Carta Marina* drawn by Olaus Magnus. I thank Ethan Dyer for permission to reuse his joke. He reminds me that the main debate is whether the dragon lives in the interior or at the singularity. It's also possible there is no interior [34, 53], in which case we'd have to draw a very different map.

show that a single precursor fills a spatial region that grows linearly in t. In a lattice system, products of such operators can be represented using tensor networks. In gauge/gravity duality, they are related to Einstein-Rosen bridges supported by localized shock waves. We find a geometrical correspondence between these two descriptions, generalizing earlier work in the spatially homogeneous case.

# 2.1 Introduction

How is the region behind the horizon of a large AdS black hole described in the dual gauge theory? A number of probes of the interior have been proposed. These include two-sided correlation functions [88–90], mutual information and entropy [85, 91–93], and the pullback-pushforward [94, 95] modification of the standard smearing function procedure [96–100]. These probes work well for young black holes, but for older black holes, or for black holes in a typical state,<sup>3</sup> they do not appear to be helpful. One way of stating the problem is as follows: the gauge theory scrambles and settles down to static equilibrium in a short time, during which two sided correlations decay and mutual information saturates. On the other hand, an appropriately defined interior geometry of the black hole continues to grow for a much longer time [86]. What types of gauge theory variables can describe this growth?

Two possible answers have been suggested. Hartman and Maldacena [85] studied the time evolution of the thermofield double state and pointed out the relationship between the growth of the interior and the growth of a tensor network (TN) description of the state. Building on this work and ideas of Swingle [84], Maldacena [86] suggested that the interior could be understood as a refined type of tensor network describing the state of the dual gauge theory. According to this picture, the overall length of the interior is proportional to size of the minimal tensor network that can represent the state.

A second suggestion focused on the evolution of the quantum state as modeled by a quantum circuit (QC). It was conjectured [102] that the length of the black hole interior at a given time is proportional to the computational complexity of the state at the same time. The computational complexity is the size of the minimial quantum circuit that can generate the state, and it is expected to increase linearly for a long time  $\sim e^S$ . In [79], two

<sup>&</sup>lt;sup>3</sup>Arguments have been made that black holes in a typical state do not have an interior geometry [34, 53, 101]. We will not address typical states in this paper.

of us checked a refinement of this relationship for a family of states, corresponding to the spherical shock wave geometries constructed in [54].

There are strong reasons to think that the tensor network and quantum circuit descriptions are essentially the same thing. A QC is a special case of a TN. It has some special features, such as time-translation invariance and unitarity of the gates, which are not shared by the most general TN. But as noted in [85], these features are also necessary for a TN to be able to describe the black hole interior. Thus we will assume that that the TN and QC representations are the same.

In this paper, we will continue exploring the relationship between tensor network geometry and Einstein geometry. We will work in the setting of two-sided black holes. Our hypothesis, following [86], is that the geometry of the minimal tensor network describing the entangled state is a coarse-graining (on scale  $\ell_{AdS}$ ) of the Einstein-Rosen bridge connecting the two sides. We will consider TN and Einstein geometry associated to products of localized precursor operators, each of the form

$$W_x(t_w) = e^{-iHt_w} W_x e^{iHt_w}.$$
 (2.4)

For  $t_w = 0$  this precursor is simply  $W_x$ , an operator local on the thermal scale. But as  $t_w$  advances, it becomes increasingly nonlocal. In a lattice system, such an operator can be represented in terms of tensor networks, and we will argue on general grounds that the characteristic TN geometry of a single precursor consists of two solid cones, glued together along their slanted faces. General products of precursor operators

$$W_{x_n}(t_n)...W_{x_1}(t_1)$$
 (2.5)

can also be represented in terms of tensor networks, with geometries that we will characterize in  $\S$  2.2.

Spatially homogeneous precursors were analyzed using gauge/gravity duality in [44, 54, 55, 79]. Their action on a thermal state corresponds to adding a small amount of energy to an AdS black hole. As  $t_w$  is increases, the stress energy is boosted and a gravitational shock wave is produced. Products of precursors create an intersecting network of shock waves behind the horizon [54]. In § 2.3, we will extend this analysis to the case of spatially localized precursor operators. We will study the spatial geometry of the two-sided black

hole dual to  $W_x(t_w)$ , and find that it has the same "glued cone" geometry we inferred for the TN. More generally, we will see that the ERB geometry dual to multiple perturbations, local at different times and positions, agrees with the expected structure of the TN. More specifically, it agrees on scales large compared to  $\ell_{AdS}$ . This generalizes [79] and provides a wide range of examples relating TN and Einstein geometry [86].

A central object in our analysis will be the size and shape of a precursor operator. In the spin chain and holographic systems that we study, precursors become space-filling, covering a region that increases outwards ballistically with respect to the time variable  $t_w$ . This behavior can be diagnosed using the thermal trace of the square of the commutator,

$$C(t_w, |x - y|) = \operatorname{tr}\left\{\rho(\beta)[W_x(t_w), W_y]^{\dagger}[W_x(t_w), W_y]\right\}$$
(2.6)

where  $W_x(t_w)$  is the precursor, and  $W_y$  is a local operator at point y. For simplicity, let us consider unitary operators, so that the maximum of this quantity is two. We will define the size  $s[W_x(t_w)]$  as the (d-1) dimensional volume of the region in y such that C is greater than or equal to one.<sup>4</sup> In the examples that we consider, this region consists of a ball centered at location x. We will define the radius of the operator  $r[W_x(t_w)]$  as the radius of this ball.<sup>5</sup>

We will see that the radius increases linearly with  $t_w$ . In the spin chain system, we will check this numerically. In the large N holographic system, we will use the geometry of the localized shock wave to determine

$$r[W_x(t_w)] \approx v_B(t_w - t_*). \tag{2.7}$$

Here  $t_* = \frac{\beta}{2\pi} \log N^2$  is the scrambling time, and the "butterfly effect speed"  $v_B$  is  $\sqrt{\frac{d}{2(d-1)}}$ [44], where *d* is the spacetime dimension of the boundary theory. This expression is negative for times  $t_w < t_*$ , indicating that the *W* perturbation has not had an order one effect on any local subsystem. However, after the scrambling time, the precursor grows outwards at

<sup>&</sup>lt;sup>4</sup>It would be more precise to optimize over all operators  $W_y$  at location y. However, for a suitably chaotic system, this step is not necessary: the butterfly effect suggests that any operator will do.

<sup>&</sup>lt;sup>5</sup>It is important to distinguish growth from movement: the growing operator is not a superposition of operators at different locations. If this were the case, the commutator C might be nonzero in a large region, but it would be numerically small. The fact that the commutator is order one indicates that the operator  $W_x(t_w)$  is sum of complicated products, each summand including nontrivial operators at most sites within the region of size  $s[W_x(t_w)]$ . This is the sense in which the operator is space-filling.

speed  $v_B$ ; this should be understood as the spread of the butterfly effect.

### 2.1.1 Some terminology

For the convenience of the reader we will list some terminology used in this paper.

- *d* refers to the space-time dimension of the boundary theory.
- $t_*$  is the scrambling time  $\frac{\beta}{2\pi} \log N^2$ , where N is the rank of the dual gauge theory.
- A precursor  $W(t_w)$  of an operator W is given by  $e^{-iHt_w}We^{iHt_w}$ . A localized precursor  $W_x(t_w)$  is a precursor of an approximately local operator  $W_x$ .
- A localized precursor  $W_x(t_w)$  is associated to a region of influence, i.e. the region in which local operators have an order-one commutator with  $W_x(t_w)$ .
- The (d 1)-volume of this region defines the size s[W<sub>x</sub>(t<sub>w</sub>)]. In a qubit model, the size indicates the number of qubits affected at t = 0 by the action of a single qubit W<sub>x</sub> a time t<sub>w</sub> in the past.
- The radius of the affected region is  $r[W_x(t_w)]$ . Size and radius are related: size is the (d-1)-volume of a ball of radius r. We will see  $r \approx v_B(t_w t_*)$ , where  $v_B$  defines the speed at which the precursor grows.
- $\Sigma_{max}$  is the spatial slice of maximal *d*-volume that passes through the ERB, from time t = 0 on the left boundary, to t = 0 on the right.  $\Sigma_{dec}$  is similarly defined, but it maximizes a functional obtained from the volume by dropping transverse gradients.

## 2.2 Qubit systems

Black holes of radius  $\ell_{AdS}$  have entropy of order  $N^2$  where N is the rank of the dual gauge group. This is the number of degrees of freedom of a single cutoff cell of the regulated gauge theory. We can think of such black holes as units out of which larger black holes are built. Technically this means that the gauge theory can be represented as a lattice of cells, and that in order to represent the thermal state at temperature T, the coordinate size of a cell should be no bigger than  $T^{-1}$ . A perturbation applied to a degree of freedom on a specific site will evolve in two ways: it will spread out through the  $N \times N$  matrix system through fast-scrambling dynamics, and it will spread out through the lattice. In previous studies of unit black holes, the focus was on operator growth in the fast-scrambling dynamics of a single cell. This paper is mostly about the complementary mechanisms of spatial growth on scales larger than  $T^{-1}$ . We can see the important phenomena by studying spatial lattices of low dimensional objects; namely qubits.

In this section, we will study precursors in a simple qubit system. In section 2.2.1, we will numerically simulate the system, and observe linear growth of the precursor as a function of time  $t_w$ . In section 2.2.2, we will use this pattern of growth to qualitatively analyze the geometry of the minimal tensor network for a product of precursor operators.

#### 2.2.1 Precursor growth

The spin Hamiltonian we will use is an Ising system defined on a one-dimensional chain

$$H = -\sum_{i} Z_{i} Z_{i+1} + g X_{i} + h Z_{i}, \qquad (2.8)$$

where  $X_i, Y_i, Z_i$  are the Pauli operators on the *i*th site, i = 1, 2, ..., n. In our numerics, we use n = 8. We will consider two choices for the couplings: one for which the system is strongly chaotic (g = -1.05, h = 0.5) [103], and one for which it is integrable (g = 1, h = 0).

We will study the size of the precursor associated to  $Z_1$ :

$$Z_1(t_w) = e^{-iHt_w} Z_1 e^{iHt_w}.$$
(2.9)

In this setting, where the operator begins at the endpoint of a one dimensional chain, size and radius are interchangeable. As a function of  $t_w$ , we define either as the number of sites *i* such that  $tr[Z_1(t_w), A_i]^2$  is greater than or equal to one, with A = X, *Y*, or *Z*. The rate at which the operator grows can be controlled using the Lieb-Robinson bound for the commutator of local operators [74–76]. This bound states that

$$\|[W_x(t_w), W_y]\| \le c_0 \|W_x\| \|W_y\| e^{c_1 t_w - c_2 |x-y|},$$
(2.10)

where  $c_0, c_1, c_2$  are constants that depend on the Hamiltonian. The norm is the operator (infinity) norm, and the bound is valid as long as the interactions decay exponentially (or faster) with distance. This bound implies that the radius of the operator can grow no faster than linearly,  $r[Z_1(t_w)] < (c_1/c_2)t_w$ .

It is not hard to see that some systems can saturate this linear behavior. A rather trivial example is the spin chain (2.8), with g = 1 and h = 0. This system is integrable for all g and can be solved by mapping to a system of free spinless Majorana fermions via the nonlocal Jordan-Wigner transformation (see e.g. [104]). This nonlocal mapping relates  $a_k$ , the fermion annihilation operator at site k, to spin operators of the form  $X_1X_2...X_{k-1}Z_k$ and  $X_1X_2...X_{k-1}Y_k$ . The fact that the free fermions propagate linearly in time corresponds, in the spin variables, to a linear growth of the operator.<sup>6</sup>

This integrable system is clearly very special: even typically diffusive quantities such as energy density move ballistically. One might naively guess that the linear precursor growth is also exceptional, and that a chaotic system will exhibit slower (perhaps diffusive) growth. To show that this is not the case, we numerically analyze a chaotic version (g = -1.05, h = 0.5) of the same spin chain alongside the integrable model, and plot the size of the precursor in the right panel of Fig. 2-2. The size, according to the commutator definition, corresponds to the staircase plots. The rate of growth is clearly linear for both the chaotic (solid blue) and integrable (dashed blue) curves. The only significant difference occurs once the operator grows to the size of the entire chain—in the chaotic system it saturates and in the integrable system it begins to shrink.

The commutator is a useful measure of the size of the operator, but having an explicit numerical representation allows us to understand the growth in other ways. It is helpful to think about expanding  $Z_1(t_w)$  in a basis of Pauli strings, e.g.  $X_1Z_2Y_3I_4X_5X_6Y_7Z_8Y_9...$ Starting from the simple operator  $Z_1$ , such strings are generated by the Baker-Campbell-Hausdorff formula

$$Z_1(t_w) = Z_1 - it_w \ [H, Z_1] - \frac{t_w^2}{2!} \ [H, [H, Z_1]] + \frac{it_w^3}{3!} \ [H, [H, [H, Z_1]]] + \dots$$
(2.11)

<sup>&</sup>lt;sup>6</sup>This is an interesting case where growth and movement are related, despite footnote 5. The reason this is possible is that the change of variables is nonlocal.



Figure 2-2: Ballistic growth of the operator  $Z_1(t_w)$ , evolved with the chaotic g = -1.05, h = 0.5 Hamiltonian (solid) and the integrable g = 1, h = 0 Hamiltonian (dotted). Left:  $c_k(t)$  is the sum of the squares of the coefficients of Pauli strings of length k in  $Z_1(t_w)$ . Notice that the integrable and chaotic behavior is rather similar until the strings grow to reach the end of the chain (n = 8 spins). Right: for both types of evolution, the size grows linearly until it approaches the size of the system. After this point, the chaotically-evolving operator saturates, while the integrably-evolving operator begins to shrink. The blue "staircase" curves show the size  $s[Z_1(t_w)]$ . The smooth black curves show  $s_2[Z_1(t_w)] \propto \sum_k k c_k(t_w)$ .

For example, suppressing coefficients and site indices, one finds the sixth order term

$$[H, [H, [H, [H, [H, Z]]]]] = X, Z, XX, XZ, YY, ZX, ZZ, IX, XXX, XXZ, XYY, XZZ, YXY, YYZ, ZXZ, XXXZ.$$
 (2.12)

As  $t_w$  increases, high order terms in the BCH expansion become important, corresponding to longer and more complicated Pauli strings.

To quantify this growth, we will group together Pauli strings according to their length. We define the length of a Pauli string as the highest site index that is associated to a nonidentity Pauli operator. For example, the length of  $X_1I_2Y_3$  is three. Given a Pauli string representation of an operator, we define  $c_k$  as the sum of the squares of coefficients of all Pauli strings of length k. We plot the  $\{c_k\}$ , as a function of  $t_w$ , in the left panel of Fig. 2-2. Again, the chaotic and integrable systems track each other well until the operator becomes roughly as large as the entire system. Using these coefficients, we can define a smoother notion of size by taking

$$s_2[Z_1(t_w)] = \frac{\sum_k k \ c_k(t_w)}{\sum_k c_k}.$$
(2.13)

This is plotted as the black curves in the right panel of Fig. 2-2. It agrees fairly well with



Figure 2-3: The tensor network description of the identity operator (left) and operators  $e^{-iHt}$  with successively larger t.

the staircase definition using the commutator. The initial delay in the growth is due to the fact that  $Z_1$  needs to be converted to  $Y_1$  before the strings can grow in length. This can be thought of as the time to "scramble" a single site.

## 2.2.2 TN for multiple localized precursors

In this section, we will assume that precursor operators grow at a rate  $v_B$ . That is,  $r[W_x(t_w)] = v_B t_w$ . We will use this pattern of growth to characterize the tensor networks associated to a product of localized precursors. To begin, let us review the TN geometry associated to the time evolution operator  $e^{-iHt}$ . This is illustrated in Fig. 2-3 for a quantum system of eleven sites. Each diagram represents a formula for  $[e^{-iHt}]_{mn}$ , with different values of t. The line endpoints on the left represent the tensor decomposition of the m index into eleven indices  $m_1, ..., m_{11}$ , and the line endpoints on the right represent the n index as  $n_1, ..., n_{11}$ . Line segments correspond to Kronecker delta contractions. Thus the figure at the far left is a formula for the identity matrix:  $I_{mn} = \delta_{m_1n_1} \delta_{m_2n_2} ... \delta_{m_{11}n_{11}}$ .

Intersections of lines represent a tensor with rank equal to the number of lines. In the figure, we have only four-fold intersections, corresponding to tensors  $t_{i_1i_2i_3i_4}$ . The grids of intersecting lines represent a particular contraction of a large number of these tensors. It is known that time evolution by a local Hamiltonian can be represented in terms of such networks, using a technique known as Trotterization.

We have presented the TN as giving a formula for an operator, but we can also think about it as a formula for an entangled pure state of two quantum systems L and R. In this interpretation, the left ends correspond to indices in the L system, and the right ends correspond to indices in the R system. Contracting the tensors gives the wave function.



Figure 2-4: Left: the naive tensor network describing a precursor operator  $e^{-iHt_w}We^{iHt_w}$ . The red network represents backwards time evolution, the black dot represents the local operator W, and the green network represents forwards time evolution. Shading indicates the region affected by the linearly growing W insertion. In the unshaded region, the forwards and backwards evolutions cancel. **Right:** the network after removing tensors that cancel. The dotted lines indicate contractions; their endpoints should be identified.

This is an example of a general correspondence: an operator  $A_{mn}$  can be understood as the wave function for a state given by acting with the operator A on one side of a maximally entangled state:  $|A\rangle = A \sum_{i} |i\rangle |i\rangle = \sum_{mn} A_{mn} |m\rangle |n\rangle.$ 

Thought about either way, the basic feature of this network is that it grows linearly with time. The work of [85] pointed out the relationship between this linear growth and the geometry of the Einstein-Rosen bridge of the time-evolved thermofield double state  $e^{-iHt}|TFD\rangle$ .

#### One precursor

Let us now understand the tensor network for a single precursor operator  $W_x(t_w)$ . The structure is illustrated in Fig. 2-4. Concatenation of tensor networks represents multiplication of operators, and in the left panel, we have a naive TN for the operator, in which we simply concatenate the networks for  $e^{-iHt_w}$ ,  $W_x$ , and  $e^{iHt_w}$ . The tensor network for the Woperator is the identity on all sites except for the central one, represented by the dot. This naive tensor network is an explicit representation of a time-fold [95, 105], where we evolve backwards in time, insert the operator, then evolve forwards again. As in [79], we can simplify the network by considering the partial cancellation between  $e^{-iHt_w}$  and  $e^{iHt_w}$ . If we had not inserted the  $W_x$  operator, the cancellation would be complete. This means that we



Figure 2-5: Fibering the position-dependent time fold over the x space gives the minimalized TN. The geometry at right is equivalent to the right panel of Fig. 2-4.

can remove the tensors in the region that is not affected by the linearly growing precursor. After doing so, we obtain the simpler network shown in the right panel of Fig. 2-4.<sup>7</sup>

This "minimalized" tensor network can be understood in terms of a position-dependent time-fold. Because the region of influence of the operator grows outwards at a rate  $v_B$ , at distance |x| from the insertion of the operator, we only need to include a time fold of length  $t_w - |x|/v_B$ ; the rest of the fold cancels. The minimalized TN geometry in Fig. 2-4 can be constructed as the fibration of this position-dependent fold over the x space, as shown in Fig. 2-5. This procedure extends to higher dimensions, where one finds a geometry consisting of two solid cones, glued together along their slanted faces.

#### Multiple precursors

Now, consider a product of n localized precursor operators<sup>8</sup>

$$W_{x_n}(t_n)\dots W_{x_1}(t_1) = e^{-iHt_n} W_{x_n} e^{iHt_n} \dots e^{-iHt_1} W_{x_1} e^{iHt_1}.$$
(2.14)

As before, we can form a naive TN by simply concatenating the networks for each of the Wand  $e^{\pm iHt}$  operators on the RHS. To form the minimalized TN, we cancel adjacent regions of forward and backward evolution outside the influence of any insertion. This procedure is explicit, but it becomes complicated as the number of operators increases, as their insertion positions are varied, and for systems with spatial dimension larger than one. Even with

<sup>&</sup>lt;sup>7</sup>D.S. is grateful to Don Marolf for pointing out an error in a previous version of this argument.

<sup>&</sup>lt;sup>8</sup>We emphasize that the times  $t_i$  are not necessarily in time order. In fact, the most interesting case is the one in which the differences between adjacent times alternate in sign.



Figure 2-6: The minimal tensor network describing a product of three precursor operators  $\mathcal{O} = W(t_3)W(t_2)W(t_1)$  inserted at the same spatial point and with  $0 < -t_2 < t_3 < t_1$ .



Figure 2-7: Slices of the x-dependent time-fold for the product of three local precursor operators inserted at x = 0 with  $0 < -t_2 < t_3 < t_1$ . As |x| increases, the folds are pulled inwards. Eventually the second and third folds merge, leaving a single fold.

three operators, the geometry can be rather nontrivial, as shown in Fig. 2-6.

In order to compare with holography, it will be useful to emphasize the representation of the minimalized TN using position-dependent time-folds. Let us begin with the case where all operators are localized at x = 0. We define the position-dependent time fold as follows. At x = 0, we simply have the folded time axis defined by the times  $t_1, ..., t_n$ , as in the spherically symmetric case studied in [79]. Depending on the configuration of times, this will have  $k \leq n$  folds, and (n - k) through-going insertions (see [79]). As |x| increases, each fold gets pulled inwards, by an amount  $|x|/v_B$ ; that is, the time location of the fold associated to operator j becomes  $t_j \pm |x|/v_B$ , with a sign that depends on the direction of the fold. This reflects the fact that as we move farther from the insertion, there is additional cancellation between forwards and backwards time evolution. At certain special values of |x|, pairs of folds will merge and annihilate, leaving behind two fewer folds (Fig. 2-7). This procedure defines a folded time axis as a function of x. Fibering this geometry over the x space gives the geometry of the minimalized tensor network.

In the case where the operators are localized at general positions  $\{x_j\}$ , the procedure is slightly more complicated. To define the position-dependent fold at location x, we begin with the position-independent time fold, passing through insertions at times  $\{t_j\}$ . We then replace each  $t_j$  by a new variable  $\tau_j$ , and we minimize the length of the folded time axis subject to two constraints: first, the time fold must pass through each  $\tau_i$  in order; second, the  $\{\tau_i\}$  must satisfy  $|\tau_j - t_j| \leq |x - x_j|/v_B$ . The minimalized TN is the fibration of this time fold over the x space. In general, it consists of flat regions glued together at the loci of the folds.

## 2.3 Holographic systems

In this section, we will use holography to study the dynamics of precursors. The analysis will be highly geometric: the action of a precursor on the thermofield double state generates a gravitational shock wave that distorts the Einstein-Rosen bridge, as in [44]. By analyzing the transverse profile of this shock wave, in section 2.3.2, we will be able to estimate the commutator with other operators, and thus the size of the precursor as a function of time. In section 2.3.3, we will study the geometry dual to a product of local precursor operators. We will find a detailed match with the geometry of the corresponding tensor networks, characterized in section 2.2.2.

## 2.3.1 Localized shock waves

Let us start by reviewing the AdS black hole dual to the thermofield double state  $|TFD\rangle$  of two CFTs L and R [106].<sup>9</sup> Planar AdS black holes have only one scale, namely  $\ell_{AdS}$ , and

<sup>&</sup>lt;sup>9</sup>We should clarify the role of the thermofield double state and the background black hole geometry. Why couldn't we study the same problem with vacuum AdS as the background? The answer is that matrix elements of the precursor, and its commutators with other operators, depend on the energy. Near the vacuum, the dynamics is integrable. The relevant commutators might be large as operators, but they have small expectation value in the vacuum state (this can be seen explicitly in the spin model numerics). The black hole geometry allows us to study nontrivial matrix elements.

we can write the metric in terms of dimensionless coordinates as

$$ds^{2} = \ell_{AdS}^{2} \left[ -f(r)dt^{2} + f^{-1}(r)dr^{2} + r^{2}dx^{i}dx^{i} \right]$$
(2.15)

$$f(r) = r^2 - r^{2-d}, (2.16)$$

where *i* runs over (d-1) transverse directions. It will convenient to use the smooth Kruskal coordinates, *u* and *v*, which are defined in terms of *r* and *t* by

$$uv = -e^{f'(1)r_*(r)}, \qquad u/v = -e^{-f'(1)t},$$
(2.17)

where the tortoise coordinate is  $r_*(r) = \int_{\infty}^r dr' f^{-1}(r')$ . In terms of these Kruskal coordinates, the black hole metric can be written

$$ds^{2} = \ell_{AdS}^{2} \left[ -A(uv)dudv + B(uv)dx^{i}dx^{i} \right]$$

$$(2.18)$$

$$A(uv) = -\frac{4}{uv} \frac{f(r)}{f'(1)^2} \qquad B(uv) = r^2.$$
(2.19)

The horizon of the black hole is at r = 1, or uv = 0. The inverse temperature is  $\beta = 4\pi/d$ .

Following [44], we will act on the thermofield double state with a precursor of an operator in the L CFT:

$$W_x(t_w)|TFD\rangle = e^{-iH_L t_w} W_x e^{iH_L t_w} |TFD\rangle, \qquad (2.20)$$

where the  $W_x$  operator is an approximately local, thermal scale operator acting near location x on the left boundary. We would like to understand the geometry dual to this state.<sup>10</sup> If the time  $t_w$  is not large, then the geometry will not be substantially affected by the perturbation. However, Schwarzschild time evolution acts near the horizon as a boost, and as we make the Killing time  $t_w$  of the peturbation earlier, its energy in the t = 0 frame gets boosted  $\propto e^{2\pi t_w/\beta}$ . When this exceeds the  $G_N$  suppression from the gravitational coupling,  $t_w \sim t_* = \frac{\beta}{2\pi} \log N^2$ , the geometry will be affected.<sup>11</sup>

If  $G_N$  is small, then the boost must be large in order to overcome the suppression. The

<sup>&</sup>lt;sup>10</sup>In order to have a well defined notion of geometry, we will should consider a coherent operator W built out of a large but fixed number (e.g. 100) of quanta. However an operator corresponding to even a single quantum will lead to similar effects on the commutator that we estimate below; the metric should be understood in the single-particle case as giving the eikonal phase [107].

<sup>&</sup>lt;sup>11</sup>This potential important of this time scale was first noticed by [108]. The connection to scrambling was made in [38], and the connection to gauge/gravity duality was made in [39]. This connection was made precise in [44].

associated stress energy distribution is highly compressed in the u direction and stretched in the v direction. We can replace this by a stress tensor localized on the u = 0 horizon,

$$T_{uu} = \frac{E}{\ell_{AdS}^{d+1}} e^{2\pi t_w/\beta} \delta(u) a_0(x),$$
(2.21)

where E is the dimensionless asymptotic energy of the perturbation, and  $a_0$  is a function concentrated within  $|x| \leq 1$ , with integral of order one. The precise form of  $a_0$  depends on details of the perturbation, as well as the propagation to the horizon.



Figure 2-8: A constant x slice through the one-shock geometry. The geometry consists of two halves of the eternal AdS black hole, glued together with a null shift of magnitude h(x) in the v direction. The shock lies along the surface u = 0. In the coordinates (2.22), the surface v = 0 is discontinuous.

The backreaction of this matter distribution is extremely simple. It was worked out first in [109] for the case of a flat space Schwarzschild black hole, and by [44, 110] for AdS black holes. The idea is to consider a shock wave ansatz of the form<sup>12</sup>

$$ds^{2} = \ell_{AdS}^{2} \Big[ -A(uv) dudv + B(uv) dx^{i} dx^{i} + A(uv) \delta(u) h(x) du^{2} \Big].$$
(2.22)

This metric can be understood as two halves of the AdS black hole, glued together along u = 0 with a shift of magnitude h(x) in the v direction (see Fig. 2-8). By evaluating the curvature of this metric (setting  $u\delta'(u) = -\delta(u)$  and  $u^2\delta(u)^2 = 0$ ) and plugging into Einstein's equations, we find a solution if

$$\left(-\partial_i \partial_i + \mu^2\right) h(x) = \frac{16\pi G_N}{A(0)\ell_{AdS}^{d-1}} E e^{\frac{2\pi}{\beta}t_w} a_0(x), \qquad (2.23)$$

where  $\mu^2 = \frac{d(d-1)}{2}$ . For  $|x| \gg 1$ , the solution to this equation will depend only on the integral of  $a_0(x)$ , which we can replace with a delta function. The differential operator in

<sup>&</sup>lt;sup>12</sup>For finite  $t_w$ , the metric will not be exactly of this type. Corrections are analyzed in appendix 2.5.1.

Eq. (2.23) can then be inverted in terms of Bessel functions. Expanding for large |x|, and assuming a thermal-scale initial energy E, we find the solution

$$h(x) = \frac{e^{\frac{2\pi}{\beta}(t_w - t_*) - \mu|x|}}{|x|^{\frac{d-2}{2}}} , \qquad (2.24)$$

where the scrambling time  $t_* = \frac{\beta}{2\pi} \log \frac{c\ell_{AdS}^{d-1}}{G_N} \approx \frac{\beta}{2\pi} \log N^2$  has been defined with *c* chosen to absorb certain order-one constants. For  $|x| \leq 1$ , the approximation of  $a_0$  by a delta function is incorrect; the power-law singularity in the denominator should be smoothed out.

The strength of the shock wave is exponentially growing as a function of  $t_w$ , reflecting the growing boost of the initial perturbation. However, it is exponentially suppressed as a function of x. In the next section, we will use the interplay of these exponentials to determine the growth of the precursor operator  $W_x(t_w)$ .

## 2.3.2 Precursor growth

To measure the size of the precusor, let us consider the expectation value of the commutatorsquared

$$C(t_w, |x - y|) = \operatorname{tr}\left\{\rho(\beta)[W_x(t_w), W_y]^{\dagger}[W_x(t_w), W_y]\right\},$$
(2.25)

where  $\rho(\beta)$  is the thermal density matrix. To simplify the analysis, we will consider the setting in which (i)  $W_x$  and  $W_y$  are both unitary operators, so that  $0 \leq C \leq 2$ , and (ii) they correspond to different bulk fields. As before, we will define the radius of the operator at time  $t_w$  as the maximum distance |x - y| such that  $C(t_w, |x - y|)$  is equal to one.

To calculate the above expectation value we will follow the procedure used by Ref. [54] in the spherical shock case. Similar calculations were also previously done by [111, 112]. Using the purification by the thermofield state, we can write

$$C(t_w, |x - y|) = \langle TFD | [W_x(t_w), W_y]^{\dagger} [W_x(t_w), W_y] | TFD \rangle$$

$$(2.26)$$

$$= 2 - 2Re\langle \psi | \psi' \rangle, \tag{2.27}$$

where  $|\psi\rangle = W_x(t_w)W_y|TFD\rangle$  and  $|\psi'\rangle = W_yW_x(t_w)|TFD\rangle$ . It is helpful to think about this inner product in the t = 0 frame, in which the  $W_x$  perturbation is highly boosted, and the  $W_y$  perturbation is not. Let us suppose that  $t_w$  is not too large, so the relative boost is a small fraction of the Planck scale. Then the  $W_x$  operator creates a mild shock wave, and the  $W_y$  operator creates a field theory disturbance that propagates on this background. The difference between applying the  $W_y$  operator before or after the  $W_x$  operator is a null shift in the v direction of magnitude h(y-x) [54]. The inner product of the states  $|\psi\rangle$ ,  $|\psi'\rangle$ then reduces to the following question: if we take the field theory state corresponding to the  $W_y$  perturbation and shift it by h(x-y) in the v direction, what is its overlap with the original unshifted state?

If the strength of the shock is sufficiently weak, the shift is small and the overlap is close to one; we recover a small commutator. However, once  $h(x - y) \sim 1$  the shift becomes of order the typical wavelength in the field theory excitation created by  $W_y$ , and the inner product begins to decrease. Because the strength of the shock is exponential in  $t_w$ , the overlap will be quite small within a few thermal times. At order one precision, we can therefore determine the size of the operator by setting the formula (2.24) equal to one. We find (for perturbations with order one energy E)

$$r[W_x(t_w)] = \frac{2\pi}{\beta\mu} t_w - \frac{1}{\mu} \log N^2 - O(\log t_w)$$
(2.28)

$$= v_B(t_w - t_*) - O(\log t_w), \qquad (2.29)$$

where  $t_* = \frac{\beta}{2\pi} \log N^2$  and [44]

$$v_B = \frac{2\pi}{\beta\mu} = \sqrt{\frac{d}{2(d-1)}}.$$
 (2.30)

Here, d is the space time dimension of the boundary theory. The radius is negative for  $t_w$  less than the scrambling time  $t_*$ , indicating that the commutator is small everywhere. However, for larger values of  $t_w$ , the region of influence spreads ballistically, with the velocity  $v_B$ . This is the speed at which precursors grow, or equivalently, the speed at which the butterfly effect propagates.

It is interesting to compare this speed with  $v_E$ , the rate at which entanglement spreads. On rather general grounds, entanglement should spread no faster than the commutator of local operators.<sup>13</sup> The speed  $v_E$  was recently computed by [85, 92] for holographic systems

 $<sup>^{13}</sup>$ We are grateful to Sean Hartnoll for emphasizing this point to us.

dual to Einstein gravity, with the result

$$v_E = \frac{\sqrt{d}(d-2)^{\frac{1}{2}-\frac{1}{d}}}{[2(d-1)]^{1-\frac{1}{d}}}.$$
(2.31)

One can check that  $v_B \ge v_E$ , with equality at d = 1 + 1.

The speed  $v_B$  will be corrected in bulk theories that differ from Einstein gravity. Stringy effects are considered in [57]. In appendix 2.5.2, we work out the shock solutions in Gauss-Bonnet gravity<sup>14</sup>. The result for  $d \ge 4$  is that the velocity is corrected to

$$v_B(\lambda_{GB}) = \frac{1}{2}\sqrt{1 + \sqrt{1 - 4\lambda_{GB}}}\sqrt{\frac{d}{d - 1}}$$

$$(2.32)$$

$$= \left(1 - \frac{\lambda_{GB}}{2} + \dots\right) v_B. \tag{2.33}$$

This speed increases for negative  $\lambda_{GB}$ , and it exceeds the speed of light (in d = 4) for  $\lambda_{GB} < -3/4$ . In fact, Gauss-Bonnet gravity is known to violate boundary causality for  $\lambda_{GB} < -0.36$  [113]. (In fact the recent work of [114] shows that maintaining causality requires an infinite number of massive higher spin fields for any nonzero value of  $\lambda_{GB}$ .)

