

Quantum Indefinite Spacetime

By
Ding Jia (贾丁)

A thesis
presented to the Independent Studies Program
of the University of Waterloo
in fulfilment of the
thesis requirements for the degree
Bachelor of Independent Studies (BIS)



Waterloo, Canada

2017

I hereby declare that I am the sole author of this research paper.

I authorize the University of Waterloo to lend this research paper to other institutions or individuals for the purpose of scholarly research.

(signature)

(date)

I further authorize the University of Waterloo to reproduce this research paper by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

(signature)

(date)

Abstract

We combine the principle of superposition from quantum theory and the principle of dynamical causal structure from general relativity to attack fundamental questions in quantum gravity. We generalize the concept of entanglement to parties whose causal relation is quantum indefinite. The generalized notion of entanglement gives meaning to timelike and more generally spacetime entanglement in quantum theory both with and without indefinite causal structure. Using this generalization, we identify quantum gravitational fluctuations of causal structure as a possible mechanism that regularizes the otherwise divergent entanglement. We give the name “quantum indefinite spacetime” to the model of spacetime incorporating quantum gravitational causal fluctuations.

Quantum indefinite spacetime sheds new light on black hole information problem as we argue that quantum gravitational causal fluctuations allow positive information communication capacity to the outside of the black holes. The new generalized notion of entanglement offers additional support from the black hole thermodynamics perspective.

Towards the end of the thesis we make a preliminary proposal that the quantum fluctuating entanglement regularization may explain the apparent accelerated expansion of the universe without introducing dark energy or cosmological constant.

All these results and proposals are obtained on the basis of jointly applying quantum and general relativistic principles, but without making tentative postulates about the microscopic degrees of freedom of quantum spacetime. We hope to convey the message that this more conservative approach can offer firm answers to several questions in quantum gravity. Moreover, other approaches to quantum gravity should incorporate features of quantum indefinite spacetime if they assume the same principles.

Acknowledgements

I would like to thank my dear supervisors Lucien Hardy and Achim Kempf for their continuous feedback, insightful questions and valuable suggestions that proved to be crucial to the development of the thesis, and for the conducive research environment they created for me during the of completion of this thesis.

I would like to thank the colleagues in my group, Mikhail Panine, Maria Papageorgiou, Jason Pye, Nayeli Azucena Rodriguez-Briones, and Guillaume Verdon-Akzam for interesting discussions. Special thanks goes to Jason Pye, who clarified many questions throughout the development and completion of the thesis, and read through the draft to offer lots of good suggestions. I would like to thank Eduardo Martin-Martinez and the participants of his group meeting for learning things together in entertaining ways. I would like to thank Ognyan Oreshkov, Alex Smith, Daniel Grimmer, Sam Leutheusser, TC Fraser, Yidong Chen, Nitica Sakharwade, Fabio Scardigli and Robert Oeckl with whom I had useful discussions in developing the ideas of the thesis.

I would like to thank people who helped me through the Independent Studies Program at University of Waterloo, Susan Gow, Linda Carson, Bill Abbott, Ted McGee, Anne Dagg, Steve Furino, Sarah Tolmie, Mavis Fenn, Barbara A. Moffatt, Gordon Stublely and Devon Hutchinson. Independent studies is a wonderful program with wonderful people.

I would like to thank my parents Hui Jia (贾辉) and Li Jin (金丽), and my partner Yicen Lu (卢亦岑) for their generous support.

Dedication

To my grandparents Lian'an Jia and Huarong Yu.

献给我的爷爷奶奶贾连安和于化荣

Table of Contents

Notation	viii
1 Introduction	1
1.1 Indefinite causal structure as a consequence of principles . . .	2
1.2 The overall picture of quantum indefinite spacetime	3
1.3 Outline of the thesis	4
2 Review of the process framework	5
2.1 Basic assumptions	5
2.2 Mathematical formulation	6
2.3 Normalization constraints	12
2.4 Purification postulate	14
3 Models of indefinite causal structure	15
3.1 Superposition of $A \nearrow B$ and $A - B$	16
3.1.1 Process with ancilla	16
3.1.2 Process with partial swap	17
3.1.3 Communication capacity	18
3.2 Superposition of $A \nearrow B$, $A \nwarrow B$ and $A - B$	19
3.3 Superposition of $A \nearrow B$ and $A \nwarrow B$	19
4 Entanglement measures	20
4.1 Process entanglement measures	21
4.2 Coherent information for processes	23
4.3 Coherent-information-based measures	25
5 Entanglement in spacetime	26
5.1 Quantum channel	27
5.2 Quantum network	29
6 Entanglement regularization	30
6.1 Entanglement regularization for $A \nearrow B$	31
6.2 Entanglement regularization for three causal relations in superposition	34
7 Black holes	36
7.1 Do black holes exist?	37
7.2 Entropy of black holes	39

7.3	Information transmission through evaporation	42
8	Cosmology	47
8.1	Postulates	47
8.2	Modified travelling distance	49
8.3	Questions	51
9	Conclusions	52
9.1	Summary	52
9.2	Outlook	53
9.2.1	Incorporating indefinite causal structure to quantum field theory	53
9.2.2	Developing a theory of quantum gravity	55
9.2.3	Generalizing communication theory	56
	APPENDICES	58
A	Standard theorems from quantum information theory	58
A.1	Quantum relative entropy	58
A.2	Entropic inequalities	59
A.3	State and channel coherent information	60
A.4	Continuity of coherent information	64
B	Results about process coherent information	64

Notations

$A_{ab\dots}^{xy\dots}$ ($A^{ab\dots xy\dots}$)	A process A with input systems a, b, \dots and output systems x, y, \dots (the form in the bracket is also used when no ambiguity arises).
$A_{ab\dots}^{xy\dots}$ ($A^{ab\dots xy\dots}$)	The process operator for the process A with input systems a, b, \dots and output systems x, y, \dots (the form in the bracket is also used when no ambiguity arises).
$A_{ab}^{xy} B_{xc}^{az}$, ($A_{ab}^{xy} B_{xc}^{az}$)	Composition of processes (process operators) A and B by joining inputs and outputs of the same letters, resulting in a new process (process operators) with input systems b, c and output systems y, z after composing A with B .
A^a of A^{ab} , (A^a of A^{ab})	The reduced process (process operator) when system b is discarded.
ω	The maximally mixed state.
ρ_+	The canonical maximally entangled state.
$S(A_a^x)$ or $S(A_a^x)$	The von Neumann entropy of the process operator of A_a^x .
S^x	Short hand for $S(A^x)$ or $S(A^x)$, when it is clear from context which A is referred to.
$S^{a b}(A^{ab})$ or $S^{a b}(A^{ab})$	The conditional entropy (conditioning on b) of the operator of A^{ab} .
$I_R^x(A^{xy})$ or $I_R^x(A^{xy})$	The coherent information of the process A with target system x and supplemental resource R .
$I^{a:b}(A^{ab})$ or $I^{a:b}(A^{ab})$	The mutual information of the operator A^{ab} .
$A \nearrow B$	A causally precedes B .
$A \nwarrow B$	A causally succeeds B .
$A - B$	A causally disconnected with B .
$A \overline{\nearrow} B$	Superposition of $A \nearrow B$ and $A - B$.
$A \overline{\nwarrow} B$	Superposition of $A \nwarrow B$ and $A - B$.
$A \overline{\nearrow \nwarrow} B$	Superposition of $A \nearrow B$ and $A \nwarrow B$.
$A \overline{\nearrow \nwarrow} B$	Superposition of $A \nearrow B$, $A \nwarrow B$ and $A - B$.

1 Introduction

In classical general relativity (GR), spacetime is completely specified by the causal structure plus the conformal factor [1, 2, 3]. This can be taken as a starting point for a theory of quantum spacetime. Instead of quantizing the metric as a field, we can apply quantum principles to the dynamical causal structure and the conformal factor. In particular, the prospect of combining quantum probabilistic features with dynamical causal structure led to the so-called indefinite causal structure approach to quantum gravity initiated in [4, 5].

In this thesis we follow the indefinite causal structure approach and adopt the more recent “process framework” [6] to obtain the following new results.

- We propose an axiom for entanglement measures for parties whose causal relation is indefinite (Section 4).
- We use the new axiom to give meaning to entanglement of two parties in arbitrary causal relation, in a context with or without indefinite causal structure (Section 5).
- We identify quantum gravitational fluctuations of causal structure as a possible concrete mechanism that regularizes entanglement in the ultraviolet (Section 6).

Studies along these lines lead us to make the following speculative and preliminary proposals.

- Quantum black holes in the strict sense do not exist. Quantum gravitational fluctuations allow information to transmit out of black holes (Section 7).
- Observations used as evidence for dark energy may be explained by quantum gravitational fluctuations of the area scale of spacetime without introducing extra energy density (Section 8).

A message we hope to convey about the subject of quantum gravity is that several questions can be answered without making tentative postulates about the microscopic degrees of freedom of spacetime/gravity. These questions can be addressed directly by extracting consequences of combining fundamental principles of quantum theory and of general relativity. In

this thesis, in particular, we obtain the above results by combining the quantum principle of superposition and the general relativistic principle of dynamical causal structure.

The key to these results is to incorporate what we call “causal fluctuations” into spacetime. Causal fluctuations are quantum fluctuations of causal structure expected to take place when as gravity quantum fluctuates at the Planck scale. We assign the name **quantum indefinite spacetime** to the spacetime model incorporating such causal fluctuations. Apart from spontaneous causal fluctuations at the Planck scale, quantum indefinite spacetime may also incorporate indefinite causal structure arising from superposition of matter. We take “quantum indefinite spacetime” as the title and theme of this thesis because of the central role it plays in obtaining our results.

In the rest of this section we give a conceptual-level overview of quantum indefinite spacetime to familiarize the readers with its motivations and implications. We then outline the rest of the thesis where we get into technical details.

1.1 Indefinite causal structure as a consequence of principles

Two major expectations on a theory of quantum gravity are to account for the microscopic degree of freedom of spacetime, and make quantum and general relativistic principles compatible. In our view, the second one is more fundamental. First of all it is possible that spacetime does not have any underlying microscopic degree of freedom to be discovered, and the search for it may be doomed to be fruitless. Moreover, even if an underlying structures exists, without much empirical guidance, it is not easy to limit the possibilities of the underlying structure to make firm statements. Although more effort has been made in quantum gravity phenomenology in recent years, there is still a lack of decisive data. Any conviction to a single possibility runs a high risk of having the wrong starting point. It would be better to base a theory of quantum gravity on general grounds not sensitive to any tentative postulate about the microscopic degrees of freedom.

A natural approach is to start with combining principles of quantum theory and general relativity. Such an effort may already offer firm answers to some fundamental questions without the need to commit to any particular postulate on the microscopic degrees of freedom. Admittedly, if

quantum spacetime has yet unknown underlying structure, there may be questions one can only address in a deeper theory. However, this deeper theory must also tell us how to combine quantum and general relativistic principles, so the study at the principle level can be regarded as forming the basis of the deeper theory.

In this thesis, we approach quantum gravity combining the principle of superposition from quantum theory and the principle of dynamical causal structure from general relativity. In quantum theory, the superposition principle applies to dynamical degrees of freedom. In general relativity, causal structures of spacetime are dynamical degrees of freedom. If these principles still hold in quantum gravity, spacetime causal structure can be in superposition. This conclusion is reached without any new postulates, and its negation would imply changes of quantum theory or general relativity at the fundamental level. This thesis assumes that the conclusion holds and study its consequences.¹

1.2 The overall picture of quantum indefinite spacetime

A main feature of quantum indefinite spacetime is that it does not presume a differentiable manifold structure. The description of spacetime as a differentiable manifold with a metric is standard in classical GR, but it is unclear how to apply quantum superposition principle to spacetime in this description. We take the alternative perspective to introduce quantumness in the causal structure of spacetime with the quantum process framework. The process framework description is more akin to quantum circuit networks that specifies how different systems compose with each other but does not presume any differentiable structure. Our outlook is that the classical spacetime description in terms of differentiable manifold emerges from the quantum process description of spacetime in the classical limit of no superposition of causal structure.

Rather than logically starting with a description of spacetime itself and build further structures and concepts on this spacetime, the framework starts with local parties where agents (or nature itself as a special case of

¹It is possible that the foundations of quantum theory or general relativity needs modification in quantum gravity. We approach this possibility conservatively by first extracting consequences of combining the original principles. If problems arise, they give clues on how to modify the foundations.

agents) can perform operations. The processes that carry information about quantum spacetime are defined as sets of probabilities correlating operations performed in the local parties. In other words, the description of quantum spacetime arises from the description of correlations of local operations. In this sense, this framework of quantum spacetime has an operational flavor. This is a welcoming feature since usually operational approaches in physics are subject to fewer risky assumptions. Additionally, it is relatively easy to discuss information processing in the present framework of quantum spacetime because of the close interconnections between operational approaches to quantum theory and information theory. In particular, this facilitates studies of black hole information problems.

As quantum indefinite spacetime is described in terms of processes correlating local parties, the theory does not commit to either continuous or discrete spacetime picture. In either case, the theory can be set up as long as the notion of local party is meaningful. A point worth emphasizing is that the entanglement regularization mechanism we identify in Section 6 exists independently of whether spacetime is continuous or discrete.² We speculate that the same mechanism regularizes quantum field theory 9.2.1. If this turns out to be true, it would to some extent suppress the motivation for debating whether spacetime is digital or analog.

For this thesis, the most important feature of quantum indefinite spacetime is that spontaneous quantum gravitational fluctuations of causal structure, i.e., causal fluctuations, take place around the Planck scale. Causal fluctuations have significant implications on fundamental questions of quantum gravity such as ultraviolet finiteness, black hole information, and possibly also cosmology. We devote Sections 6 to 8 to study these implications.

1.3 Outline of the thesis

The mathematical framework we adopt, the so-called “process framework” [6] which generalizes ordinary quantum theory, is reviewed in Section 2. In Section 3 we present a few explicit models that describe indefinite causal structure of different kinds. In Section 4 we extend the notion of entanglement measures to apply to quantum networks and quantum processes with indefinite causal structure. In Section 5 we demonstrate through examples how the new notion allows us to study entanglement in more general

²This is reminiscent of the regularization from Shannon sampling theory that allow spacetime to be “simultaneously continuous and discrete” [7, 8, 9]

settings such as timelike separated parties. In Section 6 we identify causal fluctuation as a possible mechanism that regularizes entanglement in the ultraviolet. In Section 7 we examine the definition of black holes in quantum indefinite spacetime, show that strict black holes do not exist, and offer a possible solution to the information paradox. In Section 8 we present a preliminary proposal that explains “dark energy” using quantum fluctuation of entanglement. Finally, in the concluding Section 9 we summarize the logical structure of the thesis and discuss some topics for future research. Appendix A reviews some standard results in quantum information theory to be used in the thesis, while Appendix B proves some elementary new results about process coherent information.

2 Review of the process framework

2.1 Basic assumptions

The process framework first proposed in [6] is an extension of ordinary quantum theory to allow indefinite global causal structure. The main idea is to assume that ordinary quantum theory with fixed causal structure holds locally, while globally the causal structure can be indefinite. This is reminiscent of classical GR, where locally spacetime is flat and special relativity holds, while globally spacetime can be curved and general relativistic effects appear. Specifically, the process framework builds on the following assumptions³, whose conceptual meaning is explained in this section and whose mathematical meaning will be explained in the next section:

1. **Local quantum mechanics:** ordinary quantum theory with fixed causal structure holds locally.
2. **Linearity of probability:** probabilities of randomized and coarse-grained operation outcomes obey linearity.
3. **Non-negative normalized probability:** probabilities are non-negative and sum to 1.

The idea of the first assumption is that there is a set of local parties within each of which ordinary quantum theory with definite causal structure holds. **Local parties** (or simply “parties”) are “local” first because

³An implicit additional assumption is that the joint probabilities are non-contextual, i.e., equivalent local operations lead to the same probabilities.

operations performed in one party is separated from those performed in another. In terms of information transmission, each party is associated with an input and an output system, which are different from the input and output systems of other parties. Furthermore, the parties are “local” in the sense that all information transmission between a local party and the outside is through the specified input and output systems, i.e., information can transmit to the party only through the specified input system, and can transmit out of the party only through the specified output system. Intuitively, one can picture local parties as closed laboratories with specified input and output systems.

The “local quantum mechanics” assumption states that physical operations conducted within a local party are described by ordinary quantum theory. In particular, the causal structure is ordinary within the party such that information propagates from the input to the output system but not the other way around.

The next assumption is about the probabilities the theory assigns to outcomes of local operations. In ordinary quantum theory a state can be viewed as assigning probabilities to measurement outcomes. In the theory incorporating indefinite causal structure a generalized state assigns probabilities correlating outcomes of local operations of different parties. Such a generalized state can carry non-trivial correlation that amounts to indefinite causal structure among its subsystems, and is called a **process**. The framework itself is called the process framework.

In the next section we discuss the mathematical representation of processes and the implications of the basic assumptions.

2.2 Mathematical formulation

In this section we review the mathematical language of the process framework, mainly following [6, 10] with minor adaptations⁴.

As mentioned in the last section, a local party A is associated with an input system a_1 with Hilbert space \mathcal{H}^{a_1} and an output system a_2 with Hilbert space \mathcal{H}^{a_2} . In accordance with the “local quantum mechanics” assumption, a local party A can apply operations represented by ordinary quantum instruments. A quantum instrument is a set of completely positive

⁴In the original articles, processes are represented by unnormalized operators. We adapt the framework to use normalized operators, because this is more convenient when we study entanglement measures in later sections. The adapted representation is equivalent to the original one and does not make an essential difference.

(CP)⁵ trace-non-increasing maps $\{\mathcal{E}_i\}_i$ that sum up to a CP trace-preserving map (channel) $\mathcal{E} = \sum_i \mathcal{E}_i$. These maps take in density matrices on \mathcal{H}^{a_1} and output (possibly unnormalized) density matrices on \mathcal{H}^{a_2} . The label i represents the classical outcome of the operation. For example, a quantum instrument can represent a “measure-and-prepare” operation, where the party measures the input quantum state in some basis to obtain the classical outcome i , and prepares output states according to i .

It is convenient to represent the CP maps of a quantum instrument by the so-called Choi states [11]. Given a CP map $\mathcal{M} : L(\mathcal{H}^{a_1}) \rightarrow L(\mathcal{H}^{a_2})$ from system a_1 to system a_2 ($L(\mathcal{H})$ stands for linear operators on Hilbert space \mathcal{H}), its Choi state is defined as

$$\mathbf{M}_{a_1}^{a_2} := \mathcal{M} \otimes \mathbb{1}(|\phi_+\rangle\langle\phi_+|) \in L(\mathcal{H}^{a_2} \otimes \mathcal{H}^{a_1}). \quad (1)$$

The subscript a_1 (inheriting changed font of the operator) represents the input system, while the superscript a_2 (inheriting changed font of the operator) represents the output system. $\mathbb{1}$ is the identity channel on system a_1 . $|\phi_+\rangle = \sum_i d_{a_1}^{-1/2} |ii\rangle \in \mathcal{H}^{a_1} \otimes \mathcal{H}^{a_1}$ is a normalized⁶ maximally entangled state on two copies of system a_1 , and d_{a_1} is the dimension of \mathcal{H}^{a_1} . The Choi states and the CP maps are in one-to-one correspondence [11], so we can identify the original maps with their Choi states. Finally, in the following we use sans serif font letters to denote the Choi state operators corresponding to the maps in the normal font, e.g., we use $\mathbf{M}_{a_1}^{a_2}$ to denote the Choi state operator of the map $M_{a_1}^{a_2}$.

