An Operational Road towards Understanding Causal Indefiniteness within Post-Quantum Theories

by

Nitica Sakharwade

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Časlav Brukner
Professor, Faculty of Physics, University of Vienna

Supervisor(s): Lucien Hardy
Faculty, Perimeter Institute for Theoretical Physics
Adjunct Associate Professor, Dept. of Physics and Astronomy
University of Waterloo

Achim Kempf
Professor, Dept. of Applied Mathematics
University of Waterloo

Internal Member: Kevin Resch
Professor, Dept. of Physics and Astronomy
University of Waterloo

Internal-External Member: Doreen Fraser
Associate Professor, Dept. of Philosophy
University of Waterloo

Other Member: Eduardo Martin-Martinez
Associate Professor, Dept. of Applied Mathematics
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contribution

I would like to detail the research contributions presented in this thesis.

Chapter 2 is a continuation from the my Perimeter Scholar International (PSI) Essay (part of my Master's degree with U Waterloo with PSI certification from Perimeter Institute). The idea of the project was born from a discussion with Ognyan Oreshkov and Lucien Hardy. The progress on the research came from the many discussions with Lucien Hardy and also benefited from discussions with Achim Kempf. I took the lead on this project and the Chapter is written solely by me. It contains unpublished material, presented at various conferences including Quantum Networks Oxford 2017, ICQF India 2017, and YQIS Vienna 2018.

Chapter 3 is work I co-authored with Ding Jia published in PRA [66, 67]. The project was proposed by Ding Jia and research was done together. The Chapter has been adapted from the paper to best fit the thesis.

Work from Chapters 4, 5 and 6 belong to a research program, contain unpublished work presented at various conferences (QISS HK 2020, Q-turn 2020 [90], APS 2021 [91], Quantizing Time 2021 [89], QPL 2021) as the work developed.

Chapter 4 reviews the Causaloid Framework based on earlier papers by Lucien Hardy [41][42]. We present new diagrammatics largely worked on by myself and the next two Chapters build upon this. The Chapter is written by me, and are based on mentioned sections the papers by Lucien Hardy mentioned.

Chapter 5 The project was initiated by me and the problem statement of characterising Meta Compression was presented by Lucien Hardy. The hierarchy came from the many discussions with Lucien Hardy and the results are my contribution.

Chapter 6 is related to Chapters 4, 5 focusing on the bridge between the Causaloid Framework and Duotensor framework by Lucien Hardy [46]. The project and research was done together with Lucien Hardy. The Chapter is written solely by me.

All figures (created using TikZ) as well as illustrations presented in this thesis were solely created by me, barring Figure 3.2 in Chapter 3 which is taken from [66, 67].

Note that a statement of contribution per Chapter is also provided at the end of each Chapter 2 to 6.

Abstract

A theory, whatever it does, must correlate data. We commit ourselves to operational methodology, as a means towards studying the space of Generalised Probability Theories that are compatible with Indefinite Causality. Such a space of theories that combines the radical aspects of Quantum Theory: its probabilistic nature and of Relativity: its dynamic causal structure, is expected to house Quantum Gravity. In this thesis, we ask how we may understand *Indefinite Causality*.

In the first half, we explore the consequences of (Indefinite) Causality on Communication tasks. Motivated by the game with one way signalling that provided the Causal Inequality [82], we study competing two-way signalling and provide protocols for Bidirectional Teleportation and Bidirectional Dense-Coding. Further, we provide a theorem for when tensor products of processes are valid. This result has consequences for setting up of a theory of process communication.

In the second half, we revisit the Causaloid Framework by Hardy [41, 42], a framework that studies this space of theories and prescribes how to recover the correlations within a theory from operational data to calculate probabilities through three levels of physical compression – Tomographic, Compositional and Meta. We present a diagrammatic representation for the Causaloid Framework and leverage it to study Meta-Compression through which we characterise a Hierarchy of theories. The rungs of this Hierarchy are differentiated by the nature of the Causal Structure of the theory. We apply the Causaloid Framework to the space of Generalised Probability Theories pertaining to circuits, through the Duotensor Framework [49, 46] and show that finite dimensional Quantum Theory as well as Classical Probability Theory belong to its second rung.

To summarise, we work towards a better understanding of communication tasks when the underlying causal structure is indefinite, and characterise the space of Generalised Probability Theories with Indefinite Causality, through the nature of their causal structures.

Acknowledgements

I am deeply grateful to Lucien Hardy for his mentorship through the years that encouraged me to think independently, come up with my own problems and tackle them, which was challenging but also helped me become the researcher I am today. Thank you for the many discussions, conversations, anecdotes, presence and support through which not only did I learn a lot about academia and research but also got to grow as a person and these teaching will continue to guide me through my life.

I am grateful to Achim Kempf, for his guidance, support and hospitality for having me in the Physics of Information Lab, the many discussions we had would leave me inspired. I would like to thank Doreen Fraser, for training me in Philosophy (of Physics), for her mentorship, constant support and for inspiring to be articulate around technical nuances. I would like to thank Ding Jia for his warmth, collaboration and ideas, thank you for inspiring me to ask hard fundamental questions and to learn to embrace scientific critique.

I am grateful to my PhD Advisory Committee and to my PhD Defense Committee – Časlav Brukner, Doreen Fraser, Eduardo Martin-Martinez, Kevin Resch, Achim Kempf and Lucien Hardy – for their timely guidance through my time as a PhD student and well as the very interesting questions that led to an enjoyable and illuminating discussion during my PhD Defense.

The scientific atmosphere at Perimeter Institute, IQC and U Waterloo, encouraged the birth of new ideas and were central to the work of this thesis. While it is difficult to exhaustively capture the people and discussions that in some way informed my work I would like to extend my gratitude to a multitude of groups:

The Foundations of QFT reading group that I co-hosted with Doreen Fraser was vital to expanding my understanding of subtleties around infinities in quantum theory and I thank all those who came and contributed their thoughts, including – Maria Papageorgiou, Tomas Gonda, Ding Jia, Adrian Franco-Rubio, Ravi Kunjwal, Angella Yamamoto, Rafael Sorkin.

I would like to thank the Quantum Foundations group at Perimeter for the many seminars, group meetings (which were my favourite), discussions in the Bistro and companionship, including – Rob Spekkens, Markus Muller, Ana Belen Sainz, Elie Wolfe, Ravi Kunjwal, John Selby, Andrei Shieber, Tobias Fritz, Jamie Sikora, Alvaro Alhambra, Denis Rosset, Flaminia Giacomini, Tom Galley, Tomas Gonda, David Schmid, TC Fraser.

I would like to thank the members of the Physics of Information Group – Aida Ahmadzadegan-Shapiro (for hosting the ML reading group), Maria Papageorgiou, Jose de Ramon Rivera, Marco Letizia, Jason Pye, Nayeli Rodriguez-Briones, Julia Amoros Binefa, Allison Sachs.

I am grateful that I was able to travel and discuss Physics. I would like to thank all the researchers I met at conferences, schools and workshops with whom I had stimulating discussions that left a strong impression on me. A non-exhaustive list in no particular order includes – Rob Spekkens, Matt Leifer, Rafael Sorkin, Justin Dressel, Ognyan Oreshkov, Caslav Brukner, Paulo Perinotti, Giulio Chiribella, Lidia del Rio, Bob Coecke, Aleks Kissinger, Philipp Hohn, Sandu Popescu, Artur Ekert, Carlo Rovelli, Sumati Surya, Ian Durham, Nuriya Nurgalieva, V Vilasini, Patricia Contreras Tejada, Hippolyte Dourdent, Alastair Abbott, Debashis Saha, Tian Zhang, Marco Tulio Quintino, Nicholas Gisin, Fabio Costa, Eduardo Martin-Martinez, Anne-Catherine de la Hamette, Alok Kumar Pan, Ralph Silva, Tony Shorts, Markus Huber, Juani Bermejo Vega, Wayne Myrvold, Laura Ruetsche, Some Sankar Bhattacharya, Manik Banik. I am sure some names evade me at the moment, please forgive me if I have missed your name.

I am grateful to my academic mentors from my undergraduate years, in particular – Shasanka M. Roy, Sarah Croke, Thomas Busch, Albert Benseny Cases and Saikat Ghosh, for giving me opportunities to learn and research and for their encouragement and faith. I am grateful to my PSI academic mentors – Dan, Gang, Aggie, David, Tibra, Denis. I am also grateful to Pawel Horodecki for his patience as I finished my thesis.

The journey of a PhD student, requires a strong support system and I am grateful to my fellow PhD students, colleagues and friends for their support. Thank you Christophe, Florian, Tomas, Andrei, Fiona, Ravi, Bel, Rath, Jaini, Leilee, Alvaro, Maria, Qingwen, Aditya and Navya for listening to and sharing in this journey. Thank you Angela, Melissa, Sayali, Sonali, Sayak, Jayesh, Venkat and Amrita for always having my back. Thanks to the Waterloo Community – Janice, Bashar, Eve, KWPS, Emily, Faryal, Alysha, Hannah for letting me belong. I would like to thank my Yoga teachers Leilee Chojnacki, Desiree Dawson, Alysha Brilla and Adrienne for their teaching that kept me sane while writing my thesis during the ongoing pandemic. Thank you Supranta, Soham, Adrian, Namrata, Ashley, Tomas for being my chosen family. And finally, thank you Adrian for your companionship, being my home and office mate for four years, for entertaining my late night dumb mathematics questions, Waterloo would be colder without you.

I would not be able to be where I am without the unconditional love and support of my family. I thank my sister Sanica for her companionship, my father Chandrakant for encouraging me to dream big, and my mother Susmita for teaching me to be fearless. I thank the Chowdhury family for their love and support. Lastly, I am deeply grateful to my partner Shreyan for the many zooms over which you supported my daily thesis writing, for believing in me and for always being there. A special shout-out to my father and Shreyan for their intense and thorough proof-reading of my thesis!

Dedication

Aai	*Baba*
Three decades ago	All my childhood
you wrote	you told me
your thesis	to dream
as I grew	as big as
in your womb	the sky

This is for you

Table of Contents

List of Figures **xiii**

List of Tables **xv**

1 This Thesis in a Nutshell **1**

1.1 Operational Methodology 2

1.2 Generalised Probability Theories 5

1.3 Indefinite Causal Structure 6

1.3.1 Frameworks for Quantum Causality 7

1.3.2 Communication in Theories with Indefinite Causality 9

1.4 An Outline of this Thesis 11

2 Bidirectional Quantum Teleportation and Dense-Coding **14**

2.1 From Quanta-net to Post-Quantum theories 15

2.2 Problem Statement and Setup 16

2.3 Diagrammatics and Notation 19

2.3.1 Diagrammatic conventions 19

2.3.2 Notation for Resources 20

2.3.3 Summary of Protocols 20

2.4 Teleportation and Dense-Coding 21

2.4.1 Quantum Teleportation 21

2.4.2 Quantum Dense-Coding . . . 23
2.4.3 Resource Theory of Unit Quantum Protocols . . . 24
2.5 Butterfly Networks . . . 26
2.5.1 Classical Butterfly Network . . . 26
2.5.2 Quantum Butterfly Network . . . 28
2.6 Bidirectional Quantum Teleportation . . . 31
2.6.1 Bidirectional Quantum Teleportation (Version I) . . . 31
2.6.2 Bidirectional Quantum Teleportation (Version II) . . . 33
2.7 Bidirectional Quantum Dense-Coding . . . 37
2.7.1 Bidirectional Quantum Dense-Coding (Version I) . . . 37
2.7.2 Bidirectional Quantum Dense-Coding (Version II) . . . 39
2.7.3 Masked Encoding and Security . . . 42
2.7.4 Bidirectional Quantum Dense-Coding (Version III) . . . 43
2.8 Discussion . . . 45
2.8.1 Duality within Bidirectional Teleportation and Dense-Coding . . . 45
2.8.2 Quanta-Net and the Boundary Rule . . . 47
2.8.3 Some final thoughts . . . 49

3 Tensor Products of Process Matrices **51**
3.1 Towards a Theory of 'Process Communication'.. . . . 52
3.2 Processes . . . 53
3.3 Conditions for forming Process Products . . . 55
3.3.1 Restriction on Process Product: a Theorem . . . 58
3.4 Conclusions . . . 61

4 Revisiting the Causaloid Framework **64**
4.1 Why the Causaloid Approach? . . . 66
4.1.1 Towards Quantum Gravity... . . . 66

4.1.2 Letting go of states evolving in time 66
4.1.3 Operational Methodology, Data and Probabilities 68
4.1.4 Diagrammatics: Old Framework, New Clothes 69
4.2 Causaloid Framework 70
4.2.1 Thinking inside the box 70
4.2.2 Organising Stacks of Cards 71
4.2.3 Operational Regions 72
4.2.4 Statement of objective 73
4.3 Three levels of Physical Compression of Data 74
4.3.1 Zeroth level: Pre-Compression 75
4.3.2 First level: Tomographic Compression 77
4.3.3 Second level: Compositional Compression 81
4.3.4 Third level: Meta Compression 88
4.3.5 Synopsis: Three Levels of Physical Compression 93