#### 2.3.3 ERB dual to multiple localized precursors

In this section, we will characterize the geometry dual to a product of localized precursor operators,

$$W_{x_n}(t_n)...W_{x_1}(t_1)|TFD\rangle.$$
 (2.34)

Following [115], Ref. [54] showed how to construct states of this type in the spatially homogeneous case, building the geometry up one shock at a time. If the masses of the perturbations are small and the times are large, the geometry consists of a number of patches of AdS-Schwarzschild, glued together along their horizons, with null shifts  $h_1...h_n$  determined by the times  $t_1...t_n$ . For spatially localized perturbations, there are two differences: first, the null shifts depend on the transverse position  $h_1(x)...h_n(x)$  and second, the geometry after shocks collide is not generally known. These regions will not substantially affect our analysis, for reasons explained below.

The point we will emphasize is that the intrinsic geometry of the maximal spatial slice<sup>15</sup>

<sup>&</sup>lt;sup>14</sup>We are grateful to Steve Shenker for suggesting this.

<sup>&</sup>lt;sup>15</sup>The timelike interval inside the ERB remains of order  $\ell_{AdS}$ , even as the shocks make the total spatial

through the ERB,  $\Sigma_{max}$ , agrees with the geometry of the corresponding TN on scales large compared to  $\ell_{AdS}$ . In general, the geometries dual to (2.34) do not have any symmetry, and finding the maximal surface exactly would require the solution of a nonlinear PDE. In order to understand the large-scale features of  $\Sigma_{max}$ , we use the following fact: in the exact AdS-Schwarzschild geometry, maximal surfaces within the ERB are attracted to a spatial slice defined by a constant Schwarzschild radius  $r = r_m$ . This was first pointed out (for co-dimension two surfaces) in [85], and we will refer to the attractor surface as the Hartman-Maldacena (HM) surface.

Within each patch of a multi-shock wormhole, the maximal surface hugs the HM surface of that patch. As  $\Sigma_{max}$  passes through a shock connecting adjacent patches, it transitions from one HM surface to the next. Due to the lack of symmetry, we will not be able to determine  $\Sigma_{max}$  in this transition region. However, the transition takes place over an intrinsic distance of order  $\ell_{AdS}$ . Each HM surface is intrinsically flat, so the surface  $\Sigma_{max}$ consists of approximately flat regions, joined by curved regions of size  $\sim \ell_{AdS}$ . The flat regions grow large in proportion to the time between W insertions. We will calculate their size and shape, in the limit that they become large, and match to the geometry of the TN associated to (2.34).

Before we begin, let us make one more technical comment. In the analysis below, we will focus on the geometry of the "decoupled maximal surface"  $\Sigma_{dec}$ , which maximizes a volume-like functional,  $V_{dec}$ , obtained from the true volume by dropping terms involving x gradients. Finding this surface is technically simpler, but we argue in appendix 2.5.3 that  $\Sigma_{dec}$  and  $\Sigma_{max}$  agree at the level of an  $\ell_{AdS}$  coarse-graining, because both are attracted to the same HM surfaces.

#### A simple example

We will start by demonstrating agreement with the TN geometry in a case in which (i) all shocks are centered near  $x_j = 0$ , (ii) all odd numbered times  $t_1, t_3, ...$  are positive and all even nubered times  $t_2, t_4, ...$  are negative, and (iii) all shocks are strong, i.e. adjacent time differences are large,  $|t_{j+1} - t_j| - 2t_* \gg 1$ . The x = 0 slice through the geometry dual to a configuration of this type is shown in Fig. 2-9. Across shock j (counting from the right),

volume very large. The conjecture of [86] was that the TN geometry reflects a coarse-graining of the ERB on scale  $\ell_{AdS}$ . At this level, the maximal spatial surface represents the entire ERB.



Figure 2-9: A slice through a three-shock geometry. The white pre-collision regions are cut and pasted from the displayed white regions of AdS black holes. The pale blue curves are HM surfaces of the separate patches. The dark blue curve is the maximal surface. Dark grey indicates post-collision regions. As the shocks grow stronger, the dark grey regions shrink and the maximal surface tracks more of each HM surface.

we have a null shift of magnitude

$$h_j(x) = \frac{e^{\frac{2\pi}{\beta}(\pm t_j - t_*) - \mu |x|}}{|x|^{\frac{d-2}{2}}}.$$
(2.35)

The upper sign is appropriate for the v shifts associated to odd-numbered perturbations (right-moving shocks), and the lower sign is appropriate for u shifts associated to the evennumbered perturbations (left-moving shocks). This form is accurate for  $|x| \gg 1$ . At smaller values of |x|, the singularity in the denominator should be smeared out over the thermal scale  $\sim \ell_{AdS}$ .

Because the defining functional for  $\Sigma_{dec}$  does not contain x gradients, the surface can be constructed independently at each x, by solving a maximal surface problem in a spatially homogeneous shock geometry with x-independent shifts  $h_i = h_i(x)$ . This problem was studied in [79], following [85]. For a configuration with n shocks, the intersection of  $\Sigma_{dec}$ with x = 0 is a curve made up of n + 1 segments. Two of these connect to the asymptotic boundaries, and n-1 of them pass between shocks. All but an order-one contribution to the length of these segments comes from regions near the HM surface. Following the analysis in [79] one finds that the large t behavior of the length of the j-th segment is

$$\log h_{j+1}h_j \propto |t_{j+1} - t_j| - 2t_*. \tag{2.36}$$

These segments can be identified with pieces of the folded time axis at x = 0.

So far, this is identical to the homogeneous case in [79]. But now, keeping the same configuration of shocks, we consider a slice at nonzero x. As |x| increases, the only difference will be that the shocks are weaker, according to the transverse profile in Eq. (2.35), and the segments will be correspondingly shorter,

$$|t_{j+1} - t_j| - 2t_* - 2|x|/v_B - O(\log|x|).$$
(2.37)

Apart from the  $t_*$  and the logarithm, this agrees with the |x| dependence of the positiondependent folds from section 2.2.2. The  $t_*$  represents further cancellation of  $e^{-iHt}$  and  $e^{iHt}$ during single-site fast scrambling [79]. The logarithm in |x| indicates a slight modification of linear growth, but it is subleading in the limit that the segments are large. An important point is that, even as x is varied, all but an order one contribution to the length will come from the region near the flat HM surfaces, thus all but an order-one pieces of  $\Sigma$  will be approximately flat. Varying x, the geometry of  $\Sigma_{dec}$  is a fibration of the collection of segments, that is, a fibration of the folded time axis.

This analysis is enough to cover the regions of x over which the length of the folds change, but their number remains constant. We also need to understand the merger and annihilation of folds, as in Fig. 2-7. This takes place when the length of a given segment vanishes. In terms of the shock wave profiles, this corresponds to  $h_j(x)h_{j+1}(x) \leq 1$ . The transition is sketched in Fig. 2-10. The essential point is that two of the shocks become very weak, so the resulting geometry effectively has two fewer shocks, in keeping with the TN.

There is an interesting point regarding nonlinearities during this transition. Classical GR nonlinearities in the collision of shocks are proportional to  $h_j(x)h_{j+1}(x)$ . When this product is large, the effects are strong, but the null shifts ensure that the surface  $\Sigma$  passes far from the post-collision regions. However, when  $h_j(x)h_{j+1}(x)$  is of order one, nonlinear effects are still important, but the shifts are small enough that  $\Sigma$  can pass through the post-collision region (Fig. 2-10). We do not know the geometry in this region, and we are unable to characterize the shape of the maximal surface. Fortunately, this corresponds to a small piece of  $\Sigma$ , of characteristic size  $\ell_{AdS}$ . The reason is that as we vary x, the strength of the shocks changes exponentially, and the product  $h_j(x)h_{j+1}(x)$  rapidly becomes much



Figure 2-10: The ERB implementation of the transition in Fig. 2-7. At left, both  $h_1h_2$ and  $h_2h_3$  are large, so the shaded post-collision regions are far from  $\Sigma$ . Increasing |x|, the shocks become weaker. Eventually  $h_2h_3$  becomes order one (middle) and the surface passes through a nonlinear post-collision region. At yet larger |x|,  $h_2h_3 \ll 1$ ; nonlinear effects are small and the geometry has effectively one shock (right). The transition occurs over an order  $\ell_{AdS}$  region in  $\Sigma$ .

smaller than one.

## The general case

Given a general time-configuration of homogeneous shocks, the only subtlety in constructing  $\Sigma$  comes from the distinction between "through-going" and "switchback" operators [79]. There, as here, one can ignore through-going operators. In the localized case, shocks can also be ignorable because they are sufficiently far away that their profile at the x location of interest is very weak.

Let us therefore begin by determining which operators in a general product (2.34) are relevant. Let  $j_1$  be the first index j such that  $|t_j| - |x - x_j| > t_*$ . This corresponds to the first shock with appreciable strength in the frame of the maximal volume surface anchored at t = 0. As a recursive step, let  $j_{k+1}$  be the least index greater than  $j_k$  such that either the insertion makes  $j_k$  through-going:

$$\operatorname{sgn}(t_{j_{k+1}} - t_{j_k}) = \operatorname{sgn}(t_{j_k} - t_{j_{k-1}})$$
(2.38)

$$|t_{j_{k+1}} - t_{j_{k-1}}| - |x - x_{j_{k+1}}|/v_B > |t_{j_k} - t_{j_{k-1}}| - |x - x_{j_k}|/v_B$$
(2.39)

or the insertion makes  $j_k$  a switchback:

$$\operatorname{sgn}(t_{j_{k+1}} - t_{j_k}) = -\operatorname{sgn}(t_{j_k} - t_{j_{k-1}})$$
(2.40)

$$|t_{j_{k+1}} - t_{j_k}| - 2t_* - |x - x_{j_{k+1}}|/v_B - |x - x_{j_k}|/v_B > 0.$$
(2.41)

The first line of the first condition ensures that the time fold is through-going at  $j_k$ , and the second condition ensures that the shock  $h_{j_{k+1}}(x)$  is stronger than  $h_{j_k}(x)$ . In this case, shock  $j_k$  becomes ignorable for reasons similar to those explained in [79]. The first line of the second condition ensures that the fold switches back, and the second line ensures that the product of the strengths of the shocks,  $h_{j_k}(x)h_{j_{k+1}}(x)$ , will be at least order one. Operators failing both conditions correspond to weak shocks that only mildly affect  $\Sigma$  at this transverse position x. After extracting this subset, we further discard all through-going shocks. To simplify notation, we re-index the remaining shocks by j.

Each shock in this set now corresponds to a switchback of the time fold, so the geometry is nearly identical to the simple case considered above. The only difference is that the profiles are replaced by

$$h_j(x) = \frac{e^{\frac{2\pi}{\beta}(\pm t_j - t_*) - \mu |x - x_j|}}{|x - x_j|^{\frac{d-2}{2}}}.$$
(2.42)

and the length of segment (j + 1) of the folded time axis is generalized to

$$|t_{j+1} - t_j| - 2t_* - |x - x_{j+1}| / v_B - |x - x_j| / v_B.$$
(2.43)

Although we will not give a formal proof, one can check that this set of intervals is the solution to the minimization problem that defines the cross section of the minimal TN at location x, as described at the end of § 2.2.2.<sup>16</sup> This implies that both the TN and the ERB geometry are fibrations of the same collection of intervals over the x space, and therefore that they agree.

<sup>&</sup>lt;sup>16</sup>The  $t_*$  was not present in section 2.2.2, because the single-site "scrambling time" is order one for a qubit system. Its appearance in this equation is consistent with the interpretation of extra cancellation between U(t) and U(-t) in a large N system with a single-site perturbation [79].

# 2.4 Discussion

Connections have been found between quantum mechanics and geometry, most notable the connection between spatial connectivity and entanglement [82, 83, 116–118]. But entanglement may not be enough; the growth of entanglement saturates after a very short time, while geometry continues to evolve for a very long time. Evidently there is need for quantities more subtle than entanglement entropy to encode this evolution. In lattice quantum systems (in contrast to classical systems) computational complexity evolves for an exponentially long time, as does the form of the minimal tensor network describing a state. It has been conjectured that these quantities are related to the geometry behind the horizon [86, 102].

States obtained through the action of precursor operators provide a setting to test these conjectures. In [79], spherically symmetric precursors were studied. Because of the symmetry, it was only possible to relate a single geometric quantity—ERB-volume—to a single information-theoretic quantity—complexity. By contrast, the spatially localized precursors studied in this paper provide an opportunity to relate a much wider range of local geometric properties of ERBs and TNs.

Our basic hypothesis, following [82], was that the structure of the minimal tensor network, encoding the instantaneous state of the holographic boundary theory, directly reflects the Einstein geometry of the ERB. We emphasize that it is the minimal TN—the one that defines complexity—that seems to be picked out by general relativity. There are many TNs that can describe a given state. For example the naive TN on the left side of figure 2-4 generates the same state as one on the right. For a general product of precursor operators, we found a match between the "minimalized" TN and the large-scale spatial geometry implied by general relativity.

More precisely, we found a match between the TN associated to a product of precursors in a strongly interacting lattice system, and the large-scale features of the ERB geometry characterizing an analogous product in the holographic theory. We do not mean to imply that spin systems are dual to black holes, or that we know precisely how to describe states of a continuum quantum field theory using tensor networks.<sup>17</sup> The idea is one of universality: the pattern of growth of precursor operators leads to a geometric description that is shared

<sup>&</sup>lt;sup>17</sup>But see [119, 120] for a description of continuous-time complexity.

by a wide collection of quantum systems.

It would be interesting to understand how wide this collection is. Strong coupling plays an important role, since precursors do not grow in free theories. A natural problem is to understand corrections to the pattern of growth at finite coupling. In gauge/gravity duality, the effects of finite gauge theory coupling translate to stringy corrections in the bulk. The role of inelastic and stringy physics in the context of precursors and shock waves is the subject of [57].

# 2.5 Appendix

## 2.5.1 Localized shocks at finite time

In the main text of the paper, we presented a localized shock solution proportional to  $\delta(u)$ . This solution is appropriate in a particular limit, where we take  $t_w$  to infinity and the mass of the perturbation to zero, with  $T_{uu}$  held fixed. If  $t_w$  is finite, there are corrections to this solution. For example, the shock must be confined within the future lightcone of the source point. Exact solutions with this property in planar BTZ can be obtained from [121]. However, those solutions are precisely localized on the light cone, corresponding to a carefully tuned high energy insertion at the boundary. In our setting, the boundary operator should be low energy, and the solution will not be precisely localized on the light cone. In principle, the exact linearized backreaction can be calculated using retarded propagators. Unfortunately, these functions are not known exactly for the black hole background.

Without constructing the exact solution explicitly, we will show in this appendix that the shock profile is accurate for  $x < v_B t$ . If  $G_N$  is small, the perturbation can be very small in this region, but it will be large compared to  $G_N$ .

We begin with a planar black hole metric with an arbitrary perturbation h:

$$ds^{2} = -A(uv)dudv + B(uv)dx^{i}dx^{i} + h_{\mu\nu}dx^{\mu}dx^{\nu}.$$
(2.44)

We are interested in the response to an order one source at time  $t_w$  in the past, with  $t_w$ 

large. This solution can be constructed by applying a boost to a reference solution

$$h_{uu}(u, v, x^{i}) = \gamma^{2} H_{uu}(\gamma u, \gamma^{-1} v, x^{i})$$
(2.45)

$$h_{ui}(u,v,x^i) = \gamma H_{ui}(\gamma u,\gamma^{-1}v,x^i)$$
(2.46)

$$h_{ii}(u, v, x^{i}) = H_{ii}(\gamma u, \gamma^{-1}v, x^{i}), \qquad (2.47)$$

where  $\gamma = e^{\frac{2\pi}{\beta}t_w}$ , and H is the backreaction of a perturbation at time t = 0.

As we increase  $\gamma$ , the  $h_{uu}$  and  $h_{ui}$  components of the metric become larger, but the entire profile is compressed towards the u = 0 horizon. Physically, it is clear that the disturbance to the geometry at an order one value  $u_0$  of the u coordinate should remain finite as  $\gamma$  grows large. That is, the effect on a late infaller of an early perturbation (on the same side) does not grow with the time separation. We will use this fact below, in the form  $h_{\mu\nu}(u_0, 0, x^i) \sim G_N O(1)$ , where the O(1) refers to  $\gamma$  dependence.

Let us consider the quantity

$$\delta v(x^{i}) = \int_{0}^{u_{0}} du \, \frac{h_{uu}(u,0,x^{i})}{A(0)}.$$
(2.48)

For small h, this is the shift in the v direction of a null curve near v = 0, traveling from u = 0 to  $u = u_0$ . We would like to show that this quantity is approximately equal to the shock wave profile h(x). To do so, we consider the u-u component of Einstein's equations. We define  $E_{uu} = R_{uu} - \frac{1}{2}g_{uu}R + g_{uu}\Lambda$ . To linear order in h, this is

$$E_{uu} = \frac{D-2}{2} \frac{B'}{AB} (-2 + v\partial_v - u\partial_u)h_{uu} - \frac{1}{2B}\partial_i^2 h_{uu} + \frac{A}{2B} \left[ 2\partial_u \partial_i \frac{h_{ui}}{A} - \partial_u \left( \frac{B}{A} \partial_u \frac{h_{ii}}{A} \right) \right] - \frac{vB'}{B} \frac{h_{uv}}{A}.$$
(2.49)

where the first term was simplified using the background equations of motion for A, B. Einstein's equations set this equal to  $8\pi G_N T_{uu}$ .

We will take this component of Einstein's equations, and integrate it du along v = 0. Integrating by parts, we find

$$(a_0 - a_1\partial_i^2)\delta v(x^i) + a_2\partial_i h_{ui}(u_0, 0, x^i) - a_3\partial_u h_{ii}(u_0, 0, x^i) = 8\pi G_N \int_0^{u_0} du \, T_{uu}(u, 0, x^i), \quad (2.50)$$

where the constants  $a_i$  are related to A, B, B' evaluated at the horizon. Since the metric

components at  $u_0$  are O(1) at large  $\gamma$ , we therefore have

$$(a_0 - a_1 \partial_i^2) \delta v(x^i) = 8\pi G_N \int_0^{u_0} du \, T_{uu}(u, 0, x^i) + G_N O(1).$$
(2.51)

This is the same equation satisfied by the shock profile. We see that it is accurate at large  $\gamma$ , up to a correction that is O(1) in  $\gamma$ . The solution  $\delta v(x^i)$  will therefore agree with the shock profile  $G_N e^{\frac{2\pi}{\beta}t_w - \mu|x|}$  in the region that this profile is large compared to  $G_N$ . That is, for  $|x| < \frac{2\pi}{\beta\mu}t_w = v_Bt_w$ .

#### 2.5.2 Localized shocks in Gauss-Bonnet

In Gauss-Bonnet gravity [122] with a negative cosmological constant, the action is

$$S = \frac{1}{16\pi G_N} \int d^{d+1}x \sqrt{-g} \left\{ R + \frac{d(d-1)}{\ell_{AdS}^2} + \alpha (R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2) \right\}, \quad (2.52)$$

with AdS radius  $\ell_{AdS}$ , and Gauss-Bonnet coefficient  $\alpha$ . For d < 4, the Gauss-Bonnet term is topological. It will be convenient to rewrite this coefficient as

$$\alpha = \frac{\lambda_{GB} \ \ell_{AdS}^2}{(d-2)(d-3)},\tag{2.53}$$

with  $\lambda_{GB}$  a dimensionless parameter. The planar black hole solution [123–125] is

$$ds^{2} = \ell_{AdS}^{2} \Big[ -f(r)N_{\sharp}^{2}dt^{2} + f(r)^{-1}dr^{2} + r^{2}dx^{i}dx^{i} \Big], \qquad (2.54)$$

with

$$f(r) = \frac{r^2}{2\lambda_{GB}} \left[ 1 - \sqrt{1 - 4\lambda_{GB} \left(1 - r^{-d}\right)} \right]$$
(2.55)

$$N_{\sharp}^{2} = \frac{1}{2} \Big( 1 + \sqrt{1 - 4\lambda_{GB}} \Big). \tag{2.56}$$

The horizon is at r = 1, and one can check that f'(1) = d, as in Einstein gravity.

Following § 2.3.1, we pass to Kruskal coordinates and assume a shock wave ansatz as the same form as (2.22). Plugging into the GB equation of motion, with a stress tensor

 $T_{uu} \propto \delta(u) \delta^{d-1}(x)$ , we find the condition

$$(1+2\lambda_{GB})\left(-\partial_i\partial_i+\mu^2\right)h(x)\propto\delta^{d-1}(x),\tag{2.57}$$

where again  $\mu^2 = \frac{d(d-1)}{2}$ . The factor  $(1+2\lambda_{GB})$  rescales the source, effectively changing the scrambling time by a small amount of order  $\log(1+2\lambda_{GB})$ , but the transverse dependence of h(x) is unchanged. The important difference is the presence of  $N_{\sharp}$  in the metric. This changes the temperature such that now  $\beta = 4\pi/N_{\sharp}f'(1)$ . The strength of the perturbation still grows with the boost factor  $e^{\frac{2\pi}{\beta}t_w}$ , but because the relationship between  $\beta$  and f'(1) has been rescaled, one finds  $v_B(\lambda_{GB}) = N_{\sharp}v_B$ .<sup>18</sup>

## 2.5.3 Maximal volume surface and decoupled surface

In this appendix, we will give some details about  $\Sigma_{max}$ ,  $\Sigma_{dec}$  and the relationship between them. First, let us recap the definitions:

- $\Sigma_{max}$  is the maximal volume codimension one surface crossing the wormhole, anchored at t = 0 on the two asymptotic boundaries.
- $\Sigma_{dec}$  is the "decoupled maximal surface." Specifically, it maximizes a modified volume functional,  $V_{dec}$  obtained by dropping all gradients in the x direction.

The coarse-grained features of these surfaces are very similar to each other, because most of the volume comes from a region where x gradients are small. However,  $\Sigma_{dec}$  is much easier to work with, because the defining equation is decoupled in the x coordinate. Here, we will study the relationship between the surfaces in the example setting of a single shock wave.

We can use the symmetry of the shock wave geometry to reduce the problem to one in the unperturbed black hole geometry. We illustrate this in Fig. 2-11. The left panel is a representation of a constant x slice through the shock wave geometry, with a slice of the maximal volume surface shown in blue. On the right, we display only the region on one side of the shock (the other is related by a symmetry). The portion of the surface in this region is a maximal volume surface in the unperturbed black hole geometry, with boundary conditions t = 0 at the left boundary, and v = h(x)/2 at the horizon u = 0.

<sup>&</sup>lt;sup>18</sup>D.S. is grateful to Juan Maldacena for pointing out a mistake in v1 of this appendix, which stated that the  $e^{\frac{2\pi}{\beta}t_w}$  relationship between boosts and  $t_w$  was modified. The error did not propagate elsewhere in the paper.



Figure 2-11: Left: a constant x slice through the shock wave geometry, with the corresponding slice of the maximal volume surface shown in blue. Right: the portion of the surface to the left of the shock is a maximal volume surface in the unperturbed black hole geometry, with the boundary conditions shown. The full surface is obtained by gluing two such pieces together.

We can parametrize the surface using  $v(u, x^i)$ . The pullback metric is then

$$G_{ab}dy^a dy^b = \ell_{AdS}^2 \Big[ -A(uv)\partial_u v \ du^2 - A(uv)\partial_i v \ du dx^i + B(uv) \ dx^i dx^i \Big],$$
(2.58)

and the volume of both pieces is

$$V = 2\ell_{AdS}^{d} \int d^{d-1}x du \sqrt{-AB^{d-1}\partial_{u}v - \frac{A^{2}B^{d-2}}{4}(\partial_{i}v)^{2}}.$$
 (2.59)

It is useful to keep in mind that  $\partial_u v$  is negative, so the first term is positive. The decoupled volume  $V_{dec}$  is given by dropping the second term inside the square root.

Using the surface  $\Sigma_{dec}$ , we can find upper and lower bounds on the volume of the maximal surface,

$$V(\Sigma_{dec}) \le V(\Sigma_{max}) \le V_{dec}(\Sigma_{dec}).$$
(2.60)

The first inequality follows from the fact that  $\Sigma_{max}$  is maximal. The second follows from the fact that  $V \leq V_{dec}$  for any surface, and that  $\Sigma_{dec}$  maximizes  $V_{dec}$ . In Fig. 2-12, we plot  $V_{dec}(\Sigma_{dec})$  and the gap in the bounds for a shock in the BTZ geometry, as a function of the strength  $t_w - t_*$ . Numerically, the gap between the bounds is quite small, and appears to be proportional to the  $t_w$ -derivative of the volume.

We can explain this as follows. At a fixed value of x, the decoupled surface v(u) is given by finding a maximal surface in a sptially homogeneous shock background. This problem was studied in [79] using techniques from [85]. For large h(x), the surface tends to hug a fixed radius in the interior,  $r_m$ , that maximizes the function  $r^{d-1}\sqrt{|f(r)|}$ . In terms of u, v



Figure 2-12: The blue curve (left axis) shows the decoupled volume of the decoupled surface,  $V_{dec}(\Sigma_{dec})$ , as a function of  $(t_w - t_*)$ , for the BTZ setting of d = 1 + 1 theory on a spatial circle. Initially, the volume grows quadratically, but after a  $(t_w - t_*) \sim \pi$ , the size of the precursor saturates, and the volume grows linearly. The red curve (right axis) shows the gap between the upper and lower bounds in Eq. (2.60). The gap is quite small, and is roughly proportional to the  $t_w$ -derivative of the volume itself.

coordinates, this special radius corresponds to a surface given by vu = const, independent of x.

In the spatially homogeneous case, the contribution to the volume coming from the region near this surface is proportional to log h. The surfaces  $\Sigma_{max}$  and  $\Sigma_{dec}$  will be very similar in this region, because x gradients are small. To put it differently,  $r = r_m$  is an attractor for both  $\Sigma_{dec}$  and  $\Sigma_{max}$ . For large h, most of the surface is near this radius, and the surfaces will therefore agree at a coarse-grained level. Away from the special surface  $r = r_m$ , the surface  $\Sigma_{max}$  and  $\Sigma_{dec}$  will differ, but the regularized volume in this region is subleading at large h.

Integrating over x, the regularized decoupled volume of  $\Sigma_{dec}$  will be proportional to

$$V_{dec}(\Sigma_{dec}) \propto \int_{h(x)\ge 1} d^{d-1}x \log h(x), \qquad (2.61)$$

while the difference will be proportional to

$$V_{dec}(\Sigma_{dec}) - V(\Sigma_{max}) \propto \int_{h(x) \ge 1} d^{d-1}x.$$
(2.62)

It follows that the difference is subleading at large h, and in fact proportional to the  $t_w$ derivative of the volume.

# Chapter 3

# Agere sequitur credere

(Complexity equals action)

Last chapter explored the relationship between tensor networks and black holes for a complicated many-parameter family of states. While not directly emphasized due to the focus on tensor networks, these examples provide nontrivial evidence for a conjecture of Stanford and Susskind. In [79], they conjecture that the computational complexity of a holographic state is dual to the volume of a maximal co-dimension one spacelike slice through the bulk

$$\mathcal{C}(t_L, t_R) \sim \frac{V(t_L, t_R)}{G_N \ell_{Ads}},\tag{3.1}$$

where C is the complexity, and V is the volume of the maximal surface anchored at times  $t_L$  and  $t_R$  on the left and right boundaries, respectively.<sup>1</sup> On the tensor network side of the duality, the complexity is essentially the number of tensors in the network—the number of simple gates in the minimal quantum circuit.

In this chapter, we explore an alternative (and improved) conjecture: "complexity equals action." This conjecture reproduces all the nontrivial checks of "complexity equals volume" (as will be explained below), but also improves on some of the notable confusions. In particular, it suggests that one can actually fix the constant of proportionality between complexity and action.

This leads to a corollary conjecture that black holes are the fastest complexifiers in

 $<sup>{}^{1}\</sup>mathcal{C}$  is also a function of the overall state, e.g. where and when shock wave perturbations  $W_{x}(t)$  are applied to the thermofield double state.

nature; that black holes are nature's fastest computers. In some sense, this seems reasonable given that black holes are the densest hard drives [21–23] and the fastest scramblers [39, 44, 58]. Scrambling has a natural interpretation in terms of the "processing" of quantum information (one could analogize this as black holes have the fastest CPUs) and begs to be related to some notion of computation.<sup>2</sup>

Regardless of whether either conjecture is exactly correct, it seems that there's an important sense in which the black hole interior is a geometric representation of a quantum circuit.<sup>3</sup> While it may be unclear whether the volume or the action are exactly the right geometric invariant to consider, my opinion is that we should keep studying them (even if only to try and break them). My hope is that by pursing these conjectures we can make more precise the definition of computational complexity in conformal field theory and ultimately learn more about the underlying puzzles centered on understanding the black hole interior.

#### \*\*\*\*

This chapter was first presented as [2] under the title "Complexity Equals Action." Upon submission, the title was immediately changed by the editor to "Computational complexity of a holographic quantum state equals action of a spacetime patch?" After much back and forth (and our insistence that a "title" is not an "abstract"), we finally landed on "Holographic complexity equals bulk action?" (They wouldn't give up on the question mark.) Regardless of what you write at the top of the article or chapter, it is never true that *action follows belief*: conjectures require evidence. Unfortunately, that evidence came some months later in a obsencely long follow-up [49]. Both works are a collaboration with Adam R. Brown, Leonard Susskind, Brian Swingle, and Ying Zhao, but (thankfully for the reader) this chapter based only on the short paper,

*Phys. Rev. Lett.* **116** no. 19, (2016) 191301, arXiv:1509.07876 [hep-th].

#### Abstract

We conjecture that the quantum complexity of a holographic state is dual to the action of a certain spacetime region that we call a Wheeler-DeWitt patch. We illustrate and test

<sup>&</sup>lt;sup>2</sup>This is the subject of future work.

<sup>&</sup>lt;sup>3</sup>This idea is elaborated more in Chapter 5.

the conjecture in the context of neutral, charged, and rotating black holes in AdS, as well as black holes perturbed with static shells and with shock waves. This conjecture evolved from a previous conjecture that complexity is dual to spatial volume, but appears to be a major improvement over the original. In light of our results, we discuss the hypothesis that black holes are the fastest computers in nature.

#### \*\*\*\*

The interior of a black hole is the purest form of emergent space: once the black hole has formed, the interior grows linearly for an exponentially long time. One of the few holographic ideas about the black hole interior is that its growth is dual to the growth of quantum complexity [102, 126]. This duality is a conjecture but it has passed a number of tests.

In the context of AdS/CFT duality, the conjecture has taken a fairly concrete form: the volume of a certain maximal spacelike slice, which extends into the black hole interior, is proportional to the computational complexity of the instantaneous boundary conformal field theory (CFT) state [79]. The conjecture is an example of the proposed connection between tensor networks and geometry—the geometry being defined by the smallest tensor network that prepares the state. (See also [1, 84, 85, 102, 119, 126, 127].)

For the case of the two-sided AdS black hole the conjecture is schematically described by

Complexity 
$$\sim \frac{V}{G\ell_{AdS}},$$
 (3.2)

where V is the volume of the Einstein-Rosen bridge (ERB),  $\ell_{AdS}$  is the radius of curvature of the AdS spacetime, and G is Newton's constant. Multiplying and dividing Eq. 3.2 by  $\ell_{AdS}$  suggests a new perspective on the identification of complexity and geometry,

Complexity 
$$\sim \frac{\mathcal{W}}{G\ell_{AdS}^2},$$
 (3.3)

where  $\mathcal{W} \equiv \ell_{AdS} V$  now has units of spacetime volume and represents the world volume of the ERB. Further noting that  $1/\ell_{AdS}^2$  is proportional to the cosmological constant of the AdS space inspires a new conjecture which we suspect may have deep implications for the connection between quantum information and gravitational dynamics. We propose:

CA-conjecture: Complexity = 
$$\frac{\text{Action}}{\pi\hbar}$$
. (3.4)

The detailed calculations are presented in [49]; here we will only attempt to explain the conjecture. The systems we will consider are those whose low-energy bulk physics is described by the Einstein-Maxwell action

Action = 
$$\frac{1}{16\pi G} \int_{\mathcal{M}} \sqrt{|g|} \left(\mathcal{R} - 2\Lambda\right) - \frac{1}{16\pi} \int_{\mathcal{M}} \sqrt{|g|} F_{\mu\nu} F^{\mu\nu} + \frac{1}{8\pi G} \int_{\partial\mathcal{M}} \sqrt{|h|} K,$$
 (3.5)

with the usual conventions [128]. The three terms in Eq. 3.5 representing the action of a region  $\mathcal{M}$  are the Einstein-Hilbert (EH) action including a (negative) cosmological constant, a Maxwell term, and a York-Gibbons-Hawking (YGH) surface term constructed from the extrinsic curvature tensor K.

In AdS/CFT, the spacetime region dual to the boundary state is the "Wheeler-DeWitt (WDW) patch" whose action, according to the conjecture, gives the complexity of the state. The WDW patch, plotted in Fig. 3-1, is given by the union of all spatial slices anchored at a given boundary time t (or pair of times  $(t_L, t_R)$  for the two-sided case).

### \*\*\*\*

Black holes are known to excel at information theoretic tasks: they are the densest memory [21–25]; they are the fastest scramblers [38, 39, 58]. Here we explore the possibility that black holes also saturate a universal bound on complexity growth. Computational complexity is the minimum number of quantum gates required to prepare the boundary state from a simple state [129, 130]. The reference state may be taken to be an unentangled state so that all non-trivial correlations are accounted for, or simply the initial state if we are interested only in complexity growth [49]. The use of discrete gates can be justified by lattice regulating the field theory of interest and making a renormalization group argument [49].

Interpreting complexity growth as computation, and normalizing complexity using Eq. 3.4, we find our proposals are compatible with the saturation of Lloyd's conjectured bound<sup>4</sup> on

<sup>&</sup>lt;sup>4</sup>The numerical coefficient in Eq. 3.6 is not fixed by our considerations. Indeed the normalization of complexity depends on precise details of the quantum circuits used to prepare the state. What our claim



Figure 3-1: The two-sided eternal AdS black hole (left) and a one-sided AdS black hole that forms from a collapsing shockwave (right). The two-sided AdS black hole is dual to an entangled (thermofield double) state of two CFTs that live on the left and right boundaries; the one-sided black hole is dual to a single CFT. Our complexity/action conjecture relates the complexity of the CFT state to the action of the Wheeler-DeWitt patch (shown shaded).

the rate of computation for a system of energy M[47]

$$\frac{d\text{Complexity}}{dt} \le \frac{2M}{\pi\hbar}.$$
(3.6)

The bottom left panel of Fig. 3-1 shows the WDW patch for a neutral two-sided black hole in AdS. There are two boundary times, one on each side of the wormhole, and the symmetry ensures that the action is only a function of the sum  $t_L + t_R$ . Calculating the total action of the WDW patch requires a regulator, since the relevant integrals diverge at the asymptotic AdS boundaries. However, the divergences are time-independent and do not affect the rate of change of action.