Suppose there exist several local parties⁷ A, B, \dots, C that can choose to apply local operations represented by quantum instruments $\{\mathcal{E}_i^A\}_i, \{\mathcal{E}_j^B\}_j, \dots, \{\mathcal{E}_k^C\}_k$. These can be identified with their Choi states in $L(\mathcal{H}^{a_1} \otimes$

⁵A linear map $T : L(\mathcal{H}^1) \rightarrow L(\mathcal{H}^2)$ is positive if it takes positive semi-definite operators to positive semi-definite operators. It is completely positive if $T \otimes \mathbb{1} : L(\mathcal{H}^1 \otimes \mathcal{H}^3) \rightarrow L(\mathcal{H}^2 \otimes \mathcal{H}^3)$ is positive for all $\mathcal{H}^3 = \mathbb{C}^n$ where n is a positive integer. In particular, a completely positive map is positive. We require a physical operation to be positive because it should take states to states. We require it to be completely positive because the same should hold when it acts on a subsystem of a bipartite state.

⁶Many articles use the unnormalized state $\sum_i |ii\rangle$ for the Choi state. We find it more convenient to use the normalized state. This is just a convention and does not make a real difference for the content of the framework. As for the basis of the maximally entangled state, it is a matter of convention and different choices lead to the same predictions of probabilities.

⁷In this thesis we restrict attention to finite-dimensional systems, and consider only finite-dimensional Hilbert spaces.

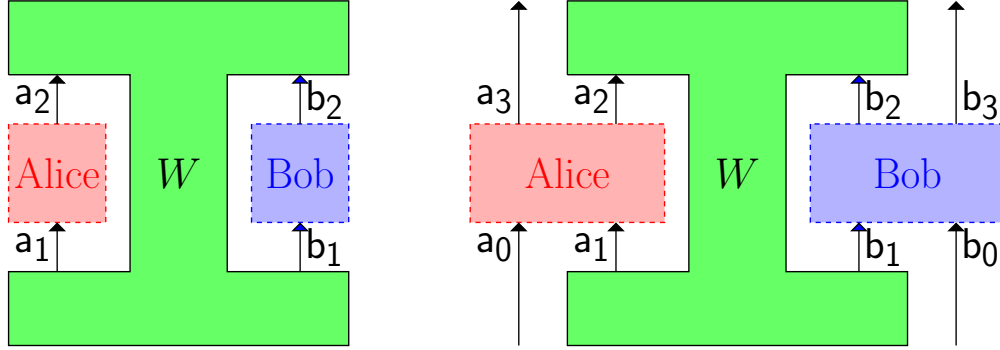


Figure 1: Left: process W as a probability assigning function that maps quantum operations from Alice and Bob to a probability. Right: process W as a map that acts partially on Alice’s and Bob’s operations to yield an object awaiting further operations to obtain a probability.

\mathcal{H}^{a_2} , $L(\mathcal{H}^{b_1} \otimes \mathcal{H}^{b_2}), \dots, L(\mathcal{H}^{c_1} \otimes \mathcal{H}^{c_2})$. The “linearity of probability” assumption says that the probability a process assigns to a particular set of outcomes i, j, \dots, k is multilinear in the CP maps, and hence in their Choi states.

Physically the linear sums can arise from randomization and coarse-graining of local outcomes. Suppose A can choose from two instruments $\{\mathcal{E}_i^A\}_i$ and $\{\mathcal{F}_i^A\}_i$ with the same number of outcomes. Then A can randomize over the two instruments to perform the first with probability p and the second with $1-p$. The new instrument is $\{\mathcal{G}_i^A\}_i$ with $\mathcal{G}_i^A = p\mathcal{E}_i^A + (1-p)\mathcal{F}_i^A$ for all i , where each new CP map is a convex linear sum. Now suppose A starts with only one instrument $\{\mathcal{E}_i^A\}_{i=1}^n$, but coarse grain over two outcomes, say $i = n-1$ and $i = n$. Then the new instrument is $\{\mathcal{F}_i^A\}_{i=1}^{n-1}$, with $\mathcal{F}_i^A = \mathcal{E}_i^A$ for $i < n-1$ and $\mathcal{F}_{n-1}^A = \mathcal{E}_{n-1}^A + \mathcal{E}_n^A$. The linearity of probability assumption says that the probability assignment of the process commutes with these linear sums. This is physically reasonable, as probability assignment should commute with randomization or coarse-graining.

The “non-negative and normalized probability” assumption implies that the probabilities are non-negative for any set of outcomes and they add to 1 when we sum over all outcomes. More generally, local operations can be extended to include ancillary systems and the multilinear probability function of the process can be extended to a multilinear map acting only on subsystems of the local parties (Figure 1). The further requirement is that the resulting object is physical – when its open slots are closed by feeding

in physical operations, the probability is non-negative and normalized (this is analogous to moving from requiring positivity to complete positivity for quantum transformations).

It can be shown⁸ [6, 10] that such probability maps can be represented by operators $W \in L(\mathcal{H})$, where $\mathcal{H} := \mathcal{H}^{a_1} \otimes \mathcal{H}^{a_2} \otimes \mathcal{H}^{b_1} \otimes \mathcal{H}^{b_2} \otimes \dots \otimes \mathcal{H}^{c_1} \otimes \mathcal{H}^{c_2}$, such that the probability for a set of local classical outcomes i, j, \dots, k is given by (T denotes operator transpose)

$$P(\mathbf{E}_i^A, \mathbf{E}_j^B, \dots, \mathbf{E}_k^C) = \text{Tr}\{[\mathbf{E}_i^A \otimes \mathbf{E}_j^B \otimes \dots \otimes \mathbf{E}_k^C]^T W\}. \quad (2)$$

Here W is to be understood as a generalized density matrix that assigns probabilities to measurement outcomes. The “non-negative normalized probability” assumption of the process framework imposes the following conditions on W :

$$W \geq 0, \quad (3)$$

$$\text{Tr } W = d_{\mathcal{H}}, \quad (4)$$

$$W = L_V(W), \quad (5)$$

where $d_{\mathcal{H}}$ is the dimension of Hilbert space \mathcal{H} and L_V is a projector onto a linear subspace⁹

$$L_V(W) := [1 - \prod_A (1 - a_2 + a_1 a_2) + \prod_A a_1 a_2] W. \quad (6)$$

The first condition follows from non-negativity of probability, and the latter two conditions follow from the normalization of probability. The third condition may look complicated. In the next section we illustrate it in the concrete two-party case and give intuitive interpretations.

In the original articles, W is not a normalized operator (as seen from (4)). However, one of the goals of this thesis is to study entanglement measures on W . It is convenient to define entanglement measures on normalized operators. Therefore we adapt the convention of the original articles and normalize W as follows. One may wonder why the operator transpose shows up in (2). This has to do with the way Choi states model operation composition. Given operations¹⁰ A_{ya}^x and B_v^{au} , we can compose them by feeding

⁸If a definite causal order among the local parties is imposed, the problem falls into the “quantum-comb” formalism of quantum networks [12].

⁹We use the shorthand notation that ${}_x W := \omega^x \otimes \text{Tr}_x W$, where ω^x is the maximally mixed state on system x , and that ${}_{[1-x]} W := W - {}_x W$. \prod_A denotes the product over all different parties A, B, C, \dots each with an input and an output system.

¹⁰We introduce the notation to use normal font capital letters to denote maps with input systems in the subscript and output systems in the superscript.

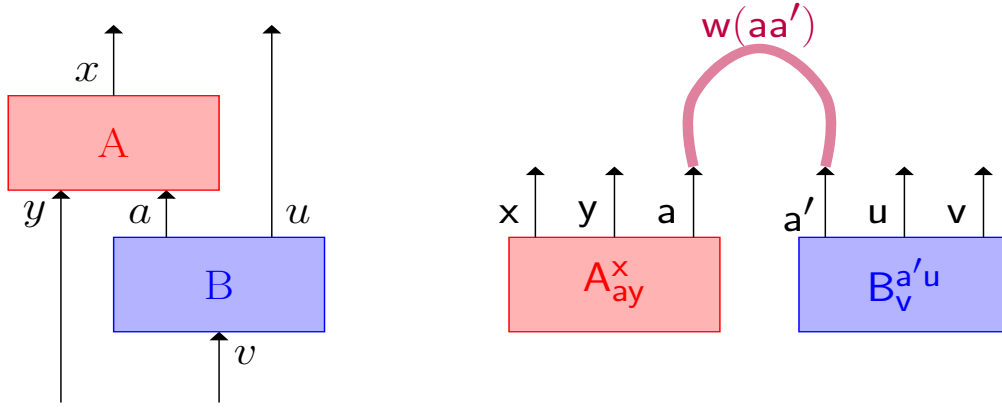


Figure 2: Left: composition of operations. Right: composition of operations at the Choi state level. An unnormalized wire operator is introduced to connect the two normalized Choi states.

the output a of B_v^{au} into the input a of A_{ya}^x to obtain a new operation¹¹ $C_{yv}^{xu} := A_{ya}^x B_v^{au}$ (Left picture of Figure 2). At the Choi state level, we have the composition rule [12]

$$C_{yv}^{xu} = d_a \operatorname{Tr}_a(A_{ya}^x B_v^{aTu}) = d_a \operatorname{Tr}_a(A_{ya}^x B_v^{au}), \quad (7)$$

where $d_a = \dim \mathcal{H}^a$, and \mathbf{a}^T denotes operator transpose on \mathcal{H}^a . The second equality says that under composition, input transpose and output transpose are equivalent, and either choice works. We can view (2) as implementing a composition rule. Each \mathbf{E} composes with the positive semi-definite operator \mathbf{W} twice, once at the input and once at the output. If we choose the convention to apply transposes on \mathbf{E} rather than on \mathbf{W} , then every \mathbf{E} is fully transposed. The composition rule has a pictorial interpretation that the two Choi states are composed through an entangled measurement [15] (Right picture of Figure 2).

Defining the unnormalized “wire” operator (a' is a copy of a)¹²

$$w(aa') := d_a \sum_{ij} (|ii\rangle\langle jj|)^{aa'}, \quad (8)$$

¹¹We introduce the “tensor-type” convention that repeated super- and subscripts denote composition. This symbolism comes from the so-called “operator tensor” formulation [13, 14].

¹²The “composition by wire” trick is also used in [16].

we can rewrite the composition rule as

$$C_{yv}^{xu} = \text{Tr}_{aa'}(A_{ya}^x w(aa') B_v^{a'u}). \quad (9)$$

Now we can rewrite (2) as the composition of the multiple E 's with a rescaled W as

$$P(E_i^A, \dots, E_k^C) \quad (10)$$

$$= \text{Tr}\{[E_i^A \otimes \dots \otimes E_k^C] [w(a_1 a'_1) \otimes w(a_2 a'_2) \otimes \dots \otimes w(c_1 c'_1) \otimes w(c_2 c'_2)] W\}. \quad (11)$$

The wire composition formula (9) appears redundant, so we introduce the shorthand notation

$$C_{yv}^{xu} = \text{Tr}_{aa'}(A_{ya}^x w(aa') B_v^{a'u}) \quad (12)$$

$$=: A_{ya}^x B_v^{au}, \quad (13)$$

where repeated index a in operators at the superscript and the subscript signify composition through a wire.

Then conditions (3) to (5) upgrade to (only the second condition is affected)

$$W \geq 0, \quad (14)$$

$$\text{Tr} W = 1, \quad (15)$$

$$W = L_V(W). \quad (16)$$

By introducing the wire operator, we normalized W to be a trace-one operator. We call W a **process operator**, representing the original process W . Processes incorporate information about quantum correlations and quantum causal structures among local parties, and will be the central objects we study in this paper. Ordinary trace-preserving quantum operations are special cases of processes, as they assign probabilities to operations and obey the three basic assumptions. The process framework is therefore a generalization of ordinary quantum theory with definite causal structure.

Finally, note that inputting maximally entangled states in the canonical basis to the processes does not change the process operator. Up to normalization, composing with the wire or the maximally entangled state in the canonical basis is equivalent to a partial transpose at the system of composition. Up to normalization, the operator obtained by inputting the

maximally entangled state is therefore equivalent to twice partial transposing the same system (composing with both the wire and the maximally entangled state), which is equivalent to doing nothing. Since the resulting operator is normalized, the whole procedure simply yields the original operator. An important implication is that any process operator can be obtained by inputting canonical maximally entangled states at all the inputs of the original process map.

2.3 Normalization constraints

We are particularly interested in processes of the form $W_{a_2 b_2}^{a_1 b_1}$ with two local parties A and B , where A has input system a_1 and output system a_2 , and B has input system b_1 and output system b_2 ¹³. Recall (15) and (16) that follow from the normalized probability assumption:

$$\text{Tr } W = 1, \tag{17}$$

$$W = L_V(W). \tag{18}$$

The first condition says that the process W is represented as a trace-one operator, while the second imposes the non-trivial constraints [10]:

$${}_b W = {}_{a_2 b} W, \tag{19}$$

$${}_a W = {}_{b_2 a} W, \tag{20}$$

$$W = {}_{a_2} W + {}_{b_2} W - {}_{a_2 b_2} W. \tag{21}$$

We specify some notational conventions to be used throughout the thesis. We already introduced the shorthand notation ${}_x W := \omega^x \otimes \text{Tr}_x W$ with ω the maximally mixed state. When no ambiguity arises, we sometimes omit super- and subscripts of a process or a process operator (e.g., write W), and sometimes do not distinguish super- and subscripts to write all systems as superscripts (e.g., write $W^{a_1 a_2 b_1 b_2}$ for $W_{a_2 b_2}^{a_1 b_1}$). We sometimes group the input and the output of some local party together as a single letter (e.g., write \mathbf{a} for $a_1 a_2$). Finally, the reduced operators of $W^{a_1 a_2 b_1 b_2}$ are denoted by labelling the remaining subsystems (e.g., $W^{\mathbf{a}} = \text{Tr}_b W^{\mathbf{a}b}$).

¹³The reason input systems appear as superscripts and output systems appear as subscripts in the process $W_{a_2 b_2}^{a_1 b_1}$ is that inputs of local parties become outputs of the process, and *vice versa*.

Using these conventions, the above constraints imply that

$$W^a = W^{a_1} \otimes \omega^{a_2}, \quad (22)$$

$$W^b = W^{b_1} \otimes \omega^{b_2}, \quad (23)$$

$$W^{a_2 b_2} = \omega^{a_2 b_2}, \quad (24)$$

$$W^{a b_2} = \omega^{a_2} \otimes W^{a_1 b_2}, \quad (25)$$

$$W^{a_2 b} = \omega^{b_2} \otimes W^{a_2 b_1}. \quad (26)$$

These conditions all have intuitive interpretations. The operator W^{ab} can be expanded in the Hilbert-Schmidt basis of the subsystems as

$$W^{a_1 a_2 b_1 b_2} = \sum_{i,j,k,l} W_{ijkl} \sigma_i^{a_1} \sigma_j^{a_2} \sigma_k^{b_1} \sigma_l^{b_2}, \quad W_{ijkl} \in \mathbb{R}. \quad (27)$$

For example, for qubit systems the Pauli basis contains four elements $\sigma_0 = \mathbb{1}$, and the Pauli operators $\sigma_i, i = 1, 2, 3$. For general systems, we always take $\sigma_0 = \mathbb{1}$.

The condition of normalized probability excludes components of certain types. We refer to terms of the form $\sigma_i^x \otimes \mathbb{1}^{\text{rest}}$ for $i \geq 1$ as of type x , $\sigma_i^x \otimes \sigma_j^y \otimes \mathbb{1}^{\text{rest}}$ for $i, j \geq 1$ as of type xy etc. It can be shown that type

$$a_1, b_1, a_1 b_1, a_1 b_2, a_2 b_1, a_1 a_2 b_1, a_1 b_1 b_2 \quad (28)$$

terms are allowed, while type

$$a_2, b_2, a_2 b_2, a_1 a_2, b_1 b_2, a_1 a_2 b_2, a_2 b_1 b_2, a_1 a_2 b_1 b_2 \quad (29)$$

terms are not. One can heuristically interpret the constraints as allowing and excluding certain types of correlations [6].

Allowed types:

- States: $a_1, b_1, a_1 b_1$
- Channels: $a_1 b_2, a_2 b_1$
- Channels with memory: $a_1 a_2 b_1, a_1 b_1 b_2$

Excluded types:

- Postselection: $a_2, b_2, a_2 b_2$
- Local loops: $a_1 a_2, b_1 b_2$

- Channels with local loops: $a_1 a_2 b_2, a_2 b_1 b_2$
- Global loops: $a_1 a_2 b_1 b_2$

Finally, for the process W to be represented by a trace-one density matrix, the coefficient W_{0000} for component $\mathbb{1}^{a_1 a_2 b_1 b_2}$ must be $1/d_{a_1 a_2 b_1 b_2}$, as all other components are traceless.

2.4 Purification postulate

The three basic assumptions of the process framework are quite general and they allow a large family of processes. There remains the question whether all of them are physical. One proposal is that physical processes should further obey a “purification postulate”, which generalizes the fact for states that all mixed states can be (mathematically) purified [17]¹⁴. Apart from fundamental considerations, process purification, like state purification, is very useful in deriving technical results. To keep the theory general, we will not impose this postulate at a fundamental level for the process framework. However, we assume the postulate in Section 6 for the technical convenience it offers in proving results.

Consider a two party process $W_{a_2 b_2 p}^{a_1 b_1 f}$ extended to have a global past p and a global future f . It can act on extended local operations $A_{a_1 a'_1}^{a_2 a'_2}$ and $B_{b_1 b'_1}^{b_2 b'_2}$ to create

$$G_{a'_1 b'_1 p}^{a'_2 b'_2 f} := W_{a_2 b_2 p}^{a_1 b_1 f} A_{a_1 a'_1}^{a_2 a'_2} B_{b_1 b'_1}^{b_2 b'_2}. \quad (30)$$

The process W is said to be **pure** if for all unitaries A and B , G is a unitary. It can be proved that [17]

Theorem 1. A process W is pure if and only if W equals the Choi state for some unitary channel.

A two-party process $W_{a_2 b_2}^{a_1 b_1}$ with trivial p and f systems is **purifiable** if it can be recovered from a pure process $W_{a_2 b_2 p}^{a_1 b_1 f}$ by inputting the state $|0\rangle$ in p and tracing out f , i.e., if

$$W_{a_2 b_2}^{a_1 b_1} = W_{a_2 b_2 p}^{a_1 b_1 f} |0\rangle\langle 0|^p. \quad (31)$$

The **purification postulate** states that a process is physical only if it is purifiable. We call $W_{a_2 b_2 p}^{a_1 b_1 f}$ or $W_{a_2 b_2 p}^{a_1 b_1 f} |0\rangle\langle 0|^p$ the **purification** of $W_{a_2 b_2}^{a_1 b_1}$. By Theorem 1, the purifications have pure Choi states.