5 A Hierarchy through Meta-Compression 97

5.1 Defining the Hierarchy 98
5.2 Proof of concept *a la Hardy* 100
5.3 Toolkit for Working with Omega Sets 101
5.3.1 Projections 102
5.3.2 Filling 102
5.3.3 Permutations and Order 103
5.3.4 Visualising Omega Sets 103
5.4 $\Lambda_{\mathcal{O}_1}$-Sufficiency 105
5.5 $\Lambda_{\mathcal{O}_2}$-Sufficiency 105
5.5.1 Tripartite Region: Causal Relations Results 108
5.5.2 Going beyond Tripartite Regions: $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$, $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$. 112
5.6 $\Lambda_{\mathcal{O}_d}$-Sufficiency 113
5.7 The Road Ahead... 114

6 The Hierarchy's Second Rung: Duotensors meet the Causaloid 118
6.1 Applying the Causaloid Framework 119
6.1.1 Formalism-local and Operational Formulations 120
6.1.2 Outline of the Chapter 121
6.2 Duotensor Framework 121
6.2.1 Operational descriptions 122
6.2.2 Probabilities 125
6.2.3 Duotensors 127
6.3 Causaloid meets Duotensors 134
6.3.1 Operational Data 135
6.3.2 Probabilities 136
6.3.3 Level Zero: Pre-Compression 137
6.3.4 Level One: Tomographic Compression 140
6.3.5 Level Two: Compositional Compression 143
6.3.6 Level Three: Meta Compression 150
6.4 Populating Higher Rungs 158
6.4.1 Hyper$_d$wired Circuits 158
6.4.2 Beyond the Hierarchy 161

7 Conclusion 163

References 164

List of Figures

1.1 A Visual Outline of this Thesis 12

2.1 Diagrammatic conventions for parties, channels and entanglement 19
2.2 Quantum Teleportation 22
2.3 Quantum Dense-Coding 23
2.4 Classical Butterfly Network 27
2.5 Trivial protocol for BQT (left) and BQDC (right) 28
2.6 Quantum Butterfly Network 29
2.7 Bidirectional Quantum Teleportation (Version I) 32
2.8 Bidirectional Quantum Teleportation (Version II) 34
2.9 Bidirectional Quantum Dense-Coding (Version I) 37
2.10 Bidirectional Quantum Dense-Coding (Version II) 40
2.11 Bidirectional Quantum Dense-Coding (Version III) 43
2.12 Coherent Classical Butterfly Network 46
2.13 An example for a Quantum Network 47
2.14 Spatial Perspective: BQDC II (left) and BQT II (right) 48

3.1 Bipartite Process 53
3.2 Process Product for two Bipartite Processes 56

4.1 Operational Region R_1 72

5.1 Visual form of $\Omega_{1,2} \subseteq \Omega_{1,\not{2}} \times \Omega_{\not{1},2} \subseteq \Omega_1 \times \Omega_2$ 104

5.2 The space of Generised Probability Theories with respect to the two Hierarchies 115

List of Tables

2.1 Summary of Communication Protocols 20

2.2 Takeaway for Bidirectional Quantum Teleportation Protocols 36

2.3 Takeaway for Bidirectional Quantum Dense-Coding Protocols 45

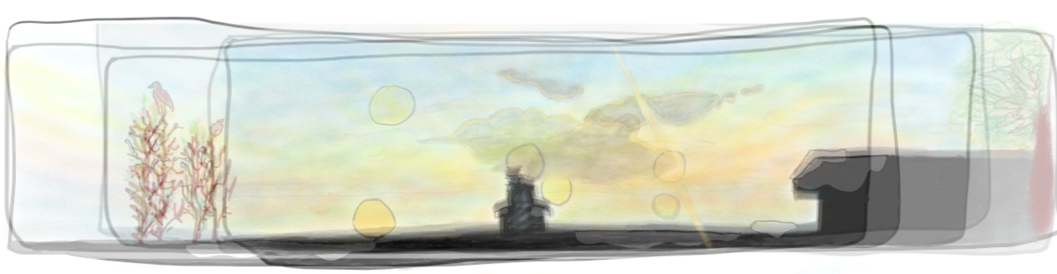
3.1 Types of Process Terms 55

5.1 Form for $\Lambda_{\mathcal{O}_2}$ given Causal Relations between R_1, R_2, R_3 108

5.2 $\Omega_{1,2,3}$ and $\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ given Causal Relations between R_1, R_2, R_3 . . 111

I measured
the passage of time
looking at sunsets
through my window...

when the world in effect
was uncertain time wrapped
in quantum clothes
waiting to be observed



Chapter 1

This Thesis in a Nutshell

in Ojibwe languages there are no words for the past or the future,
there are three words for now— already, at this time, then now;
today, nowadays now; and after a while, eventually, finally just now

—Paraphrasing of an Elder Teaching[1],[2]

Causality permeates our life deeply; in ways we experience the world and the passage of time, in ways we perceive and make inferences, in ways languages are structured and cultures think and in the way you are reading this sentence. And humans have long pondered, theorised and studied causation and time as philosophers in Metaphysics, and more recently (in time-scales of human history) as physicists. Till before the 20$^{\text{th}}$ century the notion of linear or Newtonian time was prevalent until Relativity radically shifted how we understand time and causality in two ways. Firstly, we had to reject notions of space and time as independent entities as they were in Newtonian space and time and accept dynamic causality. Secondly, we adopted a more operational approach to describing physical theories, we gave up the notion of simultaneity for two space-like separated events; since the temporal order became dependent on the observer's reference-frame and for events that cannot signal to each other the temporal order becomes operationally meaningless.

In this thesis, we take inspiration from the space-time revolution and take the two ideas a quantum further... Quantum Theory is formulated on background Newtonian space-time,

[1] from the poem, *Black Snake* by Ashley Hynd (*Contemporary Verse 2, Vol. 42 No. 02*)

[2] This thesis in part was written on the traditional territory of the Attawandaron (Neutral), Anishinaabe and Haudenosaunee peoples, situated on the Haldimand Tract, the land promised to the Six Nations; and in part in the author's homeland, India.

within which temporal order of events are definite. In the fundamental physical theory of Quantum Gravity we expect for space-time itself to admit a quantum nature allowing for indefinite temporal order of events. Towards constructing such a theory, we consider *causally neutral approaches* where space and time are treated on an equal footing *a priori*. Further, we adopt operational methodology, that allows us to avoid making assumptions about the structure of space and time, and helps us find a mathematical framework that can incorporate *Indefinite Causality*.

We are interested in the broad road-map towards Quantum Gravity and approach this through operational frameworks that may accommodate theories with indefinite causal structures. The standpoint of Quantum Foundations will heavily inform the approach we undertake, the operational methodology we employ and the tools we utilise. This will help us motivate the field of *Indefinite Causal Structures* that has gained momentum in the last decade, which will be discussed in Section 1.3 in greater detail.

In this introductory chapter, we explain what we mean by operational methodology in Section 1.1 followed by Section 1.2, where we briefly discuss the reconstruction program, that was motivated by the desire to find a principled reconstruction of quantum theory, which led to studies of Generalised Probability Theories. We then motivate the field of Indefinite Causal Structure in Section 1.3, largely focused on the Causaloid Framework by Hardy [41, 42], which is central to the latter half of this thesis. Here we also provide other approaches to Indefinite Causal Structure, which start from quantum theory and generalise them to allow for indefinite causality. These approaches may be termed as pursuing Quantum Causality. We discuss how Quantum Communication theory is contingent on definite causal structures, and ask how one may extend them to the setting of indefinite causal structures, that motivate the first half of this thesis. Finally we provide an outline for this thesis in Section 1.4, and share in which order the reader may approach the Chapters and what they may expect from this thesis.

1.1 Operational Methodology

What do we mean by an operational methodology and why do we adopt it? In order to answer this meaningfully, we first look at the broad picture by briefly discussing some metaphysical terms at the heart of the foundations of physics.

Broadly, the question physicists are asking is what is the nature of reality, by which we mean *what* it is that exists, let's loosely call this *stuff* (although physical objects, particles, matter are the terms more often used by physicists); and further physicists are concerned with how this stuff relates to each other and how it moves through space and time. This is what constitutes *Ontology*. The study of Ontology predates Physics and is part of the Metaphysics branch of Philosophy. The ontology that shows up in Physics has been shaped over centuries and till the 19^{th} century had been heavily influenced by Classical Physics, in terms of rigid bodies, trajectories and deterministic evolution of these bodies in time, as well as in terms of an objective understanding of nature, both of which have been challenged by Quantum Physics (as well as by General Relativity).

To answer the ontological questions posed, physicists use experiments to probe nature and to test their hypothesis through the scientific method thereby expanding their knowledge of what it is there is and how stuff behaves, which is coded into physical theories. This knowledge of reality is what constitutes *Epistemology*. The study of Epistemology also predates Physics and is studied within Philosophy. Within Physics, through experimentation and theory construction we continue to move closer to the ontic answers sought by physicists and in an ideal physical theory the epistemology and ontology would coincide (in the limits the theory is valid within).

> *Therefore, a successful physical theory should be consistent with and be able to predict the outcome of experiments (epistemology); as well as provide us a picture of the nature of reality in the approximations it applies to (ontology).*

While Classical Physics provided a crisp distinction between epistemology and ontology (for example, Thermodynamics is the epistemological counterpart of the ontologically sound Statistical Physics), Quantum Theory obfuscates such a distinction and the question of whether the quantum state is ontic or epistemic leaves the research community picking sides and different interpretations in varying degrees between realist and anti-realist positions, positions which are all consistent with the predictively-successful mathematical framework of quantum theory.

Among this, the operational approach has yielded a means to progress by allowing concrete steps forward. No-go theorems, that are theory-independent and show incompatibility of some assumptions help us make some definite statements. Three important no-go theorems pertaining to Quantum Theory, which can be recasted as inequalities, are Bell's theorem, Kochen–Specker theorem, and Reeh–Schlieder theorem (in the Quantum Field Theoretic setting); these tell us about the non-local and contextual nature of Quantum theory and the issue of the creation of non-localised states by the action of local operators in QFT. Further, the confluence of the operational approach of information theory to quantum theory has led to advances in quantum information theory.

Back to the purposes of this thesis, we are interested in a road-map towards Quantum Gravity (QG), a theory more fundamental than General Relativity and Quantum theory which QG should reduce to in the appropriate limits. With the already present issues around ontology and epistemology in the predictively-successful Quantum Theory, and the lack of a predictively-successful framework for Quantum Gravity, the operational methodological route – based on the assertion that *"any physical theory, whatever it does, must correlate recorded data"* – lends us a stable footing.

Herein, we find a pragmatic meeting point between differing philosophical positions. Agnostic of a physicist's philosophical tendencies towards (broadly speaking) realism and anti-realism within physics, we expect that a physical theory will *at least* be consistent with experiments that produce certain classical data.

Given the difficulties around finding a mathematical framework where Quantum theory and General Relativity meet, let alone finding a coherent ontology between the two, adopting an operational methodology as a basis is a safe route, perhaps even a desperate attempt to revisit the blank canvas before clouding it with assumptions that may not hold. The upside is that whatever we can recover operationally won't be wrong even if incomplete, and hopefully would lead towards a point where we are ready to tack on an ontology, should it be possible and desirable. Hardy discusses this position through the following quote:

> "Operationalism played a big role in the discovery of both relativity theory and QT. There are different ways of thinking about operationalism. We can either take it to be fundamental and assert that physical theories are about the behaviour of instruments and nothing more. Or we can take it to be a methodology aimed at finding a theory in which the fundamental entities are beyond the operational realm. In the latter case operationalism helps us put in place a scaffolding from which we can attempt to construct the fundamental theory. Once the scaffolding has served its purpose it can be removed leaving the fundamental theory partially or fully constructed. The physicist operates best as a philosophical opportunist (and indeed as a mathematical opportunist). For this reason we will not commit to either point of view for the time being noting only that the methodology of operationalism serves our purposes. Indeed, operationalism is an important weapon in our armory when we are faced with trying to reconcile apparently irreconcilable theories. A likely reason for any such apparent irreconcilability is that we are making some unwarranted assumptions beyond the operational realm... The operational methodology is a way of not making wrong statements. If we are lucky we can use it to make progress." (Section 3, [41])

1.2 Generalised Probability Theories

When we study special relativity we learn two physical principles that govern it:

1. The laws of physics look the same in all inertial frames of reference
2. The speed of light (in vacuum) is constant for all observers

The above principles do not invoke the mathematical objects that are used to perform calculations. In contrast, typically one's first encounter with quantum theory's axioms requires one to commit to the mathematics – complex Hilbert space structures – used to perform calculations (for example, "a state is represented by a vector in a Hilbert space"). This obfuscates the physics. The question of if a reconstruction of quantum theory through a principled approach through some reasonable postulates is possible was answered by Hardy [40], which was followed by more reconstructions through different choice of postulates [19, 28, 75, 47]. (The reconstruction program can be traced back to Mackey [73] building on work by John von Neumann.)