We have found that the late-time rate of change of action of the WDW patch of the neutral AdS black hole is [49]

$$\frac{d\text{Action}}{d(t_L + t_R)} = 2M. \tag{3.7}$$

This result is simple, but the derivation is non-trivial. It involves a complicated cancellation between EH volume term, and the YGH surface term in Eq. 3.5. Remarkably, this result holds for black holes of any size—small, intermediate, or large compared to the AdS radius and in any number of spacetime dimensions. In the previous proposal of [79], the coefficient in the rate of growth of complexity depended both on the size of the black hole and on the number of dimensions, which made it impossible to saturate a universal bound of the form of Eq. 3.6. The universality of the rate of growth of action means that our CA-duality implies that all neutral black holes of any size and in any number of dimensions saturate the same bound with the same coefficient.

Additional evidence for our conjecture is provided by black holes with conserved charge Q. With a conserved charge, the system is more restricted and should complexify slower. In [49] we will argue that the bound should generalize to

$$\frac{d\text{Complexity}}{dt} \le \frac{2}{\pi\hbar} \left[ (M - \mu Q) - (M - \mu Q)_{\text{ground state}} \right].$$
(3.8)

For a given chemical potential  $\mu$ , the ground state is the state of minimum  $M - \mu Q$ . No new coefficient is required in this equation—the coefficient is fixed by requiring that Eq. 3.8

means is that once the normalization is fixed by Eq. 3.6 for any particular black hole, the same coefficient determines the complexity-action relation for all black holes, and indeed all systems.

We find it interesting that by adopting the normalization of Lloyd [47] (see also [48, 131]), an increase of complexity by one gate advances the phase of  $e^{i\operatorname{Action}/\hbar}$  from 1 to -1.
reduce to Eq. 3.6 in the limit  $\mu \to 0$ . For rotating black holes the charge is the angular momentum J and the chemical potential is the angular velocity  $\Omega$ .

We will now study a number of special cases, choosing the dimensionality to make the calculations easy. Our conclusions should apply in any number of dimensions.

For rotating black holes in 2+1-dimensional AdS, the ground state has  $M, J \rightarrow 0$ , and we have found that the bound of Eq. 3.8 becomes [49]

$$\frac{d\text{Complexity}}{dt} \le \frac{2}{\pi\hbar} (M - \Omega J) = \frac{2}{\pi\hbar} \sqrt{M^2 - \frac{J^2}{\ell_{AdS}^2}}.$$
(3.9)

The rate of change of action of the WDW patch of a rotating black hole in 2+1 dimensions has also been calculated [49]. Assuming action and complexity are related by our conjecture, Eq. 3.4, the bound is again saturated. As in the static case, the action calculation involves nontrivial cancellations between EH volume term and the YGH surface term.

For electrically charged black holes in 3+1 dimensional AdS that are much smaller than  $\ell_{AdS}$ , the minimum of  $M - \mu Q$  at fixed  $\mu$  is at  $M, Q \rightarrow 0$ . The bound of Eq. 3.8 becomes

$$\frac{d\text{Complexity}}{dt} \le \frac{2}{\pi\hbar}(M - \mu Q) = \frac{2}{\pi\hbar}\sqrt{M^2 - \frac{Q^2}{G}}.$$
(3.10)

At late times, the rate of change of action of the WDW patch of a small charged black hole in 3+1 dimensions can be calculated [49] and shown, assuming our conjecture, Eq. 3.4, to precisely saturate this bound. The action calculation involves intricate cancellations, this time between all three terms in Eq. 3.5.

For electrically charged black holes in 3+1 dimensional AdS that are much larger than  $\ell_{AdS}$  (shown in Fig. 3-2), the situation is more complicated. Large highly-charged black holes have large  $\mu$ ; for large enough  $\mu$ , the quantity  $M - \mu Q$  has a nontrivial minimum less than zero. Naively taking this minimum to be the large extremal black hole at fixed  $\mu$ , near extremality the bound of Eq. 3.8 becomes

$$\frac{d\text{Complexity}}{dt} \le \frac{4}{\pi\hbar} (M - M_Q) + O\left[ (M - M_Q)^2 \right], \quad \text{[naive]} \quad (3.11)$$

where  $M_Q \equiv \frac{2}{3} \left(\frac{G}{3}\right)^{-\frac{1}{4}} \sqrt{\frac{Q^3}{\ell_{AdS}}}$ . The bound goes linearly to zero at extremality. On the other hand, the late-time rate of change of action of the WDW patch of a large charged black



Figure 3-2: For a charged AdS black hole, the Wheeler-DeWitt patch does not extend all the way to the singularity, and instead ends when the ingoing lightsheets self-intersect just outside the inner horizon at  $r_{-}$ .

hole in 3+1 dimensions can be calculated [49] to be

$$\frac{dAction}{dt} = \sqrt{6M_Q(M - M_Q)} + O[M - M_Q].$$
 (3.12)

Near extremality this is much larger than Eq. 3.11, and so large charged black holes apparently violate the complexity bound. (This problem also afflicts the complexity-volume conjecture of [79].)

However, the naive bound of Eq 3.11 would not apply if the extremal black hole were not the state of lowest  $M-\mu Q$  and fixed  $\mu$ . Were there another even lower state, then the correct application of Eq. 3.8 would use that as the ground state, and the large near-extremal AdS black hole would satisfy the correct bound by a wide margin. We suspect that this is what happens for large  $\mu$ . This would imply that large near-extremal Reissner-Nordstrom-AdS black holes must be unstable.

Given the absence of a no-hair theorem this does not seem impossible. General arguments have been given in this direction based on the weak gravity conjecture [132], and several explicit classes of instabilities are known, including superconducting condensation of charged fields [133–135] (such fields are typically present in UV completions of Eq. 3.5). CA-duality and the complexity growth bound demand that we make the stronger conjecture

that large near-extremal Reissner-Nordstrom-AdS black holes are always unstable. Turning this around, the apparent violation of the complexity-bound can be used as a tool to diagnose instability.

This subtlety does not arise in the other cases we have considered. There are good reasons to be believe that neutral, rotating, and small charged AdS black holes are stable. There are no known instabilities of neutral AdS black holes, large or small. Extremal rotating (2 + 1)-dimensional black holes have UV completions which are supersymmetric BPS states and are rigorously known to be stable. Finally small black holes are essentially the same as their flat space counterparts and are protected by no-hair theorems. The only case which we expect to be unstable is the one where the naive bound fails.

Taking the example of a superconducting instability, we may ask whether fluctuations about the true superconducting ground state obey the complexity growth bound. The finite temperature solution corresponds to a far-from-extremal charged black hole surrounded by shell of condensed charge [136]. The zero-temperature limit of this solution has all the charge residing in the condensate shell and has vanishing horizon area [137], and hence does not complexify. This fact along with the known power-law heat capacity of the hairy black hole [136] imply that the complexity growth-bound is qualitatively obeyed (see also the discussion of static shells below).

#### \*\*\*\*

A strong test of the relationship between geometry and complexity is provided by perturbing the black hole. In [44] (see also [1, 3, 54, 55, 57]), the field theory was perturbed with a small thermal-scale operator. In the boundary, this small perturbation grows due to the butterfly effect, and increases the complexity. In the bulk, the perturbation gives rise to an ingoing null shock wave that increases the volume and action of the Einstein-Rosen bridge. Both the complexity-volume duality of [1, 79] and the complexity-action duality of this paper successfully have these two growths match.

The match is remarkably detailed. Not only does the bulk shock wave calculation successfully reproduce the chaotic growth of complexity in the boundary state, it also reproduces the partial cancellation that occurs during the time it takes the perturbation to spread over the whole system (the "scrambling time" [38, 39]). That the action calculation is sensitive to this cancellation is evidence that it counts the gates of the *minimal* circuit.

The shock wave tests can be made even more stringent. As was shown for the spatial volume of the ERB in [1, 79], and as will be shown for its action in [49], we can add more than one shock wave [54] in more than one location [1], and the dual calculations continue to match. (The action calculations that will be presented in [49] are much easier than the volume calculations of [1, 79], since there are now no differential equations to be solved.) These multiple shock wave states provide detailed evidence for the duality between complexity and geometry because the correspondence continues for all possible times and locations of the perturbations.

Another test of the relationship between complexity and geometry is provided by a different kind of perturbation. Rather than sending a null shockwave into the black hole, we will instead surround the black hole with a static shell held aloft by compressive strength. In calculating the rate of change of action, the shell itself does not contribute since it is static. The only effect is indirect—the shell places the black hole in a gravitational well and so gravitational time dilation slows the black hole's rate of action growth. This action calculation fits with our complexity expectation. First, not all energy computes—the static shell is computationally inert. Second, gravitational time dilation makes computers placed in gravitational wells run slow.

(One way of understanding superconducting black holes is as non-extremal black holes surrounded by a computationally inert superconducting shell.)

#### \*\*\*\*

We have introduced a new conjecture in this paper: that the complexity of a holographic state is dual to the action of the associated Wheeler-DeWitt patch. Although motivated by the older complexity/volume duality of [79], the new conjecture subsumes the old. One may ask in what way is it an improvement?

The original conjecture had some ad hoc features, most notably the introduction of a new length scale with each new configuration to be studied. For large black holes in a given AdS background the length scale was chosen to match the AdS radius of curvature. For small AdS black holes or black holes in flat spacetime the scale was the Schwarzschild radius. No such arbitrary scale is needed in the duality relating complexity and action.

The complexity-action duality has been subjected to a number of non-trivial tests. In [49] we will subject it to a broader battery of tests, perform more detailed calculations, and discuss the connection with tensor networks. A number of open questions remain. We would like to better understand the implications of the conjecture not just for the rate of change of complexity, as we have in this paper, but for the ground state complexity already present at t = 0. Finally, as a means of testing our complexity bound, we would like to extend our proposal to less strongly-coupled boundary theories; this will introduce higher-derivative terms in the bulk action, and we will need to be careful to understand their contribution near the singularity.

The coarse-grained geometry of the Einstein-Rosen bridge is given by the circuit (or tensor network) of least complexity that connects the two sides. One wonders if there is a connection between the principle of least action and this principle of least computation.

The complexity-action conjecture relates the geometry of the bulk to the computational complexity of the boundary. The CA-duality provides a tool for diagnosing when horizons are transparent [35], and also for diagnosing when the state does not belong to a consistent truncation of a UV-complete theory. The WDW patch is the natural bulk spacetime region to associate with a boundary state, and is robust against small perturbations. The action is a natural quantity associated with the Wheeler-DeWitt patch that generalizes to higher dimensions, to more general theories, and to more complicated semiclassical states, without having to make arbitrary choices. Using CA-duality, we saw that neutral AdS black holes in any number of dimensions and of any size all saturate the same bound on the rate of computation with the same coefficient. (The same coefficient also applies to charged and rotating black holes.)

If our complexity-action conjecture is correct, then black holes saturate Lloyd's proposed limit on the rate of computation [47]. CA-duality thus provides a natural framework in which to think about black holes as the fastest computers in nature.

## Chapter 4

# Ordo ab chao<sup>1</sup>

## (Two-dimensional conformal field theory and the butterfly effect)

In this chapter we move away from holography to focus on chaos purely in terms of conformal field theory. This is a slight lie since we focus on two-dimensional CFT in the large central charge limit—theories that are expected to have holographic duals—and we even include matching bulk calculations!

Nevertheless, the point of this chapter is to serve as a model calculation of how to compute "chaotic" observables in quantum field theory: out-of-time-order (OTO) correlation functions such as

$$\langle \psi | W(t) V(0) W(t) V(0) | \psi \rangle.$$
 (4.1)

These OTO correlation functions are simply related to the commutators of operators separated in space and time that we studied in Chapter 2. As emphasized there, it is the growth of operators in time (outwards as a "ball") that is synonymous with the butterfly effect.

Therefore, such calculations were being done implicitly in holography (as in [44, 54, 55] and in Chapter 2), but this work was among the first to really emphasize the importance of such out-of-time-order observables (see also this talk by Kitaev [45]). Additionally, we demonstrate a central property of computing such observables in CFT: they are determined

<sup>&</sup>lt;sup>1</sup>It is often said that it's hard to get *order out of chaos*, but it turns out the converse is really more appropriate: *chao, ab ordo*. Consistent with our intiuition, it is by taking things *out of order* that we discover *chaos*.

by a particular analytic continuation of the Euclidean-time correlation function.

As one hopes with a model calculation, this work has inspired a large body of direct follow-ups [63, 66, 67, 69]. These calculations focus on quantum (or 1/c) corrections [63], rational CFTs [66, 67], and higher-spin theories [69]. For our model calculation, we focus on large c CFTs with a particularly sparse low-lying spectrum. These theories are expected to be chaotic since their thermal states are described by BTZ black holes. We contrast this with a well-known integrable system, the two-dimensional Ising model.

As is appropriate for a study of out-of-time order observables, the work in this chapter actually preceded the previous chapter. Whether the intention was to make a point about time-ordering or to fit in with some other arbitrary (non-science-based) constraint/theme— or whether this is the natural presentation order—I leave this question as an exercise to the reader.

#### \*\*\*\*

This chapter was first presented as [3] under the title "Two-dimensional conformal field theory and the butterfly effect." The title was subsequently changed during the review process to "Diagnosing Chaos Using Four-Point Functions in Two-Dimensional Conformal Field Theory." (For some reason the editors and referees did not agree with the authors that the phenomenon being studied had anything to do with the butterfly effect.) Regardless of what you call it, it was a collaboration with Douglas Stanford,

Phys. Rev. Lett. 115 no. 13, (2015) 131603, arXiv:1412.5123 [hep-th].

#### Abstract

We study chaotic dynamics in two-dimensional conformal field theory through out-of-time order thermal correlators of the form  $\langle W(t)VW(t)V\rangle$ . We reproduce bulk calculations similar to those of [44], by studying the large c Virasoro identity block. The contribution of this block to the above correlation function begins to decrease exponentially after a delay of  $\sim t_* - \frac{\beta}{2\pi} \log \beta^2 E_w E_v$ , where  $t_*$  is the scrambling time  $\frac{\beta}{2\pi} \log c$ , and  $E_w, E_v$  are the energy scales of the W, V operators.

### 4.1 Introduction

Chaos is a generic feature of thermal systems, but it is difficult to study directly. This is both because chaotic systems tend to be poorly suited to perturbation theory, and because e.g. the butterfly effect is not easily visible in time-ordered vacuum correlation functions. In recent work [1, 44, 45, 54, 55], gauge/gravity duality was used to expose chaos in the gauge theory in terms of shock waves on the horizons of AdS black holes.

In this paper, we will reproduce part of that analysis without using holography directly. We will instead focus on a particular contribution to the four point function given by the large c Virasoro identity block. The close relationship between 2+1 gravity and the identity Virasoro block has been demonstrated recently in [138–140]; our work should be understood as an application of the techniques in these papers to the problem studied in [1, 44, 45, 54, 55].

The central object of our study will be a non-time-ordered correlator of the  $form^2$ 

$$\langle W(t)VW(t)V\rangle_{\beta}.$$
 (4.2)

Here, the subscript  $\beta$  indicates a thermal expectation value. W, V are approximately local operators, smeared over a thermal scale, and with one-point functions subtracted. In a suitably chaotic system, this type of correlation function becomes small as t becomes large. We believe that this happens for practically any choice of W and V (subject to the above conditions) and that, in fact, this behavior is a basic diagnostic of quantum chaos. To motivate this point, it is helpful to regard Eq. (4.2) as an inner product of two states: one given by applying V then W(t), and one given by applying W(t) then V. Intuition from classical chaos suggests that these states should be rather different if the time t is sufficiently long, hence a small overlap. This was confirmed for holographic theories in the gravity analysis of [1, 44, 54]. Such correlation functions have also been studied recently, using similar techniques, by Kitaev [45], who interpreted the initially exponential decrease in terms of Lyapunov exponents, following [52]. Kitaev also pointed out the unusual quantization of these exponents in a theory dual to gravity.

The discussion above should be contrasted with the behavior of correlation functions

<sup>&</sup>lt;sup>2</sup>In this paper, we use the standard convention  $W(t) = e^{iHt}We^{-iHt}$ .

with the orderings

$$\langle W(t)W(t)VV \rangle_{\beta}, \qquad \langle VW(t)W(t)V \rangle_{\beta},$$

$$(4.3)$$

both of which tend to an order-one value  $\langle WW \rangle_{\beta} \langle VV \rangle_{\beta}$  as t becomes large. In the case of the WWVV configuration, this is because the correlator is time-ordered, and the connected contribution decays with t. The VWWV ordering can be understood as an expectation value of WW in a state given by acting with V on the thermal state. If the energy injected by V is small, the state will relax and the expectation value will approach the thermal value,  $\langle WW \rangle_{\beta}$ , multiplied by the norm of the state,  $\langle VV \rangle_{\beta}$ .

Each of the Lorentzian correlators in (4.2) and (4.3) can be obtained by analytic continuation of the same Euclidean four point function. The sharp difference in behavior arises because the continuation defines a multivalued function, with different orderings corresponding to different sheets. To see the buttefly effect, one has to move off the principal sheet.

Although what we have said up to now is completely general, most of the paper will be restricted to two dimensional conformal field theory. There are two reasons for this. The first is that conformal symmetry in two dimensions relates thermal expectation values (in a spatially infinite system) to vacuum expectation values. This makes it possible to study the thermal correlation functions described above using standard conformal blocks on the plane. A key feature of the conformal mapping is that long time t in (4.2) and (4.3) translates to small values of the cross ratios  $z, \bar{z}$ . This suggests an OPE limit, and indeed, for the orderings in Eq. (4.3), the large t behavior is dominated by the identity operator in the operator product of WW. However, for the ordering in Eq. (4.2), z has moved to a second sheet of the Euclidean correlation function, where the ordinary OPE does not apply.

A very similar situation arose in the study of high energy AdS scattering by [141–143]. As pointed out in those references, a formal expansion in continued conformal blocks is still possible on the second sheet of the four point function. The conformal block associated to global primaries with spin J and energy E give a contribution proportional to (in our variables)

$$e^{\frac{2\pi}{\beta}(J-1)|t|}e^{-\frac{2\pi}{\beta}(E-1)|x|}, (4.4)$$

where x is the spatial separation of the operators W, V. Here, the contribution of high-spin operators increases with t. In purely CFT terms, this is the reason the ordering in Eq. (4.2) is more interesting than those in Eq. (4.3).

This brings us to the second advantage of 2d CFT: we can sum a particular class of global primaries by studying the Virasoro conformal block of the identity operator. As we will see, this reproduces an AdS<sub>3</sub> bulk calculation in the style of [1, 44, 54], and it provides a purely field-theoretic window into fast scrambling [38–41]. However, we emphasize that contributions of operators with high spin are important. Even with the usual assumption of a large gap in operator dimensions, restricting to Virasoro descendants of the identity becomes a bad approximation at sufficiently large t. In the bulk, what is left out are stringy corrections to the Einstein gravity computation [57]. Still, the Einstein gravity results provided a useful starting point for the investigation of chaos using holography, and we hope that the Virasoro identity block will provide a similarly useful starting point in conformal field theory.

The plan of the paper is as follows. In § 4.2.1, we will set up the configuration of operators and review the Euclidean conformal block expansion. In § 4.2.2, we will discuss the analytic continuations required to obtain Eq.s (4.2) and (4.3) from the Euclidean correlator. We will also review the "second sheet OPE," following [142]. In § 4.2.3, we will study the contribution of the stress tensor (and its Virasoro module) in detail. By analytically continuing Virasoro conformal blocks, we will reproduce a special case of the gravity calculations that demonstrated the butterfly effect in [1, 44, 54]. For the convenience of the reader, this calculation is spelled out in detail in appendix 4.4.1.

Before we begin, we will clarify one further point. We have advertised the behavior of the correlation function (4.2) as a diagnostic of quantum chaos. We should contrast this with the expected behavior in an integrable system. In such a system, certain out-of-time order correlation functions of the type (4.2) might tend to zero for large time t. However, this should not happen for *all* pairs of operators W, V. In appendix 4.4.2 we demonstrate this in a familiar integrable system: the two dimensional Ising model.

## 4.2 CFT calculations

#### 4.2.1 Conventions and review

In this paper, we will study thermal four-point correlation functions of W and V, Eq.s (4.2) and (4.3). Eventually, these operators will be arranged in the timelike configuration shown



Figure 4-1: Left: the spacetime arrangement of the W and V operators. Right: their locations after the conformal mapping, viewed in the Rindler patch on the boundary of AdS<sub>3</sub> (grey) covered by x, t. The union of the grey and yellow regions are the Poincare patch covered by  $z, \bar{z}$ .

in Fig. 4-1, where V is at the origin, and W is at position t > x > 0. However, we will obtain these correlation functions by starting with the Euclidean correlator and analytically continuing. We will work exclusively in the setting of two-dimensional conformal field theory. This allows us to map the thermal expectation values to vacuum expectation values through the conformal transformation

$$z(x,t) = e^{\frac{2\pi}{\beta}(x+t)}, \qquad \bar{z}(x,t) = e^{\frac{2\pi}{\beta}(x-t)}.$$
 (4.5)

Here, x, t are the original coordinates on the spatially infinite thermal system and  $z, \bar{z}$  are coordinates on the vacuum system. Explicitly,

$$\langle \mathcal{O}(x,t)...\rangle_{\beta} = \left(\frac{2\pi z}{\beta}\right)^{h} \left(\frac{2\pi \bar{z}}{\beta}\right)^{\bar{h}} \langle \mathcal{O}(z,\bar{z})...\rangle, \tag{4.6}$$

where  $h, \bar{h}$  are the conformal weights of the  $\mathcal{O}$  operator, related to the dimension and spin by  $\Delta = h + \bar{h}$  and  $J = h - \bar{h}$ . On the left hand side, we have a thermal expectation value, at inverse temperature  $\beta$ , and on the right hand side we have a vacuum expectation value on the  $z, \bar{z}$  space. It is common to work with units in which  $\beta = 2\pi$ , but we prefer to keep the  $\beta$  dependence explicit.

It will be essential in this paper to study correlation functions with operators at complexified times  $t_i$ . In our convention, real t corresponds to Minkowski time, and imaginary t corresponds to Euclidean time. Notice from (4.5) that  $\bar{z}$  is the complex conjugate of z only if the time t is purely Euclidean. In order to make contact with standard CFT formulas for the four point function, we will begin with a purely Euclidean arrangement of the operators. This means a choice of  $z_1, \bar{z}_1, ..., z_4, \bar{z}_4$  with  $\bar{z}_i = z_i^*$ . With such a configuration, the ordering of the operators is unimportant, and global conformal invariance on the  $z, \bar{z}$  plane implies that the four point function can be written

$$\langle W(z_1, \bar{z}_1) W(z_2, \bar{z}_2) V(z_3, \bar{z}_3) V(z_4, \bar{z}_4) \rangle = \frac{1}{z_{12}^{2h_w} z_{34}^{2h_v}} \frac{1}{\bar{z}_{12}^{2\bar{h}_w} \bar{z}_{34}^{2\bar{h}_v}} f(z, \bar{z}).$$
(4.7)

in terms of a function f of the conformally invariant cross ratios

$$z = \frac{z_{12}z_{34}}{z_{13}z_{24}}, \qquad \bar{z} = \frac{\bar{z}_{12}\bar{z}_{34}}{\bar{z}_{13}\bar{z}_{24}}.$$
(4.8)

According to the general principles of CFT, we can expand f as a sum of global conformal blocks, explicitly [144]

$$f(z,\bar{z}) = \sum_{h,\bar{h}} p(h,\bar{h}) \ z^{h} \bar{z}^{\bar{h}} \ F(h,h,2h,z) F(\bar{h},\bar{h},2\bar{h},\bar{z})$$
(4.9)

where F is the Gauss hypergeometric function, the sum is over the dimensions of global SL(2) primary operators, and the constants p are related to OPE coefficients  $p(h, \bar{h}) = \lambda_{WW\mathcal{O}_{h,\bar{h}}} \lambda_{VV\mathcal{O}_{h,\bar{h}}}.$ 

#### 4.2.2 Continuation to the second sheet

In order to apply the above formulas to the correlators (4.2) and (4.3), we need to understand how to obtain them as analytic continuations of the Euclidean four point function. That this is possible follows from the fact that all Wightman functions are analytic continuations of each other.<sup>3</sup> The procedure involves three steps: first, one starts with the Euclidean function, assigning small and different imaginary times  $t_j = i\epsilon_j$  to each of the operators. Second, with the imaginary times held fixed, one increases the real times of the operators to the desired Lorentzian values. Finally, one smears the operators in real time and then takes the imaginary times { $\epsilon_i$ } to zero.<sup>4</sup> The result will be a Lorentzian correlator ordered such that the leftmost operator corresponds to the smallest value of  $\epsilon$ , the second operator

<sup>&</sup>lt;sup>3</sup>Theorem 3.6 of [145].

<sup>&</sup>lt;sup>4</sup>In fact, we will omit this final step in this paper. However, we will omit it consistently on both sides of the bulk and boundary calculations that we are comparing.



Figure 4-2: The paths taken by the cross ratio z during the continuations corresponding to (from left to right)  $\langle WVWV \rangle$ ,  $\langle WWVV \rangle$ , and  $\langle WVVW \rangle$ . Only in the first case does the path pass around the branch point at z = 1. Our choice of t > x > 0 breaks the symmetry between z and  $\bar{z}$ , and the corresponding figures for  $\bar{z}$  are very boring: the  $\bar{z}$  coordinate never circles one.

corresponds to the second smallest, and so on.

This elaborate procedure is necessary because Eq. (4.7) is a multivalued function of the independent complex variables  $\{z_i, \bar{z}_i\}$ . The interesting multivaluedness (for our purposes) comes from  $f(z, \bar{z})$ . By crossing symmetry, this function is single valued on the Euclidean section  $\bar{z} = z^*$ , but it is multivalued as a function of independent z and  $\bar{z}$ , with branch cuts extending from one to infinity. Different orderings of the W, V operators correspond to different sheets of this function. To determine the correct sheet, we must assign  $i\epsilon$ 's as above, and follow the path of the cross ratios, watching to see if they pass around the branch loci at z = 1 and  $\bar{z} = 1$ .

To carry this out directly, we write

$$z_1 = e^{\frac{2\pi}{\beta}(t' + i\epsilon_1)} \qquad \bar{z}_1 = e^{-\frac{2\pi}{\beta}(t' + i\epsilon_1)}$$
(4.10)

$$z_2 = e^{\frac{2\pi}{\beta}(t' + i\epsilon_2)} \qquad \bar{z}_2 = e^{-\frac{2\pi}{\beta}(t' + i\epsilon_2)}$$
(4.11)

$$z_{3} = e^{\frac{2\pi}{\beta}(x+i\epsilon_{3})} \qquad \bar{z}_{3} = e^{\frac{2\pi}{\beta}(x-i\epsilon_{3})}$$
(4.12)

$$z_4 = e^{\frac{2\pi}{\beta}(x+i\epsilon_4)} \qquad \bar{z}_4 = e^{\frac{2\pi}{\beta}(x-i\epsilon_4)}$$
(4.13)

as a function of the continuation parameter t'. When t' = 0, we have a purely Euclidean correlator, on the principal sheet of the function  $f(z, \bar{z})$ . When t' = t > x, we have an arrangement of operators as shown in Fig. 4-1.

The cross ratios  $z, \bar{z}$  are determined by these coordinates as in Eq. (4.8). Their paths, as a function of t', depend on the ordering of operators through the associated  $i\epsilon$  prescription. Representative paths for the three cases of interest are shown in Fig. 4-2. The variable  $\bar{z}$  never passes around the branch point at one, and the z variable does so only in the case corresponding to WVWV.

In the final configuration with t' = t, the cross ratios are small. For  $t \gg x$ , we have

$$z \approx -e^{\frac{2\pi}{\beta}(x-t)} \epsilon_{12}^* \epsilon_{34} \qquad \bar{z} \approx -e^{-\frac{2\pi}{\beta}(x+t)} \epsilon_{12}^* \epsilon_{34}, \tag{4.14}$$

where we introduced the abbreviation

$$\epsilon_{ij} = i(e^{\frac{2\pi}{\beta}i\epsilon_i} - e^{\frac{2\pi}{\beta}i\epsilon_j}). \tag{4.15}$$

For the orderings WWVV and WVVW, no branch cuts are crossed, so the limit of small cross ratios can be taken on the principal sheet of Eq. (4.9). The contribution from the identity operator dominates, verifying our statement in the Introduction that both  $\langle W(t)VVW(t)\rangle_{\beta}$  and  $\langle W(t)W(t)VV\rangle_{\beta}$  approach  $\langle WW\rangle\langle VV\rangle_{\beta}$  for large t.

The case corresponding to WVWV is more interesting. Here, z passes around the branch point at one. The hypergeometric function F(a, b, c, z) has known monodromy around z = 1, returning to a multiple of itself, plus a multiple of the other linearly independent solution to the hypergeometric equation,  $z^{1-c}F(1 + a - c, 1 + b - c, 2 - c, z)$ . For small  $z, \bar{z}$  we then have

$$f(z,\bar{z}) \approx \sum_{h,\bar{h}} \tilde{p}(h,\bar{h}) z^{1-h} \bar{z}^{\bar{h}}, \qquad (4.16)$$

where  $\tilde{p}$  has been defined to absorb the transformation coefficient. On this sheet, as  $z, \bar{z}$  become small, global primaries with large spin become important. As a function of x, t, individual terms in this sum grow like  $e^{(h-\bar{h}-1)t}e^{-(h+\bar{h}-1)x}$ . For sufficiently large t, this sum diverges, and it must be defind by analytic continuation. In other words, we must do the sum over  $h, \bar{h}$  before we continue the cross ratios. In a CFT dual to string theory in AdS<sub>3</sub>, we expect this divergence even at a fixed order in the large c expansion, because of the sum over higher spin bulk exchanges. The implications of this stringy physics for boundary CFT chaos is under investigation in [57].

#### 4.2.3 The Virasoro identity block

We will make a few remarks about the inclusion of other operators in the Discussion. However, the primary focus of this paper is to reproduce the Einstein gravity calculation of the correlation function. In the bulk, this calculation is done by studying free propagation on a shock wave background, which implicitly sums an infinite tower of ladder exchange diagrams. In the CFT, these diagrams are related to terms involving powers and derivatives of the stress tensor in the OPE representation of the four point function. In a two dimensional CFT, all such terms can be treated simultaneously using the Virasoro conformal block of the identity operator, itself an infinite sum of SL(2) conformal blocks. Including these terms (and only these!) in the OPE amounts to replacing

$$f(z,\bar{z}) \to \mathcal{F}(z)\bar{\mathcal{F}}(\bar{z}).$$
 (4.17)

where  $\mathcal{F}$  is the Virasoro conformal block with dimension zero in the intermediate channel.

The function  $\mathcal{F}$  is not known explicitly, but there are several methods for approximating it. We will use a formula from [139], which is valid at large c, with  $h_w/c$  fixed and small, and  $h_v$  fixed and large. In our notation, the formula reads

$$\mathcal{F}(z) \approx \left(\frac{z}{1 - (1 - z)^{1 - 12h_w/c}}\right)^{2h_v}.$$
(4.18)

This function has a branch point at z = 1, as expected. Following the contour around z = 1and taking z small, we find

$$\mathcal{F}(z) \approx \left(\frac{1}{1 - \frac{24\pi i h_w}{cz}}\right)^{2h_v}.$$
(4.19)

The trajectory of  $\bar{z}$  does not circle the branch point at  $\bar{z} = 1$ , so for small  $\bar{z}$ , we simply have  $\bar{\mathcal{F}}(\bar{z}) \approx 1$ , the contribution of the identity operator itself. Substituting (4.19) in (4.17) and then in (4.7), we find

$$\frac{\langle W(t+i\epsilon_1)V(i\epsilon_3)W(t+i\epsilon_2)V(i\epsilon_4)\rangle_{\beta}}{\langle W(i\epsilon_1)W(i\epsilon_2)\rangle_{\beta} \langle V(i\epsilon_3)V(i\epsilon_4)\rangle_{\beta}} \approx \left(\frac{1}{1+\frac{24\pi i\,h_w}{\epsilon_{12}^*\epsilon_{34}}}e^{\frac{2\pi}{\beta}(t-t_*-x)}\right)^{2h_v}$$
(4.20)

where we define the scrambling time  $t_*$  with the convention

$$t_* = \frac{\beta}{2\pi} \log c. \tag{4.21}$$

Eq. (4.20) is the main result of our paper. The essential point is that as t increases beyond  $t_* + x$ , the correlation function starts to decrease, in keeping with expectations from [1, 44, 54]. In Appendix 4.4.1, we calculate the above correlation function using the gravitational shock wave techniques of the above-cited papers, finding precise agreement. We will finish this section with several comments about this formula and extensions.

#### 4.2.4 Various comments

#### The $\epsilon_{12}\epsilon_{34}$ factors

If we take the  $\epsilon_j$  parameters to zero, the denominator in (4.21) diverges. This because, in this limit, we have two pairs of coincident operators on the LHS, and the expectation value is dominated by the high energy components of the operators. The function tends to zero because the high energy components disturb the state violently and decorrelate faster. We can avoid this divergence by smearing W and V over an interval of Lorentzian time Lbefore taking the  $\epsilon$  parameters to zero. This procedure will replace the  $\epsilon_{12}\epsilon_{34}$  factors with  $\sim L^2$ . Instead of doing this explicitly, we will leave the  $\{\epsilon_j\}$  finite. This is another way of cutting off the high energy components, essentially inserting  $e^{-\epsilon H}$  between each operator. Either way,  $\epsilon_{12}\epsilon_{34}h_wh_v$  should be understood as proportional to  $\beta^2 E_w E_v$ , the product of the energy scales of the operators.

#### Validity of the block formula

The derivation of (4.18) in [139] is valid for large but fixed  $h_v$ , in the limit of large c with  $h_w/c$  fixed and small. It is accurate uniformly in z, including on the second sheet. With this scaling, W is a heavy operator. In the context of quantum chaos, we are interested in a somewhat different limit, with light operators and long times:  $h_w, h_v$ , and cz fixed (on the second sheet) as  $c \to \infty$ . The distinction is subtle but important. With  $h_w/c$  fixed, the correlation function (4.20) becomes affected at a time  $t \sim \frac{\beta}{2\pi} \log c/h_w$ , which is order one as a function of c. On the other hand, with  $h_w$  fixed, this timescale grows with c.