¹⁴The purification postulate for general probabilistic theories is studied in [18].

3 Models of indefinite causal structure

In this section, we present explicit two-party processes exhibiting superposition of causal relations. With fixed causal structure, two local parties must be in one of the three causal relations: $A \nearrow B$, $A \nwarrow B$, and $A - B$, denoting A causally preceding B , A causally succeeding B and A causally disconnected with B , respectively. With indefinite causal structure, there are the possibilities of $A \nwarrow \nearrow B$, $A \nearrow \nwarrow B$, and $A \nwarrow \nearrow B$, denoting the respective superpositions of two or three causal relations. In this section, we give example processes of each kind. These serve as concrete examples to gain intuition from, and are good to keep in mind in understanding the general results in later sections. To prepare for the study of black hole information in Section 7, we also discuss communication capacity of relevant processes in this section.

One useful overall strategy to realize superposition of different causal structure we use several times in this section is to adjoin ancilla systems. Given a set of processes $\{|w_i\rangle\langle w_i|\}_i$ describing different causal relations, we can superpose them as $|w\rangle := \sum_i a_i |i\rangle \otimes |w_i\rangle$ by introducing an ancilla system. The process $|w\rangle\langle w|$ then describes a coherent superposition of the different causal relations. This method requires the original processes $W_i = |w_i\rangle\langle w_i|$ to be described by pure vectors. To incorporate originally mixed processes one can start with the purification of the original processes and trace out the purifying systems in the end.

The examples we give in this section exhibit superposition of causal relation in the operational sense. Operational causal relation is to be distinguished from spacetime causal relation. $A \nearrow B$ in the operational sense implies $A \nearrow B$ in the spacetime sense, but not *vice versa*. Event A can causally precede event B in spacetime without being able to influence B , e.g., communication from A to B may be blocked by other objects. Conversely, $A \nearrow B$ in the spacetime sense implies $A \nearrow B$ in the operational sense, but not *vice versa*. Processes are defined operationally with respect to local parties where local operations can be applied. Therefore processes naturally describe operational causal relations. Nevertheless when applying the theory, as we do in later sections, we can choose processes whose spacetime causal relation coincides with the operational causal relation, and have the processes describe quantum spacetime.

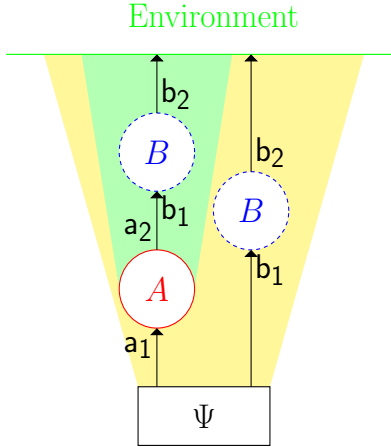


Figure 3: Heuristic illustration of the superposition of $A \nearrow B$ and $A - B$. Information from the initial state Ψ propagates through the region filled with yellow color, while information from party A propagates through the region filled with green color. Party B is in a superposition of being in the causal future of A and of being out of the causal future of A . All information not collected by A or B ends up collected by the environment system.

3.1 Superposition of $A \nearrow B$ and $A - B$

The situation we consider is illustrated in Figure 3. We superpose two causal relations, A causally disconnected with B where they share a bipartite state, and A causally precede B where they share a channel.

There are previous works studying descriptions of the same superposition of causal relation. In [19], this is called “coherent mixture” of “cause-effect” and “common-cause” relations, while in [20], this is called “coherent superposition” of “direct-cause” and “common-cause” causal structures. Below we present the processes constructed in these previous works with a slight generalization to allow arbitrary probability amplitudes for the two causal relations. In addition, we discuss quantum communication capacity of the processes, which is not considered in the previous works.

3.1.1 Process with ancilla

The first strategy to implement the superposition is to introduce an ancilla system, the general idea of which we discussed above. The process is given

by

$$W_{a_2 b_2}^{a_1 b_1} := \text{Tr}_{c e_1 e_2} |w\rangle\langle w|, \quad (32)$$

where

$$|w\rangle := \alpha |0\rangle^c |\Psi\rangle^{a_1 b_1} |I\rangle^{a_2 e_1} |I\rangle^{b_2 e_2} + \beta |1\rangle^c |\Psi\rangle^{a_1 e_2} |I\rangle^{a_2 b_1} |I\rangle^{b_2 e_1}. \quad (33)$$

Here α and β are complex probability amplitudes such that $|\alpha|^2 + |\beta|^2 = 1$. $|I\rangle^{xy}$ is the (normalized) Choi state of the identity channel. The systems c, e_1, e_2 are subsystems that altogether purify $W_{a_2 b_2}^{a_1 b_1}$.

Comparing the systems and terms with Figure 3 gives some intuitions about why the process takes this form. Systems e_1 and e_2 can be regarded as subsystems of the environment the left and right arrows point to, respectively. The ancilla system c is not shown in the figure, but its basis states $|0\rangle$ and $|1\rangle$ indicate the two different places of B in the figure. The Choi states $|I\rangle$ correspond to arrows in the figure representing information transmission between systems. The whole process $W_{a_2 b_2}^{a_1 b_1}$ describes the superposition of B at the two places shown in the figure.

The special case with $\alpha = \beta = 1/\sqrt{2}$ is was constructed in [20] to describe equal amplitude superposition of “direct-cause” and “common-cause” relations between A and B .

3.1.2 Process with partial swap

The second strategy to implement the superposition is to use a partial swap unitary. This was used in [19] to construct an equal probability amplitude superposition of a maximally entangled initial state and a noiseless channel. Below, we generalize it to allow arbitrary probability amplitudes, plus noisy initial states and channels. The main idea is to coherently mix the initial state and the channel through the partial swap unitary.

Define the “partial swap” channel $P(p)$ ($0 \leq p \leq 1$) on two subsystems of equal dimension to correspond to the following partial swap unitary:

$$\sqrt{1-p} \mathbb{1} + \sqrt{p} i U_{SW}. \quad (34)$$

Here $\mathbb{1}$ is the identity operator and U_{SW} is the swap operator. Define

$$W_{a_2}^{a_1 b_1} := \text{Tr}_e P(p)_{a'_1 a'_2}^{b_1 e} \rho^{a_1 a'_1} N_{a_2}^{a'_2}, \quad (35)$$

where a'_1 and a'_2 are copies of a_1 and a_2 , ρ is an initial state, and N is a channel.

The $\mathbb{1}$ part of P transmits a'_1 to b_1 and a'_2 to the environment e , while the $i U_{SW}$ part of P does the opposite, sending a'_2 to b_1 and a'_1 to e . The whole setting puts $\mathbb{1}$ and $i U_{SW}$ parts into a “coherent superposition”, such that A and B partially share a channel from a_2 to b_1 , and partially share a bipartite state on a_1 and b_1 . The output b_2 of B is not correlated with a_1, a_2 or b_1 . Its information goes directly to the environment through a unitary channel such that the reduced Choi state on b_2 is maximally mixed. The whole two-party process on $a_1 a_2 b_1 b_2$ is

$$W_{a_2 b_2}^{a_1 b_1} = W_{a_2}^{a_1 b_1} \otimes W_{b_2}, \quad (36)$$

where $W_{b_2} = \omega$. This strategy superposes noisy states and channels directly, in contrast to the ancilla system strategy which starts with pure resources.

3.1.3 Communication capacity

In this section we discuss classical and quantum communication capacities for $A \not\prec B$ processes. The crucial result for applications to quantum indefinite spacetime is that under reasonable assumptions two “almost causally disconnected” parties can communicate both quantum and classical information, even for the tiniest probability amplitude of $A \not\prec B$. We use this to argue in Section 7 that in quantum indefinite spacetime black holes do not exist, and information can transmit out of regions classically considered to be within black holes.

It is not hard to obtain a lower bound on the communication capacity for a $A \not\prec B$ process. If A conducts a trivial measurement to trace out system a_1 of the process, then the two parties share a channel from a_2 to b_1 (b_2 is not correlated with a_2 or b_1 so it is irrelevant). The classical and quantum communication capacities are lower bounded by those of this channel.

We consider the case with noiseless initial state $\rho^{a_1 b_1}$ and noiseless channel $N_{a_2}^{b_1}$. The general case with noisy initial resources can be regarded as arising from this case by dropping subsystems. We make the further simplifying assumption that $\rho^{a_1 b_1}$ is maximally entangled. This is reasonable in the context of quantum black holes in the vacuum, which is what we want to apply the results in this section for. The channel to be shared after A performs the trivial measurement on a_1 is unitarily equivalent to the depolarizing channel $D_{a_2}^{b_1}$ parametrized by p , which is the $A \not\prec B$ amplitude squared:

$$D(\rho) = p\rho + (1 - p)\omega. \quad (37)$$

The classical channel capacity of the depolarizing channel is known to be [21]:

$$\log d + \left(1 - p + \frac{p}{d}\right) \log \left(1 - p + \frac{p}{d}\right) + (d - 1) \frac{p}{d} \log \left(\frac{p}{d}\right), \quad (38)$$

where d is the dimension of the system. Unitarily equivalent channels have same capacities, so this is a lower bound to the classical capacity. It is positive as long as the probability amplitude for $A \nearrow B$ is positive. We arrive at the profound implication that for initially causally disconnected A and B , any tiny bit of causal fluctuation can induce positive classical communication capacity.

Even further, the tiny causal fluctuation also induces quantum communication capacity. When the probability amplitude for $A \nearrow B$ is small, A can trace out a_2 by sending in a maximally mixed state to a_2 . A and B then share a bipartite state on $a_1 b_1$ that is close to maximally entangled, from which maximally entangled state can be distilled. In the quantum Shannon setting A and B share many copies of the process, so they can use some to communicate classical communication and some to distill maximal entanglement. Quantum teleportation is possible through this, so the quantum communication capacity is positive.

To obtain the exact values of the classical and quantum capacities is a more complicated task. For this thesis these lower bounds already suffice for us to draw some definite conclusions on black holes in quantum indefinite spacetime in Section 7.

3.2 Superposition of $A \nearrow B$, $A \nwarrow B$ and $A - B$

A superposition of the three causal relations illustrated in Figure 4 can be realized using a qutrit ancilla. Define

$$\begin{aligned} |w\rangle := & \alpha |0\rangle^c |\Psi\rangle^{a_1 e_2} |I\rangle^{a_2 b_1} |I\rangle^{b_2 e_1} + \beta |1\rangle^c |\Psi\rangle^{a_1 b_1} |I\rangle^{a_2 e_1} |I\rangle^{b_2 e_2} \\ & + \gamma |2\rangle^c |\Psi\rangle^{e_1 b_1} |I\rangle^{b_2 a_1} |I\rangle^{a_2 e_2}. \end{aligned} \quad (39)$$

Then $W_{a_2 b_2}^{a_1 b_1} := \text{Tr}_{c e_1 e_2} |w\rangle\langle w|$ is a process with the designated causal structure.

3.3 Superposition of $A \nearrow B$ and $A \nwarrow B$

Although we will not make further use of it in this thesis, we mention that a superposition $A \nearrow B$ and $A \nwarrow B$ can be realized using an ancilla qubit.

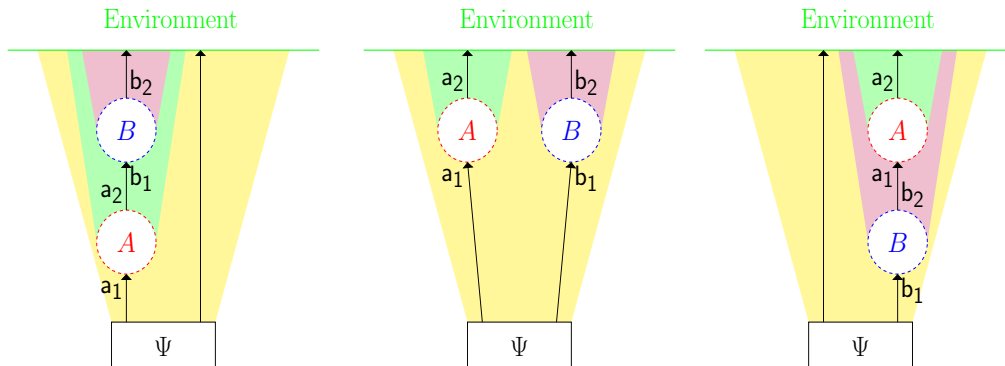


Figure 4: The three causal relations $A \nearrow B$, $A - B$ and $B \prec A$ with the same initial state Ψ . All information not collected by A or B ends up collected by the environment.

Define

$$|w\rangle := \alpha |0\rangle^c |\Psi\rangle^{a_1} |I\rangle^{a_2 b_1} |I\rangle^{b_2 e_1} + \beta |1\rangle^c |\Psi\rangle^{b_1} |I\rangle^{b_2 a_1} |I\rangle^{a_2 e_1}. \quad (40)$$

Then $W_{a_2 b_2}^{a_1 b_1} := \text{Tr}_{ce_1} |w\rangle\langle w|$ is of the causal structure $A \not\bowtie B$. In the context of quantum computational circuit processes with this causal structure have been studied under the name of “quantum switch” and are shown to offer advantages in certain information processing tasks [22, 23, 24, 25].

4 Entanglement measures

Expectations on entanglement to play an important role in quantum gravity is high, as it is a quantum property that also connects to spacetime structure (for some examples exploiting this connection, see [26, 27, 28, 29, 30]). Ordinarily, entanglement is a property of bipartite states. It is usually assumed that the two parties have fixed causal relation (causally disconnected at spacelike separated regions) in order that they share a state. However, in quantum gravity with the presence of indefinite causal structure the causal relation between two parties is generically indefinite, and a state can be “superposed” with a transformation. There is a need to redefine entanglement in quantum gravity. In this section, we propose a new axiom for entanglement measures in the process framework¹⁵, and study some particular

¹⁵The related question of how to define entropy in the causaloid framework for indefinite causal structure was studied in [31]. See [32] for an early proposal to define

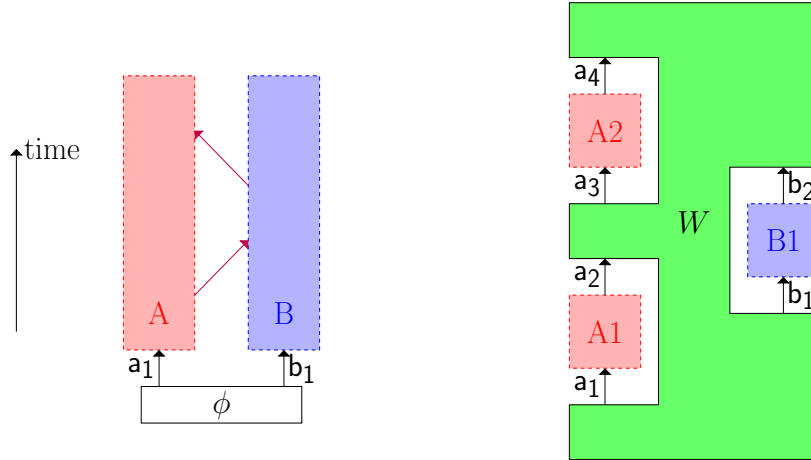


Figure 5: Left: The bipartite state LOCC paradigm assumes that two parties can freely communicate classically (signified by purple arrows in the left picture). This implies that they can apply operations in extended periods of time – the parties are not “localized” in time.

Right: To have non-trivial causal structure, local parties for processes are “localized” in time (if they are “extended” in time to enable two-way classical communication, the causal structure is trivial). In the process LO paradigm only local operations are free.

The LOCC paradigm can be reproduced in the process LO paradigm by endowing the processes with classical communication resources. For example, setting $W^{a_1 a_2 a_3 a_4 b_1 b_2} = \phi^{a_1 b_{11}} \otimes C_{a_{22}}^{b_{12}} \otimes C_{b_{22}}^{a_{32}} \otimes I_{a_{21}}^{a_{31}}$ with classical channel C and quantum identity channel I in the right picture reproduces the situation in the left picture ($a_2 = a_{21} \otimes a_{22}$, $b_1 = b_{11} \otimes b_{12}$ etc.).

measures.

4.1 Process entanglement measures

For general bipartite states (including mixed states) there is no unique entanglement measure, and the choice of measure depends on the operational tasks of interest. It was proposed that all state entanglement measures should obey the monotonicity axiom (entanglement measures cannot increase under local operation and classical communication (LOCC)) [33].

The naturalness of the monotonicity axiom relies on the LOCC paradigm,

entanglement for unitary channels.

where local operations and classical communications are assumed to be free resources. For processes designed to incorporate indefinite causal structure, however, we propose that the suitable paradigm is instead the LO (local operation) paradigm, where classical communications are not free. This is for the simple reason that otherwise the local parties are always causally related and the causal structure is trivial (Figure 5).

In the LO paradigm for processes, resource for classical communication is to be carried by the process itself, and the communication is conducted by the parties applying local operation at the input and output of the channel to encode and decode messages. One can view the process framework with the LO paradigm as a fine-grained version of the LOCC paradigm such that a party localized in space (such as A in the first picture of Figure 5) is dissected into several parties localized in spacetime (such as $A1$ and $A2$ in the second picture of Figure 5), and the encoding and decoding local parties of each use of classical communications is recorded. In the LOCC paradigm one distinguishes among classes of resources (C1 - class of local operations, C2a - class of one-way forward LOCC, C2b - class of one-way backward LOCC, C3 - class of two-way LOCC). These correspond to ending processes with different classical communication resources in the LO paradigm.

Note that for processes each local party has an input and an output system. Local operations of a party can apply a joint operation on the input and the output systems. A party A with input a_1 and output a_2 is allowed to compose a channel $N_{a_1 a_1}^{a_2 a_2}$ with the process at a_1 and a_2 as a local operation (applying a channel to a_1 and inputting some state to a_2 at the same time is a special case). This differs from state local operations where local operations on one system cannot be correlated with those on another system.

We ask **process entanglement measures** to obey the following monotonicity under local operations axiom:

- **Monotonicity under LO.** Process entanglement measures do not increase under local operations.

On states, this requirement reduces to the monotonicity under LOCC axiom. Suppose in the bipartite state LOCC paradigm we are given a state ρ and are free to apply some class (C1, C2a, C2b, C3) of classical communication. In process LO paradigm this translates into being given a process $W = \rho \otimes R \otimes N$, where R is the corresponding classical communication resources and N is the quantum identity channels induced by breaking parties

with time extension into parties “localized” in time (e.g., I in the caption of Figure 5). Consider an arbitrary operation O in the LOCC paradigm taking ρ to $\chi = O(\rho)$. Suppose E is a process entanglement measure, we show that $E(\rho) \geq E(\chi)$. We make the implicit assumption that in the process LO paradigm, as far as entanglement measures are concerned, it is equivalent to applying no operation by applying the identity map inside local parties to pass the initial state ρ through the induced identity channels N and discarding the resource R (analogously, in the state LOCC paradigm it is equivalent to applying nothing by applying the identity channels as local operation and discarding the classical communication resources). This takes W to ρ , and we have $E(W) = E(\rho)$. Moreover, since O can always be reproduced by some local operation O' on W , we have $\chi = O'(W)$. By the monotonicity axiom,

$$E(\rho) = E(W) \geq E(O'(W)) = E(\chi). \quad (41)$$

Since O is an arbitrary LOCC operation, E is monotonic under LOCC.