The recent reconstructions of the current century often work with finite-dimensional quantum theory and operational or information-theoretic postulates. Under these reconstructions one could find common postulates for Classical Probability Theory and Quantum Theory, where some additional postulate which differentiates between them. We will refer to this in our discussion of the Causaloid Framework in Chapter 5.

Further, these reconstructions led to the space of *Generalised Probability Theories*, theories governed by operational notion of settings and measurements, and associated probabilities. This space of theories, often outside the bounds of nature, were able to shed light on understanding quantum theory. For example, PR-boxes belonging to the space of GPTs are able to maximally violate Bell's inequality as well as the Tsirelson's bound and further is shown to violate Information Causality, while quantum theory violates Bell's inequality and obeys Tsirelson's bound as well as Information Causality.

In this thesis, we will revisit the Causaloid Framework (Chapter 4), the space of GPTs that may admit indefinite causal structure which we discuss in the coming sections. We will also apply the Causaloid Framework to the Duotensor Framework, pertaining to probabilistic circuits (GPTs expressible through circuits) (Chapter 6).

1.3 Indefinite Causal Structure

Modern studies of indefinite causal structure in generalised probabilistic theories began with Hardy's Causaloid framework [41, 42]. The main motivation to study indefinite causal structures is the expectation that in Quantum Gravity causal structures (dynamical degrees of freedom of gravity) should subject to indefiniteness (as other dynamical degrees of freedom do in ordinary quantum theory). Hardy motivates this through the following quote:

> "Quantum theory is a probabilistic theory with fixed causal structure. General relativity is a deterministic theory but where the causal structure is dynamic. It is reasonable to expect that quantum gravity will be a probabilistic theory with dynamic causal structure." [41]

Causaloid Framework

The Causaloid framework is based on operational methodology – it is based on the assertion that any physical theory, whatever it does, must correlate recorded data. Imagine a person inside a closed space, having access to stacks of cards with recorded data (procedures, outcomes, locations); and the person is tasked with inferring (aspects of) the underlying physical theory that governs the data. The correlation within recorded data due to the physical theory implies that the stacks of cards are filled with (some) redundancy. The person in the box distils away the redundancy by *compressing* the data. We call this *Physical Compression* as it is governed by the nature of the underlying physical theory. In this framework there are three levels of compression: 1) Tomographic Compression, 2) Compositional Compression and 3) Meta Compression.

The Causaloid Framework considers regions in a causally neutral manner, and through the levels of compression, causal structure emerges. What a physical theory does among other things is help us predict quantities, in particular outcome probabilities, and the Causaloid shares this objective. Therefore, the Causaloid Framework can be regarded as a "general probability theory for theories with indefinite causal structure"; it may also be regarded as a "theory to study correlated data". The Causaloid framework includes complex Hilbert space theory and other probabilistic theories such as real Hilbert space theory as special cases.

Apart from interest in fundamental physics, Hardy also pointed out that the new framework suggests useful applications for practical information processing – a quantum computer that takes advantage of indefinite causal structure may outperform a quantum computer that doesn't [44].

In Chapter 4, we present new terminology and diagrammatic language for the three levels of physical compression to facilitate exposition, while also providing a review of the Causaloid Framework. Further, we study the third level of physical compression to characterise theories through a hierarchy pertaining to causal structures (Chapter 5).

1.3.1 Frameworks for Quantum Causality

The program Hardy initiated more than ten years ago has since then seen a boom across different areas of physics and information science. While the Causaloid Framework is very general (perhaps "too general" depending upon the purpose), other frameworks have emerged, that begin from a complex Hilbert space perspective and relax the temporal order of events by allowing for higher order maps that can be applied to Completely Positive Trace Preserving (CPTP) channels.

Choi-Jamiołkowski Isomorphism

One of the obstacles in relaxing the definite causality in quantum theory is being able to express quantum channels (mathematically represented by complete positive maps) and quantum states (mathematically described by density matrices) in a causally neutral manner. The Choi–Jamiołkowski isomorphism [22, 63] enables one to do this within quantum information and is a mathematical correspondence between positive operators and the complete positive maps (or gives a equivalence between density matrices and Completely Positive Trace Preserving maps):

$$|i\rangle \otimes |j\rangle \equiv |i\rangle \langle j| \tag{1.1}$$

Consider a CPTP map $\Phi : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$ where the dimensions of the Hilbert spaces are $\dim(\mathcal{H}_1) = d_1, \dim(\mathcal{H}_2) = d_2$. The Choi–Jamiołkowski isomorphism [22, 63] gives us the Choi state for this CPTP map as follows, [96]:

$$\text{Choi State:} \quad \rho_\Phi = (\mathbb{1} \otimes \Phi) \sum_{i,j=1}^{d_1} |ii\rangle \langle jj| = \sum_{ij} |i\rangle \langle j| \otimes \Phi(|i\rangle \langle j|) \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2) \tag{1.2}$$

The Choi state of a CPTP map is a density matrix. The Choi state is equivalent to the CPTP map and one can find the action of the CPTP map Φ on some density matrix σ in terms of the CPTP map's Choi state ρ_Φ as follows:

$$\Phi(\sigma) = \Phi(\sum_{i,j} \sigma_{i,j} |i\rangle \langle j|) = \sum_{i,j} \sigma_{i,j} \Phi(|i\rangle \langle j|) \tag{1.3}$$

$$= Tr_1 \left[(\sigma^T \otimes \mathbb{1}) \rho_\Phi \right] \tag{1.4}$$

The frameworks studying indefinite causality in generalised quantum theories often utilise the Choi–Jamiołkowski isomorphism.

Quantum Combs

Chiribella, D'Ariano and Perinotti [18] developed an important framework for quantum networks in complex Hilbert space for general purposes from both a constructive and a neat axiomatic perspective (a framework based on similar mathematical content was developed to study quantum games previously [38]). Although the framework of [18] still assumes a definite causal order (represented by a directed acyclic graph) among the elementary circuits, the mathematical elements that enable indefinite causal order are already present. The framework is causally neutral in the sense that all objects (including both channels and states) are represented through the Choi isomorphism [22] as an operator, and a composition rule (the link product) is given to specify how the operators compose and offer predictions of probabilities. Given this general setup one could already talk about an operational probabilistic theory without specifying a definite causal order for the elementary circuits. Indeed, in [20] (see further developments in [85]), the original framework of [18] was developed to include indefinite causal order, and a computation protocol that cannot be reproduced by definite causal order computation is given. This explicit protocol confirms Hardy's previous suggestion that "quantum gravity computers" outperform ordinary quantum computers [44], and has attracted much attention from both theoreticians and experimentalists.

Process Matrix Framework

The Process Matrix Framework by Oreshkov, Costa and Brukner [82] (see also [2, 83]) is an important framework in the study of indefinite causal structure. This framework is devised from the outset to incorporate indefinite causal structure into complex Hilbert space quantum theory. As in [18], the Process Matrix Framework represents objects such as channels and states as operators through the Choi isomorphism. A series of works based on this framework were carried out to study the new features indefinite causal structure brings to quantum theory (e.g., [3, 31, 36]), and we are gathering an increasingly better understanding of "quantum causality" (see [13] for a review of related works and other frameworks including the duotensor [47] and quantum conditional states [70]).

In the paper [82], analogous to Bell's inequality for non-signalling scenario, a one-way signalling scenario was considered that gave the *Causal Inequality*, violated by indefinite causal orders and this motivated the work done in Chapter 2 (See next Subsection). We will provide more details on the Process Matrix Framework in Chapter 3.

1.3.2 Communication in Theories with Indefinite Causality

Communication or signalling is the operational signature of causally related regions. In this thesis we also were motivated to study extensions of quantum communication theory in the context of indefinite causal structures (Chapters 2,3). We provide the context for these Chapters below.

Guess-Your-Neighbour's Input (GYNI) game

Games provide for studying the possibility of phenomenon in a theory-independent way. Bounds on non-signalling resources between two parties are known through Bell's inequality and Tsirelson bound which can be recast as probabilities of success under given resource in a *Non-Local Game*.

Similarly one can study one-way signalling between two parties through the *Guess-Your-Neighbour's Input (GYNI) game* as done in [82] by Oreshkov et al. The GYNI game helps characterise one-way signalling within different indefinite causal structures. In the process framework [82] one can cast this game to obtain a causal inequality which is obeyed by definite causal orders and violated by indefinite causal orders.

The setting considered by Oreshkov et al [82] has Alice and Bob in two closed laboratories. In the beginning of the game Alice (and Bob) receives a physical system and performs local quantum operations on it (the causal structure inside the lab is definite) after which Alice (and Bob) sends the physical system out of the lab. It is assumed that the labs are isolated, only open once for physical systems to enter and once for physical systems to leave. This assumption, and that of a causally ordered world, restricts the way in which Alice and Bob can communicate during the game. As an example, consider that Alice sends a signal to Bob. Since Bob can only receive a signal through the system entering his lab, this means that Alice must act on her system before. Therefore, Bob cannot send a signal to Alice since they both receive a system only once, and two-way signalling is forbidden. The constraint on signalling then only allows one-way signalling under different causal structures.

Alice (and Bob), after receiving the system toss a coin to obtain random bits a (and b). In addition, Bob also generates a random 'choice" bit c. If $c = 0$, Alice must guess Bob's bit b and encode it in her output bit x, and if $c = 1$ Bob must guess Alice's bit a and encode it in his output bit y, to win the GYNI game. The aim of the game is to maximise the probability of success, given by:

$$P_{GYNI} = \frac{1}{2} \sum_{a,b=\{0,1\}} (P(x = b|c = 0) + P(y = a|c = 1)) \tag{1.5}$$

where $P(x = b|c = 0)$ is the probability of Alice correctly guessing Bob's bit given $c = 0$, and similarly, $P(y = a|c = 1)$ is the probability of Bob correctly guessing Alice's bit given $c = 1$. It can be shown that P_{GYNI} is bounded by 0.75 with definite causal structures (called the **Causal inequality** $P_{GYNI} \leq 0.75$), and is bounded by ~ 0.85 with indefinite causal structures as shown in the paper [82].

We were motivated by the GYNI game. In Chapter 2, we extend it to a two-way signalling scenario with restricted channel capacity on a channel which has both tasks competing on it, setup in a Quantum Butterfly Network. This setup can be modified to play the GYNI game in a quantum theoretic setting. Through this setup, we are able to provide protocols for Bidirectional Teleportation and Bidirectional Dense-Coding.

Tensor products of Processes

The introduction of indefinite causal structure into quantum theory not only brought in new understandings, but also posed some new questions. In particular, the subtlety in taking tensor products of objects with indefinite causal structure was known to some researchers in the community and was explicitly mentioned in [7] and implicitly encompassed in the construction of the combs formulation in [85]. We will look at this in Chapter 3, where we provide a necessary and sufficient condition (Theorem 2) to characterise objects whose tensor products need qualification. This poses a challenge to extending communication theory to indefinite causal structures, as the tensor product is the fundamental ingredient in the asymptotic setting of communication theory. Further, Guérin et al. [37] considered if maps other than the tensor product could circumvent the restriction provided by Theorem 2, through a no-go theorem, to find it would not be possible. We discuss a few options to evade this issue. In particular, we show that the sequential asymptotic setting does not suffer this issue.

1.4 An Outline of this Thesis

Chapter 2 and Chapter 3 can be read independently and all relevant material is presented within each Chapter.

In Chapter 2 we provide protocols for bidirectional generalisations for quantum teleportation and dense-coding with an improvement over the communication resources required. These results may have an impact on quantum communication theory; and were motivated by the Guess-Your-Neighbour's-Input game which was presented in the Process Matrix Formalism [82] within the context of indefinite casual structures.

Chapters 3, 4, 5 and 6 are relevant to the field of indefinite causal structures.

An important result regarding the composition via tensor product of two process matrices is presented in Chapter 3. This result is useful for the extension of quantum communication protocols to post-quantum communication protocols specifically to define asymptotic channel capacities within the process matrix formalism.

Section 6.3 is the Culmination of Chapters 4, 5 and 6 and requires sequential reading, though Section 6.2 can be read on its own.

We revisit the Causaloid framework by Hardy [41, 42] in Chapter 4. We present some new terminology as well as a diagrammatic representation for the Causaloid Framework.

In Chapter 5, we utilise the diagrammatics introduced in Chapter 4 and focus on studying Meta-Compression (the third level of physical compression) to establish a hierarchy of causal theories pertaining to causal structure.

In Chapter 6 we apply the Causaloid Framework to the Duotensor Framework, to populate the second rung of the hierarchy introduced in Chapter 5.

In the final Chapter, we briefly summarise our work and discuss future directions.