Based on agreement with bulk calculations, it seems that formula (4.20) is actually still

valid in this scaling, provided that  $h_w \gg h_v \gg 1$ . It would be good to derive this fact directly. In the meantime, we note the SL(2) blocks are enough to show that (for fixed  $h_w, h_v$ , large c) the time until the correlator is affected is of order  $t_* = \frac{\beta}{2\pi} \log c$ . This criterion comes from the contribution of the stress tensor to (4.16), which is proportional to  $h_v h_w z/c \sim h_v h_w e^{\frac{2\pi}{\beta}(t-t_*-x)}$ .

#### Convergence in 1/c

Notice that the power series in 1/c diverges when the second term in the denominator becomes equal to one. This is for  $t \approx x + t_*$ . However, the correlation function is perfectly well behaved at that point. This illustrates the importance of doing the sum over  $h, \bar{h}$  before continuing onto the second sheet.

#### Asymmetry in $h_w$ and $h_v$

The expression on the LHS is basically symmetric under interchanging W and V, but our answer on the RHS is definitely not. This is because we have assumed that  $h_w \gg h_v$ . We assumed this for purely technical reasons—the Virasoro identity block is not known in the limit of interest if  $h_w$  and  $h_v$  are comparable. Using holography, we can write the answer for general  $h_w, h_v$  as an integral over momenta and transverse coordinates. We give this expression in (4.51) of appendix (4.4.1). This constitutes a prediction for the behavior of the Virasoro identity block at small values of z on the second sheet. It would be interesting to check this directly in conformal field theory.

#### A two-sided configuration

Although the correlation function (4.20) is a purely 'one-sided' quantity, defined with respect to the thermal expectation value, its behavior is closely related to the disruption of 'twosided' entanglement of the thermofield double state, as studied in [44, 54, 55].

The thermofield double state is a purification of the thermal ensemble we have been considering, given by

$$|TFD\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_{i} e^{-\beta E_i/2} |i\rangle_L |i\rangle_R, \qquad (4.22)$$

where two noninteracting copies of the system are denoted L and R, and  $Z(\beta)$  is the thermal partition function of an individual system. Given an operator O, we define  $O_R = 1 \otimes O$  acting on the R system and  $O_L = O^T \otimes 1$  acting on the L system. The thermofield double state has a high degree of local correlation between these pairs. As shown in [44], a small perturbation to one of the systems at an early time will disrupt the local entanglement. We can diagnose this by studying correlation functions in the perturbed state  $|\psi\rangle = W(t)_R |TFD\rangle$ 

$$\langle \psi | V_L V_R | \psi \rangle = \langle TFD | W(t)_R V_L V_R W(t)_R | TFD \rangle.$$
(4.23)

It is easy to check that this correlation function is related to a one-sided correlation function of the type (4.2) by analytic continuation,

$$\langle \psi | V_L V_R | \psi \rangle = \langle W(t) V W(t) V(i\beta/2) \rangle_{\beta}. \tag{4.24}$$

We can make this continuation directly in (4.20) by setting  $\epsilon_4 = \beta/2$  and  $\epsilon_3 = 0$ . Then  $\epsilon_{34} \rightarrow 2i$ . Again, as the time t increases past  $t_* + x$ , the two-sided correlation begins to decrease exponentially.

#### Commutators and growth of operators

The correlation function (4.20) is closely related to the square of the commutator

$$\langle [W(t), V]^2 \rangle_{\beta}.$$
 (4.25)

As the correlation function (4.20) becomes small, the commutator becomes large, of order  $\langle WW \rangle \langle VV \rangle$ . In [1], the behavior of this commutator was used to determine the growth of a precursor operator such as W(t). Here, we think about V as a probe to determine the size of W(t). Translated in terms of the correlator (4.20), the definition of [1] is that the size of W(t) is the volume of the region (as a function of the location of V) in which the correlator is less than one half. This volume is a ball, and radius of the operator r[W(t)] is defined as the radius of that ball. For d dimensional CFTs dual to Einstein gravity, [1, 44] found that the ball has radius

$$r[W(t)] \approx v_B(t - t_*), \qquad v_B = \sqrt{\frac{d}{2(d-1)}}.$$
 (4.26)

Specializing to d = 2, we find  $v_B = 1$ , in agreement with (4.20).

#### Two-sided entropy

Another two-sided quantity of interest is the entropy  $S_{AB}$  between two matching semiinfinite intervals, one on each boundary. Apart from the usual UV divergence near the boundary of the intervals, this entropy is finite. In other words, there is no IR divergence despite the infinite length of the intervals. As pointed out in [85], this is related to the local entanglement of the two sides of the thermofield state.

As discussed above, applying a small local perturbation at an early time will disrupt this entanglement and cause the entropy  $S_{AB}$  to grow. Using the gravity techniques of [1, 44], one can check that the mutual information decreases linearly with the time of this perturbation, t, as (for large t)

$$\Delta S_{AB} = \frac{\pi c}{3\beta} (|t| - t_* - |x|) + (const), \qquad (4.27)$$

where the constant depends on the smearing and dimension of the perturbing operator, and where x is the distance outside the interval that the perturbation was applied. In this paper, we chose to focus on four point functions instead of entropy, but the Virasoro identity block can also be used to calculate the above, following very closely the techniques in [140] and the conformal mappings used in this section. We have verified that the answer matches.<sup>5</sup>

#### Scrambling in the vacuum

In this paper, we have focused on thermal expectation values. It is interesting to compare to similar correlation functions, evaluated in the vacuum. Taking limit of (4.10)-(4.13) for infinite  $\beta$ , the conformal cross ratio for  $t \gg x$  becomes  $z \approx -(\epsilon_1 - \epsilon_2)(\epsilon_3 - \epsilon_4)/(t - x)^2$ . In this configuration, the contour still crosses the branch cut, and we find

$$\langle WVWV \rangle \propto \left( \frac{1}{1 + \frac{24\pi i h_w (t-x)^2}{c(\epsilon_1 - \epsilon_2)(\epsilon_3 - \epsilon_4)}} \right)^{2h_v},$$
(4.28)

where in this case the expectation is taken in the vacuum state.

We can understand this result as a type of "slow scrambling" that happens in an extremely low temperature state on an infinite line, with  $\beta \gg \sqrt{cL}$ , where L is the smearing

 $<sup>{}^{5}</sup>$ A similar problem was recently studied by [146], and a linear increase in the entropy was found, but the formula differs significantly from (4.27).

scale of the operators. In that case, for  $L\sqrt{c} < t < \beta$ , the correlation will decrease with power law dependence as  $\sim t^{-4h_v}$ . After t becomes comparable to  $\beta$ , the dependence will be given by (4.20) and will decrease exponentially.

### 4.3 Discussion

Several manifestations of chaos in quantum field theory have recently been discussed: disruption of atypical correlations by small perturbations, [44, 55], large commutators [45, 54], linear growth of precursor operators [1], —all of which are ultimately connected to an outof-time-order correlation function of the form  $\langle W(t)VW(t)V\rangle_{\beta}$ . Previous studies have relied on holography. In this paper, we showed how to obtain the relevant correlation function starting from the Euclidean correlator, following [141] and [138]. We then used an approximation to the Euclidean correlation function, the Virasoro identity block [138, 139], to compute the out-of-time order correlation function. We found agreement with a generalization of the original holographic calculation of [44].

However, it is clear both from the perspective of holography and conformal field theory that there is more to the correlation function than our main result, (4.20). In the bulk, we see this in the form of stringy corrections to the correlator (which are important at large t) [57]. In conformal field theory, we have to consider the effect of operators that are not Virasoro descendants of the identity. We can make some progress from the CFT side by considering additional operators in the channel of the four-point function.

To explore these contributions, let us compare the second-sheet contribution of the stress tensor,  $\propto e^{t-x}$ , to the contribution of a general SL(2) primary,  $e^{(h-\bar{h}-1)t-(h+\bar{h}-1)x}$ . At leading order in 1/c, the most important contributions should be from single-trace operators. In a theory with a weakly curved holographic dual, the lowest-dimension single trace operators with spin greater than two should have parametrically large dimension. A useful example to keep in mind is a low n string state on the leading Regge trajectory, with  $h - \bar{h} = n$  and  $h + \bar{h} \sim \sqrt{n} m_s \ell_{AdS}$ . The contribution of this type of operator is suppressed as a function of x, but at sufficiently large t, such that

$$(h - \bar{h} - 2)t > (h + \bar{h} - 2)x, \tag{4.29}$$

it will dominate over the stress tensor. If we fix x, then at  $t_*$ , the above criterion will be

satisfied. This is in keeping with the bulk analysis of [57], which shows that a Regge-type resummation of an infinite number of stringy operators is necessary.

At sufficiently large x, this criterion will not be satisfied (here, we are assuming large  $m_s \ell_{AdS}$ ), even at the relevant scrambling time  $t = t_* + x$ . This indicates a region in which identity block will dominate. We would like to suggest that there is another: very large  $t \gg t_* + x$ . Our reason for this suspicion is the bulk geodesic 'bending' discussed in appendix 4.4.1. For simplicity, let us set  $x \approx 0$ . Then for  $t > t_*$ , the important correlation comes from bulk geodesics that pass through the horizon at transverse coordinate  $|x'| \sim (t - t_*)/2$ . If this x' is large enough to violate (4.29), we expect the extra operator to be subleading compared to the identity.

To put a little flesh on this idea, let us consider the Virasoro block of the  $h, \bar{h}$  operator, rather than the SL(2) block. Using, again, the results from [139] and continuting to the second sheet, we find

$$\frac{\mathcal{F}_h(z)}{\mathcal{F}(z)} \approx \left(\frac{16}{z - \frac{24\pi i h_w}{c}}\right)^h \tag{4.30}$$

where  $\mathcal{F}(z)$  is the identity block (4.20). For x = 0 and  $t > t_*$ , this ratio is proportional to  $e^{\frac{2\pi}{\beta}ht_*}$ . Together with the contribution from the antiholomorphic block, which is  $e^{-\frac{2\pi}{\beta}\bar{h}t}$ , we find that the  $h, \bar{h}$  operator will dominate only if

$$\bar{h}t < ht_*. \tag{4.31}$$

For large  $E = h + \bar{h}$  and to leading order in  $J/E = (h - \bar{h})/(h + \bar{h})$ , this is consistent with the criterion (4.29) for the SL(2) blocks, but evaluated at the 'bent' location  $x' = (t - t_*)/2$ .

Finally, we will comment on two possible extensions that were not indicated above. First, using methods outlined in [138], it seems possible to extend our analysis to multipoint functions dual to the multiple shock wave backgrounds studied in [1, 54]. Second, we note that shock wave geometries and de-correlation have been studied in relation to tensor networks [79] and computational complexity [87, 102, 119, 126]. It would be interesting to translate the Virasoro identity block approximation (and corrections) into the language of those papers.

## 4.4 Appendix

#### 4.4.1 Bulk calculations

In this section we will compute the correlation function (4.20) using the gravitational shock wave methods of [1, 44]. The idea is as follows. If  $h_w \gg h_v \gg 1$ , we can calculate the correlation function by treating the W operator as creating a shock wave, and calculating the two point function of the V operator on that background. Although the calculation can be done for general  $\epsilon_j$ , we will focus on the case where  $\epsilon_1 = -\tau$ ,  $\epsilon_2 = \tau$ ,  $\epsilon_3 = 0$  and  $\epsilon_4 = \beta/2$ , because this corresponds to a two-sided expectation value

$$\frac{\langle \psi | V_L V_R | \psi \rangle}{\langle \psi | \psi \rangle \langle V_L V_R \rangle} \tag{4.32}$$

in the state

$$|\psi\rangle = W_R(t+i\tau)|TFD\rangle. \tag{4.33}$$

Here, the normalization factor  $\langle V_L V_R \rangle$  is evaluated in the unperturbed thermofield double state. Eq. (4.32) is the correlation function studied in [44]. The analysis breaks into two parts: (i) finding the geometry of the shock sourced by W and (ii) computing the correlation function of the V operators in that background.

#### The shock geometry

The metric of a localized shock wave<sup>6</sup> in 2+1 dimensional AdS Rindler space, is [1, 44, 109, 121, 141]

$$ds^{2} = -\frac{4}{(1+uv)^{2}}dudv + \frac{(1-uv)^{2}}{(1+uv)^{2}}dx^{2} + 4\delta(u)h(x)du^{2},$$
(4.34)

where h will be defined below. Schwarzschild (r, t) are related to the Kruskal (u, v) by

$$r = \frac{1 - uv}{1 + uv}, \qquad t = \frac{1}{2}\log\left(-\frac{v}{u}\right).$$
 (4.35)

<sup>&</sup>lt;sup>6</sup>Although we refer to these geometries as shock waves, the terminology is somewhat misleading in 2+1 dimensions, since the geometry is locally pure AdS<sub>3</sub> away from the source.

The inverse temperature  $\beta$  is  $2\pi$ . We will consider a shock sourced by a stress tensor

$$T_{uu}(u', v', x') = P\,\delta(u')\delta(x' - x), \tag{4.36}$$

appropriate for a particle sourced by the W operator, traveling along the u = 0 horizon at transverse position x. The constant P is related to the momentum  $p^v$  of the particle. The metric (4.34) can be understood as two halves of AdS-Rindler, glued together at u = 0 with a shift

$$\delta v(x) = h(x) \tag{4.37}$$

in the v direction. Plugging in to Einstein's equations, we determine h as

$$h(x') = 2\pi G_N P e^{-|x'-x|}.$$
(4.38)

To relate this geometry to the state  $|\psi\rangle$ , we need to fix P. In other words, we need to evaluate

$$\frac{\langle \psi | \int dx' du' \ T_{uu} | \psi \rangle}{\langle \psi | \psi \rangle},\tag{4.39}$$

where we take the integral to run over the slice v = 0. We will assume that W is dual to a single-particle operator in the bulk, so that the state  $|\psi\rangle$  can be described by a Klein-Gordon wave function K. This wave function is a bulk-to boundary propagator from the location (x, t) of the W operator. It is given in terms of the regularized geodesic distance dfrom the boundary point, as  $K \propto (\cosh d)^{-2h_w}$ . At v = 0, we find

$$K(t, x; u', x') = \frac{\mathcal{N}}{[e^t u' + \cosh(x - x')]^{2h_w}}.$$
(4.40)

 $\mathcal{N}$  is a normalization that will ultimately cancel, so we will set it to one.

We will evaluate the denominator of (4.39) first. The norm  $\langle \psi | \psi \rangle$  is a Klein-Gordon inner product

$$\langle \psi | \psi \rangle = 2i \int dx' du' \ K(t + i\tau, x; u', x')^* \partial_{u'} K(t + i\tau, x; u', x').$$
(4.41)

The u' integral can be done using contour integration, and the x' integral can be done in

terms of  $\Gamma$  functions:

$$\langle \psi | \psi \rangle = \frac{4\pi^{3/2}}{(2\sin\tau)^{4h_w}} \frac{\Gamma[4h_w]}{\Gamma[2h_w]\Gamma[2h_w + \frac{1}{2}]}.$$
(4.42)

For the numerator, the stress tensor for the Klein-Gordon field is given by the expression  $T_{uu} = \partial_u \varphi \partial_u \varphi$ . Contracting bulk operators with boundary operators using K, we have

$$\langle \psi | \int dx' du' \ T_{uu} | \psi \rangle = 2 \int dx' du' \ \partial_{u'} K(t_w + i\tau, x; u', x')^* \partial_{u'} K(t_w + i\tau, x; u', x'), \quad (4.43)$$

where the factor of 2 comes from the two different ways of doing the contractions. The integrals can be done the same way as before:

$$\langle \psi | \int dx' du' \ T_{uu} | \psi \rangle = \frac{8\pi^{3/2} e^{t_w}}{(2\sin\tau)^{4h_w+1}} \frac{\Gamma[4h_w]\Gamma[2h_w + \frac{1}{2}]}{\Gamma[2h_w]^3}.$$
 (4.44)

Taking the ratio at large  $h_w$ , we find

$$P = \frac{2h_w e^t}{\sin \tau}.\tag{4.45}$$

#### The two point correlator

The second step, following [44], is to compute the two-sided correlation function of the V operators in this shock background. We will do this using the geodesic approximation

$$\langle \psi | V_L V_R | \psi \rangle \propto e^{-md},$$
 (4.46)

where d is the regularized geodesic distance, and the mass m is approximately  $2h_v$ , the conformal weight of V.

The relevant geodesic passes from t = 0 and x = 0 on the left boundary to t = 0 and x = 0 on the right. As in [44], we first compute the length of a geodesic from the left boundary to an intermediate point on the u = 0 surface (v', x'), and then we add it to the length from that point to the right boundary. These distances can be worked out from the embedding coordinates:

$$d = 2\log 2r_{\infty} + \log\left[\cosh x' - v'\right] + \log\left[\cosh x' + v' + h(x')\right],$$
(4.47)

where  $r_{\infty}$  is the cutoff at the boundary of AdS, and h(x) is given by (4.38). The true

geodesic distance is given by extremizing this sum over the intermediate point. This gives

$$d = 2\log 2r_{\infty} + \log\left[1 + h(0)\right].$$
 (4.48)

Note the difference from the behavior  $2\log(1 + h/2)$  in the spherically symmetric setting of [44]. At large h(0), the distance is half of what one might naively expect based on the spherical problem. The reason is that, for  $t > t_* + x$ , the geodesic bends away from the source of the shock, passing through at transverse position

$$x' \approx \frac{t_* + x - t}{2} < 0, \tag{4.49}$$

here x > 0 is the transverse position of the source of the shock.

After subtracting the divergent distance in the unperturbed thermofield double state,  $d_{TFD} = 2 \log 2r_{\infty}$ , we plug the distance into (4.46), finding

$$\frac{\langle \psi | V_L V_R | \psi \rangle}{\langle \psi | \psi \rangle \langle V_L V_R \rangle} = \left(\frac{1}{1 + \frac{4\pi G_N h_w}{\sin \tau} e^{t-x}}\right)^{2h_v}.$$
(4.50)

This agrees with (4.20) after (i) using  $G_N = 3/2c$  to express the gravitational constant in terms of the central charge, (ii) plugging in  $\epsilon_1 = -\tau$ ,  $\epsilon_2 = \tau$ ,  $\epsilon_3 = 0$ ,  $\epsilon_4 = \beta/2$  and (iii) using  $\beta = 2\pi$ .

#### Generalization

The case with general  $\epsilon_j$  and general  $h_w$ ,  $h_v$  is analyzed in [57]. The calculation is similar to the above, but includes an integral over the momentum of the shock, weighted by wave functions associated to the W operator. We will not give the details of this calculation here, but we will note that it can be used to give a prediction for the behavior of the Virasoro conformal block  $\mathcal{F}$ 

$$\mathcal{F}(z) = c_0^2 \int_0^\infty dp \, dq \int_{-\infty}^\infty dx \, dy \, p^{4h_w - 1} q^{4h_v - 1} e^{-p \cosh x} e^{-q \cosh y} e^{3\pi i p q \, e^{x + y}/cz}.$$
 (4.51)

where the normalization constant ensures that  $\mathcal{F}(z) = 1$  when  $c = \infty$ :

$$c_0 = \frac{2^{1-2h_w - 2h_v}}{\Gamma[2h_w]\Gamma[2h_v]}.$$
(4.52)

We remind the reader that this expression is on the second sheet, where z has passed around the branch point at z = 1. For  $h_w \gg h_v \gg 1$ , one can check that this reduces to (4.20). However, the formula should be accurate for all  $h_w, h_v$  in the large c limit with cz fixed. It would be interesting to check this directly in CFT.

#### 4.4.2 Ising model

It is interesting to contrast the behavior of out-of-time-order correlators in a chaotic theory to those in an integrable theory. In a chaotic theory, we expect all thermal correlators of the form  $\langle W(t)VW(t)V\rangle_{\beta}$  should become small for sufficiently large t, regardless of which operators V and W are correlated. This will not be the case in an integrable theory.

As an example, we consider the two-dimensional Ising CFT. This theory has c = 1/2and three Virasoro primary operators: I,  $\sigma$ , and  $\epsilon$ , corresponding to the identity, 'spin,' and 'energy' operators. The different combinations of four-point correlators of these primaries are well known [147–149]. We will present these by giving the function  $f(z, \bar{z})$  in Eq.(4.7) for the three cases:

$$f_{\sigma\sigma}(z,\bar{z}) = \frac{1}{2} \Big( |1 + \sqrt{1-z}| + |1 - \sqrt{1-z}| \Big), \tag{4.53}$$

$$f_{\sigma\epsilon}(z,\bar{z}) = \left|\frac{2-z}{2\sqrt{1-z}}\right|^2,\tag{4.54}$$

$$f_{\epsilon\epsilon}(z,\bar{z}) = \left|\frac{1-z+z^2}{1-z}\right|^2,\tag{4.55}$$

where the subscripts in  $f_{WV}$  denote W and V, and the operators are ordered WVWV with the configuration specified by (4.10)–(4.13). z and  $\bar{z}$  still given by (4.14). Since we treat  $z, \bar{z}$ as independent, the notation |h(z)| should be interpreted as  $\sqrt{h(z)}\sqrt{h(\bar{z})}$ . It's easy to see that the  $\langle \sigma\sigma\sigma\sigma \rangle$  and  $\langle \sigma\epsilon\sigma\epsilon \rangle$  correlators have branch points at z = 1, while  $\langle\epsilon\epsilon\epsilon\epsilon\rangle$  does not. Following the contour across the branch cut for the two correlators that do have a second sheet and taking z small, we find

$$\frac{\langle \sigma \sigma \sigma \sigma \rangle_{\beta}}{\langle \sigma \sigma \rangle_{\beta}^{2}} = 0, \qquad \frac{\langle \sigma \epsilon \sigma \epsilon \rangle_{\beta}}{\langle \sigma \sigma \rangle_{\beta} \langle \epsilon \epsilon \rangle_{\beta}} = -1, \qquad \frac{\langle \epsilon \epsilon \epsilon \epsilon \rangle_{\beta}}{\langle \epsilon \epsilon \rangle_{\beta}^{2}} = 1.$$
(4.56)

Only  $\langle \sigma \sigma \sigma \sigma \rangle$  vanishes at large t.

## Chapter 5

# Scientia ipsa potentia est — 🛷

(Chaos in quantum channels)

In the last chapter we discussed a diagnostic for quantum chaos: out-of-time-order (OTO) correlation functions that decay for almost all choices of operators. However, the phrase "almost all" leaves a little bit to be desired. Does one have to check every correlator to determine whether a theory is chaotic? Shouldn't there be a more information-theoretic definition of chaos?

The goal of this chapter is the pursuit of such a definition. To study chaos in an information-theoretic sense, we will precisely connect it to the related phenomena of "scrambling" [38, 39]. Scrambling is usually phrased in terms of states: can the pure state  $\rho_{AB} = |\psi\rangle\langle\psi|$  on AB be distinguished from a maximally mixed state  $\mathbb{I}_A/|A|$  by local measurement on A alone [37]? For  $\rho_A = \operatorname{tr}_B\{\rho_{AB}\}$ , if

$$\|\rho_A - \frac{\mathbb{I}_A}{|A|}\|_1 \le 0 + O(2^{(S_A - S_B)/2}), \tag{5.1}$$

then the state is said to be "scrambled" [37, 39], where  $\|\cdot\|_1$  is the 1-norm or trace distance. The upshot is that for any subsystem A that contains less than half of the total degrees of freedom  $(S_A < S_B)$  a scrambled state appears "random" and cannot be distinguished by local measurements on A alone.

Intuitively, scrambling captures the idea that the butterfly effect ought to lead to the delocalization of information. However, as phrased it's not obvious how to directly connect

this definition of scrambling to chaos, operator growth, or the decay of OTO correlators. To make the connection, we want to think about scrambling in terms of operators rather than states. Operators that scramble take states with localized information to states with multipartite entanglement. To characterize a unitary operator U(t) with this property, we need to understand the entanglement properties of U(t); we need to think of it as a state

$$U(t) \to |U(t)\rangle.$$
 (5.2)

This mapping will be explained in detail in the chapter. The key point is that by considering the state  $|U(t)\rangle$  we can use machinery of entanglement entropy to study unitary operators—that is, we can study entanglement in "time." Operators that scramble tend to maximize the tripartite information  $I_3$ , which is a known measure of multipartite entanglement. By considering an average over OTO correlators, we are able to find a precise and quantitative connection between our previous definition of chaos and our new understanding of scrambling.

In holography, the state interpretation of a unitary operator discussed in this chapter is not a new idea but a natural extension of ideas that have been around for a while [82– 86, 106]. Interpreted in terms of holographic complexity conjectures such as those discussed in Chapter 3, this suggests a deep relationship between chaos and computation. *Knowledge itself is power*, and the computational "power" of time evolution is related to how efficiently it scrambles or entangles information (or "knowledge").

This chapter serves to bring together the central ideas of this thesis: chaos, computation, and holography. *Summa summarum*, chaos is essential to computation. As extreme of both, black holes are an obvious testing grounds for the pursuit of this relationship between chaos and computation. The question of backing off the extreme limit and making this connection precise is the subject of ongoing research.

#### \*\*\*\*

This chapter was first presented as [4] under the title "Chaos in quantum channels," and is a collaboration with Pavan Hosur, Xiao-Liang Qi, and Beni Yoshida,

JHEP 02 (2016) 004, arXiv:1511.04021 [hep-th].

#### Abstract

We study chaos and scrambling in unitary channels by considering their entanglement properties as states. Using out-of-time-order correlation functions to diagnose chaos, we characterize the ability of a channel to process quantum information. We show that the generic decay of such correlators implies that any input subsystem must have near vanishing mutual information with almost all partitions of the output. Additionally, we propose the negativity of the tripartite information of the channel as a general diagnostic of scrambling. This measures the delocalization of information and is closely related to the decay of out-oftime-order correlators. We back up our results with numerics in two non-integrable models and analytic results in a perfect tensor network model of chaotic time evolution. These results show that the butterfly effect in quantum systems implies the information-theoretic definition of scrambling.

### 5.1 Introduction

Quantum information is processed in quantum circuits, or more generally, quantum channels. A useful way to characterize fault-tolerance and computational power of such channels is by whether input information remains localized or is spread over many degrees of freedom. This delocalization of quantum information by a quantum channel over the entire system is known as scrambling [38–40]. Scrambling implies that information about the input cannot be deduced by any local measurement of the output [37].

The information-theoretic phenomenon of scrambling is closely related to the notion of chaos in thermal systems. A vivid feature of quantum chaos is the butterfly effect: simple localized operators grow under time evolution to have large commutators with almost all other operators in the system. Consider a pair of local Hermitian operators W and Vsupported on non-overlapping subsystems such that [W, V] = 0. Under a chaotic time evolution with Hamiltonian H, a local operator W(0) will evolve into a complicated operator  $W(t) = e^{iHt}We^{-iHt}$  which has an expansion as a sum of products of many local operators

$$W(t) = W + it [H, W] - \frac{t^2}{2!} [H, [H, W]] - \frac{it^3}{3!} [H, [H, [H, W]]] + \dots$$
(5.3)

For a generic H with local interactions, the kth-order nested commutator of H with W can

lead to a product of as many as k local operators that acts non-trivially on a large volume of the system [1]. This implies that  $[W(t), V] \neq 0$  and will generically be a large operator of high weight.

The degree of non-commutativity between W(t) and V can be measured by their group commutator: W(t) V W(t) V. In fact, the generic decay of out-of-time-order (OTO) correlators of the form

$$\langle W(t) V W(t) V \rangle_{\beta} \equiv Z^{-1} \operatorname{tr} \left\{ e^{-\beta H} W(t) V W(t) V \right\},$$
(5.4)

is a distinguishing feature of quantum chaos [1, 3, 44, 54, 57, 58], where  $\beta$  is the inverse temperature, and  $Z = \text{tr} e^{-\beta H}$ . While this definition of chaos is very direct in terms of operators and observables, it should be possible to understand chaos solely as a property of the system itself.

In this paper, we will consider unitary quantum channels as states in order to characterize their ability to process quantum information in terms of entanglement. A unitary quantum channel is simply a unitary circuit where we allow the inputs to be mixed states. To use standard quantum-information measures, we introduce a mapping from a unitary quantum channel to a quantum pure state in a doubled Hilbert space. Using this map, we propose that the tripartite information of a subsystem of the input A and a division of the output CD into two subsystems C and D

$$I_3(A:C:D) = I(A:C) + I(A:D) - I(A:CD),$$
(5.5)

provides a basic measure of scrambling in such channels. The negativity of the tripartite information is a natural measure of multipartite entanglement, and in particular, channels that scramble will have near maximally negative tripartite information for all possible input/output subsystem combinations.

Next, we will use this definition of scrambling to make a direct connection between the OTO correlator diagnostic of chaos (5.4) and the information-theoretic definition of scrambling (5.5). It is intuitive that chaotic time evolution should correspond to scrambling: the growth of the operator W(t) in time means that in order to recover or reconstruct the simple operator W one will need to measure a growing fraction of the degrees of freedom of the system. By averaging OTO correlators of the form (5.4) over a complete basis of operators on an input subsystem A and output subsystem D, we can directly relate these correlators to the second Rényi entropy between subsystems of the input and output

$$|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle| \sim 2^{-S_{AC}^{(2)}}, \tag{5.6}$$

where  $\mathcal{O}_A$  represents an operator in the input A,  $\mathcal{O}_D$  represents an operator in the output D, and  $S_{AC}^{(2)}$  is the second Rényi entropy of the input/output system AC. With this result, we can show that the butterfly effect implies scrambling

$$|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle_{\beta=0}| = \epsilon \implies I_3(A:C:D) \le I_{3,\min} + 2\log_2 \frac{\epsilon}{\epsilon_{\min}}.$$
 (5.7)

Here,  $I_{3, \min}$  is the minimum of  $I_3$ , and  $\epsilon_{\min}$  is the minimum of averaged OTO correlations. We will support all of these results with numerics in two non-integrable models, the onedimensional Ising spin chain with parallel and transverse field, and a four-Majorana-fermion model introduced by Kitaev [77] known to be a fast scrambler. Our results show that for chaotic systems, all of the information-theoretic quantities relevant to scrambling approach their Haar-scrambled values.

Relatedly, we will comment on the relationship between the butterfly effect and the ballistic spreading of information in a channel. For systems with a notion of spatial locality, we will show that the information contained in an input subsystem A will expand ballistically in the output with a characteristic velocity  $v_B$  (usually denoted the "butterfly velocity" [1, 44]). This supports the idea that  $v_B$  is the velocity of information in strongly-chaotic systems. We will also comment on the conceptual differences between the butterfly velocity and the entanglement or tsunami velocity  $v_E$  of [85, 92, 93].

Finally, we will use our enhanced understanding of the relationship between scrambling, chaos, and entanglement in time to propose a solvable model of a unitary quantum channel that exhibits scrambling. Building on the work of [85] and [150], we discuss a perfect tensor model of a chaotic Hamiltonian time evolution. This can be thought of as a toy model for an Einstein-Rosen bridge that connects the two sides of the eternal black hole in AdS.

The plan of this paper is as follows. In §5.2 we discuss unitary quantum channels and elaborate on the notion of entanglement in time. There, we consider the entanglement properties of such channels and introduce the tripartite information as a measure of scrambling. In §5.3, we show that the decay of OTO correlation functions implies strong bounds

on information-theoretic quantities, directly connecting chaos to scrambling. We provide evidence for our claims via numerical studies of qubit systems in §5.4 and with a perfect tensor model of chaotic time evolution in §5.5. We conclude in §5.6 with a discussion of the relationship between chaos and computation. Some extended calculations, tangential discussions, and lengthy definitions are left to the Appendices.

## 5.2 Unitary quantum channels

To study the scrambling properties of different unitary operators U by using informationtheoretic quantities, we will interpret them as states. To be concrete, let us assume that the quantum system consists of n qubits with a time independent Hamiltonian H. We will be interested in a particular unitary operator, the time evolution operator  $U(t) = e^{-iHt}$ . This will let us study a one parameter family of unitary operators indexed by t.

A unitary operator U(t) that acts n qubits is described by a  $2^n \times 2^n$  dimensional matrix

$$U(t) = \sum_{i,j=0}^{2^n - 1} u_{ij} |i\rangle \langle j|.$$
 (5.8)

which we usually choose to think of in terms of a tensor with n input and n output legs, as shown in Fig. 5-1(a). However, it is also natural to map this to a 2n-qubit state by treating the input and output legs on equal footing

$$|U(t)\rangle = \frac{1}{2^{n/2}} \sum_{i,j=0}^{2^n - 1} u_{ji} |i\rangle_{in} \otimes |j\rangle_{out}.$$
(5.9)

This is depicted graphically in Fig. 5-1(b). Clearly  $|U(t)\rangle$  encodes all the coefficients  $(u_{ij})$  necessary to represent the unitary operator U(t).

When t = 0, the unitary operator is simply the identity operator, and (5.8) reduces to a state consisting of n EPR pairs

$$|I\rangle = \frac{1}{2^{n/2}} \sum_{i,j=0}^{2^n - 1} |i\rangle_{in} \otimes |i\rangle_{out}, \qquad (5.10)$$

where each input leg is maximally entangled with each output leg, and there is no entanglement between different EPR pairs. For finite t, inserting a complete set of states into



Figure 5-1: Interpretations of a unitary channel: (a) a unitary operator U with input and output legs. (b) state interpretation  $|U\rangle$  of the unitary operator U. By bending the input legs down, we treat input/output equally. (c) the state interpretation is equivalent to the creation of a maximally entangled state followed by acting with U on half the EPR pairs, which gives  $|U\rangle$ .

(5.9) and using (5.8), we can rewrite it as

$$|U(t)\rangle = \frac{1}{2^{n/2}} \sum_{i,j=0}^{2^n-1} |i\rangle_{in} \otimes U(t)|i\rangle_{out} = \mathbb{I} \otimes U(t)|I\rangle,$$
(5.11)

with I the identity operator acting on the incoming states. This offers the following interpretation to the state  $|U(t)\rangle$ : a maximally entangled state is created between a reference system "in" and a system of interest "out." Next, the operator U acts on "out," giving  $|U(t)\rangle$ . This perhaps offers a more physical interpretation of this operator-state mapping, as shown in Fig. 5-1(c).

**Note:** in this paper we will adopt the perspective of the mapping shown in Fig. 5-1(b) and use language that refers to the entanglement properties of channels (in time) between subsystems of inputs and outputs. Additionally, we will always draw our channels from the operator perspective as in Fig. 5-1(a).