Since product processes of the form $\otimes_A W_A$ (W_A are processes that can be created locally by parties A) can be created freely from local operations, any process entanglement measure must reach its minimum value on them. As a convention, we usually set this minimum value to be zero. When two parties can freely communicate classically (one-way is enough), any separable state can be created from local operations. Hence process entanglement measures are zero on separable states when at least one-way (it does not matter which way) classical communication is free.

4.2 Coherent information for processes

In this section and next, we define coherent information for processes, and consider two process entanglement measures arising from this definition. Like for ordinary state entanglement measures, we are particularly interested in process entanglement measures with operational meanings. For quantum gravity and spacetime considerations we are especially interested in information communication, as this is the task that closely relates to causal structure. The celebrated quantum channel capacity theorem [21] says that the quantum channel capacity is given by the regularized coherent information, suggesting that we study coherent information for processes.

Processes generalize states and channels. The nice feature of coherent information with respect to processes is that it has been defined in a standard way for both bipartite states and channels. For a bipartite state ρ^{ab} ,

the coherent information with target system b is defined as

$$I^b(\rho^{ab}) := S(\rho^b) - S(\rho^{ab}) \quad (42)$$

$$= S^b - S^{ab}. \quad (43)$$

ρ^b is the reduced state on system b , $S(\rho^x)$ is the von Neumann entropy for ρ^x , and in the second line we introduced the shorthand notation that $S(\rho^x) = S^x$. $I^b(\rho^{ab})$ is the negative of the conditional entropy $S^{a|b}(\rho^{ab}) := S^{ab} - S^b$, so $I^b(\rho^{ab})$ can be positive, zero, or negative. For a pure state, the coherent information coincides with the entanglement entropy. Intuitively, the coherent information measures the quantum correlation of a bipartite state.

The coherent information for a quantum channel N_a^b with target system b is defined as

$$I^b(N_a^b) := \sup_{\rho^{aa'}} I^b(N_a^b \rho^{aa'}), \quad (44)$$

where the supremum is over input states $\rho^{aa'}$ with arbitrary auxiliary system a' ¹⁶. Operationally, it measures the maximal state coherent information of the final state the input and output parties can share. Alice at a prepares a state $\rho^{aa'}$ with arbitrary auxiliary system a' and sends the a part of it to Bob's side through the channel such that they share $N_a^b \rho^{aa'}$ on $a'b$ in the end. Alice is allowed to choose $\rho^{aa'}$ to maximize the coherent information of the final state. Intuitively, the coherent information for a channel measures its ability to set up quantum correlation.

A natural question is why is Bob not allowed to operate to optimize as well. The answer is that any local operation Bob applies cannot increase the coherent information of the final state (Theorem A.5). Therefore we can equivalently define the coherent information for quantum channel N_a^b with target system b as

$$I^b(N_a^b) := \sup_{LO} I^b(LO(N_a^b)), \quad (45)$$

where LO stands for local operations by Alice and Bob. Alice can input part of a bipartite state to a , and Bob can apply a channel to b .

¹⁶An equivalent, more commonly seen definition is to take supremum only over pure states (Theorem A.11).

This motivates us to define a generalized notion of **coherent information for processes**. The coherent information of the process W with target systems x and supplemental resource R is defined as

$$I_R^x(W) := \sup_{O_R} I^x(O_R(W)). \quad (46)$$

The target systems x is a subset of all systems $a_1, a_2, b_1, b_2, \dots$ of the local parties A, B, \dots of W . The supplemental resource R enables optimizing operations $O_R(W)$ to maximize the final state coherent information (we assume that the $O_R(W)$ is turned into a state by inputting maximally entangled states to all open input systems in the end). Usually R is taken to be local operations of the local parties. More generally R can include other resources such as classical communications.

As an example, the coherent information for a channel (45) can be recovered as a special case of the new definition if we take $W = N$, $x = b$ and $R = LO$.

In Appendix B we prove some elementary results about process coherent information.

4.3 Coherent-information-based measures

In this section we present two process entanglement measures built out of process coherent information. As the example of recovering channel coherent information suggests, a very natural way is to take R to be local operations for process coherent information:

$$I_{LO}^x(W) = \sup_{LO} I^x(LO(W)). \quad (47)$$

The most general local operation a party A with systems a_1 and a_2 can apply to a process is a channel $N_{a'_1 a_1}^{a'_2 a_2}$ that extends to auxiliary systems a'_1 and a'_2 . For coherent information optimization we restrict local operations to channels of the form $N_{a_1}^{a' a_2}$ to ensure that the final process is a shared state. $N_{a'_1 a_1}^{a'_2 a_2}$ can be turned into the form $N_{a_1}^{a' a_2}$ by inputting a bipartite state $\rho^{a'_1 a'_1}$ to a'_1 .

The **coherent information over local operation** we just defined is a process entanglement measure, as it trivially obeys the monotonicity axiom. As setting R to be local operations is so natural, we take this to be the “canonical” coherent information for processes, omit the subscript LO

and call it “the coherent information for the process W ” when no ambiguity arises.

Many state entanglement measures such as distillable entanglement are defined in the asymptotic limit. In the Shannon communication setting capacities are defined in the asymptotic limit. The quantum channel capacity is given by the regularized coherent information calculated in the asymptotic limit. It is then natural that we define a regularize measure out of process coherent information. Given $I_R^x(W)$, denote the regularized coherent information as

$$\mathfrak{I}_R^x(W) := \lim_{k \rightarrow \infty} \frac{1}{k} I_R^x(W^{\otimes k}). \quad (48)$$

Then it is easy to see that $\mathfrak{I}_{LO}^x(W)$ obeys the monotonicity axiom. This **regularized coherent information over local operation** is hence an entanglement measure. When no ambiguity arises, we also call it the **regularized coherent information**, and write it as $\mathfrak{I}^x(W)$.

Note that $I^x(W)$ without optimizing over local operation is not monotonic under local operation, and thus is not a process entanglement measure. For example, $I^b(\rho^{ab})$ for $\rho^{ab} = (1-p)\omega^{ab} + p\rho_+^{ab}$ is negative for small enough p . As ρ^{ab} can be annihilated by local operations and product states can be created for free using local operations, $I_{LO}^b(\rho^{ab}) \geq I^b(\sigma^a \otimes \tau^b) = 0$ for pure states σ^a and τ^b . Therefore local operation can increase $I^x(W)$.

Even though $I^x(W)$ is not an entanglement measure it still usefully provides a lower bound on $I_{LO}^x(W)$, $\mathfrak{I}_{LO}^x(W)$, and $\mathfrak{I}^x(W)$, which can be hard to compute.

5 Entanglement in spacetime

Traditionally, entanglement is considered as a property of states but not of channels. Both states and channels are special cases of processes, so process entanglement applies to them, and more general quantum networks and processes. The introduction of process entanglement measure generalize the concept of entanglement to a large extent. We devote this section to illustrate this point through examples.

We emphasize that although we defined generalized entanglement in the process framework, the construction applies to theories *without* indefinite causal structure, too. For example, the coherent-information-based measures defined through von Neumann entropy can be used in quantum field

theory with classical spacetime. Given some differentiable manifold where a state, transformation, or quantum network [12] lives, we are now allowed to take entangling surfaces to be any codimension 1 surface that separates the system into two parts – the surfaces are allowed to extend in a combination of spacelike, timelike and null directions. This gives us a meaningful way to talk about entanglement in time and more generally spacetime, rather than just in space,¹⁷ and can be potentially useful in “non-conventional approaches” to quantum theory, quantum field theory and quantum gravity such as boundary approaches (e.g. [36, 37, 38, 39, 40]).

For concreteness we focus on entanglement measured by coherent-information-based measures in this section. For generality we also keep causal structure definite.

5.1 Quantum channel

Consider a quantum channel with two input systems a, b and two output systems c, d . Figure 6 illustrates different ways an “entangling surface” separates the systems into two parts – let one party A have access to the shaded region and another party B have access to the rest region. The coherent-information-based measures can be calculated to yield measures of entanglement with either A ’s or B ’s systems as the target.

Of particular interest are unitary channels, for which we can draw some general conclusions to express the entanglement in terms of the sizes of the systems. The first simplification of unitaries is that its Choi state is pure and its von Neumann entropy vanishes. Consequently,

$$I^A(N_{ab}^{cd}) = S^A - S^{abcd} = S^A = S^B = I^B(N_{ab}^{cd}), \quad (49)$$

as for a pure state ρ^{AB} , $S^A = S^B$. In this case it does not make a difference which party we set as the target – without loss of generality let it be A . The second simplification of unitaries is that

$$\text{Tr}_{ab} \mathbf{N}^{abcd} = \text{Tr}_{cd} \mathbf{N}^{abcd} = \omega, \quad (50)$$

¹⁷There have been works on timelike entanglement for massless fields [34, 35]. These exploit the special property of massless field that the propagator is supported only *on* the lightcone (due to violation of the strong Huygen’s principle, this only holds for a special class of spacetimes), which allows one to quantize timelike separated oscillators analogously to spacelike separated ones. Our study of entanglement in time and spacetime, on the other hand, does not rely on such special properties.

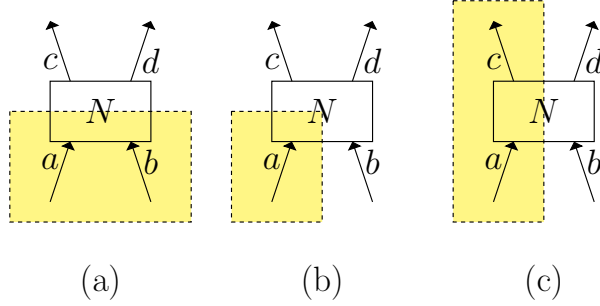


Figure 6: Entanglements of a channel

where ω is the maximally mixed state on \mathcal{H}^{ab} or \mathcal{H}^{cd} (these two spaces have the same dimension because N is unitary). Therefore,

$$S(\mathbf{N}^{ab}) = S(\mathbf{N}^{cd}) = \log |ab| = \log |cd|, \quad (51)$$

where $|x| := \dim \mathcal{H}^x$, and

$$S(\mathbf{N}^x) = \log |x| \quad \text{for } x = a, b, c, d. \quad (52)$$

Taking these into account, we have for case (a) in Figure 6,

$$I^A(N_{ab}^{cd}) = S^{ab} - S^{abcd} = S^{ab} = \log |ab|. \quad (53)$$

For case (b),

$$I^A(N_{ab}^{cd}) = S^a - S^{abcd} = S^a = \log |a|. \quad (54)$$

Case (c) is a bit more complicated.

$$I^A(N_{ab}^{cd}) = S^{ac} - S^{abcd} = S^{ac}. \quad (55)$$

Without further specification the value of S^{ac} is undetermined. For example, if $N_{ab}^{cd} = H_a^c \otimes G_b^d$ factors into two unitaries H and G , then $S^{ac} = 0$. If instead $N_{ab}^{cd} = H_a^d \otimes G_b^c$ factors into two unitaries H and G in another way, then $S^{ac} = S(\omega^{ac}) = \log |ac|$. Nevertheless, if N_{ab}^{cd} is drawn uniformly randomly from all unitaries and the “entangling surface” separate the channel into two systems of equal size, i.e., $|ac| = |bd|$, then one can show that [41]

$$I^A(N_{ab}^{cd}) > \log |ac| - 1. \quad (56)$$

be viewed as a toy model whose continuous limit is a quantum field theory [16, 40]. It is assumed that the network is foliated according to a global time such that each horizontal set of channels act at the same time. We break the network into two parts of the shaded and unshaded regions and ask for the coherent information with subsystems inside the shaded region as the target. The entangling surface is codimension 1 that is not a timelike, spacelike or lightlike extended surface, but a combination of more than one kind.

Since every box is a noiseless unitary, the whole network has a pure Choi state. To calculate the coherent information of the bipartition, we only need to trace out all the open subsystems of the unshaded region and calculate the entropy of the reduced network. According to the global time foliation, information in the output subsystems b_3b_4 of B must flow to the global future in the end and be traced out. Therefore in the reduced network on subsystem b_1b_2 we have the maximally mixed state $\omega^{b_1b_2}$, like case (a) of the last section. Similarly on g_3g_4 we have $\omega^{g_3g_4}$.

Similarly, information in a_3a_4 of A must flow to the global future and be traced out to yield a maximally mixed state. Whether or not information in a_1 originates from the global past, the reduced state on a_2 is the maximally mixed state, like in case (b) of the last section. The same conclusion holds for c_1, f_4 and h_3 .

Information in d_3 must also flow to the global future as there is no way it gets to the shaded party. Similarly information in d_1 must originate in the global past. Hence like in case (c) of the last section, if D is a random unitary and $|d_2d_4| = |d_1d_3|$, the reduced state on d_2d_4 is approximately maximally mixed. The same conclusion holds for e_1e_3 under the same assumptions.

Altogether, we have that the reduced network of the shaded party is a global maximally mixed state, and the coherent information, along with the local-operation-optimized and asymptotically-regularized ones, is approximately maximal. In other words, this bipartition yields maximal entanglement.

6 Entanglement regularization

In this section, we identify causal fluctuation as a possible concrete mechanism that regularizes the otherwise divergent entanglement in quantum field states.

To identify the concrete mechanism that regularizes quantum field entanglement is an important long-standing question [42]. In various profound proposals it is crucial that entanglement be regularized in the ultraviolet (e.g., the proposals that Einstein’s equation can be derived from thermodynamics [28], and that black hole entropy is entanglement entropy [26, 27]). It is widely believed that some quantum gravitational effect does the job, but the exact mechanism remains unclear [43]. One might believe that the mechanism will remain unclear until we find a satisfactory description of the microscopic degrees of freedom of gravity at Planck scale. However, we show below that combining quantum and general relativistic principles already yields the regularization.

The main result is presented in Section 6.2. To build up some intuition of why causal fluctuation regularizes entanglement, we prove the regularization for the restricted case of superposing two causal relations in Section 6.1 first. In this case, the regularization is a consequence of the no-cloning theorem [44, 45]. To prove these results we assume the purification postulate (Section 2.4).

The proofs below deal with coherent-information-based entanglement measures defined in the last section. We expect that the same mechanism of causal fluctuation regularizes a much larger family of entanglement measures. We leave the study of more general proofs to the future, when the general theory of process entanglement measures become more developed and more process entanglement measures with clear operational meaning are obtained.

6.1 Entanglement regularization for $A \not\prec B$

Observe from Figure 3 that when the two causal relations in superposition have equal probability amplitudes (which is what to be expected when causal fluctuation is maximal), the environment can simulate b_1 . If B is in the causal future of A , the environment gets the right half of the information originating from the initial correlation Ψ . If B is causally disconnected with A , the environment gets the information originating from the output of A that B would otherwise receive.

We now translate these observations into mathematical conditions. Let $W_{a_2 b_2}^{a_1 b_1}$ be a process describing A and B with maximal $A \not\prec B$ causal fluctuation.

tuation. The process must obey the following conditions:

$$W_{a_2 b_2}^{a_1 b_1} = W_{a_2}^{a_1 b_1} \otimes W_{b_2}, \quad (57)$$

$$W_{a_2 b_2}^{a_1 b_1 e_1 e_2} = W_{a_2}^{a_1 b_1 e_1} \otimes W_{b_2}^{e_2} \text{ purifies } W_{a_2 b_2}^{a_1 b_1}, \text{ where } W_{b_2}^{e_2} \text{ is a channel,} \quad (58)$$

$$W_{a_2}^{a_1 b_1} = W_{a_2}^{a_1 e_1}. \quad (59)$$

The first condition is enforced by the relation $A \not\prec B$, which requires B 's output b_2 not to be correlated with any of the three other systems. The first condition implies that $W_{a_2}^{a_1 b_1}$ and W_{b_2} can be purified separately. The second condition uses this fact and further asks $W_{b_2}^{e_2}$ to be a channel. This is reasonable as information from B 's output b_2 must transmit completely into the environment. The third condition implements *maximal* causal fluctuation, such that B and the environment share correlation with A in the same way.

Heuristically we can already expect the result of entanglement regularization from no-cloning. As B and the environment share correlation with A in the same way, if A could quantum communicate to B , the same information would be available to the environment. However, this duplicated communication of quantum information would violate the no-cloning theorem, so A 's correlation with B must be damped to kill off the possibility of quantum communication.

Formally, we have:

Theorem 2. For $W_{a_2 b_2}^{a_1 b_1}$ obeying (57) to (59),

$$I_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = 0. \quad (60)$$

It is not hard to see that without local operations $I^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = 0$. The key observation is that for an arbitrary state ρ^{ab} purified by ρ^{abe} , $I^b(\rho^{ab}) = -I^e(\rho^{ae})$ (Theorem A.7). By (57), $I^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = I^{b_1}(W_{a_2}^{a_1 b_1})$, which equals $-I^{e_1}(W_{a_2}^{a_1 e_1})$ by the above observation. On the other hand, by (59) this in turn equals $-I^{b_1}(W_{a_2}^{a_1 b_1})$. Therefore $I^{b_1}(W_{a_2}^{a_1 b_1}) = -I^{b_1}(W_{a_2}^{a_1 b_1})$, which implies $I^{b_1}(W_{a_2}^{a_1 b_1}) = I^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = 0$.

The crucial part of the proof below is that local operation of B can be avoided by exploiting (58). This allows us to finish the proof through a similar reasoning as above.