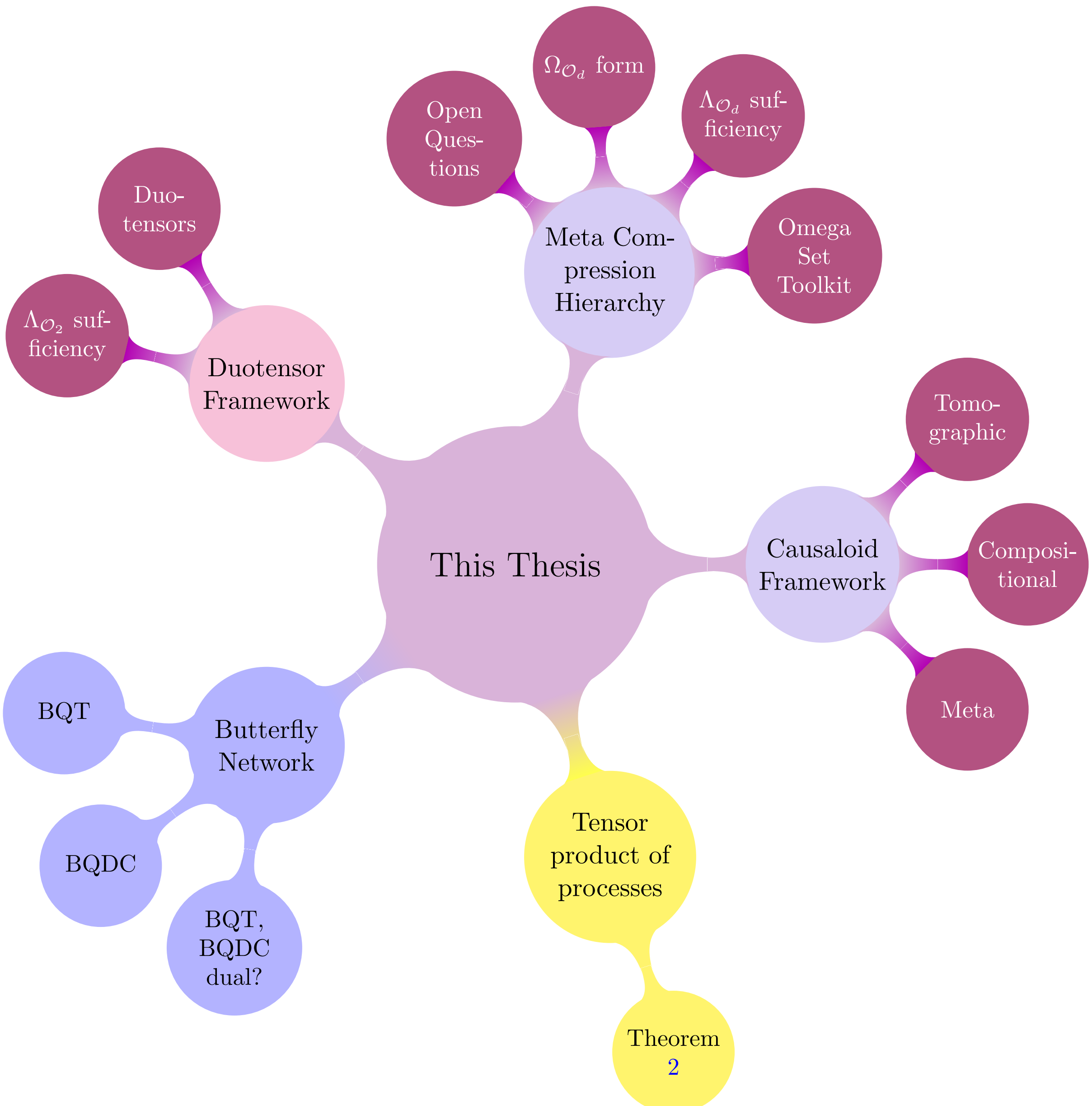
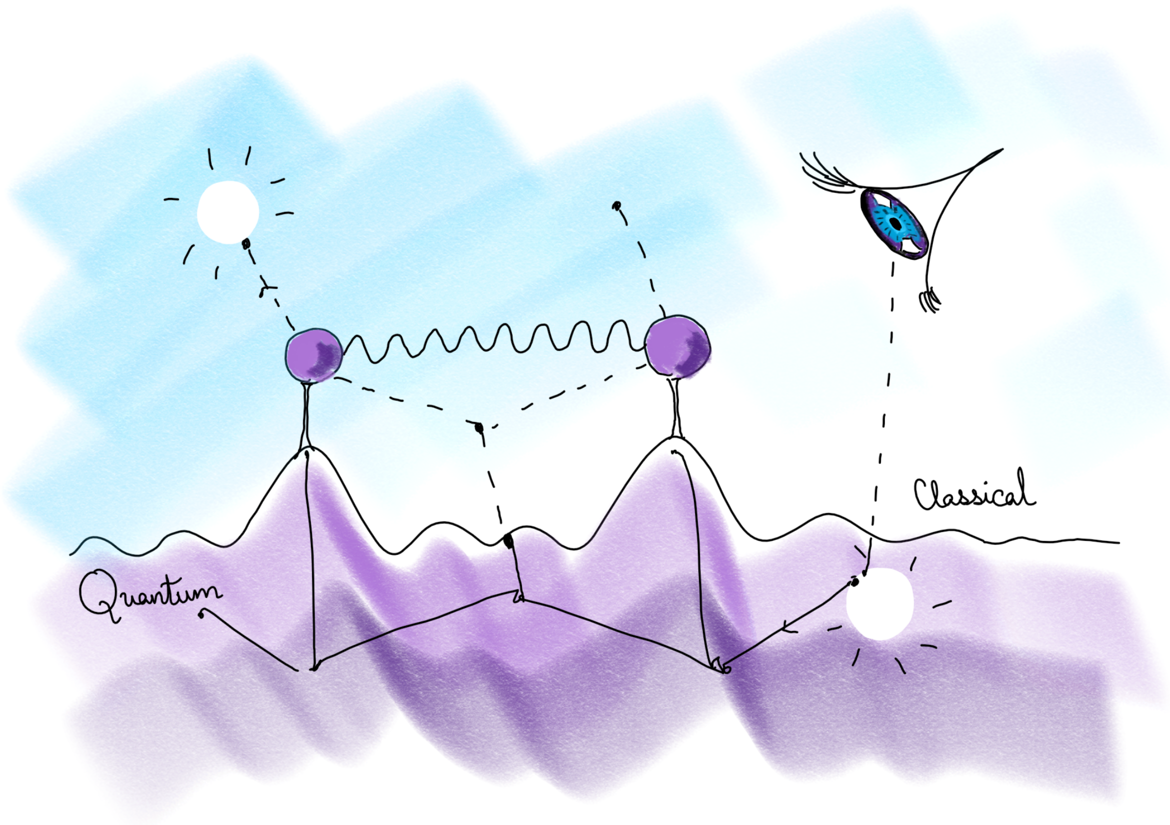


Figure 1.1: A Visual Outline of this Thesis



Teleport me to a world
where information is crisp
where coding is dense
where when we speak
we are also heard [3]

[3]Chapter 2: Conceptual illustration of the Classical-Quantum boundary

Chapter 2

Bidirectional Quantum Teleportation and Dense-Coding

Can competing quantum communication tasks on a signalling channel be optimised?
Can symmetrised communication tasks between Alice and Bob have non-trivial solutions?
Yes, we consider the Quantum Butterfly Network to answer these questions.

Entanglement can be used to aid quantum or classical communication channels to send dense-coded classical bits [11] or teleport quantum states [10] respectively; these results are at the heart of quantum communication. In this chapter, we extend these tasks to bidirectional information flow where two parties (Alice and Bob) both want to send and receive information; namely we extend it to Bidirectional Quantum Dense-Coding (BQDC) and Bidirectional Quantum Teleportation (BQT) (which is compatible with entanglement swapping similar to teleportation). We implement these protocols in a Quantum Butterfly Network, where these competing tasks are implemented over a common signalling channel, using Network Coding, in order to meaningfully study improvements in terms of resources used. As a consequence of using Network Coding, fewer communication resources are permitted at the bottleneck of the network. Hayashi et al. [58] have considered the Quantum version of Network Coding in the butterfly setup to bidirectionally send two qubits without shared entanglement. We allow shared entanglement in our butterfly setup and propose protocols for BQDC and BQT. Further, we introduce the notion of Masked Encoding, explore duality between BQT and BQDC, and discuss applications to the field of quantum communication networks (including a Quantum Internet) with possible extensions to post-quantum theories.

2.1 From Quanta-net to Post-Quantum theories

Entanglement's Role in Quantum Communication

It is known that quantum theory allows for long-distance correlations beyond classical theory through entangled states. The non-locality of quantum theory can be seen through the violation of Bell's inequality [9] and the Tsirelson's bound [23]. These correlations are non-signalling in nature, that is, they cannot be used to transfer information on their own, since that would allow for instantaneous information transfer which is physically impossible due to (special) relativistic principles. Despite being non-signalling, entangled states are an important resource for quantum communication and quantum network that are able to perform tasks impossible with classical resources. Entangled states along with communication channels can be used to enhance communication tasks, for example by enhancing quantum channel capacity in dense-coding [11], by enhancing classical channel to simulate a quantum channel in teleportation [10], or by substituting entanglement for classical communication [25].[1]

Extending to Multi-Party

While the protocols for one-qubit teleportation, two-bit dense-coding, and entanglement-substituted classical communications are optimal and present to us the undoubtedly important conceptual footings to employ entanglement, in practice to extend these protocols to bipartite multi-qubit or multi-bit information transfer is not easy given the nature of the resource theory of entanglement. However, there has been progress through several methods: Wilde derives unit resource capacity region which categorises the statistically achievable protocols using unit resources; namely quantum channel, classical channel and shared entanglement [98] which we discuss in Section 2.4.3. Further there are various protocols proposed using five-qubit cluster states, or combinations of Bell and Greenberger–Horne–Zeilinger (GHZ) states to teleport more than one-qubit across between two-parties or bidirectionally teleport qubits and Bell states [39, 72, 77, 33, 100, 92, 54, 55, 57], that peel away at the multi-party teleportation problem. Of particular interest a protocol for bidirectional teleportation for two qubits was found in [56], we will discuss how this is related to out results: namely BQT I, Protocol 4.

[1] In Section 2.4 we provide a refresher to Teleportation and Dense-coding in our notations.

Motivation I: Quantum Internet

A larger and general question remains as to how these two-party tasks can be embedded in a larger set of communication tasks between multiple parties optimally. There are two reasons to study this question, a practical and a fundamental one. Firstly, to develop quantum communications tasks for the future where conceivably these tasks are used between many parties through a *quantum network ("quanta-net")* much like how classical communication tasks were extended over time to the internet which handles many communication tasks simultaneously, including routing, broadcasting etc. This is a major project with many researchers making progress towards it.

Motivation II: Post-Quantum Communication

The second reason to study quantum multiple two-party tasks embedded in multi parties, has to do with building post-quantum communication theories. The hope is that answering these questions in the quantum realm will help provide insights with similar multi-party tasks in post-quantum theories; that explore stronger non-signalling resources in Generalised Probability Theories (GPTs) (some work has been done towards studying when teleportation in GPTs is supported [6]) or that explore Indefinite Causal Structures such as in Process theories, Quantum Combs etc. Indeed signalling or communication is intimately linked to causality in an operational manner, to the extent that only regions that are causally connected may communicate with or signal to each other.

Furthermore, a natural question to ask is can teleportation be extended within frameworks of Indefinite Causal Structures, and more importantly, is it useful to do so? Some research has been done in the literature that studies improvement in transmission for quantum teleportation that leverage the Quantum Switch [76, 15, 16].

2.2 Problem Statement and Setup

Problem Statement: Bidirectional Signalling Protocols

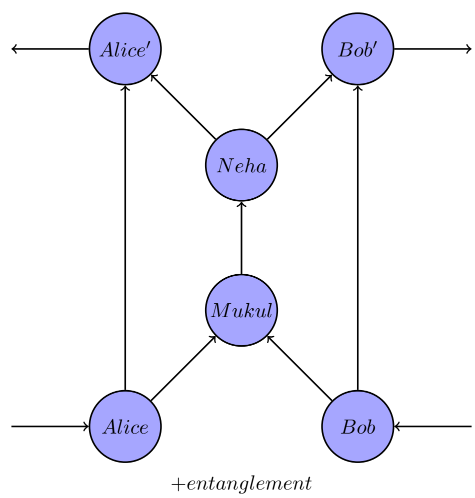
In this chapter we tackle a specific communication task, that of bidirectional signalling, which must be considered to begin addressing as to how known two-party communication tasks can be embedded in a larger network of multiple parties. The task at hand is the simple yet fundamental case with two bipartite communication tasks both between Alice and Bob which generalises teleportation (or dense-coding) to two parties who both simultaneously wish to send a qubit (or two-bits) of information across to the other. That is, we are interested in :

1. **Bidirectional Teleportation:** Alice wishes to teleport the qubit $|\psi\rangle$ to Bob and Bob wishes to teleport the qubit $|\phi\rangle$ to Alice.