It is natural to ask whether the choice of maximally entangled state  $|I\rangle$  is artificial. Although different choices of initial state  $|I\rangle$  can be made which define different mappings from U to  $|U(t)\rangle$ , all our discussions remain insensitive to the choice as long as  $|I\rangle$  is a direct product of EPR pairs. The two qubits in each pair are required to be the qubits at a given real-space position in the input and output systems, which are maximally entangled with each other. This choice guarantees that all quantum entanglement between different real-space locations in  $|U(t)\rangle$  are created by the unitary evolution U(t). The operator-state mapping can be further generalized by considering a more generic statistical ensemble as the input state. Let  $\{|\psi_j\rangle\}$  be a set of orthonormal states, and imagine that we input an initial state  $|\psi_j\rangle$  with probability  $p_j$  to a unitary quantum channel U. This means that the initial statistical ensemble is  $\rho_{in} = \sum_j p_j |\psi_j\rangle \langle \psi_j|$ . After time evolution, each input state evolves to  $|\phi_j\rangle = U|\psi_j\rangle$ , and the output statistical ensemble is given by  $\rho_{out} = \sum_j p_j |\phi_j\rangle \langle \phi_j|$ . The time evolution of a given input ensemble  $\rho_{in}$  can be mapped to the following pure state

$$|\Psi\rangle = \sum_{j} \sqrt{p_j} \, |\psi_j\rangle_{in} \otimes |\phi_j\rangle_{out} = \mathbb{I} \otimes U(t) \sum_{j} \sqrt{p_j} \, |\psi_j\rangle_{in} \otimes |\psi_j\rangle_{out}.$$
(5.12)

The isomorphic state  $|\Psi\rangle$  contains all the information required to characterize the properties of the channel. Namely, if one traces out the output system, then the reduced density matrix is the input state ( $\rho_{in} = \operatorname{tr}_{out} |\Psi\rangle\langle\Psi|$ ) while if one traces out the input system, then the reduced density matrix is the output state ( $\rho_{out} = \operatorname{tr}_{in} |\Psi\rangle\langle\Psi|$ ). The state interpretation in (5.11) corresponds to the special case of a uniform input ensemble (i.e.  $\rho_{in} = 2^{-n}\mathbb{I}$ ). In general, we will simply refer to the state given in (5.12) as a unitary quantum channel. In quantum information theory, such correspondence between quantum channels and quantum states is named as the channel-state duality.<sup>1</sup>

A familiar example of a unitary channel is the thermofield double state

$$|TFD\rangle = \frac{1}{\sqrt{Z}} \sum_{i} e^{-\beta E_i/2} |i\rangle \otimes |i\rangle, \qquad (5.13)$$

where  $E_i$  are eigenvalues of the Hamiltonian H and  $Z = \text{tr} e^{-\beta H}$ . Applying evolution for time t to the right output system, one obtains the following time evolved state

$$|TFD(t)\rangle = \frac{1}{\sqrt{Z}} \sum_{i} e^{-\beta E_i/2} e^{-iE_i t} |i\rangle \otimes |i\rangle.$$
(5.14)

One can interpret this state as a quantum channel  $U = e^{-iHt}$  whose input is given by the thermal ensemble. Note this expression reduces to (5.11) for  $\beta = 0$ .

<sup>&</sup>lt;sup>1</sup>In fact, the channel-state duality in quantum information theory extends to any quantum channels with decoherence as well as those with different sizes of the input and output Hilbert spaces [151].


Figure 5-2: Setup to study scrambling in a unitary quantum channel U. Even though we draw our channels with input and output legs, when we discuss entanglement we always mean of the state  $|U\rangle$  given by the mapping to the doubled Hilbert space as in (5.9).

## 5.2.1 Entanglement in time

Since the state as defined in (5.12) contains all the information concerning the inputs and dynamics of the channel, we would like to use it to establish a general measure for scrambling. We will do this by studying the entanglement properties of a unitary U via the state  $|\Psi\rangle$ . Our setup is as follows. The input system is divided into two subsystems A and B, and the output system is divided into two subsystems C and D, as shown in Fig. 5-2. The subsystems do not necessarily have to be of the same size (i.e. it is possible that  $|A| \neq |B|$  or  $|A| \neq |C|$ ), and at t = 0 the input and output partitions do not necessarily need to overlap (i.e. it could be that  $A \cap C = \emptyset$ ). Additionally, despite how it is drawn, there does not need to be any spatial organization to the partitions. For example, the subsystem A could be an arbitrary subset of the input qubits.

With this state interpretation of the channel (5.12), we can form a density matrix  $\rho = |\Psi\rangle\langle\Psi|$  to compute joint entropies of subsystems that include both input and output degrees of freedom. For example, the entanglement entropy  $S_{AC}$  is given by

$$S_{AC} = -\text{tr} \{ \rho_{AC} \log_2 \rho_{AC} \}, \tag{5.15}$$

where the notation  $\rho_{AC}$  means the usual partial trace

$$\rho_{AC} = \operatorname{tr}_{BD} \{\rho\}. \tag{5.16}$$

Additionally, the mutual information between A and C is given by

$$I(A:C) = S_A + S_C - S_{AC}, (5.17)$$

and we will sometimes compute Rényi entropies

$$S_{AC}^{(N)} = \frac{1}{1-N} \log \operatorname{tr} \{\rho_{AC}^N\}.$$
(5.18)

Finally, let us note that, for a uniform input ensemble (or  $\beta = 0$ ),

$$S_A = a, \qquad S_B = b, \qquad S_C = c, \qquad S_D = d,$$
 (5.19)

$$S_{AB} = S_{CD} = n \tag{5.20}$$

where a, b, c, d are the numbers of qubits on A, B, C, D respectively. These relations are true because the inputs are always maximally entangled with the outputs. Therefore, any subsystem that is only a partition of the inputs or only a partition of the outputs (including non-partitions such as AB and CD) is maximally mixed. Even if we consider the more general channel (5.12), any subsystem that does not involve both input and output systems still has an entropy that is time-independent. Therefore the scrambling effect only appears in the entropy of regions on both sides, and (thus) the mutual information terms such as I(A : C) and I(A : D). For this reason, we will primarily be interested in the mutual information between region A and different partitions of the outputs. When region A is taken to be small, such as a single lattice site, the mutual information of A with part of the output system tracks how the information about local operators in A spreads under time evolution.

# 5.2.2 Scrambling

Scrambling is usually considered as a property of a state. In [39], a reference state evolved with a random unitary sampled from the Haar ensemble is called "Haar-scrambled." A

much weaker notion of scrambling of a state (which [39] calls "Page scrambling," or usually just "scrambling") is given by a state that has the property that any arbitrary subsystem of up to half the state's degrees of freedom are nearly maximally mixed. Said another way, a state is scrambled if information about the state cannot be learned from reasonably local measurements. Naturally (and proven in [37]), Haar scrambling implies Page scrambling.

We are interested in extending the notion of scrambling to unitary quantum channels of the form (5.12). Let us try to understand the properties of scrambling channels by considering entanglement across the channel. The identity channel is just a collection of EPR pairs connecting the input to the output. An example of a channel that does not scramble is the "swap" channel, where the arrangement of the EPR pairs are simply swapped amongst the degrees of freedom in the output system. In that case, localized quantum information in the input system remains localized in the output system, though residing in a different particular location. Instead, for a channel to scramble it necessarily must convert the EPR pairs into a more complicated arrangement of multipartite entanglement between the input and output systems. Such local indistinguishability of output quantum states corresponding to simple orthogonal input states in quantum channels enables secure storage of quantum information: a realization of quantum error-correcting codes.

Let us try to formalize this idea using our setup in Fig. 5-2. If our channel is a strong scrambler, then local disturbances to initial states cannot be detected by local measurements on output states. This implies that measurements on a local region C cannot reveal much information on local disturbances applied to A. Therefore the mutual information I(A : C)must be small. From a similar reasoning, one also expects I(A : D) to be small when the channel is a good scrambler since D is a local region too. On the other hand, the mutual information I(A : CD) quantifies the total amount of information one can learn about Aby measuring the output CD jointly. Since we are interested in the amount of information concerning A which is hidden non-locally over C and D, a natural measure of scrambling would be

$$I(A:CD) - I(A:C) - I(A:D).$$
(5.21)

This quantity accounts the amount of information about A that are non-locally hidden over C and D such that any local measurements, exclusively performed on C or D, cannot know.

If the channel scrambles, we expect that this quantity is large. It is well known that the above quantity is equal to minus the tripartite information<sup>2</sup>

$$I_{3}(A:C:D) = S_{A} + S_{C} + S_{D} - S_{AC} - S_{AD} - S_{CD} + S_{ACD}$$
$$\equiv I(A:C) + I(A:D) - I(A:CD)$$
(5.22)

The tripartite information  $I_3(A : C : D)$  must be negative and have a large magnitude for systems that scramble. We propose that this is a simple diagnostic of scrambling for unitary channels.

Scrambling in unitary channels is closely related to other notions of scrambling of states. For example, if the input to the channel is fixed to be a direct product state, then tripartite scrambling implies that subsystems of the output state will be near maximally mixed. Thus, scrambling in terms of the tripartite information implies "Page scrambling" of the output state.<sup>3</sup> In Appendix 5.7.1, we analyze Haar-random channels and show that Haar scrambling also implies that the tripartite information of the channel is very negative.

Importantly, it should be noted that the tripartite information is a measure of four-party entanglement, not three-party entanglement. Namely, consider a state with three-party entanglement only,  $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_{BCD}\rangle$  where A and BCD are not entangled. Then one always has  $I_3(A:B:C) = 0$ .  $I_3$  for other choices of regions also vanish, because for pure states the tripartite information is symmetric in any partitions into four regions A, B, C, D

$$I_3(A:B:C) = I_3(A:B:D) = I_3(A:C:D) = I_3(B:C:D).$$
(5.23)

Thus, in channels  $I_3$  is really a measure of four-party entanglement between the input system and the output system. In this paper, we will often choose to write the tripartite information as  $I_3(A : C : D)$  in order to emphasize a particular decomposition. However, for unitary channels the arguments are unnecessary due to the symmetry (5.23).

Finally, note that the condition  $I_3(A:C:D) \leq I(A:D)$  is known as strong subadditivity and must always holds among entropies. On the other hand,  $I_3(A:C:D) \leq 0$  is often

<sup>&</sup>lt;sup>2</sup>In the condensed matter community, the tripartite information is referred to as the topological entanglement entropy, which measures the total quantum dimension in a (2 + 1)-dimensional TQFT [152, 153].

<sup>&</sup>lt;sup>3</sup>In fact, this operator notion of scrambling is stronger than "Page scrambling" since the latter only refers to a single state. The operator scrambling implies that the information about a local operator in the input system cannot be recovered from a subsystem C of the output system (as long as it is not very close to the entire output system) even if C is bigger than half of the system.



Figure 5-3: (a) Permutations of qubits. The isometric pure state  $|\Psi\rangle$  consists of EPR pairs between input and output qubits. (b) A swap gate and a unitary corresponding to a perfect tensor. Note the similarity to the Feynman diagrams of a  $2 \rightarrow 2$  scattering process for a free theory and interacting theory, respectively.

referred to as monogamy of mutual information and doesn't necessarily hold for arbitrary states.<sup>4</sup> However, for holographic systems the tripartite information must always be negative [154]. This result is usually only discussed for holographic states but it also applies to holographic channels (such the eternal black hole in AdS) [155]. It is natural to suggest that the negative  $I_3$  is related to the fact that such holographic systems are strongly-chaotic and fast-scramblers [44]. (See also [156] for a study of monogamy and other properties of entanglement in qubit systems.)

## 5.2.3 Examples

Here, we present a few examples of using the tripartite information of a channel to measure scrambling.

# Swap channel

Let us revisit the example discussed at the beginning of this section. Consider a system of n qubits and assume that the unitary operator is the identity operator: U = I. The channel description is given by (5.10). This is a collection of EPR pairs connecting input qubits and output qubits. Since the state consists only of two-party entanglement, the tripartite information is zero. Similarly, consider a time-evolution that consists only of permutations

<sup>&</sup>lt;sup>4</sup>e.g.  $I_3(A:C:D) = 1$  for the four-qubit GHZ state  $\frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle)$ .

of qubits, a "swap" channel. Namely, let us assume that *j*th qubit goes to  $a_j$ th qubit where  $1 \leq j, a_j \leq n$ . Then, the isomorphic state consists of EPR pairs between *j*th input qubit and  $a_j$ th output qubit, and the tripartite information is zero (Fig. 5-3(a)). Permutations of qubits can be thought of as a classical scrambling. Two initial nearby classical states may become far apart after permutations of qubits, yet it is still possible to distinguish two initial states by some local measurement on the output states.

## Perfect tensor

Next, let us look at an example where the tripartite information is maximally negative. Consider a system of two qutrits (spins with three internal states  $|0\rangle, |1\rangle, |2\rangle$ ). We consider the following unitary evolution

$$|i\rangle \otimes |j\rangle \rightarrow |-i-j\rangle \otimes |j-i\rangle$$
 (5.24)

where addition is defined modulo 3. It can be directly verified that single qutrit Pauli operators transform in the following way<sup>5</sup>

$$Z \otimes I \to Z \otimes Z$$
  $X \otimes I \to X^{\dagger} \otimes X^{\dagger}$   $I \otimes Z \to Z \otimes Z^{\dagger}$   $I \otimes X \to X^{\dagger} \otimes X.$  (5.25)

In this unitary evolution, all local operators evolve to two-body operators. As such, information concerning local disturbances to initial states cannot be detected by any single qutrit measurements on output states. We can represent this unitary evolution as the following pure state

$$|\Psi\rangle = \frac{1}{3} \sum_{i,j=0}^{2} (|i\rangle \otimes |j\rangle)_{AB} \otimes (|-i-j\rangle \otimes |j-i\rangle)_{CD}$$
(5.26)

where addition is modulo 3. It is known that this pure state is maximally entangled in any bipartition. Namely, one has  $S_A = S_B = S_C = 1$ ,  $S_{AB} = S_{BC} = S_{CA} = 2$ ,  $S_{ABC} = 1$ , and  $I_3 = -2$ , where entropy for qutrits is measured in units of log 3. This is a so-called perfect state. In general, for any pure state ABCD one can show that  $I_3 \ge -2\min(S_A, S_B, S_C, S_D)$ . Therefore, this qutrit state has minimal value of  $I_3$ .

<sup>&</sup>lt;sup>5</sup>Pauli operators for *p*-dimensional qudits are defined by  $X|j\rangle = |j+1\rangle$  and  $Z|j\rangle = \omega^j |j\rangle$  with  $\omega = e^{i\frac{2\pi}{p}}$  for all *j* where addition is modulo *p*.

Here, we note that the difference in the depiction of this qutrit perfect tensor and the swap gate  $(|i\rangle \otimes |j\rangle \rightarrow |j\rangle \otimes |i\rangle)$  resembles the difference in the Feynman diagrams of a 2  $\rightarrow$  2 scattering process between a free theory and an interacting theory (see Fig. 5-3(b)-(c)). We will comment on this in much greater detail in the context of conformal field theory in Appendix 5.7.2.

#### Black hole evaporation

Another interesting example is the thought experiment by Hayden and Preskill [38] as shown in Fig. 5-4. They considered the following scenario. Alice throws her secret (A), given in the form of some quantum state of a = |A| qubits, into a black hole (B) with the hope that the black hole will scramble her secret so that no one can retrieve it without collecting all the Hawking radiations and decoding them. Bob tries to reconstruct a quantum state of Alice by collecting some portion of Hawking radiation (D) after a scrambling unitary evolution U applied to the black hole, consisting both of Alice's secret A and the original content of the black hole B. The remaining portion of the black hole after the Hawking radiation is denoted by C. So, as usual, this channel is split into four segments A, B, C, Das shown in Fig. 5-4.



Figure 5-4: The Hayden-Preskill thought experiment.

First, assume that Bob only knows the dynamics of the black hole (i.e. the operator U). In order for Bob to successfully reconstruct Alice's system A, the mutual information between Alice's secret and the Hawking radiation must be  $I(A : D) \approx 2a$ . (Recall that

I(A:D)/2 is roughly the number of EPR pairs shared by A and D.) However this is possible only when  $c \approx 0$  since the channel is maximally entangled along any bipartition due to the assumption of U being a scrambling unitary. Namely,  $I(A:D) \approx 0$  as long as D is smaller than B. Next, let us assume that Bob not only knows the dynamics U, but also knows the initial state of the black hole B. This is possible in principle if Bob has been observing the black hole since its formation. In this case, Bob has an access to both B and D. Then for d > a, the mutual information between A and BD becomes nearly maximal;  $I(A:BD) \approx 2a$  because BD contains more than half of the entire qubits in channel ABCD. In this case, the tripartite information is given by

$$I_3(A:B:D) = I(A:D) + I(A:B) - I(A:BD).$$
(5.27)

Since I(A:D),  $I(A:B) \approx 0$ , we find  $I_3 \approx -I(A:BD) \approx -2a$  which implies that Bob can indeed learn about Alice's secret.<sup>6</sup>

#### Holographic channels

In the final example, we will consider thermal systems with Einstein gravity bulk duals. Under the holographic duality, the unitary quantum channel representing the CFT time evolution operator is geometrized as the black hole interior or Einstein-Rosen bridge [1, 85]. Such holographic states are already known to be fast scramblers [44], so here we will simply confirm that holographic channels scramble in the sense of  $I_3$ .

It's a trivial extension of the ideas in [85] and [44] to calculate the tripartite information across the eternal AdS black hole (the holographic state dual to thermofield double state [106] of two entangled CFTs) so we will be brief. For simplicity, we will take A to be aligned with C across the Einstein-Rosen bridge, and A, B, C, D to have the same size, as shown in Fig. 5-5. For simplicity, we will consider time evolution only on the right boundary  $U(t) = e^{-iHt}$ . For any finite time, the mutual information I(A : D) is always zero for any finite regions A, D. The Ryu-Takayanagi (RT) surfaces used to compute the entanglement entropy are always disconnected. The only interesting behavior is from I(A : C), which was computed in this setup in [85]. The initial RT surface extends across the Einstein-Rosen

<sup>&</sup>lt;sup>6</sup>The firewall paradox of [34, 53] is related to the fact the scrambled channel with near maximally negative  $I_3$  cannot allow bipartite entanglement I(C:D) between the evaporating black hole C and the recently evaporated Hawking radiation D. This is a consequence of the monogamy of entanglement, which is captured by the negativity of the tripartite information.

bridge, and I(A:C) begins equal to the finite part of  $S_A + S_C$ , since the finite part of  $S_{AC}$ is vanishing. Under time evolution, the finite part of  $S_{AC}$  will increase linearly in time with a characteristic entanglement velocity  $v_E$  [85], and I(A:C) will decrease linearly to zero. Since  $I(A:CD) = S_{BH}$  the Bekenstein-Hawking entropy of the black hole, after a time of at most  $O(S_{BH}/2)$ , we find

$$I_3(A:C:D) = -S_{BH}, (5.28)$$

which is its minimal possible value.



Figure 5-5: The eternal AdS black hole interior is a geometric representation of the unitary quantum channel given by the time evolution operator of the dual CFT. Left: Penrose diagram for the eternal AdS black hole geometry with a spacelike slice (blue) anchored on the left boundary at the middle of the diagram anchored at time t on the right boundary. **Right:** Geometric depiction of the spacelike slice through the Einstein-Rosen bride (ERB). The spatial coordinates on the boundary CFT are represented by  $\varphi$ . The renormalized length of the ERB is proportional to t. For small t, the RT surface used to compute the entanglement entropy  $S_{AC}$  goes across the ERB (red). After a time proportional to the size of A or C, the disconnected RT surface (blue) is preferred and the entanglement entropy is a sum of disjoint contributions ( $S_{AC} = S_A + S_C$ ).

#### Haar-random channels

In Appendix 5.7.1, we analyze Haar-random unitaries. Using those results, we can bound the tripartite information in a Haar random channel

$$I_{3,Haar} \le -2\min(S_A, S_B, S_C, S_D) + 1 + \dots$$
(5.29)

The tripartite information of a random channel is near maximally negative plus one "residual" bit of information (independent of the overall system size). As we mentioned before, Haar scrambling implies tripartite scrambling.

### 5.2.4 Chaotic channels vs. integrable channels

In this section, we will focus on the aspects of unitary quantum channels built from timeevolution operators that can be used to differentiate chaotic systems from integrable systems. A system, defined by a time independent Hamiltonian H, can be chaotic or integrable. A unitary operator  $U(t) = e^{-iHt}$  can scramble. In §5.3, we will learn that channels U(t)that scramble for most values of t must be built from chaotic systems. To be concrete, we will take integrable systems to be those that have a quantum recurrence time that is polynomial in their number of degrees of freedom n. On the other hand, chaotic systems will generally have recurrence times that are doubly exponential in n,  $O(e^{e^n})$ .

The important point about scrambling in channels is that all bipartite mutual informations between the input subsystems and output subsystems become small. As an extreme example of an integrable system, let's return to the swap channel (Fig. 5-3(b)). As a reminder, for all t this channel has I(A:C) + I(A:D) = I(A:CD) or  $I_3(A:C:D) = 0$ . Swap gates preserve bipartite entanglement and cannot create multipartite entangled output states. As such, an input localized at one lattice site can only be moved around to n possible locations. The recurrence time must scale like O(n).

In order for the tripartite information of the channel to vanish for all times, the decrease of I(A:C) must be exactly compensated by the increase of I(A:D). If we take C to contain A at t = 0, the initial values of mutual information are I(A:C) = 2a and I(A:D) = 0. If  $I_3$  vanishes for t > 0, I(A:C) must decrease in order for I(A:D) > 0. If I(A:C)returns to the initial value 2a at a later time (i.e. there is a recurrence), in the meantime we must have I(A:D) = 0. Therefore the signature of an integrable system is a sharp peak in I(A:D) = 0 (or equivalently, a dip in I(A:C)). The sharpness of the peak is determined by the relative sizes of the systems A and C. In chaotic systems for times shorter than the Poincare recurrence time  $O(e^{e^n})$ , I(A:D) and I(A:C) will asymptote to the channel's Haar-random value (see Appendix 5.7.1).

As we will see numerically in §5.4, for unequal divisions of the input  $(a \ll b)$  and output  $(c \ll d)$  the signature of an integrable system is a spike in I(A : D) which occurs at the time signals from A arrives at D. In such systems the tripartite information might become negative, but it will never become close to the Haar-scrambled value, and it will quickly return near zero. However, for equal divisions of input and output, the tripartite information of an integrable channel will tend to a constant much greater (i.e. less negative) than the Haar-scrambled value. This equal-sized subsystem configuration appears to be the most robust measure of scrambling.

This discussion of the spike in I(A : D) highlights an important point: in integrable systems, such as the transverse Ising model, it is still possible to have early time exponential decay and ballistic growth of operators [1] (see the top-middle panel of Fig. 5-7), and in integrable CFTs some (but not all!) OTO correlators can decay at late times [3]. However, for these systems the growth of operators (or the decay of I(A : C)) must always be followed by a later decrease in size (or a recorrelation in the OTO correlator) as the system exhibits a recurrence.<sup>7</sup> We will see this behavior explicitly in our numerics in §5.4.

# 5.3 Butterfly effect implies scrambling

Now, we will show that the generic decay of OTO correlators of the form  $\langle W(t) V W(t) V \rangle$ implies that the mutual information between any small subsystem in the inputs and any partition of the output should be small. We will provide an exact formula relating the operator average of OTO correlators in different size subsystems to the second Rényi entropy for a subsystem consisting of both inputs and outputs. Using tripartite information as our diagnostic of scrambling in a unitary channel, we will show the butterfly effect implies scrambling.

## 5.3.1 Average over OTO correlators

For simplicity of discussion, we consider a system consisting of qubits. We consider a complete basis of Hermitian operators  $D_i$  in subsystem D, which satisfies the orthonormal condition

$$\operatorname{tr}\left\{D_i D_j\right\} = 2^d \delta_{ij}.\tag{5.30}$$

<sup>&</sup>lt;sup>7</sup>Another relevant difference between chaotic and integrable systems is in terms of the expansion (5.3) for the time evolution of a simple operator. Due to ergodicity, such an expansion will have a number of terms exponential in the size of the system for chaotic dynamics. Integrable systems are not ergodic, so the expansion will only have a linear number of terms.

Similarly, we define an orthonormal basis  $A_i$  in subsystem A. As a reminder, our state lives in a  $2^n$ -dimensional Hilbert space, and AB and CD are two different decompositions of that Hilbert space. Hilbert space A is  $2^a$ -dimensional, and the operators  $A_i$  act on A. A similar statement holds for D. Both D and A are setup as in Fig. 5-2, and  $S_D = d$ ,  $S_A = a$ . If a = 1, then one possible basis choice for  $A_i$  the three Pauli operators X, Y, Z, and the identity I. In general, there are  $4^a$  independent operators in A. We can think of this as choosing one of the four operators I, X, Y, Z at each site. If the Hilber space decompositions A and D do not overlap, then  $[A_i, D_j] = 0$  for all i, j. However,

As a measure of OTO correlation functions for generic operator choices, we consider the following quantity

$$\begin{aligned} |\langle \mathcal{O}_D(t) \, \mathcal{O}_A \, \mathcal{O}_D(t) \, \mathcal{O}_A \rangle_\beta | &:= \frac{1}{4^{a+d}} \sum_{ij} \langle D_i(t) \, A_j \, D_i(t) \, A_j \rangle_\beta \\ &= \frac{1}{4^{a+d}} \cdot \frac{1}{Z} \sum_{ij} \operatorname{tr} \{ e^{-\beta H} D_i(t) \, A_j \, D_i(t) \, A_j \}, \end{aligned}$$
(5.31)

where the sums i, j run from 1 to  $4^d, 4^a$ , respectively. Here,  $\langle \cdot \rangle_{\beta}$  represents a thermal average, and  $|\cdot|$  represents an operator average over the complete bases  $D_i, A_j$ . Additionally, we will take the infinite temperature limit  $\beta = 0$  so we don't have to worry about the Euclidean evolution. For t = 0, every correlator is unity  $\langle D_i(0) A_j D_i(0) A_j \rangle_{\beta=0} = 1$  due to the orthonormal condition and the fact that  $A_j$  and  $D_i(0)$  commute.

Under chaotic time evolution,  $D_i(t) = e^{iHt}D_ie^{-iHt}$  will evolve into a high weight operator and cease to commute with  $A_j$ . (To see this, consider the BCH expansion for  $D_i(t)$  as in (5.3). As t increases, the later terms with high weight will become important. These terms will no longer be confined to subspace D. Under chaotic evolution the  $D_i(t)$  will grow to reach A.) This will lead to the decay of the OTO correlation functions  $\langle D_i(t) A_j D_i(t) A_j \rangle_{\beta=0}$  for generic i, j (as long as the  $D_i$  or the  $A_j$  are not the identity in which case  $\langle D_i(t) A_j D_i(t) A_j \rangle_{\beta=0} = 1$  for all t) and all such partitions A, D.

At early times, the average will be very close to unity. With chaotic time evolution, the butterfly effect will cause most of the correlation functions in the average to decay exponentially. Using standard techniques, one can relate (5.31) to the second Rényi entropies of the time evolution operator considered as a state

$$|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle_{\beta=0}| = 2^{n-a-d-S_{AC}^{(2)}}, \qquad (5.32)$$

where  $S_{AC}^{(2)}$  is the second Rényi entropy of AC defined in (5.18),  $2^n$  is the dimension of the input or output Hilbert space, and the subsystems A and D have dimension  $2^a$  and  $2^d$ , respectively. A proof of this result and its generalization to finite temperature is given in Appendix 5.7.3.

At first glance, it is a little surprising that a Rényi entropy appears here: the entropy determines mutual information, which bound two-point functions, and chaos is distinctly measured by OTO four-point functions. However, the key point is that when A and D are small (so that they only contain approximately local operators), B and C contain highly nonlocal operators covering almost the entire input and output systems, respectively. As a result,  $S_{AC}^{(2)}$  is sensitive to correlations between the few operators in A and the complete (and nonlocal and high weight) set of operators in C.

The OTO average (5.32) is an operator-independent information-theoretic quantity that is constrained by chaotic time evolution. To understand its behavior, let's consider its maximum and minimum values. The average will be the largest at t = 0, when all the correlators are unity. On the other hand, the average will be minimized when the Rényi is maximal: max  $S_{AC}^{(2)} = \min(a + c, d + b)$ . Let us assume for the rest of this section that  $a \leq d$ , so max  $S_{AC}^{(2)} = a + c$ . This means the OTO average is bounded from below by  $4^{-a}$ . (It's worth mentioning that the Haar-scrambled value of the average is generally larger than this lower bound.<sup>8</sup>) Therefore, we see the OTO average is bounded by

$$4^{-a} \le |\langle \mathcal{O}_D(t) \, \mathcal{O}_A \, \mathcal{O}_D(t) \, \mathcal{O}_A \rangle_{\beta=0}| \le 1.$$
(5.33)

Now, we will recast this result in terms of mutual information in order to make a connection to our scrambling diagnostic. Let's assume that after a long time of chaotic time evolution the OTO average asymptote to a small positive constant  $\epsilon$ . This means that the entropy  $S_{AC}$  is bounded:

$$S_{AC} \ge S_{AC}^{(2)} = n - a - d - \log_2 \epsilon,$$
 (5.34)

<sup>&</sup>lt;sup>8</sup>Since (5.31) includes  $4^d + 4^a - 1$  terms where  $A_j = I$  or  $D_i = I$  (one for each term where  $A_j = I$  or  $D_i = I$ , and minus one to prevent overcounting when they both are), if all the non-identity correlation functions decay the OTO average will be  $4^{-a} + 4^{-d} - 4^{-a-d} > 4^{-a}$ . Using the results from Appendix 5.7.1, we can show that this is larger value is exactly the Haar-scrambled value of the OTO average. To get lower value, some of the correlation functions need to cross zero so that the operators are negatively correlated.

where in the first part we used the fact that  $S_R^{(i)} > S_R^{(i+1)}$  for Rényi entropies, and in the second part we used (5.32). In terms of mutual information, we have

$$I(A:C) \le 2a + \log_2 \epsilon, \tag{5.35}$$

where here we have used the fact that  $S_A$  and  $S_C$  are always maximally mixed.

Eq. (5.35) is one of our main results. At t = 0, I(A : C) = 2a. The information about the input to the channel in A is 2a bits and that information is entirely contained in the output subsystem C. Since there are  $4^a$  linearly independent basis operators in A's Hilbert space, we can interpret these 2a bits as the information about which of the  $4^a$  operators was input into the channel. (For instance, if a = 1, it takes two bits to index the operators I, X, Y, Z.) Under chaotic time evolution,  $\epsilon \ll 1$ , and the mutual information between A and C becomes small. The smallest possible value for  $\epsilon$  is  $2^{-2a}$ , which occurs when I(A : C) = 0. In practice, there is always residual information between AC. Using the results from Appendix 5.7.1, we see that for Haar scrambling the mutual information can be bounded as

$$I(A:C)_{Haar} \le \log_2(1 + 4^{a-d} - 4^{-d}), \tag{5.36}$$

which corresponds to all the non-identity terms in the OTO average decaying to zero. If the information-theoretic quantities constructed from the second Rényi approach their Haar-scrambled value, then all the nontrivial OTO correlators  $\langle D_i(t) A_j D_i(t) A_j \rangle_{\beta=0}$  must approach zero.

Next, we note that (5.35) implies

$$I(A:C_{\alpha}) \le 2a + \log_2 \epsilon, \tag{5.37}$$

for any partitioning of  $C = C_{\alpha} \cup C_{\bar{\alpha}}$ . This can be seen from subadditivity

$$S_{AC_{\alpha}} + S_{AC_{\bar{\alpha}}} \ge S_{AC},\tag{5.38}$$

and the definition of mutual information. (This is also intuitive: any information contained about region A in region C must necessarily be more than or equal the information in a partition of C.) Therefore, we learn that in chaotic channels, local information in the input must get delocalized in the output (i.e. cannot be recovered in an output subsystem smaller in size than the total system n). Since the partitions A and C were completely arbitrary, we conclude that the decay of OTO correlators implies that all bipartite mutual informations are small. If the OTO correlator average is given by  $\epsilon$  after a long time, then we have

$$I(A:C), I(A:D) \leq 2a + \log_2 \epsilon$$
  

$$I_3(A:C:D) \leq 2a + 2\log_2 \epsilon = -2a + 2\log_2 \frac{\epsilon}{\epsilon_{\min}}$$
(5.39)

When  $\epsilon$  approaches the minimum value  $\epsilon_{\min} = 2^{-2a}$ ,  $I_3$  approaches the most negative value  $I_{3,\min} = -2a$ . From this discussion, we conclude that the butterfly effect implies scrambling.

## 5.3.2 Early-time behavior

In this section, we will attempt to connect the universal early-time behavior of OTO correlators in strongly chaotic systems with the information-theoretic quantities we use to diagnose scrambling.

In strongly chaotic systems, all OTO correlation functions of operators with nontrivial time evolution will decay to zero. However, the behavior of the OTO correlation function  $\langle W(t) V W(t) V \rangle_{\beta}$  as it asymptotes to zero is not universal. The approach will depend on the specific choices of operators W, V. For instance, in two-dimensional CFTs with large central charge and a sparse low-lying spectrum at late times the OTO correlator decays as

$$\frac{\langle W(t) V W(t) V \rangle_{\beta}}{\langle WW \rangle_{\beta} \langle VV \rangle_{\beta}} \sim e^{-4\pi h_v t/\beta}, \qquad (5.40)$$

where  $h_v$  is the conformal weight of the V operator, and it is assumed  $1 \ll h_v \ll h_w$  [3].

On the other hand, at early times the behavior of  $\langle W(t) V W(t) V \rangle_{\beta}$  usually takes a certain form. The initial decay is known to fit the form

$$\frac{\langle W(t) V W(t) V \rangle_{\beta}}{\langle WW \rangle_{\beta} \langle VV \rangle_{\beta}} \approx 1 - \# e^{\lambda_L t} + \dots, \qquad (5.41)$$

where in analogy to classically chaotic systems  $\lambda_L$  has the interpretation of a Lyapunov

exponent [45].<sup>9</sup> In [58], it was shown that quantum mechanics puts a bound on  $\lambda_L$ 

$$\lambda_L \le \frac{2\pi}{\beta},\tag{5.42}$$

with saturation for strongly-interacting conformal field theories that have holographic descriptions in terms of Einstein gravity. This Lyapunov exponent is expected to be universal independent of the choice of operators W, V—and as such any model that saturates the bound (5.42) is expected to be a toy model of holography [45, 77]. (In §5.4, we will use numerics to explore the Majorana fermion model proposed in Ref. [77] that in a certain limit is expected to have this property.)