Proof. Let $M_{a_1}^{a_2 a'}$ and $N_{b_1}^{b_2 b'}$ be arbitrary local operations conducted by A

and B . By (58),

$$\omega^{a'b'} := M_{a_1}^{a_2 a'} N_{b_1}^{b_2 b'} W_{a_2 b_2}^{a_1 b_1} \quad (61)$$

$$= M_{a_1}^{a_2 a'} N_{b_1}^{b_2 b'} W_{a_2}^{a_1 b_1} \otimes \text{Tr}_{e_2} W_{b_2}^{e_2} \quad (62)$$

$$= (\text{Tr}_{e_2} W_{b_2}^{e_2} N_{b_1}^{b_2 b'}) (M_{a_1}^{a_2 a'} W_{a_2}^{a_1 b_1}). \quad (63)$$

This is a channel $T_{b_1}^{b'}$ in the first bracket applied to a state $\omega^{a'b_1}$ in the second bracket. We can then apply the data-processing inequality of coherent information (Theorem A.5) to obtain an upper bound:

$$I^{b'}(\omega^{a'b'}) := I^{b'}(T_{b_1}^{b'} \omega^{a'b_1}) \quad (64)$$

$$\leq I^{b_1}(\omega^{a'b_1}). \quad (65)$$

Let $M_{a_1}^{a_2 a' e_3}$ be an isometric extension of $M_{a_1}^{a_2 a'}$. Since $W_{a_2}^{a_1 b_1 e_1}$ is pure, $\omega^{a'b_1 e_1 e_3} = M_{a_1}^{a_2 a' e_3} W_{a_2}^{a_1 b_1 e_1}$ is a purification of $\omega^{a'b_1} = M_{a_1}^{a_2 a'} W_{a_2}^{a_1 b_1}$. Hence

$$I^{b_1}(\omega^{a'b_1}) = -I^{e_1 e_3}(\omega^{a'e_1 e_3}) \quad (66)$$

$$\leq -I^{e_1}(\omega^{a'e_1}). \quad (67)$$

The first line uses Theorem A.7. The second uses Theorem A.5 where the post-processing channel traces out e_3 . On the other hand, by (59),

$$I^{b_1}(\omega^{a'b_1}) = I^{e_1}(\omega^{a'e_1}). \quad (68)$$

Therefore $I^{b_1}(\omega^{a'b_1}) = I^{e_1}(\omega^{a'e_1}) \leq -I^{e_1}(\omega^{a'e_1})$. It must be that $I^{b_1}(\omega^{a'b_1}) = 0$. By (65), this implies that $I^{b'}(\omega^{a'b'}) \leq 0$. Since we took arbitrary local operations M and N to define $\omega^{a'b'}$, the optimized value $I_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) \leq 0$. We also know that the value is non-negative (Corollary B.1.1), so it is 0. \square

The same proof applies to the regularized coherent information $\mathfrak{J}_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1})$. In this case consider $\tilde{W}_{\tilde{a}_2 \tilde{b}_2}^{\tilde{a}_1 \tilde{b}_1} = (W_{a_2 b_2}^{a_1 b_1})^{\otimes n}$, where the tilde systems are products of the individual systems. One can check that \tilde{W} obeys (57) to (59) for the tilde systems. The same proof works with local operations now applied to the tilde systems. Since n is arbitrary, we have:

Theorem 3. For $W_{a_2 b_2}^{a_1 b_1}$ obeying (57) and (59), the regularized coherent information with supplemental resource of local operation vanishes:

$$\mathfrak{J}_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = 0. \quad (69)$$

Based on the result that maximal causal fluctuation makes the entanglement vanish and the continuity of coherent information (Theorem A.13), we can further argue that non-maximal causal fluctuation reduces entanglement for maximally entangled state.

Consider $W_{a_2}^{a_1 b_1}$ describing a maximally entangled state without causal fluctuation, i.e., $W_{a_2 b_2}^{a_1 b_1} = \rho_+^{a_1 b_1} \otimes \omega^{a_2 b_2}$. Tracing out a_2 and b_2 by local operations, we see that $I_{LO}^b(W^{ab}) = \log \dim \mathcal{H}^{b_1}$, the maximal value any $A \nearrow B$ process can have ($A \nearrow B$ processes are of the form $W_{a_2 b_2}^{a_1 b_1} = W_{a_2}^{a_1 b_1} \otimes W_{b_2}$, so $I^b(W^{ab}) = I^{b_1}(W^{ab_1}) \leq \log \dim \mathcal{H}^{b_1}$).

We turn on causal fluctuation continuously such that in the end there is maximal causal fluctuation. Because coherent information is a continuous function of the process operator (Theorem A.13), its value varies continuously from the maximal $\log \dim \mathcal{H}^{b_1}$ to zero. In the intermediate stages with non-maximal causal fluctuation, the coherent information is reduced from the maximal value.

6.2 Entanglement regularization for three causal relations in superposition

Without causal fluctuation, entanglement diverges because regions closer to the entangling surface contributes significantly more than regions farther away, and as the entangling surface is approached the contribution grows without a bound [46]. The reasoning of the results below is that as the regions approach each other too closely around the Planck scale, causal fluctuation sets in to reduce the entanglement contribution, such that it is no longer unbounded and hence regularizes the divergence.

Technically, we want to show that coherent information is reduced to zero for large causal fluctuation of the form $A \not\bowtie B$. This is the content of Theorem 4 (without optimization over local operation) below. The relevant mathematical conditions the process is expected to obey are written in the statement of the theorem. The first condition is that $W_{a_2 b_2}^{a_1 b_1}$ is symmetric in A and B . This is expected as the strongest causal fluctuation should wash out any initial asymmetry in the causal relation of A and B . The second condition says there is a subsystem e_1 of the environment such that $S(W^{b_1}) \leq S(W^{e_1})$ and $S(W_{b_2}^{a_1}) \geq S(W^{b_1 e_1})$. Loosely speaking, these require that e_1 is a not too small (such that $S(W^{b_1}) \leq S(W^{e_1})$) subsystem that is correlated with b_1 at least as strongly as a_1 is correlated with b_2 (such that $S(W_{b_2}^{a_1}) \geq S(W^{b_1 e_1})$). To interpret the second condition we used the

intuition that the more two systems are correlated the more pure their reduced process is, e.g., a noiseless channel is purer than a noisy channel in terms of process operators.

To understanding why these hold for a process describing large $A \overset{\times}{\bowtie} B$ causal fluctuation, it is helpful to gather some intuition from Figure 4. From Figure 4 we see that a_1 is correlated with b_2 only in the right picture through a channel, whereas b_1 is correlated with the environment through Ψ in both the middle and the right picture. For cases of interest for entanglement regularization, Ψ is usually close to maximally entangled, so it can be expected that the correlation between some environmental subsystem e_1 and b_1 through Ψ is not weaker than that between a_1 and b_2 . It can also be expected that the size of e_1 is at least about the same as b_1 , the one it is close to maximally correlated with. Therefore the second condition is reasonable.

We have:

Theorem 4. Let $W_{a_2 b_2}^{a_1 b_1}$ be a two-party process with purification $W_{a_2 b_2}^{a_1 b_1 e}$. Suppose $W_{a_2 b_2}^{a_1 b_1}$ is symmetric in A and B . Suppose further that e has a subsystem e_1 such that $S^{b_1} \leq S^{e_1}$, $S^{a_1 b_2} \geq S^{b_1 e_1}$. Then

$$I^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) \leq 0. \quad (70)$$

Proof.

$$I^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) = S^b - S^{ab} \quad (71)$$

$$= S^b - (S^{ab_2} + S^{ab} - S^{ab_2}) \quad (72)$$

$$= S^b - (S^{ab_2} + S^e - S^{b_1 e}) \quad (73)$$

$$\{\text{normalization}\} = S^b - (S^{a_1 b_2} + S^{a_2} + S^e - S^{b_1 e}) \quad (74)$$

$$\{S^{a_1 b_2} \geq S^{b_1 e_1}\} \leq S^b - (S^{b_1 e_1} + S^{a_2} + S^e - S^{b_1 e}) \quad (75)$$

$$\{\text{normalization}\} = (S^{b_1} + S^{b_2}) - (S^{b_1 e_1} + S^{a_2} + S^e - S^{b_1 e}) \quad (76)$$

$$\{S^{a_2} = S^{b_2}\} = S^{b_1} - (S^{b_1 e_1} + S^e - S^{b_1 e}) \quad (77)$$

$$\{S^{b_1} \leq S^{e_1}\} \leq S^{e_1} - (S^{b_1 e_1} + S^e - S^{b_1 e}) \quad (78)$$

$$= S^{e_1} - S^{b_1 e_1} - S^{e_1 \bar{e}} + S^{b_1 e_1 \bar{e}} \quad (79)$$

$$\{\text{Strong subadd.}\} \leq 0. \quad (80)$$

The justifications for the non-trivial steps are written in the curly braces. The third line uses the fact that $S(A^a) = S(A^b)$ for pure A^{ab} . The normalization constraints of Section 2.3 imply that $W^{ab_2} = W^{a_2} \otimes W^{a_1 b_2}$ and

$W^b = W^{b_1} \otimes W^{b_2}$, from which the fourth and sixth lines follow. The condition $S^{a_2} = S^{b_2}$ used in the seventh line follows from the assumption that W is symmetric in A and B . The second to last line writes $e = e_1 \bar{e}$ to make it clear how strong-subadditivity (Theorem A.6) is applied in the last line. \square

This theorem makes it clear that causal fluctuation reduces coherent information. It would be better to prove the result for local-operation-optimized coherent information. However, at this stage we can only formulate it as a conjecture.

Conjecture 1. Under the same assumptions, $I_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) \leq 0$.

This conjecture, if true, implies that the regularized coherent information is also regularized, i.e., $\mathfrak{J}_{LO}^{b_1 b_2}(W_{a_2 b_2}^{a_1 b_1}) \leq 0$ under the same conditions. By the continuity of coherent information (Theorem A.13), if we start with a maximally entangled state and continuously turn on causal fluctuation the coherent information would vary continuously from $\log \dim \mathcal{H}^{b_1}$ to zero.

7 Black holes

In this section we study how quantum indefinite spacetime modifies our understanding of black holes. Based on the result that causal fluctuation induces positive communication capacity (Section 3.1), we argue that quantum black holes do not exist in quantum indefinite spacetime and, to the contrary of what results in classical spacetime suggest [47, 48], information will not be lost in the presence of collapsing matter.

More than 20 years ago Don Page already made the following comment in his review article of the information problem [49]:

... I think it more likely that the quantum uncertainty applied to the causal structure of the spacetime makes it impossible to define exactly an absolute horizon. The information might be taken out of the matter near what is interpreted classically as $r = 0$ and yet not be, with 100% quantum probability, within any putative absolute horizon.

There have been recent works with more detailed analysis along similar lines of reasoning to argue that information can escape when the horizon is treated as a quantum system (e.g. [50]).

Our analysis is similar in that we introduce quantum fluctuation of causal structure. The difference is that we do not presume any classical spacetime background, but set the analysis in the context of quantum spacetime. Instead of treating the horizon as a quantum state over a classical spacetime background, we study, in an operational manner, local parties in quantum spacetime. Significant convenience is offered by the quantum information theory flavored process framework such that we can study information communication capacity with ease.

It would be ideal to also calculate the black hole radiation spectrum in quantum indefinite spacetime, but this needs the development of field-theoretic tools which we currently do not have. Nevertheless we have some ideas of how to perform the field-theoretic analysis and argue towards the end of this section that the spectrum will differ from thermal to support our position that information can escape black holes.

7.1 Do black holes exist?

In classical GR black holes are specified mathematically through the notion of causal curve [51]. A future directed causal curve in spacetime (M, g_{ab}) is a differentiable curve whose tangent at each point is future directed timelike or null. A spacetime (M, g_{ab}) ¹⁸ is said to contain a black hole if M is not contained in $J^-(\mathcal{I}^+)$ (the causal past J^- of future null infinity \mathcal{I}^+). The black hole region B of this spacetime is $B := M - J^-(\mathcal{I}^+)$, i.e., the part of the spacetime that is not connected to the future null infinity through causal curves. The notion of causal curve lies at the basis of these definitions as the causal past of a subset of M is defined through connectedness of causal curves. In quantum indefinite spacetime we do not presume the existence of a differentiable manifold. The notion of causal curve loses meaning, and we need to find a new way to define quantum black holes.

We propose to define black hole through information transmission capability. In classical spacetime, two regions are connected by causal curves if and only if information can be transmitted from one region to the other.

¹⁸The spacetime is required to be strongly asymptotically predictable. This requires that (1) the spacetime is asymptotically flat such that after a conformal transformation spatial and null infinities can be represented as the boundaries of the original spacetime in an extended spacetime and (2) in the extended spacetime $(\tilde{M}, \tilde{g}_{ab})$ there is an open region $\tilde{V} \subset \tilde{M}$ with $\tilde{M} \cap J^-(\mathcal{I}^+) \subset \tilde{V}$ such that $(\tilde{V}, \tilde{g}_{ab})$ is globally hyperbolic. The first condition makes it meaningful to talk about future infinity, while the second ensures that $J^-(\mathcal{I}^+)$ is well-behaved.

Quantum indefinite spacetime with the process framework formulation is naturally suited to describe information transmission, as information transmission capacity can be inferred from the processes themselves, even in the absence of any differentiable manifold.

Set a local party F in quantum indefinite spacetime to play the role null future infinity plays in defining black holes in classical spacetime. In classical spacetime, null infinity is defined for asymptotically flat spacetimes, which represent isolated systems [51]. In this idealistic case black holes are isolated from any gravitationally non-trivial object in the rest of the spacetime such that sufficiently away from the black holes spacetime becomes asymptotically flat. Realistically, matter live around black holes if we move away far enough, and the asymptotically flat spacetime is to be viewed as modelling a subregion of the whole spacetime supporting the black hole where the presence of other matter is not felt. Null infinity should be understood to live on the boundary of the asymptotically flat subregion, so in the realistic case it should not really be regarded as living at “infinity”, but rather it represents the border of the black hole region. The function of future null infinity in defining classical black holes is that events causally preceding future null infinity also causally precede some events outside the subregion supporting the black hole. In quantum indefinite spacetime we set local party F accordingly to live at the border of this subregion. It serves the same function such that local parties that can communicate information to F can also communicate information to some local parties outside this subregion.

Without getting into the complications of global hyperbolicity, we loosely define a local party A in quantum indefinite spacetime to be inside a **quantum black hole** if A cannot transmit information to F . The quantum black hole region of a quantum indefinite spacetime consists the union of all local parties inside quantum black holes.

In quantum theory there is some ambiguity in saying that A can transmit information to F , as this can mean at least three different things. The weakest notion of information transmission is signalling, i.e., by choosing local operations A can affect the probability of classical outcomes in F . Stronger notions are to have positive classical and quantum Shannon (defined in the asymptotic limit) communication capacities.¹⁹ Positive quantum capacity

¹⁹Among the three definitions, signalling has the potential advantage of having greater generality to make sense in more theories beyond quantum theory such as general probabilistic theories. The definition through classical capacity makes sense in both quantum theory and classical probabilistic theory, while that through quantum capacity has the

implies positive classical capacity, which in turn implies signalling.

We showed in Section 3.1.3 that due to causal fluctuations, A and F have positive quantum communication capacity. Therefore even under the strongest notion of information transmission, regions whose classical limit falls within the classical black hole region are causally connected to future infinity. Classical black holes do not upgrade into quantum black holes. Rather, quantum indefinite spacetime enables information transfer when we “quantize” a classical black hole.

In the following, we use the term “black holes” in the context of quantum indefinite spacetime to denote regions that correspond to classical black holes in the classical limit when one invokes the correspondence principle. These are only “approximate” black holes because information can transmit out of the region.

7.2 Entropy of black holes

We develop the connection between entanglement entropy and black hole entropy [26, 27] made in the context of quantum field theory in curved spacetime (QFTC) and propose a new version in the context of quantum indefinite spacetime. With indefinite causal structure the entropy of entanglement used in QFTC is no longer a good measure of entanglement. Instead, the measure that connects to the entropy of the black hole should be a process entanglement measure, possibly coherent-information-based. This new understanding supports a black hole thermodynamics argument that information is not lost in the presence of quantum black holes.

In QFTC, the argument that entanglement entropy contributes to black hole entropy proceeds as follows. Suppose one wants to consider entropy for the black hole at a certain time t . It is assumed that the classical spacetime has a spacelike hypersurface Σ_t at time t and the event horizon of the black hole separates Σ_t into two regions. The state $|\Psi\rangle^{\text{in,out}}$ (assumed to be pure) of some quantum field lives on Σ_t as a bipartite state with one party inside and the other outside the horizon. The portion of the state inside the black hole is traced out to obtain ρ^{out} , and its entanglement entropy $S(\rho^{\text{out}}) = -\text{Tr} \rho^{\text{out}} \log \rho^{\text{out}}$ is proposed to contribute to the Bekenstein-Hawking entropy of the black hole. A main support for this proposal is that the entanglement entropy scales is proportional to the area of the horizon, just like Bekenstein-Hawking entropy.

greatest limitation.

We remark that the above proposal assumes a UV cutoff at the Planck scale to get the right order of magnitude for the entanglement entropy. The precise mechanism of the cutoff is unspecified in the original proposals [26, 27]. The result in Section 6 in this thesis is one candidate for the regularization mechanism.

There have been further attempts to argue that entanglement not only contributes to black hole entropy, but the entire black hole entropy is entanglement entropy. This proposal faces the “species problem”, that the total amount of entanglement entropy depend on the number of species of quantum fields, as the black hole entropy does not seem to. One way to address this problem is to show that black hole entropy actually also depends on the number of species by taking into account the effects of the quantum fields on gravity renormalization [52]. It was argued that in canonical quantum gravity matter fields and fluctuations of the gravitational field renormalize Newton’s constant such that the resulting black hole entropy diverges quadratically in the same way as the entanglement entropy [53]. Once the the black hole entropy is written in terms of the renormalized Newton’s constant, it matches the entanglement entropy of all species of matter fields, if gravity is induced by matter fields [54].

The connection between entanglement and black hole entropy and the possibility that black hole entropy *is* entanglement entropy are appealing. We want to extend these ideas from classical to quantum spacetime. The first problem we face is with indefinite causal structure the two parties sharing a process do not live within any spacelike hypersurfaces Σ_t . We need to specify the two parties differently.

In QFTC, the state is shared by the input systems of the two parties, and the output systems are irrelevant. In quantum indefinite spacetime both the input and the output systems are relevant due to causal fluctuations. Let A be the party of the (approximate) black hole (Figure 8 (1)). The output system a_2 can be separated into two subsystems, a'_2 with causal fluctuation towards outside the black hole, and a''_2 without. The process pertaining to a_1, a'_2, a''_2 is of the form

$$W_{a'_2 b_2}^{a_1 b_1} \otimes W_{a''_2}^x, \tag{81}$$

where b_1 and b_2 are the input and output systems of party B outside of the black hole, and x is some system correlated with a''_2 . While we cannot disregard a'_2 since it is coherently correlated with a_1 , we can disregard a''_2 because it is not correlated with a_1 .

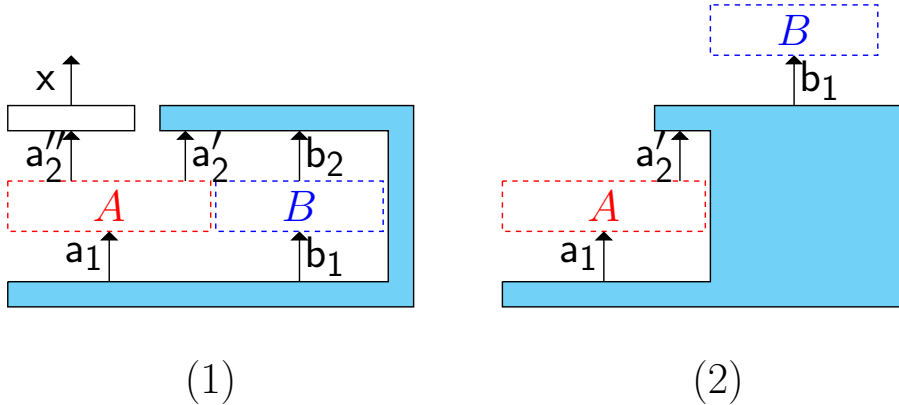


Figure 8: The black hole setting. A is the inside party and B is the outside party. In (1) B is “adjacent” to A , while in (2) B is promoted to future infinity.