2. **Bidirectional Dense-Coding:** Alice wants to communicate two classical-bits (or c-bits) (a_0,a_1) to Bob and Bob wants to communication two c-bits (b_0,b_1) to Alice.

The trivial protocol to achieve this would be to introduce two EPR states (Einstein-Podolski-Rosen states or Bell states that are maximally entangled bipartite states) between Alice and Bob and communication channels (quantum for dense-coding and classical for teleportation) going from Alice to Bob and Bob to Alice (see Section 2.4 for a review of the Teleportation and Dense-coding protocols, see Figure 2.5 fot this trivial protocol), but we learn nothing new about these tasks and essentially we simply perform teleportation or dense-coding twice consuming double the resources (two EPR pairs and two communication channels). Can we do better?

Setup: Quantum Butterfly Network



Instead we propose a more general setup where we introduce two new intermediary parties: Mukul and Neha[2] that mediate all communication between Alice and Bob. Alice and Bob

[2]Names found in India allowing for multicultural communication between Quantum Communicators.

may communicate with Mukul, further Mukul may communicate with Neha, and finally Neha may communicate back to Alice′ and Bob′ (the primes refer to the same party at a future time). In this setup to discuss optimisation we may focus on the communication channel resources between Mukul and Neha on which both signalling tasks (of Alice to Bob and Bob to Alice) are competing. This network of four parties resembles a butterfly, thus the name butterfly network, which we will utilise.

Motivation III: Extension of (Non-)Signalling Games

There is an additional motivation to this work with the butterfly setup, that has to do with theory-independent games, as a means to study bounds. For example, one can provide a non-local game setting [26] for the CHSH inequality [24] that captures bounds for non-signalling resources in different theories, such as the violation of Bell's inequality by Quantum Theory or the violation of Tsirelson's bound by some Generalised Probability Theories. The Guess-Your-Neighbour's-Input game [82] (discussed in Chapter 1) allows one-way signalling resources, where definite causal order provides the *causal inequality* which is violated in indefinite causal order scenarios. Here we wish to focus on a two-way signalling scenario. While we study constraints on communication channel capacities and resources required to perform a task, we were motivated by (non-)signalling games and the GYNI in particular. The butterfly network can be modelled to reproduce the GYNI setting if Mukul and Neha are instructed to toss a coin that allows either Alice to communicate with Bob or Bob to communicate with Alice. Indeed we came up with the butterfly network by thinking about the GYNI game.

Outline of the Chapter

Section 2.3 clarifies the diagrammatics and notation used. Section 2.4 is pedagogical, we first go through the original results on teleportation [10] and dense-coding [11] restated through our conventions and introduce the quantum communication resource theory [98]. In Section 2.5 we will learn about the Classical Butterfly Network and Network Coding results that have helped reduce communication resources on the classical internet and we will setup the Quantum Butterfly Network we use here.

We present bidirectional protocols for teleportation and dense-coding in our butterfly setup in subsequent Sections 2.6 and 2.7. We discuss the types of entanglement resources that can be shared prior to the protocol and its consequences. We also discuss the possibility of a duality between bidirectional dense-coding and bidirectional teleportation analogous to the duality between coherent teleportation and coherent dense-coding in Section 2.8.

2.3 Diagrammatics and Notation

2.3.1 Diagrammatic conventions

To begin let us share the convention that will be used in this Chapter. For spatially separated parties we give different names denoted by blue circles and for the same party at different times we use the name with a prime (such as Alice and Alice′). Different arrows are present to depict classical and quantum communication with certain channel capacity. We use squiggly lines to denote shared entanglement between parties. Time in all the diagrams goes up, similar to conventions adopted in diagrams in relativity and in contrast to the circuit diagrams prevalent in quantum information. Each vertex can be thought of as a space-time region similar to relativity diagrams as well as quantum circuit diagrams.

Note that we consider local operations to be free while shared entanglement and communication are not and therefore the channel capacity as well as the amount of entanglement are resources just as in the resource theory of quantum communication [98]. Same parties at different instances of time in the right of Fig 2.1 can be seen as having channels with memory (unrestricted channel capacity) and while these are drawn for convenience, these represent local operations and therefore are free resources.

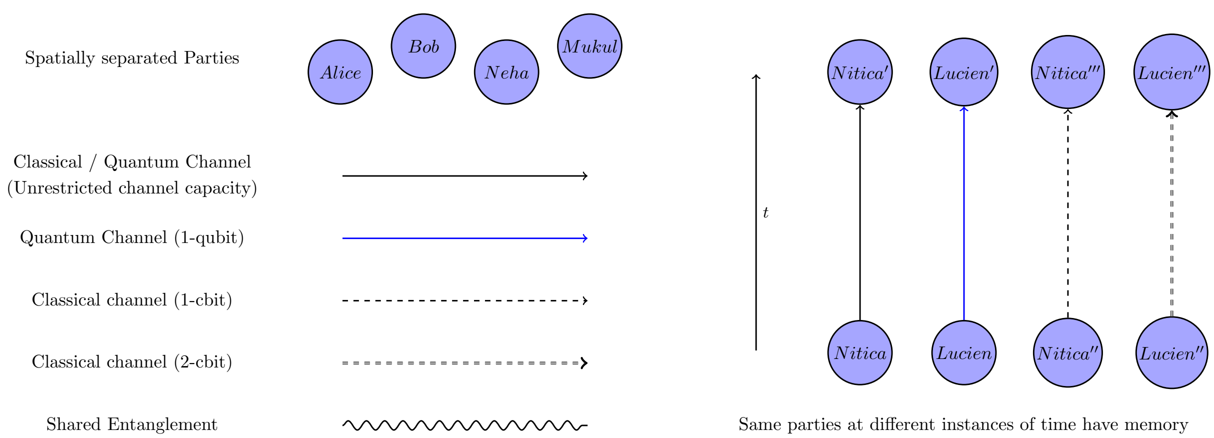


Figure 2.1: Diagrammatic conventions for parties, channels and entanglement

2.3.2 Notation for Resources

We adopt the following notation similar to that by Wilde [98] for resources between parties:

1. **Distributed / Shared entanglement**
   $[qq]_{AB}$ for Bell state shared between A and B
   $[qqq..]_{ABC..}$ for GHZ states shared between A,B,C..
2. **Classical signalling:** $n[c \to c]_{A\to B}$ n-bits sent from A to B
3. **Quantum signalling:** $n[q \to q]_{A\to B}$ n-qubits sent from A to B

2.3.3 Summary of Protocols

The following abbreviations are used : QT (Quantum Teleportation), QDC (Quantum Dense-Coding), CBNC (Classical Butterfly Network Coding), BQT (Bidirectional Quantum Teleportation)and BQDC (Bidirectional Quantum Dense-Coding)

A summary of the protocols found in this Chapter are as follows. Note that the Protocols 4-8 are the results of this work that build upon Protocols 1-3. Among the resources, we focus on the signalling resources between Mukul and Neha, which are highlighted in blue.

	Protocol	Goal	Resources
1	QT	$\|\psi\rangle_{A\to B}$	$[qq]_{AB}, 2[c\to c]_{A\to B}$
2	QDC	$(a_0, a_1)_{A\to B}$	$[qq]_{AB}, [q\to q]_{A\to B}$
3	CBNC	$a_{A\to B},\ b_{B\to A}$	$[c\to c]_{A\to M},\ [c\to c]_{B\to M},\ [c\to c]_{M\to N},$ $[c\to c]_{N\to A'},\ [c\to c]_{N\to B'}$
4	BQT I	$\|\psi\rangle_{A\to B}, \|\phi\rangle_{B\to A}$	$2[qq]_{AB}, 2[c\to c]_{A\to M}, 2[c\to c]_{B\to M},$ $2[c\to c]_{M\to N},\ 2[c\to c]_{N\to A'}, 2[c\to c]_{N\to B'}$
5	BQT II	$\|\psi\rangle_{A\to B}, \|\phi\rangle_{B\to A}$	$[qq]_{AN}, [qq]_{BN}, 2[c\to c]_{A\to M}, 2[c\to c]_{B\to M},$ $2[c\to c]_{M\to N},\ [q\to q]_{N\to A'}, [q\to q]_{N\to B'}$
6	BQDC I	$(a_0, a_1)_{A\to B},\ (b_0, b_1)_{B\to A}$	$[qq]_{MN},\ [q\to q]_{M\to N},\ 2[c\to c]_{A\to M}$, $2[c\to c]_{B\to M},\ 2[c\to c]_{N\to A'},\ 2[c\to c]_{N\to B'}$
7	BQDC II	$(a_0, a_1)_{A\to B},\ (b_0, b_1)_{B\to A}$	$[qqqq]_{ABMN},\ [q\to q]_{M\to N},\ [q\to q]_{A\to M},$ $[q\to q]_{B\to M},\ 2[c\to c]_{N\to A'},\ 2[c\to c]_{N\to B'}$
8	BQDC III	$(a_0, a_1)_{A\to B},\ (b_0, b_1)_{B\to A}$	$[qq]_{AB}, [qq]_{MN},\ [q\to q]_{M\to N},\ [q\to q]_{A\to M},$ $[q\to q]_{B\to M},\ 2[c\to c]_{N\to A'},\ 2[c\to c]_{N\to B'}$

Table 2.1: Summary of Communication Protocols

2.4 Teleportation and Dense-Coding

In 1992 the protocol for dense-coding [11] and in 1993 the protocol for teleportation were found [10]. Let us revisit them with the diagrammatic conventions provided above.

The EPR states (Einstein-Podolski-Rosen states) also known as Bell states are maximally entangled bipartite states and will be represented as follows:

$$\left|\beta^{00}\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{2.1}$$

$$\left|\beta^{01}\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \tag{2.2}$$

$$\left|\beta^{10}\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \tag{2.3}$$

$$\left|\beta^{11}\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{2.4}$$

We may write the Bell states above more succinctly as Pauli one-qubit gates applied to *second* qubit of the Bell state $|\beta^{00}\rangle$. Alternatively, it can also we represented as Pauli one-qubit gates applied to *first* qubit of the Bell state $|\beta^{00}\rangle$. This notation of the Bell states is particularly useful for the protocols ahead.

$$|\beta^{xz}\rangle = (\mathbb{1} \otimes X^x Z^z)\left|\beta^{00}\right\rangle = (Z^z X^x \otimes \mathbb{1})\left|\beta^{00}\right\rangle \tag{2.5}$$

$$\left|\beta^{00}\right\rangle = (\mathbb{1} \otimes Z^z X^x)\,|\beta^{xz}\rangle = (X^x Z^z \otimes \mathbb{1})\,|\beta^{xz}\rangle \qquad \text{(inverse)} \tag{2.6}$$

For Greenberger–Horne–Zeilinger (GHZ) states we use the notation $|\beta^{000..}\rangle = \frac{1}{\sqrt{d}}(|000..\rangle + |111..\rangle)$

2.4.1 Quantum Teleportation

The task of teleportation is to send qubits from one party (say Alice) to another party (say Bob) using shared entanglement and classical communication. This is a novel result for quantum communication theory. To appreciate this consider the following, we need two complex numbers to describe a pure non-normalised qubit, or three real numbers to describe a (mixed or pure) normalised qubit on a Bloch Sphere *if* we know the state of our qubit. In principle this information sounds impossible to send via a finite string of classical bits even if one knows the qubit state. But this task becomes possible with only two c-bits if there is shared entanglement for a qubit in an unknown state! The protocol is as follows:

Protocol 1. Quantum Teleportation

Goal: Alice wishes to send $|\psi\rangle$ to Bob.
Resources: $[qq]_{AB}, 2[c \to c]_{A\to B}$

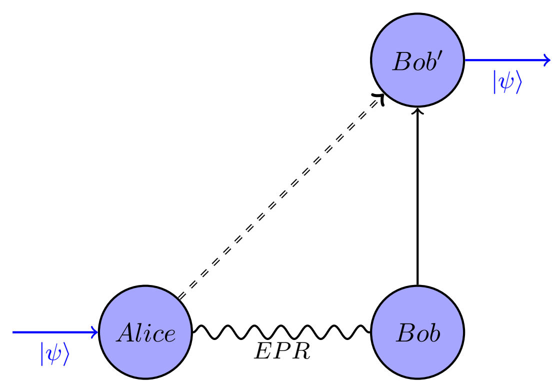


Figure 2.2: Quantum Teleportation

1. Alice has $|\psi_A\rangle$ and Alice and Bob receive Bell state

$$|\psi_A\rangle \otimes \left|\beta_{AB}^{00}\right\rangle = (\alpha_0 |0\rangle + \alpha_1 |1\rangle) \otimes \left|\beta_{AB}^{00}\right\rangle \tag{2.7}$$