Since the partitioning of the inputs into A, B and outputs into C, D was entirely arbitrary, let's first consider channels that operate on 0-dimensional systems, e.g. fast scramblers in the sense of [39], such as the Majorana fermion model of Kitaev [77] (a simplification of the Sachdev-Ye model of N SU(M) spins [78], see also [157]) or a large N strongly interacting CFT holographically dual to Einstein gravity near its Hawking-Page point.<sup>10</sup> These systems still have low-weight k-local Hamiltonians wth  $k \ll N$ , but each degree of freedom interacts with every other degree of freedom. For these systems, the OTO correlation functions decay as

$$\langle W(t) V W(t) V \rangle_{\beta} = f_0 - \frac{f_1}{N^2} e^{\lambda_L t} + O(N^{-4}),$$
 (5.43)

where  $N^2 = n$  is the total number degrees of freedom, and the constants  $f_0$  and  $f_1$  depend on properties of the W, V operators (e.g. their CFT scaling dimensions) [58]. The decay of the correlator is delayed by the large number of degrees of freedom at time  $t_* = \lambda_L^{-1} \log N^2$ . This is usually referred to as the scrambling time [39]. Plugging (5.43) into (5.35), we find that at early times the mutual information between A and C is bounded as

$$I(A:C) \le 2a - \#e^{\lambda_L(t-t_*)} + \dots$$
 (5.44)

Thus, the information between A and C must begin to decay by the scrambling time

<sup>&</sup>lt;sup>9</sup>However, the analogy is imprecise. In weakly coupled systems,  $\lambda_L$  has a semiclassical analog that does not map onto the classical Lyapunov exponent. Despite this, we will follow convention and refer to  $\lambda_L$  as a Lyapunov exponent. We are grateful to Douglas Stanford for emphasizing this point.

<sup>&</sup>lt;sup>10</sup>We require this limit so we can think of the black hole as unit sized and not yet worry about the operator growth in spatial directions.

 $t_* = \lambda_L^{-1} \log N^2$ . This inequality would be an equality if we instead considered the mutual information constructed from the second Rényi  $S_{AC}^{(2)}$ .

Now, let's consider systems arranged on a spatial lattice but do not have a large number of degrees of freedom per site, e.g. spin chains. For these systems, the butterfly effect implies ballistic growth of operators in spatial directions [1]. For local operators W and V separated by large distance  $|x| \gg \beta$ , in many known examples (such as holography and numerical investigations of one-dimensional spin chains) strong chaos implies that OTO correlation functions decay as

$$\langle W(t)VW(t)V\rangle_{\beta} = f_1' - f_2' e^{\lambda_L(t-|x|/v_B)} + O(e^{-2|x|\lambda_L/v_B}), \qquad (5.45)$$

with additional constants  $f'_1, f'_2$  that depend on the details of the operators.<sup>11</sup> In this case, the early-time decay of the correlator is suppressed by the large spatial separation between the operators. Under chaotic time evolution, the operator W(t) will grow ballistically with some characteristic "butterfly" velocity we denote  $v_B$ .<sup>12</sup> Thus, W(t) and V will commute until a time  $t > |x|/v_B$  when V enters the "butterfly" light cone of W. Let's focus on a lattice of spins in d-spatial dimensions. We will pick our subsystem A to be a ball of a sites surrounding the origin with a radius  $r_a$ . We will pick C to also be a ball surrounding the origin with a radius  $r_c$  such that  $r_c - r_a = |x|$ . Then, after a scrambling time of  $t_* = v_B t$ , the mutual information between A and C must begin to decay

$$I(A:C) \le 2a - \#e^{\lambda_L(t-|x|/v_B)} + \dots$$
(5.46)

Since we will study this quantity in §5.4, we note that we can directly equate (rather than bound) the behavior of the second Rényi entropy to the Lyapunov behavior of the OTO correlators. Let's restrict to a one-dimensional spin chain, and take A to be the first spin of the input and D to be the last spin of the output. In that case, if we assume a form

<sup>&</sup>lt;sup>11</sup>As emphasized in [58], the butterfly effect is relevant for systems with a large hierarchy of scales. For the 0-dimensional systems we just considered, the hierarchy is provided by the parametric difference between the thermal time  $\beta$  and the fast scrambling time  $\beta \log N^2$ . In the present case, the role of the large scale is instead being played by the large spatial separation |x| between the operators.

<sup>&</sup>lt;sup>12</sup>The velocity  $v_B$  can depend details of the theory that do not affect  $\lambda_L$ . For instance, it is modified in Gauss-Bonnet gravity [1] and for certain Einstein gravity theories can even acquire a temperature dependence [51].

of the correlator (5.45), plug into (5.31) and compute the average, then we find

$$S_{AC}^{(2)}(t) = S_{AC}^{(2)}(0) + \# e^{\lambda_L t} + \dots, \qquad (5.47)$$

showing that at early times the Rényi entropy can grow exponentially with characteristic Lyapunov exponent  $\lambda_L$ .<sup>13</sup> We will roughly see this behavior in Fig. 5-7.

# 5.3.3 Butterfly velocity vs. entanglement velocity

There are two nontrivial velocities relevant to the growth of information-theoretic quantities in unitary quantum channels arranged on a lattice.

The butterfly velocity  $v_B$  [1, 44] is the speed at which the butterfly effect propagates. It is the speed at which operators grow under chaotic-dynamics. Such behavior is reminiscent of the Lieb-Robinson bound on the commutator of local operators separated in time for systems with local interactions [74–76] and suggests identifying  $v_B$  with the Lieb-Robinson velocity. The butterfly velocity is often difficult to compute directly, but in holographic theories with Einstein gravity duals it is known to be [44]

$$v_B = \sqrt{\frac{d}{2(d-1)}},$$
 (Einstein gravity), (5.48)

where d is the spacetime dimension of the boundary CFT. This value is modified in Gauss-Bonnet gravity [1] and for certain theories can even acquire a temperature dependence [51].

The entanglement velocity (sometimes called the tsunami velocity) studied in [85, 92, 93, 158, 159] is often described as the rate at which entanglement spreads. It is rate of growth of entanglement entropy after a quench, and in holographic systems dual to Einstein gravity it can be computed directly [85, 92, 93]

$$v_E = \frac{\sqrt{d}(d-2)^{\frac{1}{2}-\frac{1}{d}}}{[2(d-1)]^{1-\frac{1}{d}}}, \qquad \text{(Einstein gravity)}, \tag{5.49}$$

a different nontrivial function of the CFT spacetime dimension d. In these theories,  $v_E \leq v_B$ ,

<sup>&</sup>lt;sup>13</sup>This assumes that  $\lambda_L$  is independent of the choice of operators W and V and that the ansatz (5.45) is the correct form for the initial decay of the correlator. Both of these assumptions are not necessarily true for some spin systems. Additionally, if the constant in front of the exponential is not small (for example, in holographic systems), then the expansion will not be valid and one cannot see the exponential growth behavior. We thank Tarun Grover and Douglas Stanford for emphasizing these points.

and  $v_B = v_E = 1$  for d = 1 + 1.

In the context of unitary quantum channels, these velocities have a very specific interpretation in terms of different mutual informations, see Fig 5-6. Consider a lattice of ndegrees of freedom and divide the input up such that a = b = n/2, the output such that c = d = n/2, and such that all subsystems are contiguous. If A and C are aligned such that at t = 0, I(A : C) = n, then under time evolution (for chaotic and integrable systems!) it is expected that for a long stretch of time that the mutual information will decrease linearly as

$$I(A:C) = n - v_E st, \tag{5.50}$$

until near when it saturates at I(A : C) = 0. This is often referred to as a quench, which we discuss in depth in Appendix 5.7.2 in the context of CFT. Here, s is the "entropy density," which converts  $v_E$  from a spatial velocity (with units lattice-sites/time) to an information rate (bits/time). For spin systems, the state representation of the channel requires two qubits per addition to the spatial lattice (one to represent the input leg and one to represent the output leg), so s = 2.

While this tells us about the accumulation of entanglement between AC and BD, it tells us nothing about how information propagates in the spatial directions through the channel. To study the latter, we need a different configuration of subsystems. First, let's pick A to be a small region a = O(1) center at the origin. Then, instead of considering a fixed radius C, let us let pick subsystem C to be a growing ball of radius  $r_c = v_B t$  surrounding the origin in the output system. Then, the information between the input A and output C(t)will be constant

$$I(A:C(t)) = 2a, (5.51)$$

but the output is growing ballistically as a ball of radius  $v_B t$ . Thus,  $v_B$  is the rate that information propagates spatially in unitary quantum channels.<sup>14</sup> For finite systems, eventually C(t) will grow to encompass the entire output of the channel, and the information will stay delocalized until the recurrence time. Before then, any local measurement of any subsystem of the output will be insufficient to recover the information in A. The information is scrambled.

<sup>&</sup>lt;sup>14</sup>A similar interpretation of the butterfly velocity was recently made by Mezei and Stanford in the context of holography. It was shown that the reconstruction of a particle falling into a black hole can occur for a region that grows ballistically with velocity  $v_B$  [43].



Figure 5-6: A one-dimensional tensor network model of a channel with increasing vertical network depth t. This makes it clear that  $v_E$  is related to vertical propagation and  $v_B$  is related to horizontal propagation. Left: I(A : C) decreases by 2 with every increase in circuit depth. By measuring time in terms of the network depth, the entanglement velocity is trivial ( $v_E = 1$ ). Right: the "information" or "butterfly" light cone of the input is controlled by the butterfly velocity ( $v_B = 1$ ). The information contained in this light cone is constant: I(A : C(t)) = 2a.

In Fig. 5-6, a tensor network model of a channel exhibits both of these velocities. The left side shows the linear decrease of I(A:C) with increasing circuit depth in the vertical direction. Every increase in depth is accompanied by a decrease in I(A:C) of  $v_E s \Delta t$  bits, with  $\Delta t$  the time interval corresponding to each step of the circuit. The bold tensors make it clear that this decrease is a simple consequence of the increase of circuit depth. In a gate or tensor model of a channel,  $v_E$  is simply a conversion in units between the network's depth and conventional time. More generally, it is the velocity associated with the "vertical" direction of the channel.

The right side of Fig. 5-6 shows the growth of the region C(t) required to capture all of the information from the input A. This has a natural interpretation of a light cone for information propagation and can be very different depending on the properties of the channel. This velocity is associated with the "horizontal" direction of the channel. In the simple tensor network picture and 1+1-dimensional CFT,  $v_E = v_B = 1$ , but more generally  $v_E \leq v_B$  as in holographic Einstein gravity (5.48)-(5.49).

# 5.4 Numerics in qubit channels

In this section, we will study chaos and scrambling in qubit channels. We will show explicitly that the tripartite information is a simple diagnostic of scrambling, and we will verify our main result (5.32) by directly showing that the butterfly effect implies scrambling. Finally, we will also comment on the relationship between the size of the subsystems and the expected behavior of the entropies in the channels.

We will study four different channels. First, we will directly compare two 1-dimensional Ising spin chains, one integrable and one strongly chaotic, both of which can be expressed with the Hamiltonian

$$H_{sc} = -\sum_{i} Z_i Z_{i+1} + g X_i + h Z_i.$$
(5.52)

Here, i = 1, ..., n, and  $X_i, Y_i, Z_i$  are the Pauli operators on the *i*th site. The integrable system we study is the transverse field Ising model with g = 1, h = 0, and the chaotic system has g = -1.05, h = 0.5 [103]. In our numerics, we will take n = 7 when studying subsystems of unequal sizes and n = 6 when studying subsystems of equal sizes. For these systems, we will also compute the velocities  $v_E$  and  $v_B$ .

The second model we will study is a 0-dimensional fast scrambler, the Majorana fermionic

system of Kitaev [77] (see also [78, 157]) with Hamiltonian

$$H_K = \sum_{j < k < \ell < m} J_{jk\ell m} \chi_j \chi_k \chi_\ell \chi_m, \qquad \overline{J_{jk\ell m}^2} = \frac{3!}{(N-3)(N-2)(N-1)} J^2, \tag{5.53}$$

where the  $\chi_j$  are spinless Majorana fermions, and  $j, k, \ell, m = 1, \ldots N$ ,  $J_{jk\ell m}$  are random couplings with mean zero and variance  $\overline{J_{jk\ell m}^2}$ .<sup>15</sup> We can study Majorana fermions using spin chain variables by a nonlocal change of basis known as the Jordan-Wigner transformation

$$\chi_{2i-1} = \frac{1}{\sqrt{2}} X_1 X_2 \dots X_{i-1} Z_i, \qquad \chi_{2i} = \frac{1}{\sqrt{2}} X_1 X_2 \dots X_{i-1} Y_i, \tag{5.54}$$

such that  $\{\chi_j, \chi_k\} = \delta_{jk}$ . With this representation, N Majorana fermions require N/2 qubits. In our numerics, we will take N = 14 and J = 1.

Finally, we will consider scrambling with a Haar-random unitary as a baseline for a scrambled system. For scrambling channels, at late times the quantities we use to diagnose scrambling often don't reach their extremal values, but rather asymptote to Haar-random values. In Appendix 5.7.1, we analytically compute the second Rényi entropies of the Haar-random unitary channel considered as a state. To study the mutual informations, we will numerically sample from the Haar ensemble. For the size of the channels we will study  $(\dim U = 2^6 \times 2^6 \text{ or } \dim U = 2^7 \times 2^7)$ , these properties are self-averaging and only require one sample.

# 5.4.1 Spin chains

### Unequal sized subsystems

Our first setup for the spin chains is shown in the top-left corner of Fig. 5-7, which contains 7 spins with open boundary condition. We study the  $2^7 \times 2^7$ -dimensional unitary operators as states

$$|U\rangle = \frac{1}{2^7} \sum_{i=1}^{2^7} |i\rangle_{AB} \otimes U|i\rangle_{CD}, \qquad (5.55)$$

where AB is the input to the channel and CD is the output. In the input, we take A to be the first spin, and B to be the final six spins. Similarly, we take C to be the first six

<sup>&</sup>lt;sup>15</sup>This is slightly different than in [77], because we are studying the system in the limit  $\beta J = 0$ , while the system in [77] is considered in the limit  $\beta J = \infty$ . Thus, we don't expect the Lyapunov exponent to saturate the chaos bound (5.42), and instead we expect it to be proportional to the coupling J.

spins, and D to be the last spin. This puts A and D at maximal spatial separation. For the spin chains, we take  $U = e^{-iH_{sc}t}$ , with  $H_{sc}$  given by (5.52). As a comparison, we consider a Haar channel with U a Haar random unitary operator. Since we are considering uniform input, the inverse temperature is always taken to be vanishing ( $\beta = 0$ ).

Our spin chain numerics for this unequal subsystem setup are shown in the rest of Fig. 5-7. In the top-middle panel, we show the average over OTO correlation functions  $|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle_{\beta=0}|$ . Of course, from (5.32) we know this is also equal to  $2^{n-a-d-S_{AC}^{(2)}}$ , which in this case can be verified explicitly. The chaotic spin chain asymptotes to just above the Haar-scrambled value, which can be computed from Appendix 5.7.1 and is equal to  $4^{-a} + 4^{-d} - 4^{-a-d} = 7/16$ . This is also the value given by assuming all the correlation functions in the average where neither an operator in A nor an operator in D are the identity decay to zero. (There are 9 such terms, with 16 total correlation functions in the sum.) The Haar-scrambled value is far above the absolute minimum  $4^{-a} = 1/4$  of OTO correlator average, given by plugging the maximum possible value for the second Rényi (max  $S_{AC}^{(2)} = 7$ ) into (5.32). (To reach this value, some of the correlators would have to become negative.) Additionally, we see Lyapunov behavior of the OTO correlator decay beginning around t = 2. At later times (around t = 4), the integrable system does not asymptote to the Haar-scrambled value but instead has a recurrence and recorrelates.

We see the similar behavior in the other related quantities computed on subsystem AC. In the top-right panel of Fig. 5-7, we plot the entropies  $S_{AC}$  and  $S_{AC}^{(2)}$ , and in the bottomleft panel we plot I(A:C). The entropies  $S_{AC}^{(2)}$  are directly related to the OTO average by (5.32), and  $I(A:C) = a + c - S_{AC}$ . (As a reminder a = 1 and c = 6 are constant.) At early times we see roughly Lyapunov behavior of the second Rényi entropy as suggested from (5.47). In all of these curves, the quantity is roughly unchanged until a time of order c and then begins to decay exponentially. As expected,  $S_{AC}^{(2)} < S_{AC}$ , and these entropy curves have roughly the same shape. Also as expected, localized information between A and C gets delocalized by chaos. For the chaotic channel, I(A:C) begins at 2 bits (representing the four different choices of input operators) and then decays to just above the Haar-scrambled value of roughly 0.6 bits. Using the results from Appendix 5.7.1 and (5.71), we can bound this late-time residual information to be less than 0.8 bits

$$I(A:C)_{Haar} \le 1 + \log_2(7/8) \approx 0.8$$
 bits. (5.56)



Figure 5-7: Chaos and scrambling in the integrable (blue; g = 1, h = 0 Hamiltonian) and chaotic (orange; q = -1.05, h = 0.5 Hamiltonian) spin chain unitary channels and the Haar-random channel (black). Top Left: configuration of channel (n = 7 spins; input)subsystems of a = 1, b = 6 spins; output subsystems of c = 6, d = 1 spins). For the spin chains, the channel is the time evolution operator  $(U = e^{-iHt})$ . For the Haar-channel, U is sampled from the Haar ensemble. Top Middle: the average of OTO correlators shows the butterfly effect. At later times, the chaotic system asymptotes to the Haar-scrambled value, but the integrable system doesn't remain decorrelated. Top Right:  $S_{AC}$  (solid) and  $S_{AC}^{(2)}$ (dotted) shows roughly the same behavior as the OTO average. Bottom Left: I(A:C), a trivial function of  $S_{AC}$ , shows that for the chaotic channel an initial 2 bits of information between the subsystems gets delocalized so that at late times only a small amount (0.59)bits) remains. We can also read off the butterfly velocity,  $v_B = 2.5$ . Bottom Middle: the spike in I(A:D) for the integrable channel shows that information is not delocalized by integrable time evolution. For the chaotic and Haar channels information is delocalized, and I(A:D) is always near vanishing. Bottom Right: The negative of the tripartite information normalized by its maximum value (2 bits) is a simple diagnostic of scrambling.

For the integrable channel, the information doesn't get delocalized and I(A : C) returns close to its initial value of 2 bits. The decay in both cases is delayed until roughly t = 2. The ratio of this delay to the distance between A and D lets us extract the butterfly velocity for these chains ( $v_B = 2.5$ ).

In the bottom-middle panel of Fig. 5-7, we plot I(A : D). For the chaotic channel, this quantity can never become very large: under time evolution the information in Agets spread across *all* the degrees of freedom, so there can never be significant localized information about the input A in the output subsystem D. From Appendix 5.7.1, we also note that the Haar-scrambled I(A : D) is exponentially small in the overall system size (Eq. (5.74)). However, the integrable channel has a very sharp peak after a time of order c. In Appendix 5.7.2, we explain in the context of CFT that this memory effect is the failure of integrable systems to efficiently delocalize information—i.e. scramble—due to entanglement propagation by noninteracting quasi-particles. This peak (0.28 bits) is far from the maximum (2 bits), but corresponds precisely to the point in the bottom-left panel plot of I(A : C) where the integrable channel is loosing information between A and C. This supports our hypothesis of integrable channels moving around localized information rather than actually scrambling.

Finally, we plot the negative of the tripartite information in the bottom-right panel of Fig. 5-7, normalized by its maximal value  $(-I_3(A : C : D)/2a = -I_3(A : C : D)/2)$ . As explained in §5.2.2, as a measure of multipartite entanglement this quantity is a simple diagnostic of scrambling; the chaotic channel asymptotes to the Haar-scrambled value of  $I_3$  after a time O(n). The integrable channel initially has an increase in  $I_3$ , but never reaches the Haar-scrambled value (due to the memory spike in I(A : D)) and instead has a recurrence beginning after a time of order O(n). A long time average of this quantity would make it clear that the integrable channel doesn't scramble.

As a final point, the fact that the Haar-scrambled value of the normalized negative  $I_3$  is less than unity  $(-I_3/2 \approx 0.7)$  is not a relic of small n. This "residual" information can be bounded using the results from Appendix 5.7.1 and depends only on the subsystem sizes

$$\frac{-I_3(A:C:D)_{Haar}}{2a} \ge 1 - \frac{1 + \log_2(1 - 2^{-a - d - 1})}{2a}.$$
(5.57)

Even for large systems, the residual information vanishes like 1/2a. The point is that for

 $1 \ll a, d < b, c$ , there's still always roughly one bit of information between A and C that doesn't delocalize across the entire output.

### Equal sized subsystems

For comparison, we also consider the same spin chains, but with n = 6 and equal sized subsystems of a = b = c = d = 3 spins. A and D are still taken to be opposite ends of the chain, but their boundaries are no longer separated spatially. This setup is shown in the top-left panel of Fig. 5-8. As noted before, this is the setup often used when quenches are discussed. Here, we only plot the quantities that have interesting differences from the previous configuration.

In the top-right panel, we plot the normalized negative tripartite information  $-I_3/6$ . Interestingly, for the integrable channel the tripartite information does not exhibit a recurrence, but rather saturates at a very low value (roughly at 0.24, compared to 0.7 for the chaotic channel and 0.77 for the Haar-random channel). Thus, in this configuration of subsystems it appears to be a robust measure of scrambling that doesn't require any long time average.

In the lower-left panel, we see that I(A : C) for the integrable and chaotic channels have near identical behavior until just before t = 4. This is the expected linear growth decrease in mutual information with velocity  $v_E$ 

$$I(A:C) = 6 - v_E st, (5.58)$$

where we have s = 2 for our spin chain channels, and we can read off  $v_E = 0.625$ . The relationship between  $v_E$  and  $v_B$  for these spin chains is consistent with results in holographic systems in the sense that  $v_E \leq v_B$ , but it curious that these 1 + 1-dimensional channels have  $v_E \neq v_B$ . Finally, unlike the previous subsystem configuration, the linear decrease of I(A:C) begins immediately since the distance between subsystems A and D is 0 spins.

In the lower-right panel, we see that I(A : D) for both spin chains also lay on top of each other, but only until just before t = 2. At this time, the chaotic channel saturates just above the Haar-scrambled value, while the integrable channel begins a rough pattern of linear growth to 3 bits followed by return to the Haar-scrambled value. This growth and collapse is the analog of the memory spike we discuss in Appendix 5.7.2, broadened by the



Figure 5-8: Chaos and scrambling in the integrable (blue; g = 1, h = 0 Hamiltonian) and chaotic (orange; g = -1.05, h = 0.5 Hamiltonian) spin chain channels and the Haar-random channel (black). **Top Left:** configuration of channel (n = 6 spins; input subsystems of a = 3, b = 3 spins; output subsystems of c = 3, d = 3 spins). For the spin chains, the channel is the time evolution operator ( $U = e^{-iHt}$ ). For the Haar-channel, U is sampled from the Haar ensemble. **Top Right:** The negative of the tripartite information normalized by its maximum value (6 bits) is a simple diagnostic of scrambling. **Bottom Left:** shows the linear decrease of I(A : C) in time with characteristic speed  $v_E$  known as the entanglement or tsunami velocity. **Bottom Right:** the spike of the unequal subsystem configuration (Fig. 5-7) is broadened to linear increase followed by decrease in I(A : D) for the integrable channel.

fact that the spatial separation between A and D (0 spins) is no longer larger than the size of the regions (3 spins) as required to see the spike.

## 5.4.2 Majorana fermion fast scrambler

### Unequal sized subsystems

Finally, we consider the 0-dimensional Majorana fermion fast scrambler (5.53) with N = 14fermions represented with n = 7 spins. In Fig. 5-9, the inputs and outputs are divided unequally, with inputs subsystems of a = 1, b = 6 spins and outputs subsystems of c = 6, d = 1 spins. (There is no spatial arrangement for the 0-dimensional system.) We plot the OTO average in the top-left panel, the entropies  $S_{AC}^{(1,2)}$  in the top-right panel, I(A : C) in the bottom-left panel, and the normalized tripartite information in bottom-left panel. (We do not plot I(A : D) since for both the fermion channel and the Haar-channel it never becomes greater than  $6 \times 10^{-4}$  bits.)

The main difference between the fermion channel and the spin chains in the unequal configuration is the relevant time scale of the butterfly effect. In the spin chain channels, the initially delay before the OTO correlator decay scales with c, the distance between A and D. This is because operators in A have to grow to encompass the entire spin chain so that the OTO correlators between operators in A and D can be affected. In the fermion channel, there is no notion of spatial locality so the correlators begin to decay immediately.

We don't plot the Majorana fermion channel for equal system sizes since it has essentially the same behavior: all quantities quickly asymptote to Haar-scrambled values similar to Fig. 5-9. This is in slight contrast to chaotic spin chain (plotted in Fig. 5-7), in which the relevant quantities never quite reach the Haar-scrambled values. This suggests that the Majorana fermion system has stronger scrambling power than the chaotic spin chain. Nevertheless, since the late-time values of these quantities in both channels always asymptote to very near the Haar-scrambled values, local measurements cannot differentiate the time-evolved chaotic channels from the Haar-random channel. Thus, the butterfly effect implies scrambling.



Figure 5-9: Chaos and scrambling in the 0-dimensional Majorana fermion fast scrambler unitary channel (blue) and Haar-channel (black) with n = 7 spins; input subsystems of a = 1, b = 6 spins; output subsystems of c = 6, d = 1 spins. **Top Left:** the average of OTO correlators decays immediately, showing the butterfly effect. **Top Right:**  $S_{AC}$  (solid) and  $S_{AC}^{(2)}$  (dotted) shows roughly the same behavior as the OTO average. **Bottom Left:** I(A : C), a trivial function of  $S_{AC}$ , show that for an initial 2 bits of information between the subsystems in the fermion channel gets delocalized so that at late times only a small amount (0.59 bits) remains. **Bottom Right:** The negative of the tripartite information normalized by its maximum value (2 bits) is a simple diagnostic of scrambling.

# 5.5 Perfect tensor model

Now that we understand the relationship between strong chaos and the scrambling behavior of quantum channels, we will present a tensor network model of a scrambling channel with ballistic operator growth.<sup>16</sup> This model serves two purposes.

First, it is useful as a tractable model of ballistic scrambling. The network implements the expected entanglement structure of chaotic time evolution with a (discretized) time independent Hamiltonian. It also serves as a concrete toy model to study the growth of computational complexity in scrambling quantum channels consisting of local quantum circuits.<sup>17</sup>

Second, it provides a model for the interior of the eternal AdS black hole [106]. In [85], it was proposed that the interior connecting the two asymptotic regions can be represented by a flat tensor network whose length is proportional to the total time evolution on the boundary. In [79] and [1], this proposal was explored in a larger variety of black hole states that were perturbed by shock waves.

Here, we provide a concrete model of such a network (i.e. we specify the tensors). This is in the spirit of previous work on the AdS ground state: in [84] it was suggested that the ground state of AdS can be represented by a hyperbolic tensor network (such as MERA [161]), and then an explicit tensor network model was proposed in [150] (see also [162]).<sup>18</sup>

Before we begin, let us review the proposal of [85]. The tensor network representation of the thermofield double state is shown in Fig. 5-10. At the left and right ends, we have a hyperbolic network, representing the two asymptotically AdS boundaries. This network extends infinitely from the UV into the IR thermal scale  $\beta$  at the black hole horizon. Then, the middle is flat representing the black hole interior. The entire network grows as t grows by adding more layers in the middle flat region.

We would like to further elaborate on this proposal of tensor network representation of the black hole interior. We will study networks of perfect tensors and demonstrate chaotic dynamics by finding ballistic growth of local unitary operators and the linear growth of the tripartite information until the scrambling time. For the rest of discussion, we take the

 $<sup>^{16}</sup>$ See also [159] for a similar recent tensor network model of ballistic entanglement propagation.

 $<sup>^{17}</sup>$ In fact, there is a well-defined notion of the complexity of randomness, called unitary *t*-designs, and lower bounds on the complexity growth under random quantum circuits in this sense have been rigorously established [160].

<sup>&</sup>lt;sup>18</sup>This model has the additional nice property of implementing the holographic quantum error correction proposal of [163].

infinite temperature  $\beta = 0$  limit so we can ignore the hyperbolic part and focus in on the planar tiling of tensor networks representing the interior.



Figure 5-10: Tensor network representation of the Einstein-Rosen bridge. A four-leg tensor lives at each node. We will consider a network of perfect tensors.

# 5.5.1 Ballistic growth of operators and linear growth of $I_3$

Let us review the definition of perfect tensors. Consider a tensor  $\mathcal{T}$  with 2n legs and bond dimension v. A tensor can be represented as a pure state

$$|\Psi\rangle = \sum_{i_1,\dots,i_{2n}} \mathcal{T}_{i_1,\dots,i_{2n}} |i_1,\dots,i_{2n}\rangle,$$
 (5.59)

with a proper normalization. We call a tensor  $\mathcal{T}$  perfect if it is associated with a pure state  $|\Psi\rangle$ , called a perfect state, which is maximally entangled along any bipartition. Namely,

$$S_A = n, \quad \forall A \text{ s.t } |A| = n,$$
 (5.60)

where for tensors of bond dimension v we measure entropy in units of  $\log v$ . The qutrit tensor Eq. (5.26) is an example of a perfect tensor. There are known methods for constructing perfect tensors via the framework of quantum coding theory. Also, a Haar random tensor becomes a perfect tensor at the limit of  $v \to \infty$ .

### Growth of local operators

Imagine a flat planar tiling of 4-leg perfect tensors as shown in Fig. 5-10 which may be thought of as a discretized time-evolution by a strongly-interacting Hamiltonian. We can examine time evolution of a local unitary operator V and observe linear growth of spatial profiles of operators V(t) by using a basic property of perfect tensors. Let  $|\Psi\rangle$  be a 4-spin perfect state and denote 4 legs by a, b, c, d. Consider a single-body unitary operator  $U_a \neq I$ acting exclusively on a. Since ab and cd are maximally entangled, there always exists a corresponding operator  $VU_{cd} \neq I$  acting exclusively on cd such that

$$U_a|\Psi\rangle = U_{cd}|\Psi\rangle,\tag{5.61}$$

or in the tensor representation, we have

$$= \underbrace{U_{a}}_{U_{cd}}, \qquad (5.62)$$

where a gray square represents a four-leg perfect tensor. One can prove that  $U_{cd}$  must act non-trivially both on c and d. Namely, if  $U_{cd}$  were a single-body operator acting only on c (i.e.  $U_{cd} = U_c$ ), then one would have  $U_a U_c^{\dagger} |\psi\rangle = |\psi\rangle$ . However this contradicts with the fact that ac and bd are maximally entangled. To see the contradiction, one can simply use  $U_a U_c^{\dagger} |\psi\rangle \langle\psi| = |\psi\rangle \langle\psi|$  and take a partial trace over b, d on both sides of the equation. If a, c is maximally entangled with b, d we obtain  $U_a U_c^{\dagger} = I$  is the identity operator, which is not possible. The conclusion is that, due to the perfectness of the tensors, each two-qudit unitary associated with perfect tensors always expands a single-body operator to a two-body operator. This observation is consistent with linear ballistic propagation of entanglement for single connected regions predicted for chaotic systems [164].

The implication of this ballistic expansion of unitary operators under perfect tensors is quite interesting. The size of the region of nontrivial support for V(t) increases linearly as shown in Fig. 5-11. At t = L/2, for a lattice of linear size L, a local operator will evolve into a global operator supported over the entire lattice. The growth of OTO correlation functions originates from this linear growth of spatial profiles of local operators. Namely,



Figure 5-11: Ballistic growth of local operators by perfect tensors. As a reminder, the horizontal radius of the operator V grows with the butterfly velocity as  $v_B t$ , and the vertical depth of the circuit grows as  $v_E t$ .

for a local operator W which is separated in space from V(t = 0), the commutator [V(t), W]becomes non-negligible after t = L/2 indicative of the butterfly effect.<sup>19</sup>

### Growth of tripartite information in time

Let us then compute the tripartite information for a network of perfect tensors. The entire system is split into four regions A, B, C, D of equal size as in Fig. 5-12(a). The growth of entanglement entropy can be exactly calculated by using a method developed in [150]. Recall that, for a perfect state  $|\Psi\rangle$  with four spins, there always exist a two-qubit unitary operator  $D_{ab}$  such that

$$D_{ab}|\Psi\rangle = |EPR\rangle_{ac} \otimes |EPR\rangle_{bd}.$$
(5.63)

In other words,  $D_{ab}$  disentangles a perfect state into two decoupled EPR pairs as graphically shown below

$$D_{ab} = \begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}$$
(5.64)

<sup>&</sup>lt;sup>19</sup>A qualitatively similar behavior occurs when Haar-random unitary operators are used instead of perfect tensors, which we checked numerically. For an analytical discussion a random tensor network in the context of a holographic state rather than a channel, see [165]. (See also [166].)

A key observation is that the process of finding a minimal surface by local updates can be viewed as entanglement distillation by applications of disentanglers. This led to the proof of the Ryu-Takayanagi formula for single intervals in networks of perfect tensors [150].

In general, calculation of entanglement entropies for disjoint regions is challenging even for networks of perfect tensors. Indeed, the verification of the Ryu-Takayanagi formula is given only for single intervals for a network of perfect tensors in [150]. Here we assume that the planar tensor network is translationally invariant in both time and spatial directions. To be specific, we also assume that the network consists of the qutrit perfect tensor introduced in (5.26). For such a perfect tensor tiling, an analytical calculation of the tripartite information is possible for time t shorter than the scrambling time  $t_* = L/2$ . Namely, one can prove the following:

$$I_3(A:B:C) = -2t, \qquad 0 \le t \le L/2,$$
(5.65)

where L is the linear length of the system and as a reminder for qutrits we measure entropy in units of log 3. Below, we sketch the derivation.

For a network of perfect tensors, entanglement properties can be studied by applying local disentanglers which correspond to distillations of EPR pairs. The disentanglers map each region unitarily to the minimal surface bounding it, as is shown in Fig. 5-12. For time  $t \leq L/2$ , one observes that minimal surfaces for A, C collide with each other, and similarly for B, D. Let us distill entanglement as shown in Fig. 5-12(b) by applying some appropriate local unitary transformations on each region and remove decoupled spins. Regions AC and BD possess EPR-like entanglement along the collided surface of geodesic lines. In the middle of the network, we find square regions which are responsible for four-party entanglement among A, B, C, D. Such regions, which are not included inside causal wedges of boundary regions, are referred to as residual regions [150]. These become essential in understanding entanglement properties behind the horizons of the multi-boundary black holes considered in [167]. At the time step t, there will be a pair of square residual regions with linear length t/2 as shown in Fig. 5-12(b). In Appendix 5.7.4, we study multipartite entanglement for rectangular residual regions. Namely, we show that each residual region contributes to the tripartite information by -t. We thus obtain Eq. (5.65).



Figure 5-12: (a) Computing  $I_3$  for A, B, C, D of equal size. Local moves are shown. (b) Residual tensor networks. Square-like tensor networks are responsible for the tripartite information.