In principle the output b_2 of B also separates into two subsystems correlated and not correlated with A , but we keep b_2 integrated for reasons that will be clear below. The system b_1 is clumsy, as it can contain subsystems at widely separated regions. This is because b_1 correlates with both a_1 and a'_2 . The part of b_1 correlated with a'_2 may be distant in time from the part of b_1 correlated with a_1 .

These complications can be avoided if we take party B to be the “future infinity” F . In the old QFTC setting, one can evolve the b_1 part of the bipartite state $\rho^{a_1 b_1}$ unitarily forwards to the future without affecting the entanglement entropy. We want to do something similar in the quantum indefinite spacetime setting. Assume that all the information of b_1 unitarily evolves to future infinity.²⁰ Then we can promote B to future infinity (Figure 8 (2)). Since future infinity is far to the future of A , we can assume that there is no causal fluctuation from the output b_2 towards A . Then we can dispense with b_2 because we are only interested in correlations between A and B . The process simplifies from $W_{a'_2 b_2}^{a_1 b_1}$ to $W_{a'_2}^{a_1 b_1}$, which is of the kind considered in Section 3.1.

Having fixed the two parties and the process $W_{a'_2}^{a_1 b_1}$, we propose that some entanglement measure on $W_{a'_2}^{a_1 b_1}$ such as $I^{b_1}(W_{a'_2}^{a_1 b_1})$ or $\mathfrak{J}^{b_1}(W_{a'_2}^{a_1 b_1})$ gives

²⁰This assumption holds for the case where all information that is outside but correlated with the black hole stays outside. We will argue below that even if such information falls into the black hole, it will eventually come out and reach future infinity.

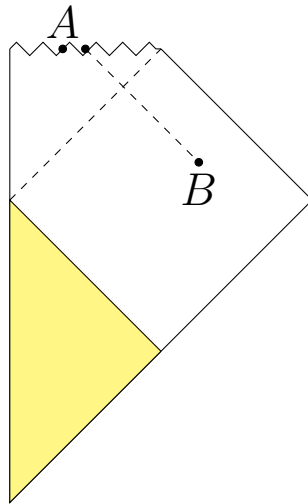


Figure 9: Information transmission from black hole singularity

the entropy of the black hole at the stage specified by A . To probe the entropy of the black hole at different stages during its evolution, we specify A at different regions. When there is no causal fluctuation, $W_{a'_2}^{a_1 b_1}$ reduces to $W^{a_1 b_1}$ and the entropy reduces to $I^{b_1}(W^{a_1 b_1})$, which is the ordinary entanglement entropy when $W^{a_1 b_1}$ is a pure bipartite state.

This interpretation of black hole entropy fits nicely with black hole thermodynamics considerations and supports an argument that information is not lost from the presence of quantum black holes. Thermodynamically, quantum fluctuation can cause the entropy of a system to decrease. In the case of black holes in quantum indefinite spacetime quantum causal fluctuations cause the black hole entropy to decrease (Section 6.1). Such fluctuation induces positive communication capacity from the black hole to the outside (Section 3.1.3), so information accompanies energy to escape the black hole in this way. In the next section we fill in some details of this general argument.

7.3 Information transmission through evaporation²¹

That quantum causal fluctuation allows information even at black hole singularity to travel out can be seen from a simple argument due to Rovelli

²¹Here by evaporation we mean any mechanism that reduces energy within the black hole, not necessarily Hawking radiation.

[55]: The situation is illustrated in Figure 9. Consider any event B outside a Schwartzchild black hole after its formation (B is outside the shaded region). There is a null curve connecting this event to some event at the singularity. Then there must be another event A at the singularity that is below one Planck distance spacelike separated from B . Allowing causal fluctuation at the Planck scale will lead to the possibility that information at A transmits to B . Therefore we expect information to transmit out through causal fluctuation not only from inside the black hole “close” to the event horizon, but also from around the singularity.

The above observation may appear counter-intuitive, because in Figure 9 A and B are not “close” to each other. It is true that they are close according to the intuition of Euclidean distance, but as far as causal fluctuation is concerned Lorentzian distance is the one to measure closeness. Lightlike separation is the critical causal relation that will become spacelike separation or timelike separation upon small perturbations of causal structure. It is then reasonable that two close to lightlike separated events are the easiest for quantum causal fluctuation to put into a superposition of different causal relations. Yet the Lorentzian proper distance is precisely zero for lightlike separated events. Hence the observation in the previous paragraph is in fact very natural if one uses intuitions from the Lorentzian proper distance.

Section 3.1.3 shows that causal fluctuation induces positive quantum communication capacity for the $A \not\prec B$ processes. This means that if A and B conduct some particular operations they can communicate quantum information. In the context of black hole this suggests that if some agent falls into a black hole he or she can act like A to communicate to the “future infinity” party B by conducting the particular operations. However, the black hole information paradox is about whether the spontaneous evaporation of black hole carries information out. Do nature’s local operations transmit information out? We present a toy model based on Section 3.1.2 to study this question.

Suppose the inside party A and the outside party B share a $A \not\prec B$ process $W_{a_2}^{a_1 b_1}$ of the form (35):

$$W_{a_2}^{a_1 b_1} := \text{Tr}_e P(p)_{a'_1 a'_2}^{b_1 e} \rho^{a_1 a'_1} N_{a_2}^{a'_2}, \quad (82)$$

where ρ is the initial state that would be shared by the two parties if no causal fluctuation happens, and N is taken as identity channel for simplicity. p parametrizes the extent of the causal fluctuation. Within the black hole nature performs a unitary evolution $U_{a_1}^{a_2}$ as the local action of A .

In the extreme case where $p = 0$, A and B share $\rho^{a_1 b_1} \otimes \omega^{a_2}$. Information of a_2 travels through the identity channel N directly to the environmental system e and stays inside the black hole, because in this case a_2 actually has no causal fluctuation towards the outside. After nature performs the unitary $U_{a_1}^{a_2}$ within the black hole, the resulting state shared by e inside and b_1 outside is

$$\sigma^{b_1 e} = U_{a_1}^{a_2} N_{a_2}^e \otimes \rho^{a_1 b_1} = U_{a_1}^e \rho^{a_1 b_1}, \quad (83)$$

which is unitarily equivalent to the initial state ρ . In the end, the dimensions of the inside and outside systems stay the same.

In the other extreme case where $p = 1$, A and B share $N_{a_2}^{b_1} \otimes \rho^{a_1}$, where ρ^{a_1} is the reduction of the initial bipartite state to a_1 , with its purifying system e lying entirely outside of the black hole but not accessible to B . There is an identity channel connecting a_2 to b_1 . After nature performs the unitary evolution $U_{a_1}^{a_2}$ within the black hole, outside the black hole b_1 and e share

$$\sigma^{b_1 e} = U_{a_1}^{a_2} N_{a_2}^{b_1} \otimes \rho^{a_1 e} = U_{a_1}^{b_1} \rho^{a_1 e}, \quad (84)$$

which is unitarily equivalent to the initial state. The information of the initial state is completely outside of the black hole and the inside system has dimension zero now. This extreme case corresponds to a complete evaporation of the black hole that carries out all information of the black hole.

In the previous extreme cases, the purifying system e is localized either within or outside the black hole, while the inside system size ends up as either the same as a_1 or as zero. Intermediate cases with $0 < p < 1$ should interpolate between the two extremes such that system e becomes delocalized and the final inside system size varies from that of a_1 to zero as p is varied from 0 to 1. In the previous cases we analyzed the size of the final inside system using the purifying system e . The subtlety in these intermediate cases is that e now contains information of the initial state and of A 's output in a coherent superposition, and one cannot find any subsystem of A that is either localized inside or outside the black hole. Therefore it is not possible to assign any definite system size to the final inside system. What one could do, instead, is to assign an effective size according to the probability amplitudes \sqrt{p} and $\sqrt{1-p}$ of A causally preceding B and A causally disconnected with B . Given a process (82) with some p , assign the

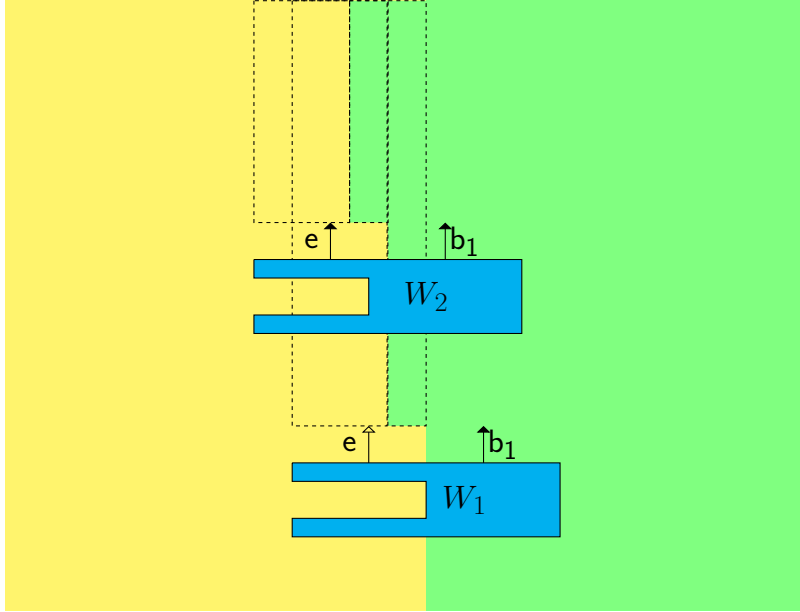


Figure 10: Heuristic illustration of information transmission with black hole evaporation. The width of the yellow region measures the effective dimension of the inside system, while that of the complementary green region measures the the effective dimension of the outside system. Causal fluctuations of the form $W_1, W_2 \dots$ reduces the yellow region.

effective dimension

$$d_{in} = (1 - p)d_{a_1} \quad (85)$$

to the final inside system, where d_{a_1} is the dimension of system a_1 . This effective dimension interpolates between the extreme cases of $p = 0$ with $d_{in} = d_{a_1}$ and $p = 1$ with $d_{in} = 0$. It captures the intuitive expectation that the larger the causal fluctuation is (the larger p is), the smaller the size of the final inside system is.

The analysis of the toy model offered us an effective dimension of the final inside system. A realistic evaporation is an accumulative effect of many radiation events. In the classical spacetime picture the size of the inside system gradually decreases until it reaches zero in the end. In the quantum spacetime picture we face the same problem as in the toy model in assigning a definite system size to some system that involves coherent superposition of inside and outside information. We could similarly use the effective dimension as a temporary tool to gain some understanding of

the general picture before a more suitable model is invented. The general picture we have about black hole evaporation is illustrated in Figure 10. Each radiation event arises from a causal fluctuation event, which is achieved through a unitary taking information from a_2 and the system correlated with a_1 through the initial state, similar to the unitary swap in 82. One subsystem of the output of the unitary belongs to the outside party B , while the other subsystem is the delocalised purifying environmental system. According to the magnitude of the causal fluctuation, an effective dimension is assigned to the inside final system. Each radiation event decreases the effective dimension of the inside system by a positive number, till in the end it reaches zero, which signifies complete evaporation. As the whole process is globally unitary, information is transmitted completely outside of the black hole in the end.

Before closing the study of black hole information, we briefly discuss the problem from the perspective of the radiation spectrum. One argument for information loss through Hawking radiation is that the spectrum is thermal. We offer some speculations on why causal fluctuation should modify the radiation spectrum away from thermality. In a field-theoretic analysis incorporating indefinite causal structure, the two-party process should play the role of the correlation function.²² The entanglement regularization mechanism of Section 6 very likely also regularizes the correlation. It has been claimed using Unruh-DeWitt detector model that ultraviolet regularization of the correlation function can drastically change the spectrum of Unruh effect²³ [57]. This suggests that causal fluctuation modifies the Unruh spectrum and, by the equivalence principle intuition, the Hawking spectrum. This argument cannot be taken seriously until a concrete field-theoretic analysis is conducted, but we find it encouraging that the thermodynamic analysis and the information-theoretic analysis of the previous sections, along with the preliminary spectrum argument here appear to converge to indicate that information is not lost.

²²In Section 9.2.1 we give some more detailed outlook on how to do field-theoretic analysis incorporating indefinite causal structure.

²³We note that this conclusion is sensitive to the particular form of the regularization [56].

8 Cosmology

It has been argued that current observational evidences for dark energy all come from distance measurements with respect to redshift [58]. The supernovae evidence can be viewed as a measurement of their luminosity distance $D_L(z)$ parametrized the redshift z . The baryon acoustic oscillation (BAO) evidence and the cosmic microwave background (CMB) anisotropy evidences can be viewed as effects on the angular diameter distance $D_A(z)$, which relates to the luminosity distance by

$$D_A = D_L/(1+z)^2. \quad (86)$$

One can challenge these as evidences for a form of energy to be present in our universe, because some travelling-distance-modifying mechanism can possibly give rise the same results, even without adding extra energy density to the universe.

In this section we propose such a mechanism. The key observation is that the entanglement regularization from causal fluctuation in Section 6 is stochastic, and entanglement can vary from region to region (to be explained below). In emergent spacetime pictures the amount of entanglement can affect distance. With a varying regularization distance can fluctuates from region to region, which in turn modifies the total distance a photon travels.

The modification of travelling distance is of the kind that accounts for the qualitative features of the above mentioned observational “evidences” for dark energy. However, at this stage the idea is still highly speculative, since we have not performed any quantitative analysis, which we hope to conduct in future works. Our goal in the following is simply to convey the idea at a qualitative level.

8.1 Postulates

We start by postulating that the length scale in emergent spacetime is given by entanglement:

- The spacetime area of an entangling surface in vacuum is proportional to the entanglement in that area.
- The entanglement fluctuates among different regions of spacetime.

To be general, we formulated the postulates independently of the process framework and causal fluctuations, as these postulates may hold for emergent spacetime proposals motivated by other considerations.

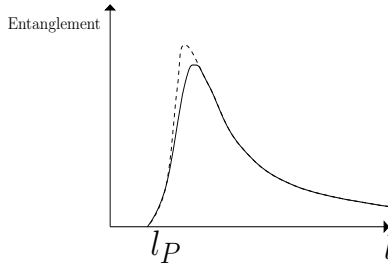


Figure 11: Schematic drawing of entanglement contribution from two parties as their separation l is varied. The solid and dashed lines represent two possible ways entanglement is regularized as l decreases to below the Planck scale l_P . The total entanglement is a sum of contributions at different l , so in this case the entanglement in the dashed line case is slightly larger.

Of course, we are led to these ideas by studying quantum indefinite spacetime. The motivation for the first postulate is to have a description of spacetime that is independent of the differentiable manifold structure of classical spacetime, which makes it hard to apply superposition principle to gravity. The two pieces of information a classical spacetime contains are the causal structure and the conformal factor [1, 2]. The processes account for quantum causal structure, and we still need to recover the conformal factor of scale. The first postulate proposes that the area scale emerges from entanglement. This is inspired by the area scaling of vacuum entanglement in quantum field theory [59].

The motivation for the second postulate is the regularization scheme of Section 6. As a quantum phenomenon, the causal fluctuation around the Planck scale does not take uniform form everywhere. There is a general trend that causal fluctuation gets larger as two parties get closer, but the detailed way processes approach the maximal causal fluctuation one differ from region to region. Such slight difference can result in different values of entanglement (Figure 11). This is the content of the second postulate.

In the next section we extract consequences of the postulates independently of the process framework. Since the postulates themselves are also formulated beyond the process framework, the upcoming proposal is general. To see that the postulates can be motivated differently, we list some examples. Related to the first postulate, Cao *et al.* made an interesting proposal to recover spacetime length scale from entanglement without presuming any classical spacetime background or the holography [30]. There the observation is that the strength of correlation between two parties A

and B measured by $I^{A:B}$ can characterize their “distance”, and so can be used to define a metric. The first postulate here differs in that the parties are required to share an entangling surface, and what we recover is a scale on the entangling surface, but not between the two parties. Related to the second postulate, there are independent works that study fluctuating UV cutoff on quantum fields [60, 61, 62]. Because of the connection between UV regularization of entanglement and quantum field, similar fluctuating cutoff on entanglement can be considered also from the perspective of these works.

8.2 Modified travelling distance

We want to analyze an expanding universe without a cosmological constant and show that the postulates in the previous section allow a photon to travel a larger comoving distance in the same amount of time. To this end we compare the cases with and without entanglement fluctuations.

Suppose a photon travels from comoving coordinate 0 at time t_0 to comoving coordinate n at t_n . The time the photon reaches each integer comoving coordinate i is denoted by t_i . We ask how these times differ in the cases with and without entanglement fluctuation.

Consider the case without entanglement fluctuation. Break each region between two integer comoving coordinates i and $i + 1$ into a certain (large) number of unit squares²⁴ as shown in Figure 12. We ask that all squares, taken as regions on some entangling surface, have the same amount of entanglement, say \bar{S} . By the first postulate, this gives a length (physical distance) of $l = \alpha\sqrt{\bar{S}}$ for the sides of the squares, which the photon passes. The universe can expand or contract such that the physical distance between adjacent integer comoving coordinates increases or decreases. Therefore the the number of squares between comoving coordinates i and $i + 1$ is a function of time. We denote it by $N(t_i)$.

Now we turn on entanglement fluctuation. The amount of entanglement in each square fluctuates, though the average value is still \bar{S} . Between comoving coordinates i and $i + 1$, the photon travels physical distance

$$L_i := \sum_{j=1}^{N(t_i)} \alpha \sqrt{S_j}, \quad (87)$$

²⁴We remark that there is no reliance on choosing *square-shaped* unit regions in the argument. The important point is that the entanglement scales quadratically in length.