2. Alice performs a Bell measurement $\{\langle\beta_{AA}^{xz}|\, |x, z \in \{0,1\}\}$ on $|\psi_A\rangle$ and her half of the Bell pair

$$|\psi_A\rangle \left|\beta_{AB}^{00}\right\rangle = \frac{1}{2}\begin{pmatrix} |\beta_{AA}^{00}\rangle |\psi\rangle_B + |\beta_{AA}^{01}\rangle Z_B |\psi\rangle_B + \\ |\beta_{AA}^{10}\rangle X_B |\psi\rangle_B + |\beta_{AA}^{11}\rangle Z_B X_B |\psi\rangle_B \end{pmatrix} \tag{2.8}$$

3. Alice communicates the two c-bits of information (x, z) to Bob which corresponds to the outcome to the Bell measurement $\{\langle\beta_{AA}^{xz}|\, |x, z \in \{0,1\}\}$

4. Bob applies the $Z^z X^x$ to correct his half of the Bell pair to get $|\psi_B\rangle$

2.4.2 Quantum Dense-Coding

The task of dense-coding is to send $2n$ classical bits from one party Alice to another party Bob using shared entanglement and n qubits. While it may seem costly to send classical information via quantum resources, it may be desired for encryption and furthermore, this is a novel result for quantum communication theory because of the *Holevo bound*. The Holevo bound [59] upper bounds the *accessible information* and restricts the capacity of a n qubit channel between two parties to n classical bits (in the absence of shared entanglement). Given this bound, Dense-coding is able to go over the Holevo bound due to the presence of shared entanglement! The protocol is as follows:

Protocol 2. Quantum Dense-Coding

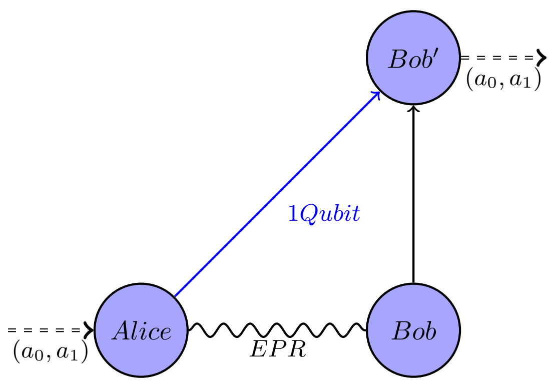


Figure 2.3: Quantum Dense-Coding

Goal: Alice wishes to send two c-bits (a_0, a_1) to Bob.
Resources: $[qq]_{AB}, [q \to q]_{A\to B}$

1. Alice and Bob receive Bell state $|\beta_{AB}^{00}\rangle$
2. Alice applies $X^{a_0}Z^{a_1}$ on her half of the Bell pair

$$X^{a_0}Z^{a_1}\left|\beta_{AB}^{00}\right\rangle = |\beta_{AB}^{a_0a_1}\rangle \tag{2.9}$$

3. Alice sends her half of the Bell pair to Bob $|\beta_{BB}^{a_0 a_1}\rangle$
4. Bob performs a Bell measurement $\{\langle\beta_{BB}^{a_0 a_1}|\, | a_0, a_1 \in \{0,1\}\}$ on the two qubits to recover the two bits of information (a_0, a_1)

2.4.3 Resource Theory of Unit Quantum Protocols

An excellent exposition of the quantum communication resource theory can be found in the book by Wilde [98]. Here the resources are quantum communication channel capacity, classical communication channel capacity and shared entanglement between spatially separated parties, while local operations for any party (quantum and classical) are free. $X > Y$ ($X \geq Y$) denotes that X is a (strictly) stronger resource than Y. Let us see what some of these resource inequalities (and equalities) look like.

Signalling is a stronger resource than shared entanglement as you can generate entanglement using a quantum channel and more specifically a Bell pair using a qubit channel:

$$[q \to q]_{A\to B} \geq [qq]_{AB} \tag{2.10}$$

Quantum signalling is a stronger resource than Classical signalling as you can simulate classical communication using a quantum channel but not the other way round. Specifically if you only encode on the computational basis of a qubit channel and measure in the same basis you can simulate a classical bit channel, therefore we have:

$$[q \to q]_{A\to B} \geq [c \to c]_{A\to B} \tag{2.11}$$

In teleportation you use a two bit classical channel and a Bell pair to simulate a qubit channel. The corresponding resource inequality looks like:

$$2[c \to c]_{A\to B} + [qq]_{AB} \geq [q \to q]_{A\to B} \tag{2.12}$$

In dense-coding you use a qubit channel and a Bell pair to simulate a two bit classical channel. The resource inequality looks like:

$$[q \to q]_{A\to B} + [qq]_{AB} \geq 2[c \to c]_{A\to B} \tag{2.13}$$

Trivially one can establish duality between teleportation and dense-coding if shared entanglement (that is $[qq]_{AB}$) was a free resource.

$$\begin{aligned} & 2[c \to c]_{A\to B} \geq [q \to q]_{A\to B} && \text{from teleportation} \\ & [q \to q]_{A\to B} \geq 2[c \to c]_{A\to B} && \text{from dense-coding} \\ \Rightarrow\quad & [q \to q]_{A\to B} = 2[c \to c]_{A\to B} \end{aligned} \tag{2.14}$$

Duality of Teleportation and Dense-Coding

Teleportation and Dense-Coding are indeed dual processes under this resource theory but it can be shown in a more sophisticated manner through coherent teleportation and coherent dense-coding. To do this we must define the Coherent bit channel as follows:

$$[q \to qq]_{A\to AB} : |i\rangle_A \to |i\rangle_A |i\rangle_B \quad \text{where} \quad i \in \{0,1\} \tag{2.15}$$

The Coherent bit channel falls between a qubit channel and shared entanglement in terms of its resource capability:

$$[q \to q]_{A\to B} \geq [q \to qq]_{A\to AB} \geq [qq]_{AB} \tag{2.16}$$

In Coherent Teleportation [12] one uses two coherent channels and a Bell pair to simulate a qubit channel. One also ends up producing two Bell states in this procedure. The resource inequality follows:

$$2[q \to qq]_{A\to AB} + [qq]_{AB} \geq [q \to q]_{A\to B} + 2[qq]_{AB} \tag{2.17}$$

$$\to \qquad 2[q \to qq]_{A\to AB} \geq [q \to q]_{A\to B} + [qq]_{AB} \tag{2.18}$$

In Coherent Dense-Coding [53] you use a qubit channel and a Bell pair to simulate two Coherent bit-channels. The resource inequality follows:

$$[q \to q]_{A\to B} + [qq]_{AB} \geq 2[q \to qq]_{A\to AB} \tag{2.19}$$

Now, one can establish duality between coherent teleportation and coherent dense-coding via the coherent bit channel using Equations 2.18 and 2.19.

$$\begin{aligned} & 2[q \to qq]_{A\to AB} \geq [q \to q]_{A\to B} + [qq]_{AB} && \text{from Coherent Teleportation} \\ & [q \to q]_{A\to B} + [qq]_{AB} \geq 2[q \to qq]_{A\to AB} && \text{from Coherent Dense-Coding} \\ \Rightarrow \quad & [q \to q]_{A\to B} + [qq]_{AB} = 2[q \to qq]_{A\to AB} && \end{aligned} \tag{2.20}$$

Notice that Coherent Teleportation is weaker while Coherent Dense-coding is stronger than their non-coherent counterparts, which is critical for this duality. Furthermore, notice that the classical task needs to be restated in quantum resources for this resource duality to work. We will explore duality for the Bidirectional protocols in Section 2.8.1.

2.5 Butterfly Networks

We have revisited known results for two-party quantum protocols and the resource theory over shared entanglement, quantum channels and classical channels in Section 2.4. How can we extend these to multi-parties and can these provide some advantage? We focus our attention to the problem of bidirectional flow of information between two parties Alice and Bob in the presence of other parties. We look at the Butterfly Network (as described in the beginning of the Chapter), wherein we have Alice and Bob wishing to communicate bidirectionally and we have two intermediary parties Mukul and Neha who mediate information between Alice and Bob.

Before considering quantum tasks we discuss an important result within Classical Networks in this setup, that is useful for tasks performed over the internet, after which we will elaborate on the Quantum Butterfly Network.

2.5.1 Classical Butterfly Network

The Classical Butterfly Network is often used to illustrate how *Network Coding* can outperform routing [1]. Network Coding refers to the (en)coding done at a node (such as Mukul's), explained by Ahlswede et al. as follows:

> "In existing computer networks, each node functions as a switch in the sense that it either relays information from an input link to an output link, or it replicates information received from an input link and sends it to a certain set of output links. From the information-theoretic point of view, there is no reason to restrict the function of a node to that of a switch. Rather, a node can function as an encoder in the sense that it receives information from all the input links, encodes, and sends information to all the output links. From this point of view, a switch is a special case of an encoder. In the sequel, we will refer to coding at a node in a network as network coding." ([1])

Here we have adapted the example from their paper [1] of six parties to a setup of four parties where Alice/Alice′ and Bob/Bob′ are nodes refering to the same party at different points in time. We have two parties Alice and Bob each trying to send a classical bit to the other (a to Bob′, b to Alice′). Instead of asking for the other's classical bit what if they ask "Is the bit with the other the same as the one I have?" The answer to this question is the same for both, the answer encoded as $m = a \oplus b$, and therefore we will see an advantage.

Protocol 3. Classical Butterfly Network Coding

Goal: Alice wishes to send a c-bit a to Bob and Bob wishes to send a c-bit b to Alice
Resources: $[c \to c]_{A \to M}, [c \to c]_{B \to M}, [c \to c]_{M \to N}, [c \to c]_{N \to A'}, [c \to c]_{N \to B'}$

1. Alice sends c-bit a and Bob sends c-bit b to Mukul
2. Mukul adds them via the boolean sum $m = a \oplus b$ and sends m to Neha
3. Neha broadcasts m to both Alice and Bob

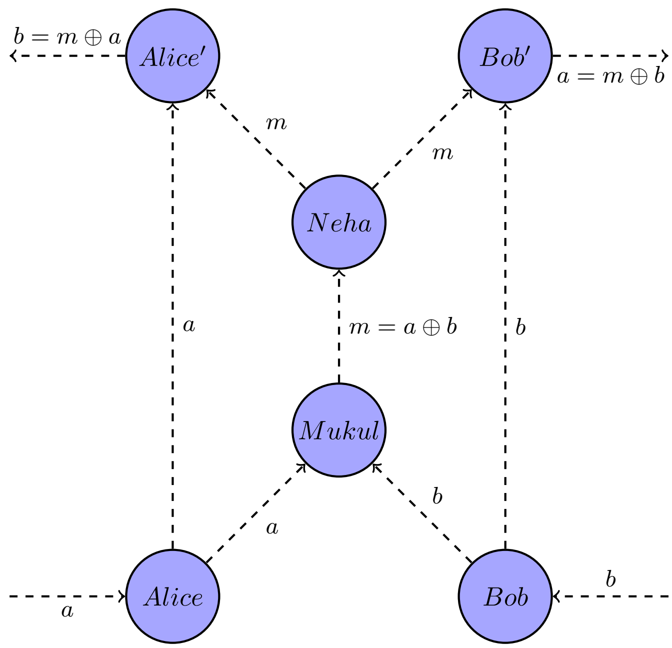


Figure 2.4: Classical Butterfly Network

4. Alice adds her c-bit a and m:

$$m \oplus a = a \oplus b \oplus a = b \tag{2.21}$$

5. Bob adds his c-bit b and m:

$$m \oplus b = a \oplus b \oplus b = a \tag{2.22}$$

Typically Mukul would require to send two c-bits to Neha, but in this particular task due to Alice and Bob having memory, one c-bit of communication ($[c \to c]_{M \to N}$) is saved. We will leverage this result[3] in our protocols in the following sections.

Notice that for this protocol to work Alice and Bob must trust Mukul and Neha. Additionally, this protocol is not secure, that is Mukul has access to a, b. Can there be a secure alternative to this protocol? We will explore these concerns in the Quantum Butterfly Network through what we will call *Masked Encoding*.

2.5.2 Quantum Butterfly Network

We wish to perform Bidirectional Quantum Teleportation (BQT) and Bidirectional Quantum Dense-Coding (BQDC). The obvious trivial protocol for BQT would involve two Bell pairs between Alice and Bob, two bits sent form Alice to Bob and two bits sent from Bob to Alice (illustrated below). Similarly, the obvious trivial protocol for BQDC would involve two Bell pairs between Alice and Bob, one qubit sent from Alice to Bob and one qubit sent from Bob to Alice (illustrated below). Can we do better? To do this meaningfully

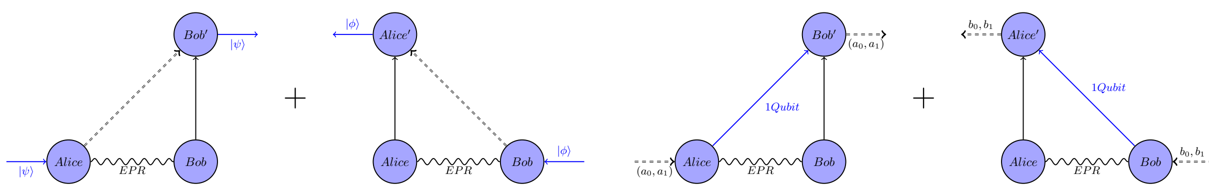


Figure 2.5: Trivial protocol for BQT (left) and BQDC (right)

we need to use a scenario where the signalling resources between Alice to Bob and Bob to Alice can come together and be compared. The butterfly network does exactly that. The channel between Mukul and Neha is a channel that has competing tasks that both Alice and Bob need for signalling. Thus, we consider a Quantum version of the Butterfly Network where channels can now be classical or quantum and shared entanglement between parties is introduced:

[3]This adaptation of the protocol was found independently by the author a couple of decades too late, before stumbling upon an Electrical Engineering Journal paper, that is largely absent in Quantum research literature.