# 5.5.2 Recurrence time

We have shown that the network of perfect tensors, as shown in Fig. 5-10 and Fig. 5-11, serves as a toy model of scrambling dynamics. A naturally arising question is whether such a system stays scrambled after the scrambling time  $t_* = L/2$ . In this section, we study the recurrence time of the planar network of perfect tensors. For concreteness, we will restrict our considerations to those with qutrit perfect tensors. We assume periodic boundary conditions in the spatial directions of the network.

Imagine that we inject some Pauli operators from the top of the tensor network and obtain output Pauli operators on the bottom. We are interested in the minimal time step necessary for a network to output the initial Pauli operators again. To find the recurrence time, we inject two-body Pauli Z operators from the top left corner of the tensor network and compute the output Pauli operator on the bottom. We define the recurrence time  $t_{rec}$ to be the minimal time step  $t_{rec}$  necessary for the network to output the initial two-body Pauli Z operators. Recall that the tensor network based on stabilizer tensors maps Pauli operators to Pauli operators. Since Pauli operators can be treated as classical variables, one can efficiently find the recurrence time via numerical methods. The recurrence time crucially depends on the system size L as shown in the plot in Fig 5-13. Note that the plot uses a logarithmic scale. When the system size is  $L = 3^m$ , the recurrence time grows only linearly:  $t_{rec} = 4L$ . This expression can be analytically obtained. The linear growth is due to the fact that the qutrit tensor can be viewed as a linear cellular automaton over  $\mathbb{F}_3$  which has scale invariance under dilations by factor of 3. For such special system sizes, the trajectories of time-evolution of Pauli operators form short periodic cycles. This is similar to the classical billiard problem where trajectories of a billiard ball are not ergodic for fine-tuned system sizes and fine-tuned angles. Yet, the billiard problem is ergodic for generic system sizes.

Likewise, the perfect tensor network has longer recurrence time for generic values of system sizes. When L is a prime number, the growth is rather fast, and seems exponential as shown in Fig. 5-13. (We do not have an analytic proof of this statement.) Assuming the exponential growth of the recurrence time for prime L ( $t_{rec} \approx e^{kL}$ ), let us find out the growth for typical values of L. For typical values of L, we expect that  $t_{rec}$  grows faster than any polynomial functions. This is because given a positive integer n, the probability for its largest prime factor to be larger than, say  $\sqrt{n}$ , is finite.<sup>20</sup> Assuming that L is not a prime number, let us decompose it as  $L = L_1L_2$ . Then, due to the translation invariance, one can show that  $t_{rec}(L) \ge t_{rec}(L_1), t_{rec}(L_2)$ . As such, the recurrence time  $t_{rec}(L)$  will be lower bounded by  $t_{rec}(p)$  where p is the largest prime factor of L. This argument implies an exponential growth of the recurrence time for typical values of L.

The recurrence time of the perfect tensor network is much longer than that of integrable systems, but is much shorter than that of chaotic systems. By construction, perfect state network based on the stabilizer formalism can have at most exponential recurrence time. This is essentially because unitary circuits implemented by stabilizer-type tensors belong to the so-called Clifford group which is a subgroup of unitary transformations that map Pauli operators into Pauli operators.<sup>21</sup> Quantum circuits solely consisting of Clifford operators are classically simulable since transformations of Pauli operators can be efficiently characterized by pairs of classical bits. In this sense, the stabilizer perfect tensor network exhibits marginally chaotic behaviors. The classical simulability enables us to study chaos

 $<sup>^{20}</sup>$  In general, the probability for the largest prime factor to be larger than  $1/n^u$  is given by the Dickman function [168].

<sup>&</sup>lt;sup>21</sup>This means that under time evolution a simple Pauli operator X, Y, or Z can only grow into a product of Pauli operators (rather than a sum of products of Pauli operators as would be generically expected by (5.3)).


Figure 5-13: The exponential growth of recurrence time for perfect tensor networks of prime size (blue) and non-prime size (black).

and scrambling behaviors in quantum channels at relatively early times in a computationally tractable manner.

This highlights an important point about perfect tensor networks as models of holography. In many cases they can exhibit key features expected of holographic systems (such as the error correcting and bulk reconstruction properties of the model presented in [150]). However, since the recurrence time of the perfect tensor network is exponential in the system size and not doubly exponential, it's clear that it fails to capture a very important feature: the possibility of exponential computational complexity. In particular, the (comparatively) quick recurrence means that the longest minimal perfect tensor networks are far less complex than the degree of complexity that generic holographic states are expected to reach.

One possible resolution is to modify the stabilizer perfect tensor by applying some single qudit non-Clifford rotation, such as a rotation around the Z axis by some angle  $\theta$ . An inclusion of a single non-Clifford operator to the full Clifford group enables us to efficiently approximate an arbitrary unitary operator, an important result known as the Solovay-Kitaev theorem [151]. As such, we speculate that non-Clifford modification of perfect tensors would create a tensor network with doubly exponential recurrence time. This resolution is along the spirit of the billiard problem since the Clifford transformations are fine-tuned operations.

Another possible resolution of this is that to reach the more complicated states (which are not at all understood holographically, see e.g. [34, 35, 53, 54]), one needs to consider superpositions of such tensor networks which do not have a geometric description and thus would not be expected to have a semiclassical bulk interpretation.

Regardless, a network of perfect tensors is very capable of scrambling. This observation leads us to envision that a certain measure of complexity can be attached to each tensor in the network, in particular,  $-I_3$  up to proper normalization. This would represent the complexity of forming the four-leg perfect tensor from a product state. It would be interesting to see if some kind of upper bound on the gate complexity can be imposed by considering an integral of  $-I_3$  over all the tensors in the network.

## 5.6 Discussion

In this paper, we have shown that the butterfly effect—as expressed by the decay of out-oftime-order (OTO) correlation functions—implies the information-theoretic notion of scrambling. The butterfly effect is manifested by the growth of simple operators under time evolution to complicated operators of high weight. These time-evolved operators will then have large commutators will all other operators in the system. If we think of the initial simple operator as an input to a unitary quantum time-evolution channel, then the output will be an operator spread over the entire system. All information associated with the input will be delocalized; the output system is scrambled.

The method of characterizing scrambling/chaos via the framework of quantum channels may also find interesting applications in studying thermalization in many-body quantum systems. We have already demonstrated the usefulness of our approach by studying the tripartite information in several different examples: numerical results in integrable/nonintegrable spin chains and the nonlocal interacting Majorana fermion model of Kitaev, and both analytical and numerical results in a perfect tensor network model of discretized time evolution. It would be interesting to study many-body/single-body localization and delocalization transitions in the setup of quantum channels. A closely related question may concern the information-theoretic formulation of the Eigenstate Thermalization Hypothesis (ETH). The state interpretation of the channel is able to consider a set of initial states as well as to probe off-diagonal elements in the Hamiltonian.

In order for quantum information to really be processed, it has to interact with the other information distributed across the system. Said another way, to process information the channel has to be capable of scrambling. This suggests that there is a strong connection between quantum chaos and computation. As a surrogate for a definition of computation, let's consider the computational complexity of the quantum circuit or channel. For tensor network models, this is simply the number of tensors in the minimal tensor network.

As a simple example, let's consider the quantum channel that only contains swap gates. The channel doesn't scramble, and information can only be moved around. As discussed multiple times, the swap channel has a quick recurrence and can never get very complex. The only output states accessible are those related to permutations of the input, all the multipartite states cannot be accessed. For a system of n qubits, the complexity can only ever reach O(n) (the complexity of swapping localized information from one end of the system to the other using local swap operations). The maximal complexity of a state of nqubits is  $O(2^n)$ ; thus, for the simple swap channel most of the possible output states are entirely inaccessible. It is essentially only capable of classical computation.

Quantum computation requires interaction, and strong chaos is a signature of a strongly interacting system. Thus, in some sense, we speculate chaos must be the capacity for a system to do computation. This suggests that strongly chaotic-systems must be fast computers. In fact, in [2, 49] it was recently hypothesized that black holes are the fastest computers in nature. Given that black holes are already known to be nature's densest hard drives [21, 22] and most chaotic systems [58], it seems reasonable to suspect that a system's computational power must be limited by its degree of chaos. It would be interesting to try and make this dependence on chaos for computation more direct.

## 5.7 Appendix

### 5.7.1 Haar scrambling

In this paper, we've generally considered scrambling by a one-parameter family of unitary operators  $U(t) = e^{-iHt}$  and found that for chaotic systems, increasing time t leads to more efficient scrambling. Instead, we will now take U to be a Haar random unitary operator. This is useful as a baseline for scrambling. We expect that the late-time values of entropies and informations computed from scrambling operators U(t) will asymptote to Haar random values. Additionally, we will see that Haar-random values of quantities such as I(A : C) and  $I_3(A : C : D)$  are not necessarily maximal, but exhibit "residual" information regardless of system size.

At its fastest, Page scrambling has a complexity of  $n \log n$  gates, while Haar scrambling is nearly maximally complex requiring ~  $e^n$  gates. In [37], it was proven that Haar scrambling implies Page scrambling. As we will now show, Haar scrambling also implies the tripartite scrambling. However, the implication does not work in the other direction: since the latetime values of I(A : C), I(A : D), and  $I_3(A : C : D)$  approach Haar-scrambled values, entanglement is not enough [87] to diagnose typicality (in the Haar-random sense). This possibly has a strong bearing on the paradoxes of [34, 53] as discussed in [35].

To proceed with this analysis, we will consider an expectation over density matrices constructed from Haar-random states. These tools were used by Page to analyze the entropy of subsystems for random states [37], and our approach will be similar to [38] and Appendix A of [44]. In fact, our calculation is very similar to that in [44]. However, beware that the results do not simply carry over; since we are pairing together input and output subsystems of possibly different size (in our notation, the fact that  $a \neq c$ ), we will find a very different result.

Our setup will be the usual division into subsystems ABCD, with the state given by (5.11), and U a random  $2^n \times 2^n$  unitary matrix taken from the Haar ensemble. The Haar average lets us consider expectations over a number of unitary matrices and is non-zero only when the number of Us equals the number of  $U^{\dagger}s$ . For instance, with two Us and two  $U^{\dagger}s$ , the formula for the average is

$$\int dU \ U_{i_1j_1} U_{i_2j_2} U_{i'_1j'_1}^* U_{i'_2j'_2}^* = \frac{1}{2^{2n} - 1} \Big( \delta_{i_1i'_1} \delta_{i_2i'_2} \delta_{j_1j'_1} \delta_{j_2j'_2} + \delta_{i_1i'_2} \delta_{i_2i'_1} \delta_{j_1j'_2} \delta_{j_2j'_1} \Big)$$
(5.66)  
$$- \frac{1}{2^n (2^{2n} - 1)} \Big( \delta_{i_1i'_1} \delta_{i_2i'_2} \delta_{j_1j'_2} \delta_{j_2j'_1} + \delta_{i_1i'_2} \delta_{i_2i'_1} \delta_{j_1j'_1} \delta_{j_2j'_2} \Big).$$

This formula will let us compute the average over the trace of the square of the density matrix  $\rho_{AC}$ 

$$\int dU \operatorname{tr} \{\rho_{AC}^2\} = \frac{1}{2^{2n}} \int dU \ U_{k\ell m o} \ U_{k'\ell m' o}^* U_{k'\ell' m' o'}^* U_{k\ell' m o'}^*, \tag{5.67}$$

where as in Appendix 5.7.3,  $k = 1 \dots 2^a$  are A indices,  $\ell = 1 \dots 2^b$  are B indices,  $m = 1 \dots 2^c$ 

are C indices, and  $o = 1 \dots 2^d$  are D indices. In applying the average (5.66) to (5.67), note that both k and  $\ell$  in (5.67) are "*i*"-type indices in (5.66), and similarly m and o are "*j*"-type indices. After a quick game of delta functions, we find

$$\int dU \operatorname{tr} \{\rho_{AC}^2\} = 2^{-a-c} + 2^{-b-d} - 2^{-a-d-n} - 2^{-b-c-n}.$$
(5.68)

As of now, we have been completely general about the size of the subsystems. For simplicity, let's now focus on the case of a+c = n and b+d = n as was often considered in our numerics. Without loss of generality, let's also take  $a \leq b$  and  $d \leq c.^{22}$  Simplifying and throwing away exponentially subleading pieces, we get

$$\int dU \operatorname{tr} \{\rho_{AC}^2\} \simeq 2^{1-n} \left(1 - 2^{-a-d-1}\right).$$
(5.69)

Using this with an appropriate caveat,<sup>23</sup> we can compute the Haar average of the Rényi entropy

$$(S_{AC}^{(2)})_{Haar} = n - 1 - \log_2(1 - 2^{-a-d-1}).$$
(5.70)

This is rather interesting: the maximal value for  $S_{AC}^{(2)}$  is *n*. Therefore, the Haar-scrambled state never reaches this maximal value. On the other hand, this is not unexpected. The corrections to Page's entropy of a subsystem formula for the divisions we are considering are expected to be O(1) [37].

Next, let us use (5.70) to put a bound on the mutual information I(A:C). We can do this using the fact that  $S_{AC} > S_{AC}^{(2)}$ , and we find

$$I(A:C) \le 1 + \log_2(1 - 2^{-a-d-1}),$$
 (5.71)

indicating the possibility of residual information between A and C that is independent of

$$\int dU \, S_{AC}^{(2)} = - \int dU \, \log_2 \operatorname{tr} \{ \rho_{AC}^2 \},$$

 $<sup>^{22}</sup>$ Ref. [44] neglects the bottom line of (5.66) as subleading. For the subsystems we consider, the first term on the second line of (5.66) is actually the same order (in n) as the terms in the first line and cannot be neglected.

<sup>&</sup>lt;sup>23</sup>To use this result to compute the Haar average of the Rényi entropy

we need to commute the Haar average with the log. This can be checked numerically and holds as long as n is sufficiently large (which in this case "large" means about n = 4). For large n, the system self-averages so that any single sample is extremely likely to be at the mean value. This lets the average commute with the log.

n, even in the Haar-scrambled limit.<sup>24</sup> By considering more equal partitions of the input  $(1 \ll a \leq b)$ , there will be more residual information between A and C, though the fraction of residual information I(A:C)/2a decreases.

Let's complete our discussion by trying to bound I(A : D). Following the approach outlined above, we find

$$\int dU \operatorname{tr} \{\rho_{AD}^2\} = 2^{-a-d} + 2^{-b-c} - 2^{-a-c-n} - 2^{-b-d-n},$$
(5.72)

which is the same as (5.68) with  $c \Leftrightarrow d$ . Taking our simplifying assumption a + c = n and b+d = n, with  $a \leq b$  and  $d \leq c$ , we see that the three latter terms are exponentially smaller in n than the first term. We find

$$(S_{AD}^{(2)})_{Haar} = a + d + O(2^{-2n+a+d}),$$
(5.73)

and we can bound the mutual information as

$$I(A:D) \le 0 + O(2^{-2n+a+d}). \tag{5.74}$$

Finally, we note that in the general case where  $a + c \neq n$  and  $b + d \neq n$ , the mutual information (5.70) is different. With  $a \leq c$ , we would instead find

$$I(A:C) \le \log_2(1 + 4^{a-d} - 4^{-d}).$$
(5.75)

Here, if a = d we still have a residual bit of information, but with a < d we do not.<sup>25</sup>

## 5.7.2 Entanglement propagation in CFT

In Fig. 5-3(b), we pointed out a strong resemblance to Feynman diagrams of a  $2 \rightarrow 2$  scattering process. The swap gate resembles a diagram that contribute to a noninteracting theory: the only allowed operation is that the particles can swap locations between the inputs and the outputs (or they can do nothing). On the other hand, the perfect tensor resembles a Feynman diagram that contributes to a scattering process in an interacting

 $<sup>^{24}</sup>$ One might have thought that it would be possible to make a better bound with the Fannes-Audenaert inequality [169, 170] and using the 2-norm to bound the 1-normal as in [38] and [44]. However, such an approach actually leads to a bound that's actually much looser than the simple one given by (5.71).

 $<sup>^{25}</sup>$ We thank Alexei Kitaev for pointing out that our first result (5.70) is not general.

theory. This is not a coincidence; the strength of chaos should be related to the strength of the coupling, see e.g. [58].

With this point of view, let us consider entanglement propagation in CFT. The general setup considered in CFT is a global quench; the system is preprepared in a groundstate of a Hamiltonian  $H_0$  and then the Hamiltonian is suddenly changed to a different Hamiltonian H such that the system is now in a finite energy configuration. The system is then evolved with the new Hamiltonian and certain entanglement entropies saturate at their thermal values. This often referred to as thermalization. For two-dimensional CFT, the entanglement entropy of a single connected region after a global quench was shown to grow linearly in time, saturating at its thermal value at a time of order half the length of the region [171–173]. To explain this, [171] proposed that entanglement is carried by pairs of entangled noninteracting quasi-particles that travel ballistically in opposite directions.<sup>26</sup> The quasi-particles would travel at the entanglement velocity  $v_E$ , which is equal to unity in two-dimensional CFT. Entanglement entropy increases as the entangle pairs are split between the region and its complement.

This model of entanglement propagation is sufficient to explain the pattern of entanglement growth after a quench of a single interval (in fact, the result is universal [171–173]), but gives puzzling results for the entanglement entropy of two separated disjoint intervals in interacting (e.g. holographic) systems [42, 140, 158, 174, 175].<sup>27</sup> Let us label the two intervals as F and G, both of size L, and the rest of the system as H. We will take F and G to be separated by a distance  $\mathcal{D}$ , and all the scales are taken to be much greater than the thermal correlation length  $L, \mathcal{D}, \gg \beta$ . Additionally, we require  $\mathcal{D} > L$ . This setup is depicted in the left-hand side of Fig. 5-14.

In the quasi-particle model, after a quench  $S_{FG}$  will grow linearly for a time D and then saturate at its thermal value. However, at time of order  $(\mathcal{D} + L)/2$  it will exhibit a dip. In the quasi-particle picture, entangled pairs created in the region between F and G are beginning to enter F and G, respectively, causing the entanglement entropy of FG with the rest of the system H to dip. For holographic systems, this is known not to happen: after  $S_{FG}$  saturates, it remains saturated [42, 140, 158, 174, 175].

This puzzle was explored in depth in the context of two-dimensional CFT in [42]. There,

<sup>&</sup>lt;sup>26</sup>See also [159] for a generalization of the non-interacting quasi-particle model to interacting systems and a discussion of the entanglement velocity.

<sup>&</sup>lt;sup>27</sup>See also [85, 92, 93] for holographic investigations of entanglement growth after a global quench.



Figure 5-14: Quench in CFT showing path of a quasi-particle EPR pair and two intervals of length L separated by distance  $\mathcal{D} > L$ . Left: for two intervals in the noninteracting quasi-particle model  $S_{FG}$  has a dip at time  $t \sim (\mathcal{D} + L)/2$ . Right: a simplified description in terms of two entangled CFT involves one partner of the EPR pair traveling in the left CFT and one partner traveling in the right CFT. This can be reinterpreted as a quantum channel. To make contact with the notation in the paper, we relabel as  $F' \to A$ ,  $G' \to D$ , and  $H' \to BC$ .

it is shown that the quasi-particle picture cannot be universal and must depend on the spectrum. Indeed, [42] concludes that for interacting CFTs "entanglement scrambles"— meaning there's no memory effect or dip in  $S_{FG}$ . Here, we would like to put these results in the context of unitary quantum channels. We will show that "entanglement scrambling" is precisely scrambling as diagnosed by the tripartite information. Furthermore, we will argue that the cause of such entanglement scrambling is chaotic dynamics. Strong chaos implies a picture of strongly interacting quasi-particles.

To make the connection, one simply has to realize that a global quench can be simply understood as the time evolution of the thermofield double state [85]. That is, a global quench is the channel  $|TFD(t)\rangle$  given by (5.14). This was also pointed out in [42], where the thermofield double state was used to simplify the setup of the two-interval calculation while retaining the basic puzzle.<sup>28</sup> In this new setup, the puzzle is a memory effect between an interval on the left CFT F' and interval on the right CFT G', where the spatial separation  $\mathcal{D}$  and interval sizes L are large. After a time of order  $(\mathcal{D} + L)/2$ , the quasi-particle model predicts a dip in the entanglement between F'G' and the rest of the system H'. This new setup is shown in the right-hand side of Fig. 5-14.

Now, let us relabel the subsystems:  $F' \to A, G' \to D$ , and  $H' \to BC$ . With the

 $<sup>^{28}</sup>$ In fact, in [42] the memory effect was diagnosed by considering properties of the second Rényi entropy for the intervals in question.

perspective of the unitary channel setup (Fig. 5-2), the memory effect is simply a question of whether I(A : D) has a spike. Integrable systems will have a spike and can be modeled by entanglement carrying quasi-particles (the swap gate in Fig. 5-3(b) provides an explicit cartoon of such noninteracting quasi-particles). Chaotic systems are strongly interacting, and the quasi-particle picture breaks down (the perfect tensor in Fig. 5-3(b) provides the cartoon for the interacting system). This memory effect was shown explicitly in the bottommiddle panel of Fig. 5-7 for the integrable spin chains we studied in section §5.4.

This connection to the work of [42] allows us to probe scrambling and chaos in particular CFTs. For instance, the results of section 4.4 in [42] suggest that the D1-D5 CFT at the "orbifold point" does not scramble (in the sense of tripartite information) as expected for a free theory. It would be interesting to make additional connections between scrambling/chaos and CFT results.

### **5.7.3 Proof of Eq.** (5.32)

The proof of the relation

$$|\langle \mathcal{O}_D(t) \, \mathcal{O}_A \, \mathcal{O}_D(t) \, \mathcal{O}_A \rangle_{\beta=0}| = 2^{n-a-d-S_{AC}^{(2)}},\tag{5.76}$$

is probably most easily understood diagrammatically as shown in Fig. 5-15. For simplicity of discussion, we assume a system consisting of qubits while our discussion straightforwardly generalizes to a system consisting of qudits by considering generalized Pauli operators. Thus, this proof applies to lattice systems with a finite-dimensional Hilbert space at each site.<sup>29</sup>

To proceed with the proof, we need to make use of an operator identity. Consider a partition of a system AB with  $A_j$  a complete basis of operators in A. Then, for any operator  $\mathcal{O}$  on the entire system AB, we have

$$\sum_{j} A_{j} \mathcal{O} A_{j} = |A| I_{A} \otimes \operatorname{tr}_{A} \{\mathcal{O}\},$$
(5.77)

 $<sup>^{29}</sup>$ In a continuum limit, we would need some notion of the operator identity (5.77), which is the completeness condition for a basis of operators. Naively, due to the infinite Hilbert space dimension, (5.32) is trivially true; the Rényi entropy is UV-divergent and the correlation function average is vanishing due to normalizing by the total number of operators. However, the connection between our results and entanglement propagation in CFT (see Appendix 5.7.2) suggests that perhaps a recasting of the relation (5.32) in terms of mutual information might lead to a sensible continuum limit.



Figure 5-15: Diagrammatic proof of the OTO average  $|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle_{\beta=0}|$ . Dotted lines indicate summation over operators (gray rectangles), and in all diagrams the normalization is suppressed. (a) The operator identity Eq. 5.77. (b)  $\rho_{AC}$  for the unitary channel. (c) Calculation of the average. In the final panel, the lines at the top and the bottom are appropriately connected.

where the sum *i* runs over the entire basis, and |A| is the size of the Hilbert space of A. The set of all qubit Pauli operators supported in A forms a complete basis of orthonormal operators, and (5.77) can be easily verified by decomposing  $\mathcal{O}$  into that basis. A diagrammatic depiction of this identity is shown in Fig. 5-15(a).

We would like to use this to evaluate the averaged correlator

$$2^{-2a-2d-n} \sum_{ij} \operatorname{tr} \{ D_i(t) A_j D_i(t) A_j \} = 2^{-2a-2d-n} \sum_{ij} \operatorname{tr} \{ U^{\dagger} D_i U A_j U^{\dagger} D_i U A_j \}.$$
(5.78)

Here the prefactor  $2^{-2a-2d}$  is the inverse of number of operators in A and D, and  $2^{-n}$  is the normalization factor such that the quantity equals to 1 if all operators are identity operators. Let's apply (5.77) to  $D_i U A_j U^{\dagger} D_i$  to do the sum over *i*. This gives

$$2^{-2a-d-n}\sum_{j} \operatorname{tr} \{ UA_{j}U^{\dagger} \operatorname{tr}_{D} \{ UA_{j}U^{\dagger} \} \otimes I_{D} \},$$
(5.79)

where note that we have made use of the cyclicity of the trace. At this point, it's useful to adopt indices. We will use  $k = 1 \dots 2^a$  for A indices,  $\ell = 1 \dots 2^b$  for B indices,  $m = 1 \dots 2^c$ for C indices, and  $o = 1 \dots 2^d$  for D indices. This lets us rewrite (5.79) as

$$2^{-2a-d-n} \sum_{j} U_{k_1\ell m o}(A_j)_{k_1k_1'} U^*_{k_1'\ell m' o} U_{k_2\ell'm' o'}(A_j)_{k_2k_2'} U^*_{k_2'\ell'm o'}, \tag{5.80}$$

where repeated indices imply summation. Now, we apply (5.77) again, specifically to  $(A_j)_{k_1k'_1}U^*_{k'_1\ell m'o}U_{k_2\ell'm'o'}(A_j)_{k_2k'_2}$ . This sets  $k'_1 = k_2$  and  $k'_2 = k_1$  (and multiplies by  $2^a$ ) to give

$$2^{-a-d-n}U_{k_1\ell m o}U^*_{k_2\ell m' o}U_{k_2\ell' m' o'}U^*_{k_1\ell' m o'}.$$
(5.81)

Now, we remember how to express the density matrix  $\rho$  of our channel (see Fig. 5-15(b))

$$\rho = 2^{-n} U_{k\ell m o} U^*_{k'\ell' m' o'}. \tag{5.82}$$

Applying this to (5.81) and then using the definition of the second Rényi entropy (5.18) gives our desired result (5.32). This whole proof, up to factors of normalization, is shown in Fig. 5-15(c).

#### Finite temperature

It is easy to generalize this formula for finite temperature  $\beta > 0$ . Define

$$Z(\beta) := \operatorname{tr}(e^{-\beta H}), \qquad |\Psi(\beta)\rangle := |TFD(\beta, t)\rangle.$$
(5.83)

To get an expression in terms of an entropy, we need to distribute the operators around the thermal circle

$$|\langle \mathcal{O}_D(t) \mathcal{O}_A \mathcal{O}_D(t) \mathcal{O}_A \rangle_\beta| \approx Z(\beta)^{-1} |\mathrm{tr} \{ \mathcal{O}_D(t - i\beta/4) \mathcal{O}_A \mathcal{O}_D(t - i\beta/4) \mathcal{O}_A \}|.$$
(5.84)

Here, rather taking a thermal expectation value we are evolving the operators in D in Euclidean time (and then renormalizing by  $Z(\beta)$ ). The trace of these Euclidean-evolved correlators is expected to be related to the thermal expectation of the original OTO correlators as long as the temperature is high enough. Following our proof Fig. 5-15 but with the time argument for the unitary operators as  $U(t - i\beta/4)$ , we find

$$=\frac{Z(\beta/2)^2}{Z(\beta)}2^{-a-d-S_{AC}^{(2)}(\beta/2)}.$$
(5.85)

where  $S_{AC}^{(2)}(\beta/2)$  is evaluated for the state  $|\Psi(\beta/2)\rangle$  defined in (5.83).

#### Higher order OTO correlators

Finally, we will briefly comment on another possible generalization. The OTO correlation functions we studied here are observables for the chaotic dynamics of a thermal system perturbed by a single operator. In [54], chaos is studied in holographic thermal systems that are perturbed by multiple operators. For two perturbations, the relevant observable is a six-point OTO correlation function of the form

$$\langle W(t_1) V(t_2) Q V(t_2) W(t_1) Q \rangle_{\beta},$$
 (5.86)

where W, V, Q are all simple Hermitian operators. (For simplicity, we will consider the case where  $\beta = 0$ .) This observable is related to the effect of simple perturbations W, V made at times  $t_1, t_2$  on measurements of Q at t = 0. This correlation can be simplified by summing over a basis of operators in three regions associated with the W, V, Q as we did for four-point functions earlier in this appendix.

However, it's easy to see that one cannot get something as simple as a Rényi entropy: since there's two explicit times  $t_1, t_2$ , we can form density matrices  $\rho(t_1), \rho(t_2)$ , and  $\rho(t_1-t_2)$ , where  $\rho(t) = |U(t)\rangle \langle U(t)|$  is the density matrix associated with the quantum channel of time evolution by  $U(t) = e^{-iHt}$ . The averaged six-point function will be related to contraction of these density matrices with a complicated permutation. This may be considered as a more generic entanglement property of the system, which is beyond Rényi entropies. In the case of finite temperature, for the four-point functions we were able to evolve the operators in Euclidean time in order to symmetrize the time arguments of the unitary operators, as is discussed around (5.84). However, cannot do that for the six-point functions since we need all three operators to be separated from each other in Lorentzian time.

#### 5.7.4 Tensor calculus

In this appendix, we provide more details about the perfect tensor calculation sketched in Fig. 5-12. We have seen that the two rectangular residual regions, which are not contained in any of causal wedges, are responsible for multipartite entanglement arising in a network of perfect tensors. To calculate  $I_3$ , we typically need to consider rectangular residual regions. In this appendix, we study multipartite entanglement in a rectangular network of perfect tensors as shown in Fig. 5-16 where tensor legs are split into four subsets P, Q, R, S. We assume that P, R contain r legs and Q, S contain t legs.

In [150], it was shown that, for any planar network of perfect tensors with non-positive curvature, the Ryu-Takayanagi formula for single intervals holds exactly. Keeping this in mind, let us summarize properties of entanglement in an arbitrary rectangular tiling of perfect tensors:

$$S_P = r, \quad S_Q = t, \quad S_R = r, \quad S_S = t,$$
  
 $S_{PQ} = r + t, \quad S_{QR} = r + t, \quad S_{RS} = r + t, \quad S_{PS} = r + t,$ 
(5.87)

where as a reminder, for tensors of bond dimension v, we measure their entropy in units of  $\log v$ . Thus, the tripartite information  $I_3$  is given by

$$I_3 = -S_{PR}.$$
 (5.88)

Note that the above statement holds for any perfect tensors and is not restricted to qutrit perfect tensors. But the value of  $S_{PR}$  is non-universal for networks of perfect tensors since PR consists of two spatially disjoint intervals. Below, we will prove that

$$I_3 = -2\min(r, t), \tag{5.89}$$

for a network of four-leg qutrit perfect tensors.

The qutrit tensor network discussed in §5.5 can be described by the stabilizer formalism [151], and analytical calculations of entanglement entropies are possible. Let us recall a



Figure 5-16: A residual region in a planar tensor network.

useful formula for entropy calculations for stabilizer states. Consider an *n*-qutrit pure state  $|\psi\rangle$  specified by a set of *n* independent Pauli stabilizer generators  $g_j$  such that  $g_j|\psi\rangle = |\psi\rangle$  for j = 1, ..., n with  $[g_i, g_j] = 0$ . The stabilizer group  $\mathcal{S}^{(stab)}$  consists of all stabilizers  $\mathcal{S}^{(stab)} = \langle \{g_j\}^{\forall j} \rangle$ . Therefore we have  $g|\psi\rangle = |\psi\rangle$  for all  $g \in \mathcal{S}^{(stab)}$ . We are interested in entanglement entropy of  $|\psi\rangle$  for some subset *A* of qutrits. A useful formula to compute  $S_A$  is the following [176, 177]

$$S_A = |A| - \log_3 |\mathcal{S}_A^{(stab)}|, \tag{5.90}$$

where |A| is the number of qutrits in A and  $\mathcal{S}_A^{(stab)}$  is the restriction of  $\mathcal{S}$  onto A (i.e. a group of stabilizer operators which are supported exclusively on A). Note,  $\log_3 |\mathcal{S}_A^{(stab)}|$  can be understood as the number of independent stabilizers supported on A.

The stabilizer generators for the qutrit tensor are given by

$$Z \otimes Z \otimes Z \otimes I, \qquad Z \otimes Z^{\dagger} \otimes I \otimes Z,$$

$$X \otimes X \otimes X \otimes I, \qquad X \otimes X^{\dagger} \otimes I \otimes X,$$

$$(5.91)$$

where  $X^{\dagger} = X^2$  and  $Z^{\dagger} = Z^2$ . Graphically, stabilizer generators are given by

where  $Z|j\rangle = \omega^j |j\rangle$  and  $X|j\rangle = |j+1\rangle$  with  $\omega = e^{\frac{i2\pi}{3}}$ . Observe that stabilizer generators commute with each other. Also observe that there is no two-body stabilizer generator. This implies, from Eq (5.90), that entanglement entropies for any subsets of two qutrits are two, and thus this stabilizer state is a four-leg perfect state.

We need to find the number of stabilizer generators which can be exclusively supported on *PR*. Let us first consider a contraction of two perfect tensors (i.e. r = 2 and t = 1). There are stabilizer generators supported only on upper and lower tensor legs



So,  $S_{PR} = |P| + |R| - 2 = 2$ . Next, let us consider the case where t = 1 and r > 2. Since X-type and Z-type stabilizers are separable, one can treat them separately. We want to find all the stabilizer operators that are supported on PR. Here we consider an input Pauli operator X(f) supported on P where f is a degree r - 1 polynomial over  $\mathbb{F}_3$ . That is, for a polynomial

$$f = c_0 + c_1 x + c_2 x^2 + \ldots + c_{r-1} x^{r-1}, \qquad c_j \in \mathbb{F}_3,$$
(5.94)

we define the Pauli-X operator as

$$X(f) = X^{c_0} \otimes X^{c_1} \otimes \ldots \otimes X^{c_{r-1}}, \tag{5.95}$$

where  $X_j$  acts on the *j*-th leg on *P* for j = 0, ..., r - 1. Given an arbitrary degree r - 1 polynomial  $f_0$  over  $\mathbb{F}_3$ , one can write  $f_0$  as follows

$$f_0 = (2+x)g_0 + h_0, (5.96)$$

where  $g_0$  is some degree r - 2 polynomial while  $h_0$  is some degree 0 polynomial (in other words, a constant). Note that X(2 + x) is the Pauli X operator on P in (5.93), whose "output" on R is given by X(1+x). Let us then look for a stabilizer operator whose action on the upper leg is given by  $X(f_0)$ . When t = 1, the output Pauli operator is supported exclusively on R if and only if  $h_0 = 0$ . Namely, the output operator can be written as

$$X(f_1) f_1 = (1+x)g_0. (5.97)$$

Similar analysis holds for Z-type stabilizers. Therefore, there are in total 2r-2 independent stabilizer generators supported on PR. Thus,  $S_{PR} = r + r - (2r - 2) = 2$ . Finally, let us consider the cases where t > 1. For this purpose, we think of decomposing  $f_j$  recursively as follows

$$f_j = (2+x)g_j + h_j, \qquad f_{j+1} = (1+x)g_j.$$
 (5.98)

The output has supports exclusively on R if and only if  $h_j = 0$  for j = 0, ..., t - 1. This implies that there are in total 2(r-t) stabilizer generators supported on AC for  $t \leq r$ , and there is no such stabilizer generator for t > r. Thus, one has  $I_3 = -2\min(r, t)$ .