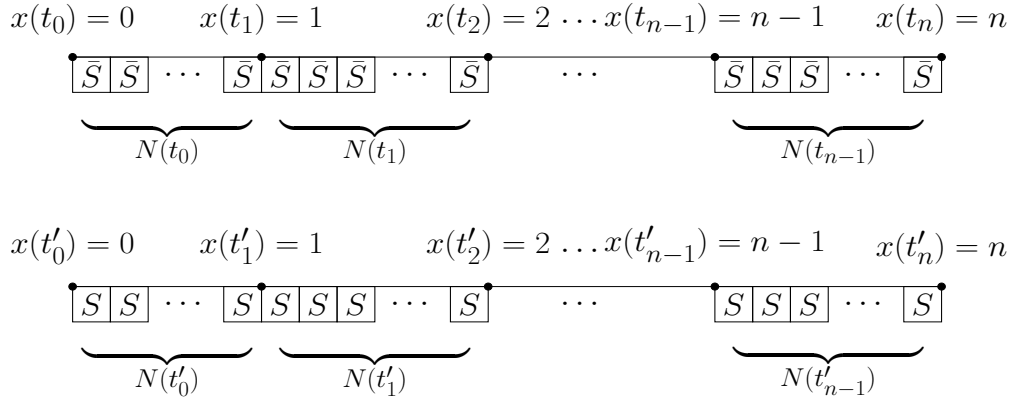


Figure 12: Up: without entanglement fluctuation. Down: with entanglement fluctuation. x is the comoving coordinate.

where S_j is the entanglement in the j -th square between comoving coordinates i and $i + 1$. When k is large enough, $\sum_{j=1}^k S_j \rightarrow k\bar{S}$. Note that the argument of N is t'_i with a prime, because the time the photon reaches comoving coordinate i in this case differs from the case without fluctuation. Generically, fluctuation in entanglement decreases the physical distance, as for samplings such as Gaussian,

$$\sum_{j=1}^N \alpha \sqrt{S_j} < \sum_{j=1}^N \alpha \sqrt{\bar{S}}. \quad (88)$$

Now we can analyze the impact of entanglement fluctuation on the travelling distance in universes with specific evolution patterns. To start with, consider the simple case of a universe that is neither expanding nor shrinking. Then N is a constant function in time. By the analysis above, between each pair of comoving coordinates, the photon travels a shorter physical distance in the fluctuation case. When N is taken to be large enough, this physical distance tends to a constant times the physical distance the photon would travel in the case without fluctuation. Entanglement fluctuation shrinks the physical distance a photon travels between any large enough comoving interval by a constant factor. Because the physical speed of a photon is fixed, fluctuation allows it to travel more comoving distance within the same amount of time.

Now consider an expanding universe. Start with $i = 0$. By (87), the photon travels a shorter physical distance to reach comoving coordinate 1

in the fluctuation case, which implies that $t'_1 < t_1$, or $\Delta t'_0 := t'_1 - t'_0 = \beta_0 \Delta t_0 := t_1 - t_0$, where $\beta_0 < 1$. This in turn implies that $N(t'_1) < N(t_1)$ as the universe is expanding. Therefore from comoving coordinate 1 to 2 the photon has to travel across fewer squares in the fluctuation case. Combined with (87), this implies that $\beta_1 < 1$ where $\beta_1, \Delta t'_1$, and Δt_1 are defined similarly. By the same reasonings, $\beta_i < 1$ for all i . Similar to the above case, in the end the photon reaches comoving coordinate n earlier, i.e., $t'_n < t_n$. If the photon in the fluctuating case is allowed to travel until time t_n , it will reach a comoving coordinate greater than n . The difference from the above case is that due to the expansion of the universe, β_i is not a constant, and how much further the photon travels within in the fluctuating case depends on the expansion history of the universe.

For supernovae, CMB, BAO observations the relevant quantity is the comoving distance $\chi(a)$ the photon travels from emission at scale factor a or redshift $z := (1 - a)/a$ to the present, as the luminosity distance and the angular diameter distance are expressed in terms of $\chi(a)$. In both cases with and without fluctuation, the photon starts at the same time $t(a)$ at scale factor a and ends at the same time $t(a_0)$ at scale factor a_0 . By the arguments above, entanglement fluctuation makes $\chi(a)$ greater. This is the same qualitative feature that dark energy or cosmological constant have on an otherwise matter dominated universe. In this sense, the two postulates may account for the observational evidences for “dark energy” without having dark energy.

8.3 Questions

We conclude the discussion on cosmology with some questions. Clearly a quantitative analysis needs to be done to turn the above proposal into a concrete one. We should also try to address the question of why the universe appears to be close to flat. In the concordance model flatness can be matched by a cosmological constant. In our proposal with modified travelling distance sharing or completely taking over cosmological constant’s (or dark energy’s) duty, it is an open question whether the energy content of the universe is still close to critical.

The first postulate states how scale emerges for the vacuum, but not excited states. The high temperature early universe is dominated by matter and radiation, so what influence will the second postulate has on the scale then? Can the postulates explain some phenomena in the early universe one would otherwise resort to dark matter or inflation to explain? More

generally, can one find some falsifiable predictions from the postulates?

9 Conclusions

9.1 Summary

In this section we summarize the thesis with an emphasis on specifying the assumptions used in different steps.

The whole work is based on the prospect of combining the principle of superposition from quantum theory and the principle of dynamical causal structure from general relativity. We assume that these principles hold also in quantum gravity.

The process framework is based on the three basic assumptions of local quantum mechanics, linearity of probability, and non-negative normalized probability, whose meanings are explained in Section 2.1. These assumptions allow us to incorporate indefinite causal structure through processes, which are represented as operators on the systems of local parties. The three basic assumptions allow a large family of not necessarily physical processes in the mathematical framework. At present it is an open question exactly what processes are physical. An optional postulate to have in the process framework is the purification postulate, which excludes non-purifiable processes as unphysical (Section 2.4). We assume this postulate only when we specify so.

We presented examples of all possible kinds of superposition of causal relation for two parties in Section 3. We refer the readers to [10, 20] for ways to certify indefinite causal structure and superposition of causal relations from the description of the processes.

We then move on in Section 4 to propose that all entanglement measures for processes obey the “monotonicity under local operations” axiom. Here local operations are trace-preserving completely positive maps (channels). In particular, they must be trace-preserving. This axiom contrasts the “monotonicity under local operations and classical communications” axiom of entanglement theory for states. The reason is that in the process framework where causal structure is important classical communication cannot be a free resource. We emphasize that the generalized notion of entanglement apply also to ordinary quantum networks without indefinite causal structure, and gives meaning to timelike and spacetime entanglement. Coherent-information-based measures have special importance as

process entanglement measures, because its fundamental status in quantum communication theory and the clear operational meanings. In Section 5 we illustrated through examples how to calculate timelike and spacetime entanglement in quantum networks which can be used to model quantum field theory in the continuous limit.

Starting with Section 6 we discuss spontaneous quantum gravitational causal fluctuations. Such fluctuations are expected to take place generically around the Planck scale. Using processes to model such fluctuations, we implicitly assumed that local parties are still meaningful at this energy scale. By manipulating entropic inequalities, we identify causal fluctuation as a possible mechanism that regularizes entanglement. The proofs assume the purification postulate.

In Section 7 we studied the impact of causal fluctuations on black holes physics. Finding that the traditional definition of black hole through causal curves in classical spacetime does not make sense in quantum spacetime, we generalized the concept of black holes to quantum spacetime using information communication capacity, assuming that there is a meaningful notion of local party at future infinity in the quantum spacetime setting. Using results from Section 3, we argue that under the generalized definition quantum black holes do not exist in quantum indefinite spacetime because of the generic causal fluctuations. This observation leads us to study the information loss paradox, offering clues from black hole thermodynamics, information processing, and radiation spectrum to argue that information is not lost through evaporation.

Finally, in Section 8 we studied the impact of causal fluctuation on cosmology based on two speculative postulates that the area scale in the vacuum derives from the amount of entanglement and the amount of entanglement quantum fluctuates. One consequence of these postulates is that the photon travelling comoving distance becomes longer. This may explain the observations related with dark energy or cosmological constant.

9.2 Outlook

9.2.1 Incorporating indefinite causal structure to quantum field theory

There are several motivations for combining the process framework with quantum field theory. First, we want to calculate the black hole radiation spectrum taking causal fluctuation into account. Second, we want to check

if the causal fluctuation regularization mechanism for entanglement also applies to quantum field regularization.²⁵ Third, we want to explore the impact of indefinite causal structure on the early universe by doing concrete calculations using field-theoretic analysis. Fourth, we want to study observable effects of indefinite causal structure in general. For this purpose concrete calculations need to be done, and it is hopeful that these can be done in connection with quantum field theory.

Our outlook is to upgrade the framework for quantum field theory in curved spacetime of Hollands and Wald [63] to incorporate processes. Hollands and Wald’s framework can be read as an improvement of algebraic quantum field theory (AQFT) [64, 65] to face the challenges such as non-existence of a unique vacuum state posed by curved spacetime background. This framework brings the already similar AQFT and process framework even closer. In the process framework the processes can be viewed as linear functionals over local operations, just as in AQFT states are linear functionals over smeared quantum fields, which represent local operations. In the process framework, processes carry essential information of indefinite causal structure of the theory which is not carried by the local operations. This contrasts AQFT, where the quantum fields representing local operations carry much of the structure of the theory, while states merely inherit the structure from the fields. In Hollands and Wald’s framework the essential structures such as constraints of the equation of motion are carried by conditions of the operator product expansion to be imposed on states and states’ interaction with fields. The same strategy can be applied to processes such that field theoretic structures are imposed as extra conditions on the processes and processes’ interaction with local operations.

A major feature to be expected in the improved field theory is the regularization of two-party correlators from causal fluctuations. Two-party processes, as linear functionals on local operations, will likely be used to express or bound two-point correlators²⁶, which are linear functionals on smeared fields representing local operations in AQFT and Hollands and Wald’s framework. The causal fluctuation regularization mechanism for entanglement very likely also regularizes the correlator in the new field the-

²⁵Indeed, it had been argued very early on that entanglement and gravity are regularized by the same mechanism [53].

²⁶Indeed we already know that mutual information $I^{a:b}(\rho^{ab}) = I^b(\rho^{ab}) + S(\rho^a)$ bounds the correlation function: $I^{a:b} \geq \frac{C(O_a, O_b)^2}{2\|O_a\|^2\|O_b\|^2}$, where $C(O_a, O_b) := \langle O_a \otimes O_b \rangle - \langle O_a \rangle \langle O_b \rangle$ for observables O_a and O_b [66].

ory, and many applications can be expected from this observation.

9.2.2 Developing a theory of quantum gravity

Currently we view quantum indefinite spacetime as a model or setting where one can address certain questions of quantum gravity. Is it possible to develop it into a theory of quantum gravity? A first step can be to set the emergence of spacetime from process operators on a firmer footing. Namely, to study in detail how the causal structure and the conformal factor can be inferred purely from the quantum description of processes.

Let us discuss causal structure first. Of course, it is possible to infer *operational* causal structure between local parties from their process. The question is whether there is an one-to-one correspondence between a process description of the universe and the *spacetime* causal structure of the universe (see the paragraph right before Section 3.1 for a discussion about the difference between operational and spacetime causal structures), which is what we want for a theory of quantum gravity. A potentially useful feature about gravity is that gravitational waves are hard to trap, so if local parties are allowed to detect small gravitational influences it is possible that operational and spacetime causal structures are equivalent. This possible equivalence needs to be examined carefully, and to exploit it in a quantum gravity theory one may need to impose extra conditions. For example, processes must describe correlations to a great level of details such that correlations as weak as mediated by gravitational waves should be included.

As for the conformal factor, we made a preliminary proposal of how the area scale in vacuum emerges from entanglement in Section 8.1. To develop it one needs to study how exactly the conformal factor can be recovered from the area scale, and how the proposal extends beyond vacuum. A goal of this theoretical development is to make the theory of quantum black holes more tenable. Currently we have a working hypothesis that a local party is inside a black hole if it cannot communicate to the future infinity party. It would be helpful to develop concepts such as future infinity party and black hole region within a mathematically more solid framework to see the impacts of quantum indefinite spacetime on singularity related topics, which are of interest to quantum gravity.

How to obtain the classical limit of general relativity from the quantum theory? In the quantum theory we allocate a local party at each “event” that is operationally meaningful. In the classical limit these events

are “small” (in terms of the space-time extension of individual events) and “dense” (in terms of the distribution of all the events) such that they can be modelled as points with a differentiable manifold. Furthermore, in the classical limit all indefinite causal structure effect disappears. Classical causal and conformal structures emerge so that a metric can be recovered on the differentiable manifold²⁷. The dynamical Einstein’s equation can then be derived à la Jacobson [28] from the now uniform scaling of entanglement (without indefinite causal structure there is no more fluctuations in entanglement, though the finiteness of field entanglement would have to come from a cutoff whose origin is now unspecified).

9.2.3 Generalizing communication theory

For obvious reasons we want to turn the conjecture of Section 6 into a theorem. The question of calculating optimized coherent information can also be formulated in the broader context of a generalized communication theory.

Ordinary quantum communication theory assumes that each party has access to either an input or an output system. On the other hand, in the process framework each party possesses two systems in order to allow superposition of causal structure. The communication theory of processes will generalize traditional communication theory to give parties access to two systems instead of one. The process communication will subsume quantum network, channel and entanglement communication theories as special cases. From a even broader perspective, other information processing tasks such as computation may be generalized to incorporate parties with both input and output systems.

The process communication theory, in particular, will have a richer structure and provide some new questions for information scientists. Tasks such as classical communication, private classical communication, and quantum communication need to be redefined and capacity theorems need to be reproved. The unification of bipartite states and channels through processes poses some new challenges. For example, for channels the capacities

²⁷The approach to recover classical spacetime from causal structure and sprinkling density is pursued in the causal set theory (see [67] and references therein). Causal set theory postulates that spacetime is fundamentally discrete, which distinguishes it from our approach as we do not make any assumption about the fundamental discreteness of spacetime. Moreover, our starting point is a quantum theory of spacetime, whereas the original causal set theory is a classical theory.

of entanglement transmission and entanglement generation are equal and is given by \mathfrak{J}_{LO}^b . However, this equivalence breaks for processes. Suppose the process is an entangled state, it certainly can have a positive entanglement generation capacity, but “no-signalling” dictates that entanglement transmission capacity is zero. Unlike for channels, the quantity \mathfrak{J}_{LO}^b no longer gives the capacity for entanglement transmission (although we think it gives an upper-bound). We will need a new expression for process entanglement transmission capacity.

On the other hand, \mathfrak{J}_{LO}^b probably still gives the entanglement generation capacity, and coherent information will in any case still play an important role for the communication theory of processes. Therefore computing or bounding coherent information is a crucial task for both high-energy physics (for ultraviolet regularization) and information science (for communication capacities).

Appendix

A Standard theorems from quantum information theory

In this section we review some standard²⁸ results from quantum information theory. References to the original literature where the results are obtained and some omitted proofs can be found in [21].

A.1 Quantum relative entropy

Quantum relative entropy is a central entropic measure, because many other measures can be written in terms of it, and many of their properties can be derived from properties of quantum relative entropy.

For a density operator ρ and a positive semi-definite operator σ , the quantum relative entropy $S(\rho||\sigma)$ is defined as

$$S(\rho||\sigma) := \begin{cases} \text{Tr}[\rho(\log \rho - \log \sigma)], & \text{supp}(\rho) \subset \text{supp}(\sigma), \\ +\infty, & \text{otherwise.} \end{cases} \quad (89)$$

Importantly, mutual information and coherent information/conditional entropy can be expressed in terms of relative entropy.

$$S(\rho^{ab}||\rho^a \otimes \rho^b) = I^{a:b}(\rho^{ab}) \quad (90)$$

$$S(\rho^{ab}||\mathbb{1}^a \otimes \rho^b) = I^b(\rho^{ab}) = -S^{a|b}(\rho^{ab}). \quad (91)$$

Note that $\mathbb{1}^a \otimes \rho^b$, having trace greater than 1 for nontrivial system a , is not a quantum state. Therefore $S(\rho^{ab}||\mathbb{1}^a \otimes \rho^b)$ does not meet the condition of Theorem A.2 below and can be negative.

We list two fundamental properties of quantum relative entropy, whose proofs can be found in [21].

²⁸Theorem A.11 is perhaps not standard as the first formula in the statement of the theorem is not usually seen in the literature. However, the proof is a straightforward application of standard results.

Theorem A.1 (Monotonicity of quantum relative entropy). For state ρ^a , positive semi-definite operator σ^a and channel N_a^b ,

$$S(\rho^a || \sigma^a) \geq S(N_a^b \rho^a || N_a^b \sigma^a). \quad (92)$$

Here the channel acts on the possibly unnormalized operator σ^a as a linear map.

Theorem A.2 (Non-negativity of quantum relative entropy). For density operator ρ and positive semi-definite σ such that $\text{Tr } \sigma \leq 1$, the quantum relative entropy $S(\rho || \sigma) \geq 0$, with equality if and only if $\rho = \sigma$.

A.2 Entropic inequalities

Theorem A.3 (Subadditivity). For bipartite state ρ^{ab} ,

$$S(\rho^{ab}) \leq S(\rho^a) + S(\rho^b), \quad (93)$$

with equal sign iff $\rho^{ab} = \rho^a \otimes \rho^b$.

Proof. The inequality is equivalent to the non-negativity of mutual information, which is a consequence of mutual information as a relative entropy of density operators (equation (90)) and the non-negativity of relative entropy (Theorem A.2). Equality holds when the relative entropy is zero, which is true iff $\rho^{ab} = \rho^a \otimes \rho^b$ in this case. \square

Theorem A.4 (Triangle inequality). For bipartite state ρ^{ab} ,

$$S(\rho^{ab}) \geq |S(\rho^a) - S(\rho^b)|.$$

Proof. Let ρ^{abe} be a purification of ρ^{ab} . Then

$$S(\rho^{be}) \leq S(\rho^b) + S(\rho^e) \quad (94)$$

$$= S(\rho^b) + S(\rho^{ab}) \quad (95)$$

where we used subadditivity (Theorem A.3) in the first line. Since $S(\rho^{be}) = S(\rho^a)$, we have $S(\rho^a) - S(\rho^b) \leq S(\rho^{ab})$. Similarly $S(\rho^b) - S(\rho^a) \leq S(\rho^{ab})$. The result follows. \square

Theorem A.5 (Data processing inequality for coherent information). For bipartite state ρ^{ab} and channel N_b^c ,

$$I^b(\rho^{ab}) \geq I^c(N_b^c \rho^{ab}).$$

Proof.

$$I^b(\rho^{ab}) = S(\rho^{ab} || \mathbb{1}^a \otimes \rho^b), \quad (96)$$

$$I^c(N_b^c \rho^{ab}) = S(N_b^c \rho^{ab} || N_b^c \mathbb{1}^a \otimes \rho^b). \quad (97)$$

The statement is then a consequence of the monotonicity of quantum relative entropy (Theorem A.1). \square

Theorem A.6 (Strong subadditivity). For tripartite state ρ^{abc} ,

$$S(\rho^{ab}) + S(\rho^{bc}) \geq S(\rho^{abc}) + S(\rho^b).$$

Proof. The inequality is equivalent to

$$I^{ab}(\rho^{abc}) \geq I^b(\rho^{bc}), \quad (98)$$

which is a consequence of Theorem A.5 when the channel N_{ab}^b traces out system a such that $N_{ab}^b \rho^{abc} = \rho^{bc}$. \square

A.3 State and channel coherent information

Theorem A.7. For bipartite state ρ^{ab} purified by ρ^{abe} ,

$$I^b(\rho^{ab}) = -I^e(\rho^{ae}). \quad (99)$$

When ρ^{ab} is pure, e is empty and the right hand side reduces to ρ^a .

Proof.

$$I^b(\rho^{ab}) = S^b - S^{ab} \quad (100)$$

$$= S^b - S^e \quad (101)$$

$$= S^{ae} - S^e \quad (102)$$

$$= -I^e(\rho^{ae}). \quad (103)$$

We used the shorthand notation that S^x denotes $S(\rho^x)$. The second and third lines use the fact that for pure ρ^{xy} , $S(\rho^x) = S(\rho^y)$. \square

Two useful lemmas below lead to the convexity of state coherent information (Theorem A.10).