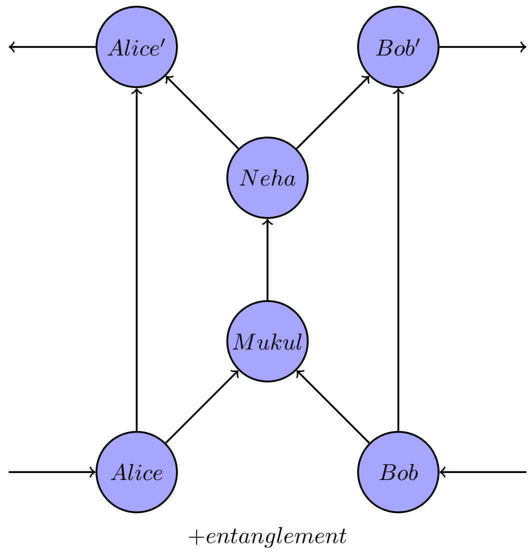


Figure 2.6: Quantum Butterfly Network

Within the Butterfly network the trivial protocols BQT corresponds to four c-bits (and BTDC corresponds to two qubits) to be sent from Mukul to Neha. Can we reduce this? We show that indeed a quantum form of the Classical Butterfly Network Coding allows us to reduce signalling resources (two c-bits for BQT and one qubit for BQDC) along the channel from Mukul to Neha.

Note that a quantum butterfly network is also studied by Leung [71] though the context and problem statement are different than what we wish to consider here. For one, they allow for backward assisted communication which we do not. Also in our butterfly network, we have only four parties, with two of these parties represented twice since we are distinguishing temporally separate events as distinct nodes.

Resources in the Butterfly Network

In the resource theory discussed in previous sections [98] local operations are free; while quantum communication, classical communication as well as shared entanglement are not free. While we largely adopt the same resource theory to present our protocols, the resource theory is more complex when considering the multi-party case, in particular the resources within shared entanglement are more involved, thus we introduce some modifications.

Given our particular interest in bidirectional flow of information for the purposes of discussion we focus on the channel between Mukul and Neha for resource counting (qubits or c-bits), since the goals compete over this channel. Note that if one imagines a time-slice at a moment *after* Mukul signals to Neha and *before* Neha receives Mukul's signal, one will notice that in fact the resources of interest are indeed the channel between Mukul and Neha (remember that the channel from Alice to Alice′ / Bob to Bob′ are not real channels, they correspond to local operations and memory over time-evolution) and shared entanglement.

Therefore, for discussion and simplicity one may consider the signalling resources over the channels from Alice to Mukul, Bob to Mukul, Neha to Alice′ and Neha to Bob′ to be *free* (but for practical purposes, for example a person with an experimental setup, all signalling resources are to be counted).

Shared Entanglement in the Butterfly Network

Here we discuss the possible shared entanglement cases one can start with in the butterfly setup to give us some structure over the shared resource. Given the signalling resources over the channels from Alice to Mukul, Bob to Mukul, Neha to Alice′ and Neha to Bob′ are considered *free*, we reduce the possible shared entanglement cases to two cases, using resource inequalities:

1. Shared entanglement between Alice and Bob: $[qq]_{AB}$

$$[qq]_{AB} \geq [qqq]_{ABM} \geq ([qq]_{AM} = [qq]_{BM} \qquad \text{free}) \tag{2.23}$$

   since Alice (and Bob) can freely send qubits to Mukul, to create Bell pairs between Mukul and Alice (Bob). If Alice and Bob have a Bell pair they can freely prepare a GHZ state with Mukul.

2. Shared entanglement between Alice, Bob and Neha: $[qq]_{AN}, [qq]_{BN}$ or $[qqq]_{ABN}$

$$[qq]_{AN} \geq [qqq]_{AMN} \tag{2.24}$$
$$[qq]_{BN} \geq [qqq]_{BMN} \tag{2.25}$$
$$[qqq]_{ABN} \geq [qqq]_{ABMN} \tag{2.26}$$
$$[qq]_{AN} \geq [qq]_{AB'} \geq [qq]_{A'B'} \tag{2.27}$$
$$[qq]_{BN} \geq [qq]_{A'B} \geq [qq]_{A'B'} \tag{2.28}$$

   where we use the fact that Alice/Bob can freely send qubits to Mukul in Equations 2.24-2.26, and Neha can freely send qubits to Alice′/ Bob′ in Equations 2.27-2.28

Therefore, we consider only shared entanglement of the form: $[qq]_{AB}, [qqq]_{ABN}$ or $[qq]_{AN}, [qq]_{BN}$ since these appear to be sufficient.

2.6 Bidirectional Quantum Teleportation

For bidirectional teleportation, we know that in the absence of the butterfly network we would need two Bell pairs, which is true for the butterfly network as well. Let us see how we may utilise the CBNC (Protocol 3) to perform bidirectional teleportation in the classical butterfly network.

2.6.1 Bidirectional Quantum Teleportation (Version I)

As a warm up, in this version we look for the easiest implementation of BQT in the Butterfly Network where we use Network Coding and Teleportation to find an advantage.

Protocol 4. Bidirectional Quantum Teleportation (Version I)

Goal: Alice wishes to send qubit $|\psi\rangle$ to Bob and Bob wishes to send qubit $|\phi\rangle$ to Alice
Resources: $2[qq]_{AB}, 2[c \to c]_{A\to M}, 2[c \to c]_{B\to M}, 2[c \to c]_{M\to N},\ 2[c \to c]_{N\to A}, 2[c \to c]_{N\to B}$.

1. Alice has qubit $|\psi_{A_1}\rangle$, Bob has qubit $|\phi_{B_1}\rangle$
2. Alice and Bob share two Bell pairs $\left|\beta^{00}_{A_2B_2}\right\rangle, \left|\beta^{00}_{A_3B_3}\right\rangle$ making the global state

$$|\psi_{A_1}\rangle \otimes \left|\beta^{00}_{A_2B_2}\right\rangle \otimes \left|\beta^{00}_{A_3B_3}\right\rangle \otimes |\phi_{B_1}\rangle \tag{2.29}$$

3. Alice does a local Bell measurement $\{\left\langle\beta^{a_0a_1}_{A_1A_2}\right| |a_0, a_1 \in \{0,1\}\}$ on A_1, A_2. We can massage the state to write it in a linear superposition over the Bell basis.

$$|\psi_{A_1}\rangle \left|\beta^{00}_{A_2B_2}\right\rangle = \frac{1}{2}\begin{pmatrix} \left|\beta^{00}_{A_1A_2}\right\rangle |\psi_{B_2}\rangle + \left|\beta^{01}_{A_1A_2}\right\rangle Z_{B_2} |\psi_{B_2}\rangle \\ + \left|\beta^{10}_{A_1A_2}\right\rangle X_{B_2} |\psi_{B_2}\rangle + \left|\beta^{11}_{A_1A_2}\right\rangle Z_{B_2}X_{B_2} |\psi_{B_2}\rangle \end{pmatrix} \tag{2.30}$$

4. Bob does a local Bell measurement $\{\langle\beta^{b_0 b_1}_{B_1 B_3}|\,|b_0, b_1 \in \{0,1\}\}$ on B_1, B_3. We can again massage the state to write it in a linear superposition over the Bell basis.

$$|\phi_{B_1}\rangle\,|\beta^{00}_{B_3 A_3}\rangle = \frac{1}{2}\begin{pmatrix} |\beta^{00}_{B_1 B_3}\rangle\,|\psi_{A_3}\rangle + |\beta^{01}_{B_1 B_3}\rangle\, Z_{A_3}\,|\psi_{A_3}\rangle \\ + |\beta^{10}_{B_1 B_3}\rangle\, X_{A_3}\,|\psi_{A_3}\rangle + |\beta^{11}_{B_1 B_3}\rangle\, Z_{A_3} X_{A_3}\,|\psi_{A_3}\rangle \end{pmatrix} \tag{2.31}$$

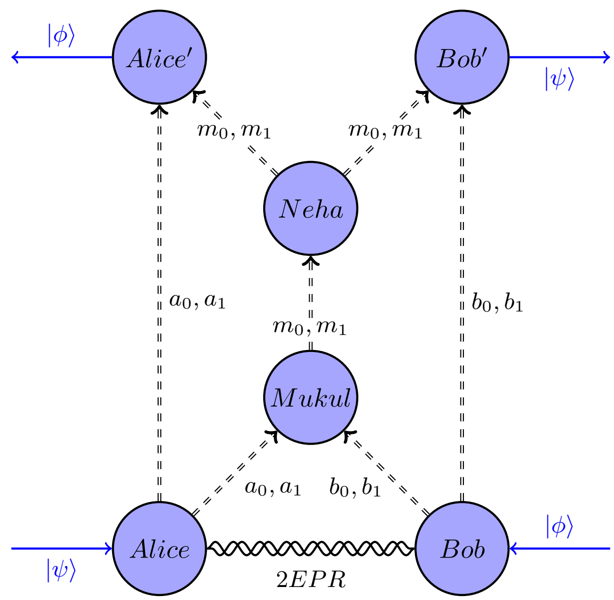


Figure 2.7: Bidirectional Quantum Teleportation (Version I)

5. At this point the quantum state of the remaining two qubits, A_3 with Alice and B_2 with Bob is:

$$Z^{a_1}_{B_2} X^{a_0}_{B_2}\,|\psi_{B_2}\rangle \otimes Z^{b_1}_{A_3} X^{b_0}_{A_3}\,|\phi_{A_3}\rangle \tag{2.32}$$

   Now all that is needed by Alice is (b_0, b_1) and Bob is (a_0, a_1), which can be sent by using the classical coding Protocol 3 twice.

6. Alice sends the outcome (a_0, a_1) of her Bell measurement $\{\langle\beta^{a_0,a_1}_{A_1 A_2}|\,|a_0, a_1 \in \{0,1\}\}$ to Mukul.

7. Bob sends the outcome (b_0, b_1) of his Bell measurement $\{\left\langle\beta^{b_0,b_1}_{B_1B_2}\right| |b_0, b_1 \in \{0,1\}\}$ to Mukul.
8. Mukul produces two c-bits (m_0, m_1) with $m_0 = a_0 \otimes b_0$ and $m_1 = a_1 \otimes b_1$ and sends them to Neha.
9. Neha broadcasts (m_0, m_1) to Alice and Bob.
10. Alice adds (a_0, a_1) from her memory to (m_0, m_1) to receive (b_0, b_1),

$$b_i = m_i \oplus a_i \quad \forall i \in \{0,1\} \tag{2.33}$$

11. Bob adds (b_0, b_1) from his memory to (m_0, m_1) to receive (a_0, a_1),

$$a_i = m_i \oplus b_i \quad \forall i \in \{0,1\} \tag{2.34}$$

12. Alice applies the gate $X^{b_0}Z^{b_1}$ to qubit A_3 and Bob applies the gate $X^{a_0}Z^{a_1}$ to qubit B_2 to complete bidirectional teleportation!

$$|\phi_A\rangle \otimes |\psi_B\rangle \tag{2.35}$$

Here, we employed the Classical Butterfly Network Coding Protocol directly over the measurement outcomes of the Bell measurements to lower the signalling costs for Mukul to send only two c-bits (m_0, m_1) (as opposed to four c-bits (a_0, a_1, b_0, b_1)). This should not be surprising and yet it is a tangible improvement.

This result is distinct from the BQT Protocol presented in by Hassanpour et al. [56]. In Hassanpour's Protocol they use two Bell pairs as one would expect in BQT though they do not focus on a classical channel on which the communication task compete such as our Mukul and Neha channel. We have focused on the Butterfly network to find a communication resource advantage not discussed in other studies.

2.6.2 Bidirectional Quantum Teleportation (Version II)

Notice that in BQT version I (Protocol 4) Alice performs a local operation on one of the Bell state qubits (A_2) while Bob′ (not Bob) performs a local operation on qubit B_2. In fact the resource $[qq]_{AB'}$ is sufficient. Similarly, Bob performs a local operation on one of the Bell state qubits (B_3) while Alice′ (not Alice) performs a local operation on qubit A_3.