In fact, the aforementioned result applies to a larger class of perfect tensors. Notice that stabilizer generators of the qutrit tensor can be written as tensor products of Pauli Z or X operators only. Such a stabilizer state is often referred to as a CSS (Calderbank-Shor-Steane) state, and a number of interesting quantum error-correcting codes belongs to this class. Let us assume that four-leg perfect tensors are based on CSS stabilizer states. Then, one is able to prove that  $I_3$  is always given by  $I_3 = -2\min(r,t)$  as long as the bond dimension v is a prime number. We reached this conclusion by finding all the possible CSStype four-leg perfect tensors with prime bond dimensions. However, this conclusion does not necessarily hold when bond dimension is not a prime number. We skip the derivation as it is similar to the one for the qutrit perfect tensor.

# Appendix A

# ? Agere sequitur complexionem

(Epilogue)

The following manuscript was submitted to Nature.

## Inaction in computation

It was all a mistake. I just wanted to know what she would say if I asked her out. Without having to ask her.

I'm really shy, and I don't have a lot of experience with this sort of thing. In fact, I've never done it before. So I have no way of *predicting* what she (or anyone, for that matter) would say.

See, my plan was to try to know what she would say, but before *she* says it. This isn't totally crazy; we (humans) do it all the time. Just usually in our head. For example:

Me Mom, can I borrow \$100 to rent a pocket-fMRI?

Mom (simulated, in my head) Didn't I just give you money for that? Money doesn't grow on trees, you know. And what's a kid your age keep needing to use an MRI-anything for? They're not toys!

So I realized that I needed a different approach:

Me Mom, I need \$100 for Mrs. Keuling's class trip.

- Mom Hmm.... Didn't I just give you money for that? And that's quite expensive! Money doesn't grow on trees, you know...can I see a permission slip explaining the project?
- **Me** Absolutely, I've got it right here. It's for a trip to the neuro-labs to learn about equipment safety and the fact that such things are dangerous and not to be used as toys.

Mom Well, okay ...

See, I simulated the second conversation in my head too and knew that I needed to make a fake permission slip. But Moms are easy, everyone knows what they're thinking! Girls are difficult, nobody knows what they're thinking ...

Hence, the idea ...

...make a better simulation! One that could never be wrong. One that would tell me exactly what she was thinking. Well, she could tell me, because I could just ask her!

But if it could never be wrong, it would have to be perfect. Hence, the Knightian-freedomviolating pocket-fMRI. And the expensive cloud-computing bill:

Me Mom, I need to make a large purchase of cycles on the TensorNetwork cloud ....

Eh, that part's boring; I'm sure you can easily simulate the rest of the conversation.

## \*\*\*\*

Anyway, as I was saying, I needed the pocket-fMRI and all the computing cycles my Mom could buy. Because even if you scan someone, what are you going to do with the data? Just stare at it? I can't *read* minds!

I needed to *run* it. And to run it, I needed her to be unaware that it was a simulation. Why would a Boltzmann brain want to date me? I've got a feeling it's not that into a long-term relationship ...

So I also needed a Replicator ("Mom, can I please had an advance on my allowance?") and an AnyonTagger ("Mom, I want to buy Dad a birthday present ..."). So now I also had a copy of the school, the class, the teacher, and the weather for the next day. (I was going to ask her out at recess tomorrow.)

But what if she tried to use her phone? Or the Internet? So I also got a Quantum-Proposal-Tester<sup>TM</sup> ("Hi, I'd like to apply for this job. You know, teen stuff. Yeah, having a girlfriend's not cheap ..."). So now—in addition to her—I had a copy of the Internet, the school, the class, the teacher, the weather for the next day, and the weather for arbitrarily into the future. You see, I copied everything on Earth.

I also made sure to perfectly copy myself. Because trying to control my avatar from the outside the simulation would be far too clunky. What if she realized something was off because I was different? What if she realized she was in a simulation and I was running it, wouldn't that change her answer? The old adage is true: the only way to *control* a sentient mind is to actually *be* that particular sentient mind.

### \*\*\*\*

It suddenly occurred to me that I had no way of knowing where I was anymore. Oh, I know where I am; I'm in my kitchen about to ask my Mom for a loan. That's not what I mean. What I mean is, am I made up of neurons or am I made up of neurons being (perfectly) simulated on the cloud?

(I try to introspect on this, but the only thing that I get back is people during my life constantly telling me to "get my head out of the *clouds*." Is that a planted memory? Is that even a real expression people say? Or is someone trying to tell me something about the metaphysics of my universe?)

I always used to wonder whether the universe is *real* or a simulation. Now I realize that's a silly question. I was probably always living in the simulation of someone else who wanted to predict the answer to a question with certainty. Like what if Mom wanted to know what Dad would say if she asked him to marry her ...?

(Somewhere in the back of my mind I wonder if my Mom's credit card that I used to pay for the TensorNetwork connection is going to run out soon. But worrying about that sort of existential risk is a lot like worrying about getting hit by a cosmic domain wall—there's no warning and then there's nothing, so why worry?)

No, that's not the mistake. That's just the way things are now. Oh wait, I have an idea:

Me Mom, I need \$100 for Mrs. Keuling's class trip.

Mom Hmm. Didn't I just give you money for that? ...

You do look, muon, in a moved sort, As if you've feared this day: be cheerful, sir. Our studies now are ended. These our physicists, As I foretold you, were all spurions and Are melted into air, into thin air: And, like the baseless fabric of this fission, The cloud-capp'd top-quark, the gorgeous valances, The soliton temples, the great SU(2) itself, Ye all which it inherit, shall dissolve And, like this insubstantial black hole faded, Leave not a charge behind. We are such quasi-particles As dreams are made on, and our little life Is rounded with a sleep. Sir, I am flux'd; Bare with my weak-force; my brane is troubled: Be not disturb'd with my infinity: If you be pleased, retire into my cell And there transpose: a turn or two I'll walk, To still my beating mind.

# Bibliography

- [1] D. A. Roberts, D. Stanford, and L. Susskind, "Localized shocks," JHEP 03 (2015) 051, arXiv:1409.8180 [hep-th].
- [2] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, "Complexity Equals Action," *Phys. Rev. Lett.* **116** no. 19, (2016) 191301, arXiv:1509.07876 [hep-th].
- [3] D. A. Roberts and D. Stanford, "Two-dimensional conformal field theory and the butterfly effect," *Phys. Rev. Lett.* 115 no. 13, (2015) 131603, arXiv:1412.5123
   [hep-th].
- [4] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels," *JHEP* 02 (2016) 004, arXiv:1511.04021 [hep-th].
- [5] K. Schwarzschild, "On the gravitational field of a mass point according to Einstein's theory," Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) 1916 (1916) 189–196.
- [6] D. Finkelstein, "Past-Future Asymmetry of the Gravitational Field of a Point Particle," Phys. Rev. 110 (1958) 965–967.
- [7] M. D. Kruskal, "Maximal extension of Schwarzschild metric," *Phys. Rev.* 119 (1960) 1743–1745.
- [8] Virgo, LIGO Scientific Collaboration, B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** no. 6, (2016) 061102, arXiv:1602.03837 [gr-qc].
- [9] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.* 27 (1948) 379–423. [Bell Syst. Tech. J.27,623(1948)].
- [10] J. W. Gibbs, "On the equilibrium of heterogeneous substances," American Journal of Science no. 96, (1878) 441–458.
- [11] J. von Neumann, Mathematische Grundlagen der Quantenmechanik. Verlag Berlin Heidelberg, 1932.
- [12] A. Einstein, "Does the inertia of a body depend upon its energy-content?," Annalen der Physik 18 (1905) 639–641.
- [13] J. D. Bekenstein, "Black holes and the second law," Lett. Nuovo Cim. 4 (1972) 737–740.

- [14] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev.* D7 (1973) 2333–2346.
- [15] S. W. Hawking, "Gravitational radiation from colliding black holes," Phys. Rev. Lett. 26 (1971) 1344–1346.
- [16] S. W. Hawking, "Black hole explosions," *Nature* **248** (1974) 30–31.
- [17] S. W. Hawking, "Particle Creation by Black Holes," Commun. Math. Phys. 43 (1975) 199–220. [,167(1975)].
- [18] J. D. Bekenstein, "Generalized second law of thermodynamics in black hole physics," *Phys. Rev.* D9 (1974) 3292–3300.
- [19] S. W. Hawking, "Black Holes and Thermodynamics," Phys. Rev. D13 (1976) 191–197.
- [20] A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy," Phys. Lett. B379 (1996) 99–104, arXiv:hep-th/9601029 [hep-th].
- [21] G. 't Hooft, "Dimensional reduction in quantum gravity," in *Salamfest* 1993:0284-296, pp. 0284-296. 1993. arXiv:gr-qc/9310026 [gr-qc].
- [22] L. Susskind, "The World as a hologram," J. Math. Phys. 36 (1995) 6377-6396, arXiv:hep-th/9409089 [hep-th].
- [23] J. D. Bekenstein, "A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems," Phys. Rev. D23 (1981) 287.
- [24] D. Marolf, D. Minic, and S. F. Ross, "Notes on space-time thermodynamics and the observer dependence of entropy," *Phys. Rev.* D69 (2004) 064006, arXiv:hep-th/0310022 [hep-th].
- [25] H. Casini, "Relative entropy and the Bekenstein bound," Class. Quant. Grav. 25 (2008) 205021, arXiv:0804.2182 [hep-th].
- [26] M. Planck, "Uber irreversible Strahlungsvorgänge," Annalen der Physik 306 no. 1, (1900) 69–122.
- [27] M. Planck, "On the Law of Distribution of Energy in the Normal Spectrum," Annalen der Physik 4 (1901) 553.
- [28] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Int. J. Theor. Phys. 38 (1999) 1113-1133, arXiv:hep-th/9711200 [hep-th]. [Adv. Theor. Math. Phys.2,231(1998)].
- [29] E. Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. 2 (1998) 253-291, arXiv:hep-th/9802150 [hep-th].
- [30] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys. Lett.* B428 (1998) 105-114, arXiv:hep-th/9802109 [hep-th].
- [31] S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse," Phys. Rev. D14 (1976) 2460–2473.

- [32] S. Arnold, "Humpty Dumpty," Juvenile Amusements 68 (1797).
- [33] L. Susskind, L. Thorlacius, and J. Uglum, "The Stretched horizon and black hole complementarity," *Phys. Rev.* D48 (1993) 3743-3761, arXiv:hep-th/9306069 [hep-th].
- [34] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black Holes: Complementarity or Firewalls?," JHEP 02 (2013) 062, arXiv:1207.3123 [hep-th].
- [35] L. Susskind, "The Typical-State Paradox: Diagnosing Horizons with Complexity," Fortsch. Phys. 64 (2016) 84-91, arXiv:1507.02287 [hep-th].
- [36] D. Harlow and P. Hayden, "Quantum Computation vs. Firewalls," JHEP 06 (2013) 085, arXiv:1301.4504 [hep-th].
- [37] D. N. Page, "Average entropy of a subsystem," Phys. Rev. Lett. 71 (1993) 1291–1294, arXiv:gr-qc/9305007 [gr-qc].
- [38] P. Hayden and J. Preskill, "Black holes as mirrors: Quantum information in random subsystems," JHEP 09 (2007) 120, arXiv:0708.4025 [hep-th].
- [39] Y. Sekino and L. Susskind, "Fast Scramblers," JHEP 10 (2008) 065, arXiv:0808.2096 [hep-th].
- [40] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, "Towards the Fast Scrambling Conjecture," JHEP 04 (2013) 022, arXiv:1111.6580 [hep-th].
- [41] W. Brown and O. Fawzi, "Scrambling speed of random quantum circuits," arXiv:1210.6644 [quant-ph].
- [42] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, "Entanglement Scrambling in 2d Conformal Field Theory," *JHEP* 09 (2015) 110, arXiv:1506.03772 [hep-th].
- [43] M. Mezei and D. Stanford. To appear.
- [44] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," JHEP 03 (2014) 067, arXiv:1306.0622 [hep-th].
- [45] A. Kitaev, "Hidden correlations in the hawking radiation and thermal noise," 2014. Talk given at the Fundamental Physics Prize Symposium, Nov. 10, 2014.
- [46] A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem," J. of Math 58 (1936) 345–363.
- [47] S. Lloyd, "Ultimate physical limits to computation," Nature 406 no. 6799, (2000) 1047-1054, arXiv:quant-ph/9908043 [quant-ph].
- [48] N. Margolus and L. B. Levitin, "The Maximum speed of dynamical evolution," *Physica* D120 (1998) 188–195, arXiv:quant-ph/9710043 [quant-ph].
- [49] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, "Complexity, Action, and Black Holes," *Phys. Rev.* D93 no. 8, (2016) 086006, arXiv:1512.04993 [hep-th].

- [50] S. Aaronson, "Why Philosophers Should Care About Computational Complexity," arXiv:1108.1791 [cs.CC].
- [51] D. A. Roberts and B. Swingle, "Lieb-Robinson and the butterfly effect," arXiv:1603.09298 [hep-th].
- [52] A. Larkin and Y. Ovchinnikov, "Quasiclassical method in the theory of superconductivity," *JETP* 28,6 (1969) 1200–1205.
- [53] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, "An Apologia for Firewalls," JHEP 09 (2013) 018, arXiv:1304.6483 [hep-th].
- [54] S. H. Shenker and D. Stanford, "Multiple Shocks," JHEP 12 (2014) 046, arXiv:1312.3296 [hep-th].
- [55] S. Leichenauer, "Disrupting Entanglement of Black Holes," Phys. Rev. D90 no. 4, (2014) 046009, arXiv:1405.7365 [hep-th].
- [56] S. Jackson, L. McGough, and H. Verlinde, "Conformal Bootstrap, Universality and Gravitational Scattering," Nucl. Phys. B901 (2015) 382-429, arXiv:1412.5205 [hep-th].
- [57] S. H. Shenker and D. Stanford, "Stringy effects in scrambling," JHEP 05 (2015) 132, arXiv:1412.6087 [hep-th].
- [58] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," arXiv:1503.01409 [hep-th].
- [59] J. Polchinski, "Chaos in the black hole S-matrix," arXiv:1505.08108 [hep-th].
- [60] D. Marolf, "Comments on Microcausality, Chaos, and Gravitational Observables," Class. Quant. Grav. 32 no. 24, (2015) 245003, arXiv:1508.00939 [gr-qc].
- [61] D. Stanford, "Many-body chaos at weak coupling," arXiv:1512.07687 [hep-th].
- [62] G. Gur-Ari, M. Hanada, and S. H. Shenker, "Chaos in Classical D0-Brane Mechanics," JHEP 02 (2016) 091, arXiv:1512.00019 [hep-th].
- [63] A. L. Fitzpatrick and J. Kaplan, "A Quantum Correction To Chaos," arXiv:1601.06164 [hep-th].
- [64] J. Polchinski and V. Rosenhaus, "The Spectrum in the Sachdev-Ye-Kitaev Model," arXiv:1601.06768 [hep-th].
- [65] B. Michel, J. Polchinski, V. Rosenhaus, and S. J. Suh, "Four-point function in the IOP matrix model," arXiv:1602.06422 [hep-th].
- [66] Y. Gu and X.-L. Qi, "Fractional Statistics and the Butterfly Effect," arXiv:1602.06543 [hep-th].
- [67] P. Caputa, T. Numasawa, and A. Veliz-Osorio, "Scrambling without chaos in RCFT," arXiv:1602.06542 [hep-th].
- [68] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, "Measuring the scrambling of quantum information," arXiv:1602.06271 [quant-ph].

- [69] E. Perlmutter, "Bounding the Space of Holographic CFTs with Chaos," arXiv:1602.08272 [hep-th].
- [70] G. Turiaci and H. Verlinde, "On CFT and Quantum Chaos," arXiv:1603.03020 [hep-th].
- [71] M. Blake, "Universal Charge Diffusion and the Butterfly Effect," arXiv:1603.08510 [hep-th].
- [72] J. Maldacena and D. Stanford, "Comments on the Sachdev-Ye-Kitaev model," arXiv:1604.07818 [hep-th].
- [73] A. Kitaev. To appear.
- [74] E. H. Lieb and D. W. Robinson, "The finite group velocity of quantum spin systems," Commun. Math. Phys. 28 (1972) 251–257.
- [75] M. B. Hastings and T. Koma, "Spectral gap and exponential decay of correlations," Commun. Math. Phys. 265 (2006) 781-804, arXiv:math-ph/0507008 [math-ph].
- [76] M. B. Hastings, "Locality in quantum systems," arXiv:1008.5137 [math-ph].
- [77] A. Kitaev, "A simple model of quantum holography," 2015. Talks at KITP, April 7, 2015 and May 27, 2015.
- [78] S. Sachdev and J.-w. Ye, "Gapless spin fluid ground state in a random, quantum Heisenberg magnet," *Phys. Rev. Lett.* **70** (1993) 3339, arXiv:cond-mat/9212030 [cond-mat].
- [79] D. Stanford and L. Susskind, "Complexity and Shock Wave Geometries," *Phys. Rev.* D90 no. 12, (2014) 126007, arXiv:1406.2678 [hep-th].
- [80] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum mechanical description of physical reality be considered complete?," *Phys. Rev.* 47 (1935) 777–780.
- [81] A. Einstein and N. Rosen, "The Particle Problem in the General Theory of Relativity," Phys. Rev. 48 (1935) 73–77.
- [82] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," Fortsch. Phys. 61 (2013) 781-811, arXiv:1306.0533 [hep-th].
- [83] M. Van Raamsdonk, "Building up spacetime with quantum entanglement," Gen. Rel. Grav. 42 (2010) 2323-2329, arXiv:1005.3035 [hep-th]. [Int. J. Mod. Phys.D19,2429(2010)].
- [84] B. Swingle, "Entanglement Renormalization and Holography," Phys. Rev. D86 (2012) 065007, arXiv:0905.1317 [cond-mat.str-el].
- [85] T. Hartman and J. Maldacena, "Time Evolution of Entanglement Entropy from Black Hole Interiors," JHEP 05 (2013) 014, arXiv:1303.1080 [hep-th].
- [86] J. Maldacena, "Spacetime from entanglement," 2015. Talk at KITP, August 20, 2013.

- [87] L. Susskind, "Entanglement is not enough," Fortsch. Phys. 64 (2016) 49-71, arXiv:1411.0690 [hep-th].
- [88] J. Louko, D. Marolf, and S. F. Ross, "On geodesic propagators and black hole holography," Phys. Rev. D62 (2000) 044041, arXiv:hep-th/0002111 [hep-th].
- [89] P. Kraus, H. Ooguri, and S. Shenker, "Inside the horizon with AdS / CFT," Phys. Rev. D67 (2003) 124022, arXiv:hep-th/0212277 [hep-th].
- [90] L. Fidkowski, V. Hubeny, M. Kleban, and S. Shenker, "The Black hole singularity in AdS / CFT," JHEP 02 (2004) 014, arXiv:hep-th/0306170 [hep-th].
- [91] I. A. Morrison and M. M. Roberts, "Mutual information between thermo-field doubles and disconnected holographic boundaries," *JHEP* 07 (2013) 081, arXiv:1211.2887 [hep-th].
- [92] H. Liu and S. J. Suh, "Entanglement Tsunami: Universal Scaling in Holographic Thermalization," Phys. Rev. Lett. 112 (2014) 011601, arXiv:1305.7244 [hep-th].
- [93] H. Liu and S. J. Suh, "Entanglement growth during thermalization in holographic systems," Phys. Rev. D89 no. 6, (2014) 066012, arXiv:1311.1200 [hep-th].
- [94] B. Freivogel and L. Susskind, "A Framework for the landscape," Phys. Rev. D70 (2004) 126007, arXiv:hep-th/0408133 [hep-th].
- [95] I. Heemskerk, D. Marolf, J. Polchinski, and J. Sully, "Bulk and Transhorizon Measurements in AdS/CFT," JHEP 10 (2012) 165, arXiv:1201.3664 [hep-th].
- [96] T. Banks, M. R. Douglas, G. T. Horowitz, and E. J. Martinec, "AdS dynamics from conformal field theory," arXiv:hep-th/9808016 [hep-th].
- [97] V. Balasubramanian, P. Kraus, A. E. Lawrence, and S. P. Trivedi, "Holographic probes of anti-de Sitter space-times," *Phys. Rev.* D59 (1999) 104021, arXiv:hep-th/9808017 [hep-th].
- [98] I. Bena, "On the construction of local fields in the bulk of AdS(5) and other spaces," *Phys. Rev.* D62 (2000) 066007, arXiv:hep-th/9905186 [hep-th].
- [99] 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.* D73 (2006) 086003, arXiv:hep-th/0506118 [hep-th].
- [100] D. Kabat, G. Lifschytz, and D. A. Lowe, "Constructing local bulk observables in interacting AdS/CFT," Phys. Rev. D83 (2011) 106009, arXiv:1102.2910 [hep-th].
- [101] D. Marolf and J. Polchinski, "Gauge/Gravity Duality and the Black Hole Interior," *Phys. Rev. Lett.* **111** (2013) 171301, arXiv:1307.4706 [hep-th].
- [102] L. Susskind, "Computational Complexity and Black Hole Horizons," Fortsch. Phys. 64 (2016) 24-43, arXiv:1402.5674 [hep-th].

- [103] M. C. Banuls, J. I. Cirac, and M. B. Hastings, "Strong and weak thermalization of infinite nonintegrable quantum systems," *Phys. Rev. Lett.* **106** no. 5, (2011) 050405, arXiv:1007.3957 [quant-ph].
- [104] S. Sachdev, Quantum Phase Transitions. Cambridge University Press, 2001.
- [105] L. Susskind, "New Concepts for Old Black Holes," arXiv:1311.3335 [hep-th].
- [106] J. M. Maldacena, "Eternal black holes in anti-de Sitter," JHEP 04 (2003) 021, arXiv:hep-th/0106112 [hep-th].
- [107] D. Amati, M. Ciafaloni, and G. Veneziano, "Classical and Quantum Gravity Effects from Planckian Energy Superstring Collisions," Int. J. Mod. Phys. A3 (1988) 1615–1661.
- [108] K. Schoutens, H. L. Verlinde, and E. P. Verlinde, "Quantum black hole evaporation," *Phys. Rev.* D48 (1993) 2670-2685, arXiv:hep-th/9304128 [hep-th].
- [109] T. Dray and G. 't Hooft, "The Gravitational Shock Wave of a Massless Particle," Nucl. Phys. B253 (1985) 173–188.
- [110] K. Sfetsos, "On gravitational shock waves in curved space-times," Nucl. Phys. B436 (1995) 721-746, arXiv:hep-th/9408169 [hep-th].
- [111] G. 't Hooft, "The black hole interpretation of string theory," Nucl. Phys. B335 (1990) 138–154.
- [112] Y. Kiem, H. L. Verlinde, and E. P. Verlinde, "Black hole horizons and complementarity," *Phys. Rev.* D52 (1995) 7053-7065, arXiv:hep-th/9502074 [hep-th].
- [113] D. M. Hofman and J. Maldacena, "Conformal collider physics: Energy and charge correlations," JHEP 05 (2008) 012, arXiv:0803.1467 [hep-th].
- [114] X. O. Camanho, J. D. Edelstein, J. Maldacena, and A. Zhiboedov, "Causality Constraints on Corrections to the Graviton Three-Point Coupling," arXiv:1407.5597 [hep-th].
- [115] T. Dray and G. 't Hooft, "The Effect of Spherical Shells of Matter on the Schwarzschild Black Hole," Commun. Math. Phys. 99 (1985) 613–625.
- [116] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," Phys. Rev. Lett. 96 (2006) 181602, arXiv:hep-th/0603001 [hep-th].
- [117] S. Ryu and T. Takayanagi, "Aspects of Holographic Entanglement Entropy," JHEP 08 (2006) 045, arXiv:hep-th/0605073 [hep-th].
- [118] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A Covariant holographic entanglement entropy proposal," JHEP 07 (2007) 062, arXiv:0705.0016 [hep-th].
- [119] L. Susskind and Y. Zhao, "Switchbacks and the Bridge to Nowhere," arXiv:1408.2823 [hep-th].

- [120] M. R. Dowling and M. A. Nielsen, "The geometry of quantum computation," arXiv:quant-ph/0701004 [quant-ph].
- [121] G. T. Horowitz and N. Itzhaki, "Black holes, shock waves, and causality in the AdS / CFT correspondence," JHEP 02 (1999) 010, arXiv:hep-th/9901012 [hep-th].
- [122] B. Zwiebach, "Curvature Squared Terms and String Theories," Phys. Lett. B156 (1985) 315.
- [123] D. G. Boulware and S. Deser, "String Generated Gravity Models," Phys. Rev. Lett. 55 (1985) 2656.
- [124] R.-G. Cai, "Gauss-Bonnet black holes in AdS spaces," Phys. Rev. D65 (2002) 084014, arXiv:hep-th/0109133 [hep-th].
- [125] M. Brigante, H. Liu, R. C. Myers, S. Shenker, and S. Yaida, "Viscosity Bound Violation in Higher Derivative Gravity," *Phys. Rev.* D77 (2008) 126006, arXiv:0712.0805 [hep-th].
- [126] L. Susskind, "Addendum to Computational Complexity and Black Hole Horizons," arXiv:1403.5695 [hep-th].
- [127] G. Evenbly and G. Vidal, "Tensor Network States and Geometry," Journal of Statistical Physics 145 (2011) 891–918, arXiv:1106.1082 [quant-ph].
- [128] E. Poisson, A Relativist's Toolkit. Cambridge University Press, 2004.
- [129] T. J. Osborne, "Hamiltonian complexity," Reports on Progress in Physics 75 (2012) 022001, arXiv:1106.5875 [quant-ph].
- [130] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, "Quantum Hamiltonian Complexity," arXiv:1401.3916 [quant-ph].
- [131] S. Lloyd, "Power of Entanglement in Quantum Communication," *Phys. Rev. Lett.* 90 (2003) 167902.
- [132] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, "The String landscape, black holes and gravity as the weakest force," *JHEP* 06 (2007) 060, arXiv:hep-th/0601001 [hep-th].
- [133] F. Denef and S. A. Hartnoll, "Landscape of superconducting membranes," Phys. Rev. D79 (2009) 126008, arXiv:0901.1160 [hep-th].
- [134] N. Iqbal, H. Liu, and M. Mezei, "Semi-local quantum liquids," JHEP 04 (2012) 086, arXiv:1105.4621 [hep-th].
- [135] S. A. Hartnoll, "Horizons, holography and condensed matter," arXiv:1106.4324 [hep-th].
- [136] S. A. Hartnoll, C. P. Herzog, and G. T. Horowitz, "Holographic Superconductors," JHEP 12 (2008) 015, arXiv:0810.1563 [hep-th].
- [137] G. T. Horowitz and M. M. Roberts, "Zero Temperature Limit of Holographic Superconductors," JHEP 11 (2009) 015, arXiv:0908.3677 [hep-th].

- [138] T. Hartman, "Entanglement Entropy at Large Central Charge," arXiv:1303.6955 [hep-th].
- [139] A. L. Fitzpatrick, J. Kaplan, and M. T. Walters, "Universality of Long-Distance AdS Physics from the CFT Bootstrap," JHEP 08 (2014) 145, arXiv:1403.6829 [hep-th].
- [140] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, "Holographic Entanglement Entropy from 2d CFT: Heavy States and Local Quenches," JHEP 02 (2015) 171, arXiv:1410.1392 [hep-th].
- [141] L. Cornalba, M. S. Costa, J. Penedones, and R. Schiappa, "Eikonal Approximation in AdS/CFT: From Shock Waves to Four-Point Functions," JHEP 08 (2007) 019, arXiv:hep-th/0611122 [hep-th].
- [142] L. Cornalba, M. S. Costa, J. Penedones, and R. Schiappa, "Eikonal Approximation in AdS/CFT: Conformal Partial Waves and Finite N Four-Point Functions," *Nucl. Phys.* B767 (2007) 327–351, arXiv:hep-th/0611123 [hep-th].
- [143] L. Cornalba, M. S. Costa, and J. Penedones, "Eikonal approximation in AdS/CFT: Resumming the gravitational loop expansion," *JHEP* 09 (2007) 037, arXiv:0707.0120 [hep-th].
- [144] F. A. Dolan and H. Osborn, "Conformal four point functions and the operator product expansion," Nucl. Phys. B599 (2001) 459-496, arXiv:hep-th/0011040 [hep-th].
- [145] R. F. Streater and A. S. Wightman, PCT, spin and statistics, and all that. 1989.
- [146] P. Caputa, J. Simon, A. Stikonas, and T. Takayanagi, "Quantum Entanglement of Localized Excited States at Finite Temperature," *JHEP* 01 (2015) 102, arXiv:1410.2287 [hep-th].
- [147] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov, "Infinite Conformal Symmetry in Two-Dimensional Quantum Field Theory," *Nucl. Phys.* B241 (1984) 333–380.
- [148] M. P. Mattis, "Correlations in Two-dimensional Critical Theories," Nucl. Phys. B285 (1987) 671.
- [149] P. H. Ginsparg, "Applied Conformal Field Theory," in Les Houches Summer School in Theoretical Physics: Fields, Strings, Critical Phenomena Les Houches, France, June 28-August 5, 1988. 1988. arXiv:hep-th/9108028 [hep-th].
- [150] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence," *JHEP* 06 (2015) 149, arXiv:1503.06237 [hep-th].
- [151] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information. Cambridge University Press, 2010.
- [152] A. Kitaev and J. Preskill, "Topological entanglement entropy," *Phys. Rev. Lett.* 96 (2006) 110404, arXiv:hep-th/0510092 [hep-th].

- [153] M. Levin and X.-G. Wen, "Detecting Topological Order in a Ground State Wave Function," Phys. Rev. Lett. 96 (2006) 110405.
- [154] P. Hayden, M. Headrick, and A. Maloney, "Holographic Mutual Information is Monogamous," Phys. Rev. D87 no. 4, (2013) 046003, arXiv:1107.2940 [hep-th].
- [155] H. Gharibyan and R. F. Penna, "Are entangled particles connected by wormholes? Evidence for the ER=EPR conjecture from entropy inequalities," *Phys. Rev.* D89 no. 6, (2014) 066001, arXiv:1308.0289 [hep-th].
- [156] M. Rangamani and M. Rota, "Entanglement structures in qubit systems," J. Phys. A48 no. 38, (2015) 385301, arXiv:1505.03696 [hep-th].
- [157] S. Sachdev, "Bekenstein-Hawking Entropy and Strange Metals," Phys. Rev. X5 no. 4, (2015) 041025, arXiv:1506.05111 [hep-th].
- [158] S. Leichenauer and M. Moosa, "Entanglement Tsunami in (1+1)-Dimensions," Phys. Rev. D92 (2015) 126004, arXiv:1505.04225 [hep-th].
- [159] H. Casini, H. Liu, and M. Mezei, "Spread of entanglement and causality," arXiv:1509.05044 [hep-th].
- [160] F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki, "Local random quantum circuits are approximate polynomial-designs," arXiv:1208.0692 [quant-ph].
- [161] G. Vidal, "Entanglement Renormalization," Phys. Rev. Lett. 99 no. 22, (2007) 220405, arXiv:cond-mat/0512165 [cond-mat].
- [162] Z. Yang, P. Hayden, and X.-L. Qi, "Bidirectional holographic codes and sub-AdS locality," JHEP 01 (2016) 175, arXiv:1510.03784 [hep-th].
- [163] A. Almheiri, X. Dong, and D. Harlow, "Bulk Locality and Quantum Error Correction in AdS/CFT," JHEP 04 (2015) 163, arXiv:1411.7041 [hep-th].
- [164] H. Kim and D. A. Huse, "Ballistic spreading of entanglement in a diffusive nonintegrable system," *Phys. Rev. Lett.* **111** (2013) 127205, arXiv:1306.4306 [quant-ph].
- [165] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, "Holographic duality from random tensor networks," arXiv:1601.01694 [hep-th].
- [166] M. B. Hastings, "Random MERA States and the Tightness of the Brandao-Horodecki Entropy Bound," arXiv:1505.06468 [quant-ph].
- [167] D. Marolf, H. Maxfield, A. Peach, and S. F. Ross, "Hot multiboundary wormholes from bipartite entanglement," *Class. Quant. Grav.* **32** no. 21, (2015) 215006, arXiv:1506.04128 [hep-th].
- [168] K. Dickman, On the frequency of numbers containing prime factors of a certain relative magnitude. Almqvist & Wiksell, 1930.
- [169] M. Fannes, "A continuity property of the entropy density for spin lattice systems," Communications in Mathematical Physics 31 no. 4, (1973) 291.

- [170] K. M. R. Audenaert, "A sharp Fannes-type inequality for the von Neumann entropy," arXiv:quant-ph/0610146 [quant-ph].
- [171] P. Calabrese and J. L. Cardy, "Evolution of entanglement entropy in one-dimensional systems," J. Stat. Mech. 0504 (2005) P04010, arXiv:cond-mat/0503393 [cond-mat].
- [172] P. Calabrese and J. Cardy, "Quantum Quenches in Extended Systems," J. Stat. Mech. 0706 (2007) P06008, arXiv:0704.1880 [cond-mat.stat-mech].
- [173] P. Calabrese and J. Cardy, "Entanglement entropy and conformal field theory," J. Phys. A42 (2009) 504005, arXiv:0905.4013 [cond-mat.stat-mech].
- [174] V. Balasubramanian, A. Bernamonti, N. Copland, B. Craps, and F. Galli, "Thermalization of mutual and tripartite information in strongly coupled two dimensional conformal field theories," *Phys. Rev.* D84 (2011) 105017, arXiv:1110.0488 [hep-th].
- [175] A. Allais and E. Tonni, "Holographic evolution of the mutual information," JHEP 01 (2012) 102, arXiv:1110.1607 [hep-th].
- [176] D. Fattal, T. S. Cubitt, Y. Yamamoto, S. Bravyi, and I. L. Chuang, "Entanglement in the stabilizer formalism," arXiv:quant-ph/0406168 [quant-ph].
- [177] B. Yoshida and I. L. Chuang, "Framework for classifying logical operators in stabilizer codes," *Phys. Rev.* A81 (2010) 052302, arXiv:1002.0085 [quant-ph].