Lemma A.8. Let $\rho^{ax} = \sum_i p_i \rho_i^a \otimes |i\rangle\langle i|^x$, where $p_i \geq 0$, $\sum_i p_i = 1$, and ρ_i^a are normalized states on a . Then

$$S(\rho^{ax}) = S(\rho^x) + \sum_i p_i S(\rho_i^a). \quad (104)$$

This can be proved by direct calculation from the definitions.

Lemma A.9. Let $\rho^{abx} = \sum_i p_i \rho_i^{ab} \otimes |i\rangle\langle i|^x$, where $p_i \geq 0$, $\sum_i p_i = 1$, and ρ_i^{ab} are normalized states on ab . Then

$$I^{bx}(\rho^{abx}) = \sum_i p_i I^b(\rho_i^{ab}). \quad (105)$$

Proof.

$$I^{bx}(\rho^{abx}) = S^{bx} - S^{abx} \quad (106)$$

$$= S^x + \sum_i p_i S(\rho_i^b) - (S^x + \sum_i p_i S(\rho_i^{ab})) \quad (107)$$

$$= \sum_i p_i (S(\rho_i^b) - S(\rho_i^{ab})) \quad (108)$$

$$= \sum_i p_i I^b(\rho_i^{ab}). \quad (109)$$

In the second line we used Lemma A.8. \square

Theorem A.10 (Convexity of coherent information). Let $\rho^{ab} = \sum_i p_i \rho_i^{ab}$, where $p_i \geq 0$, $\sum_i p_i = 1$, and ρ_i^{ab} are normalized states on ab . Then

$$\sum_i p_i I^b(\rho_i^{ab}) \geq I^b(\rho^{ab}). \quad (110)$$

Proof. By Lemma A.9, the LHS equals $I^{bx}(\rho^{abx})$, where $\rho^{abx} = \sum_i p_i \rho_i^{ab} \otimes |i\rangle\langle i|^x$. Since $\text{Tr}_x \rho^{abx} = \rho^{ab}$, the result follows from Theorem A.5. \square

An important consequence of the convexity of state coherent information is that mixed states do not improve over pure states for input optimization in the definition of channel coherent information. We can define channel coherent information in any of the following equivalent ways.

Theorem A.11 (Equivalent definitions of channel coherent information).
The following definitions of channel coherent information are equivalent:

$$I^b(N_a^b) := \sup_{\rho^{aa'}} I^b(N_a^b \rho^{aa'}), \quad (111)$$

$$I^b(N_a^b) := \sup_{|\psi\rangle^{aa'}} I^b(N_a^b |\psi\rangle\langle\psi|^{aa'}), \quad (112)$$

$$I^b(N_a^b) := \sup_{\rho^a} S(N_a^b \rho^a) - S(N_a^e \rho^a), \quad (113)$$

where ρ and $|\psi\rangle$ are arbitrary general and pure states, respectively. N_a^{be} is an isometric extension of N_a^b and $N_a^e = \text{Tr}_b \circ N_a^{be}$ is the complementary channel to N_a^b .

Proof. First we show that the second and third definitions are equivalent. Let ρ^a be a state and let $|\psi\rangle^{aa'}$ be its purification. $N_a^{be} |\psi\rangle\langle\psi|^{aa'}$ is a pure state denoted by $\psi^{a'be} = |\psi\rangle\langle\psi|^{a'be}$. We have

$$S(N_a^b \rho^a) - S(N_a^e \rho^a) = S(\psi^b) - S(\psi^e) \quad (114)$$

$$= S(\psi^b) - S(\psi^{a'b}) \quad (115)$$

$$= I^b(\psi^{a'b}) \quad (116)$$

$$= I^b(N_a^b |\psi\rangle\langle\psi|^{aa'}). \quad (117)$$

Because for any arbitrary ρ^a there is a corresponding $|\psi\rangle^{aa'}$, and conversely any $|\psi\rangle^{aa'}$ gives rise to a state ρ^a ,

$$\sup_{\rho^a} S(N_a^b \rho^a) - S(N_a^e \rho^a) = \sup_{|\psi\rangle^{aa'}} I^b(N_a^b |\psi\rangle\langle\psi|^{aa'}). \quad (118)$$

Next we show that the first and second definitions are equivalent. Obviously the first quantity is no less than the second. We use convexity of state coherent information to establish that the second quantity is no less than the first. By Theorem A.10, for any state $\sigma^{a'b}$,

$$\sum_i p_i I^b(\sigma_i^{a'b}) \geq I^b(\sigma^{a'b}), \quad (119)$$

where $\sigma^{a'b} = \sum_i p_i \sigma_i^{a'b}$ is decomposition of $\sigma^{a'b}$ into a probabilistic mixture. Then there is one particular j such that $I^b(\sigma_j^{a'b}) \geq I^b(\sigma^{a'b})$. Suppose $\sigma^{ab} =$

$N_a^b \rho^{aa'}$, where $\rho^{aa'}$ is an arbitrary input state with spectral decomposition $\rho^{aa'} = \sum_i p_i \rho_i^{aa'}$ into pure states $\rho_i^{aa'}$ such that $\sigma_i^{a'b} = N_a^b \rho_i^{aa'}$. Then

$$I^b(N_a^b \rho_j^{aa'}) \geq I^b(N_a^b \rho^{aa'}). \quad (120)$$

□

Another consequence of the above results is that supplemental forward classical communication does not increase channel coherent information.

Theorem A.12 (Forward classical communication does not increase channel coherent information). For a channel N_a^b ,

$$I_{LO \rightarrow}^b(N_a^b) = I_{LO}^b(N_a^b) = I^b(N_a^b), \quad (121)$$

where \rightarrow is classical communication from the input party to the output party.

Proof. With $LO \rightarrow$, Alice prepares a state $\rho^{acx} = \sum_i p_i \rho_i^{ac} \otimes |i\rangle\langle i|^x$, where $p_i \geq 0$, $\sum_i p_i = 1$, ρ_i^{ac} are normalized states on ac , and c is the subsystem Alice keeps. Bob gets the bx part of

$$\sigma^{bcx} = N_a^b \rho^{acx} \quad (122)$$

$$= \sum_i p_i N_a^b \rho_i^{ac} \otimes |i\rangle\langle i|^x \quad (123)$$

$$:= \sum_i p_i \omega_i^{bc} \otimes |i\rangle\langle i|^x \quad (124)$$

where the x subsystem arrives directly as classical communication. Bob applies a local operation $L_{bx}^{b'}$ on bx to finalize the shared state. Consider a set of operations that maximize the coherent information. We have

$$I_{LO \rightarrow}^b(N_a^b) = I^{b'}(L_{bx}^{b'} \sigma^{bcx}) \quad (125)$$

$$\leq I^{bx}(\sigma^{bcx}) \quad (126)$$

$$= \sum_i p_i I^b(\omega_i^{bc}) \quad (127)$$

$$\leq I_{LO}^b(N_a^b). \quad (128)$$

The first line holds because we chose a set of operations that maximize the coherent information. The second line uses Theorem A.5. The third line uses Lemma A.9. The fourth holds because $I_{LO}^b(N_a^b)$ maximizes over all input states. Finally, we showed in Section 4.2 that $I_{LO}^b(N_a^b) = I^b(N_a^b)$. □

A.4 Continuity of coherent information

Theorem A.13 (Alicki-Fannes-Winter inequality). For density matrices ρ^{ab} and σ^{ab} , if $\frac{1}{2}\|\rho^{ab} - \sigma^{ab}\|_1 \leq \epsilon$ for $\epsilon \in [0, 1]$. Then

$$|H^{a|b}(\rho) - H^{a|b}(\sigma)| = |I^b(\rho) - I^b(\sigma)| \leq 2\epsilon \log \dim \mathcal{H}^a + (1 + \epsilon)h_2(\epsilon/[1 + \epsilon]), \quad (129)$$

where $h_2(p) := -p \log p - (1 - p) \log(1 - p)$ is the binary entropy.

A proof of this theorem can be found in [21].

B Results about process coherent information

The coherent information for a state can be negative, but that of a channel cannot. This is a consequence of the following more general result.

Theorem B.1. If R contains local operations on all systems complementary to x , $I_R^x(W) \geq 0$.

Proof. Pick O_R to trace out all output subsystems complementary to x and to input product pure states to all input subsystems complementary to x . Then $I^x(O_R(W)) = 0$, from which $I_R^x(W) \geq 0$. \square

Corollary B.1.1. If R allows local operation on all systems, $I_{LO}^x(W) \geq 0$.

Theorem B.2. Local output operations cannot increase the coherent information of a process, if operations on non-target systems are restricted to *unital* channels.

Proof. Suppose we want to optimize $I^b(W^{ab}) = S(W^{ab}||\mathbb{1}^a \otimes W^b)$ by applying local channels on output subsystems of b and local unital channels on output subsystems of a .

The local output operations can be expressed as a joint channel $N_{ab}^{a'b'} := N_a^{a'} \otimes N_b^{b'}$, where $N_a^{a'}$ is a product of local unital channels, and $N_b^{b'}$ is a product of local channels (we adjoin the identity channel for subsystems

not acted on). We have

$$I^{b'}(N_{ab}^{a'b'} W^{ab}) = S(N_{ab}^{a'b'} W^{ab} \| \mathbb{1}^{a'} \otimes N_b^{b'} W^b) \quad (130)$$

$$= S(N_{ab}^{a'b'} W^{ab} \| N_a^{a'} \mathbb{1}^a \otimes N_b^{b'} W^b) \quad (131)$$

$$= S(N_{ab}^{a'b'} W^{ab} \| N_{ab}^{a'b'} (\mathbb{1}^a \otimes W^b)) \quad (132)$$

$$\leq S(W^{ab} \| \mathbb{1}^a \otimes W^b) \quad (133)$$

$$= I^b(W^{ab}). \quad (134)$$

The first and last lines use the relative entropy expression of coherent information. The second line uses the unital property of $N_a^{a'}$. The fourth line uses the monotonicity property of quantum relative entropy (Theorem A.1). \square

References

- [1] S W Hawking, A R King, and P J McCarthy. A new topology for curved space–time which incorporates the causal, differential, and conformal structures. *Journal of Mathematical Physics*, 17(2):174–181, 1976.
- [2] David B Malament. The class of continuous timelike curves determines the topology of spacetime. *Journal of Mathematical Physics*, 18(7):1399–1404, 1977.
- [3] S W Hawking and G F R Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, 1973.
- [4] Lucien Hardy. Probability theories with dynamic causal structure: a new framework for quantum gravity. *arXiv:gr-qc/0509120*.
- [5] Lucien Hardy. Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure. *Journal of Physics A: Mathematical and Theoretical*, 40(12):3081, 2007.
- [6] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092, 2012.
- [7] Achim Kempf. Information-theoretic natural ultraviolet cutoff for spacetime. *Physical Review Letters*, 103(23):231301, 2009.

- [8] Achim Kempf. Spacetime could be simultaneously continuous and discrete, in the same way that information can be. *New Journal of Physics*, 12(11):115001, 2010.
- [9] Jason Pye, William Donnelly, and Achim Kempf. Locality and entanglement in bandlimited quantum field theory. *Physical Review D*, 92(10):105022, 2015.
- [10] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. Witnessing causal nonseparability. *New Journal of Physics*, 17(10):102001, 2015.
- [11] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and Its Applications*, 10(3):285–290, 1975.
- [12] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009.
- [13] Lucien Hardy. Reformulating and reconstructing quantum theory. *arXiv:1104.2066*.
- [14] Lucien Hardy. The operator tensor formulation of quantum theory. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 370(1971):3385–3417, 2012.
- [15] Bob Coecke. Quantum picturalism. *Contemporary Physics*, 51(1):59–83, 2010.
- [16] Ognjan Oreshkov and Nicolas J Cerf. Operational quantum theory without predefined time. *New Journal of Physics*, 18(7):073037, 2016.
- [17] Mateus Araújo, Adrien Feix, Miguel Navascués, and Časlav Brukner. A purification postulate for quantum mechanics with indefinite causal order. *arXiv:1611.08535*.
- [18] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. Probabilistic theories with purification. *Physical Review A*, 81(6):062348, 2010.

- [19] Jean-Philippe W MacLean, Katja Ried, Robert W Spekkens, and Kevin J Resch. Quantum-coherent mixtures of causal relations. *arXiv:1606.04523*.
- [20] Adrien Feix and Āaslav Brukner. Quantum superpositions of “common-cause” and “direct-cause” causal structures. *arXiv:1606.09241*.
- [21] Mark M Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017.
- [22] Giulio Chiribella, Giacomo Mauro D’Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2):022318, 2013.
- [23] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Physical Review A*, 86(4):040301, 2012.
- [24] Mateus Araújo, Fabio Costa, and Āaslav Brukner. Computational advantage from quantum-controlled ordering of gates. *Physical Review Letters*, 113(25):250402, 2014.
- [25] Adrien Feix, Mateus Araújo, and Āaslav Brukner. Quantum superposition of the order of parties as a communication resource. *Physical Review A*, 92(5):052326, 2015.
- [26] Rafael D Sorkin. On the entropy of the vacuum outside a horizon. In *General Relativity and Gravitation*, pages 734–736. Springer, 1984.
- [27] Luca Bombelli, Rabinder K Koul, Joochan Lee, and Rafael D Sorkin. Quantum source of entropy for black holes. *Physical Review D*, 34(2):373, 1986.
- [28] Ted Jacobson. Thermodynamics of spacetime: The Einstein equation of state. *Physical Review Letters*, 75(7):1260, 1995.
- [29] Goffredo Chirco and Stefano Liberati. Nonequilibrium thermodynamics of spacetime: The role of gravitational dissipation. *Physical Review D*, 81(2):024016, 2010.

- [30] Chunjun Cao, Sean M Carroll, and Spyridon Michalakis. Space from hilbert space: Recovering geometry from bulk entanglement. *Physical Review D*, 95(2):024031, 2017.
- [31] Sonia Markes and Lucien Hardy. Entropy for theories with indefinite causal structure. *Journal of Physics: Conference Series*, 306(1):012043, 2011.
- [32] J Oppenheim and B Reznik. Probabilistic and information-theoretic interpretation of quantum evolutions. *Physical Review A*, 70(2):022312, 2004.
- [33] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865, 2009.
- [34] S Jay Olson and Timothy C Ralph. Entanglement between the future and the past in the quantum vacuum. *Physical Review Letters*, 106(11):110404, 2011.
- [35] S Jay Olson and Timothy C Ralph. Extraction of timelike entanglement from the quantum vacuum. *Physical Review A*, 85(1):012306, 2012.
- [36] Robert Oeckl. A “general boundary” formulation for quantum mechanics and quantum gravity. *Physics Letters B*, 575(3):318–324, 2003.
- [37] Carlo Rovelli. *Quantum Gravity*. Cambridge University Press, 2004.
- [38] Robert Oeckl. General boundary quantum field theory: Foundations and probability interpretation. *Advances in Theoretical and Mathematical Physics*, 12(2):319–352, 2008.
- [39] Robert Oeckl. A positive formalism for quantum theory in the general boundary formulation. *Foundations of Physics*, 43(10):1206–1232, 2013.
- [40] Lucien Hardy. Operational general relativity: Possibilistic, probabilistic, and quantum. *arXiv:1608.06940*.
- [41] Pavan Hosur, Xiao-Liang Qi, Daniel A Roberts, and Beni Yoshida. Chaos in quantum channels. *Journal of High Energy Physics*, 2016(2):1–49, 2016.

- [42] Eugenio Bianchi and Robert C Myers. On the architecture of spacetime geometry. *Classical and Quantum Gravity*, 31(21):214002, 2014.
- [43] Ted Jacobson. Gravitation and vacuum entanglement entropy. *International Journal of Modern Physics D*, 21(11):1242006, 2012.
- [44] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.
- [45] Dennis Dieks. Communication by EPR devices. *Physics Letters A*, 92(6):271–272, 1982.
- [46] Saurya Das, S Shankaranarayanan, and Sourav Sur. Black hole entropy from entanglement: A review. *Horizons in World Physics*, 26:8, 2008.
- [47] Stephen W Hawking. Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199–220, 1975.
- [48] Stephen W Hawking. Breakdown of predictability in gravitational collapse. *Physical Review D*, 14(10):2460, 1976.
- [49] Don N Page. Black hole information. In *Proceedings of the 5th Canadian Conference on General Relativity and Relativistic Astrophysics*, page 1, 1994.
- [50] Ram Brustein. Origin of the blackhole information paradox. *Fortschritte der Physik*, 62(3):255–265, 2014.
- [51] Robert M Wald. *General Relativity*. University of Chicago press, 1984.
- [52] Ted Jacobson. Black hole entropy and induced gravity. *arXiv:gr-qc/9404039*.
- [53] Leonard Susskind and John Uglum. Black hole entropy in canonical quantum gravity and superstring theory. *Physical Review D*, 50(4):2700, 1994.
- [54] Andrei D Sakharov. Vacuum quantum fluctuations in curved space and the theory of gravitation. In *Soviet Physics Doklady*, volume 12, page 1040, 1968.
- [55] Gerard 't Hooft, Steven B Giddings, Carlo Rovelli, Piero Nicolini, Jonas Mureika, Matthias Kaminski, and Marcus Bleicher. The good, the bad, and the ugly of gravity and information. *arXiv:1609.01725*.

- [56] Iván Agulló, José Navarro-Salas, Gonzalo J Olmo, and Leonard Parker. Two-point functions with an invariant planck scale and thermal effects. *Physical Review D*, 77(12):124032, 2008.
- [57] Piero Nicolini and Massimiliano Rinaldi. A minimal length versus the unruh effect. *Physics Letters B*, 695(1):303–306, 2011.
- [58] Ruth Durrer. What do we really know about dark energy? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1957):5102–5114, 2011.
- [59] Jens Eisert, Marcus Cramer, and Martin B Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277, 2010.
- [60] Robert Brout. The inflaton and its mass. *arXiv:gr-qc/0201060*.
- [61] Robert Brout. From inflation to dark energy. *arXiv:gr-qc/0508019*.
- [62] David Campo, Jens Niemeyer, and Renaud Parentani. Damped corrections to inflationary spectra from a fluctuating cutoff. *Physical Review D*, 76(2):023513, 2007.
- [63] Stefan Hollands and Robert M Wald. Quantum fields in curved space-time. *Physics Reports*, 574:1–35, 2015.
- [64] Rudolf Haag and Daniel Kastler. An algebraic approach to quantum field theory. *Journal of Mathematical Physics*, 5(7):848–861, 1964.
- [65] Rudolf Haag. *Local Quantum Physics: Fields, Particles, Algebras*. Springer, 2nd edition, 1996.
- [66] Michael M Wolf, Frank Verstraete, Matthew B Hastings, and J Ignacio Cirac. Area laws in quantum systems: mutual information and correlations. *Physical Review Letters*, 100(7):070502, 2008.
- [67] Rafael D Sorkin. Causal sets: Discrete gravity. In *Lectures on Quantum Gravity*, pages 305–327. Springer, 2005.