Thus, the resource $[qq]_{A'B}$ is sufficient. Further, we have discussed how in these cases it is sufficient to consider $[qq]_{AN}$ instead of $[qq]_{AB'}$ as well as $[qq]_{BN}$ instead of $[qq]_{A'B}$ (since channels from Neha to Alice′ and Bob′ are considered free). Therefore, in this second version we start with $[qq]_{AN}$ and $[qq]_{BN}$ to show a new variation for BQT wherein, the Classical Butterfly Network Coding Protocol provides a similar advantage.

Protocol 5. Bidirectional Quantum Teleportation (Version II)

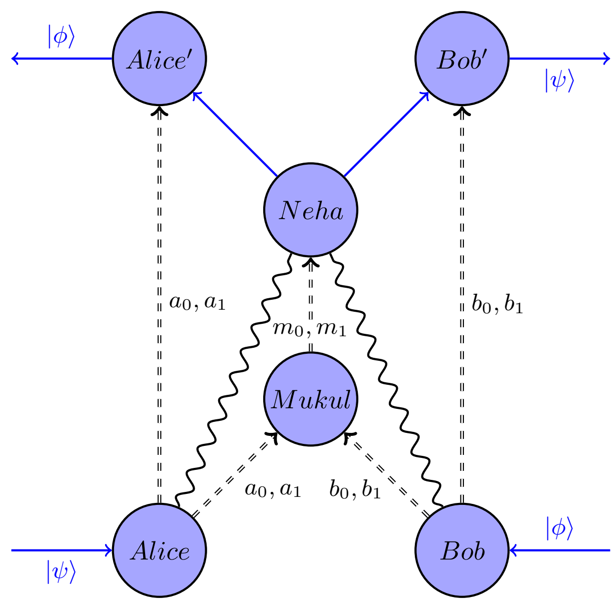


Figure 2.8: Bidirectional Quantum Teleportation (Version II)

Goal: Alice wishes to send qubit $|\psi\rangle$ to Bob and Bob wishes to send qubit $|\phi\rangle$ to Alice
Resources: $[qq]_{AN}, [qq]_{BN}, 2[c \to c]_{A\to M}, 2[c \to c]_{B\to M}, 2[c \to c]_{M\to N}$,
$[q \to q]_{N\to A'}, [q \to q]_{N\to B'}$.

1. Alice has qubit $|\psi_{A_1}\rangle$, Bob has qubit $|\phi_{B_1}\rangle$

2. Alice and Neha have a Bell pair $\left|\beta^{00}_{A_2N_1}\right\rangle$ and Bob and Neha have a Bell pair $\left|\beta^{00}_{B_2N_2}\right\rangle$

$$|\psi_{A_1}\rangle \otimes \left|\beta^{00}_{A_2N_1}\right\rangle \otimes \left|\beta^{00}_{B_2N_2}\right\rangle \otimes |\phi_{B_1}\rangle \tag{2.36}$$

3. Alice does a local Bell measurement $\{\left\langle\beta^{a_0a_1}_{A_1A_2}\right| |a_0, a_1 \in \{0,1\}\}$. We can again massage the state to write it in a linear superposition over the Bell basis.

$$|\psi_{A_1}\rangle \left|\beta^{00}_{A_2N_1}\right\rangle = \frac{1}{2}\begin{pmatrix} \left|\beta^{00}_{A_1A_2}\right\rangle |\psi_{N_1}\rangle + \left|\beta^{01}_{A_1A_2}\right\rangle Z_{N_1} |\psi_{N_1}\rangle \\ + \left|\beta^{10}_{A_1A_2}\right\rangle X_{N_1} |\psi_{N_1}\rangle + \left|\beta^{11}_{A_1A_2}\right\rangle Z_{N_1}X_{N_1} |\psi_{N_1}\rangle \end{pmatrix} \tag{2.37}$$

4. Bob does a local Bell measurement $\{\left\langle\beta^{b_0b_1}_{B_2A_1}\right| |b_0, b_1 \in \{0,1\}\}$. We can again massage the state to write it in a linear superposition over the Bell basis.

$$|\phi_{B_1}\rangle \left|\beta^{00}_{B_2N_2}\right\rangle = \frac{1}{2}\begin{pmatrix} \left|\beta^{00}_{B_1B_2}\right\rangle |\psi_{N_2}\rangle + \left|\beta^{01}_{B_1B_2}\right\rangle Z_{N_2} |\psi_{N_2}\rangle \\ + \left|\beta^{10}_{B_1B_2}\right\rangle X_{N_2} |\psi_{N_2}\rangle + \left|\beta^{11}_{B_1B_2}\right\rangle Z_{N_2}X_{N_2} |\psi_{N_2}\rangle \end{pmatrix} \tag{2.38}$$

5. At this point the quantum state of the remaining two qubits with Neha is:

$$Z^{a_1}_{N_1}X^{a_0}_{N_1} |\psi_{N_1}\rangle \otimes Z^{b_1}_{N_2}X^{b_0}_{N_2} |\phi_{N_2}\rangle \tag{2.39}$$

6. Alice sends the outcome (a_0, a_1) of her Bell Measurement $\{\left\langle\beta^{a_0,a_1}_{A_1A_2}\right| |a_0, a_1 \in \{0,1\}\}$ to Mukul

7. Bob sends the outcome (b_0, b_1) of his Bell measurement $\{\left\langle\beta^{b_0,b_1}_{B_1B_2}\right| |b_0, b_1 \in \{0,1\}\}$ to Mukul

8. Mukul produces two c-bits (m_0, m_1) with $m_0 = a_0 \otimes b_0$ and $m_1 = a_1 \otimes b_1$ and sends them to Neha

9. Neha applies the gate $X^{m_0}Z^{m_1}$ to qubit N_1 and the same gate $X^{m_0}Z^{m_1}$ to qubit N_2 as well

$$Z^{b_1}_{N_1}X^{b_0}_{N_1} |\psi_{N_1}\rangle \otimes Z^{a_1}_{N_2}X^{a_0}_{N_2} |\phi_{N_2}\rangle \tag{2.40}$$

10. Neha sends qubit N_2 to Alice and qubit N_1 to Bob.

$$Z_B^{b_1} X_B^{b_0} |\psi_B\rangle \otimes Z_A^{a_1} X_A^{a_0} |\phi_A\rangle \tag{2.41}$$

11. Alice applies the gate $X^{a_0} Z^{a_1}$ to qubit A and Bob applies the gate $X^{b_0} Z^{b_1}$ to qubit B to complete bidirectional teleportation!

$$|\psi_B\rangle \otimes |\phi_A\rangle \tag{2.42}$$

Here, we have a hybrid variation of the Classical Butterfly Network Coding Protocol employed till Neha followed by Neha sending over the qubits to Alice and Bob. This variation provides the same advantage in terms the channel from Mukul to Neha: two c-bits (m_0, m_1) (as opposed to four c-bits (a_0, a_1, b_0, b_1)); nonetheless it has its own advantages. Depending on what resources are available in a Quantum Network, if it is easier to have shared entanglement with Neha instead of between Alice and Bob this variation can be useful.

Further given BQT Protocol 4 and BQT Protocol 5, one may also easily write down protocols starting with $([qq]_{AB}$ and $[qq]_{AN})$ or $([qq]_{AB}$ and $[qq]_{BN})$. The possibility of multiple variations for BQT in the Quantum Butterfly Network that share the same advantage from network coding will no doubt allow for flexibility at the level of implementations, but also shed light on the compatibility of classical communication advantages transferring to quantum communication.

We point out that both BQT Protocols 4 and 5 are compatible with *entanglement swapping*, where instead of pure qubits if Alice or Bob were to send qubits which are entangled with other systems, the entanglement will transfer to the teleported qubits as well.

Here we share the takeaways for the BQT protocols, both of which provide the communication advantage of two c-bits being sent from Mukul to Neha as opposed to four c-bits. We proceed to BQDC in the next section.

Protocol	Resources (non-free)	Features
BQT I (Protocol 4)	$2[qq]_{AB}, 2[c \to c]_{M \to N}$	· CBNC and QT cleanly distinguished · Requires $2[c \to c]_{M \to N}$ (not 4 c-bits)
BQT II (Protocol 5)	$[qq]_{AN}, [qq]_{BN}, 2[c \to c]_{M \to N}$	· Hybrid version of CBNC and QT · Requires $2[c \to c]_{M \to N}$ (not 4 c-bits)

Table 2.2: Takeaway for Bidirectional Quantum Teleportation Protocols

2.7 Bidirectional Quantum Dense-Coding

In Bidirectional Dense-Coding Alice a_0, a_1 and Bob b_0, b_1 want to send two bits each across to the other. With classical channels one can employ the Classical Butterfly Network Coding Protocol 3 but in the absence of a classical channel between Mukul and Neha one can send information using a qubit channel and entanglement through dense-coding. We present three protocols for BQDC, each more secure than the previous, and each providing the advantage of requiring one qubit (not two) to be sent from Mukul to Neha.

2.7.1 Bidirectional Quantum Dense-Coding (Version I)

As a warm up, in this version we look for the easiest implementation of BQDC in the Butterfly Network where we use Network Coding and Dense-Coding to find an advantage.

Protocol 6. Bidirectional Quantum Dense-Coding (Version I)

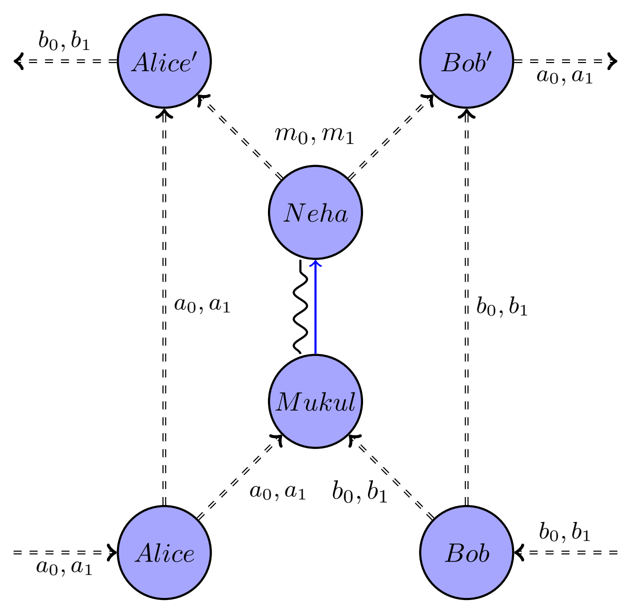


Figure 2.9: Bidirectional Quantum Dense-Coding (Version I)

Goal: Alice wishes to send two c-bits (a_0, a_1) to Bob. Bob wishes to send two c-bits (b_0, b_1) to Alice
Resources: $[qq]_{MN}, [q \to q]_{M\to N}, 2[c \to c]_{A\to M},\ 2[c \to c]_{B\to M},\ 2[c \to c]_{N\to A'},$
$2[c \to c]_{N\to B'}$

1. Mukul and Neha share a Bell pair $|\beta^{00}_{MN}\rangle$
2. Alice sends (a_0, a_1) to Mukul and Bob sends (b_0, b_1) to Mukul
3. Mukul prepares two c-bits similar to Network Coding Protocol (m_0, m_1)

$$m_0 = a_0 \oplus b_0 \quad \text{and} \quad m_1 = a_1 \oplus b_1 \tag{2.43}$$

4. Mukul encodes (m_0, m_1) in the Bell state by applying $X^{m_0}Z^{m_1}$ on his half of the Bell state

$$X^{m_0}_M Z^{m_1}_M \left|\beta^{00}_{MN}\right\rangle = |\beta^{m_0m_1}_{MN}\rangle \tag{2.44}$$

5. Mukul sends his half of the Bell pair to Neha $|\beta^{m_0m_1}_{NN}\rangle$
6. Neha performs a Bell measurement $\{\langle\beta^{m_0m_1}_{NN}|\,|m_0, m_1 \in \{0,1\}\}$ on the two qubits to recover the two bits of information (m_0, m_1)
7. Neha broadcasts (m_0, m_1) to Alice and Bob
8. Alice adds (a_0, a_1) from her memory to (m_0, m_1) to receive (b_0, b_1)

$$b_i = m_i \oplus a_i \quad \forall i \in \{0,1\} \tag{2.45}$$

9. Bob adds (b_0, b_1) from his memory to (m_0, m_1) to receive (a_0, a_1)

$$a_i = m_i \oplus b_i \quad \forall i \in \{0,1\} \tag{2.46}$$

10. Alice receives (b_0, b_1) and Bob receives (a_0, a_1). BQDC complete.

Here, we employed dense-coding within the Classical Butterfly Network Coding Protocol which requires us to send two c-bits (m_0, m_1) (as opposed to four c-bits $(a_0, a_1), (b_0, b_1)$). The c-bits (m_0, m_1) through dense-coding can be sent over one qubit (as opposed to two qubits for $(a_0, a_1), (b_0, b_1)$) sent from Mukul to Neha. This should not be surprising and provides us with a simple protocol for BQDC that provides an advantage.

2.7.2 Bidirectional Quantum Dense-Coding (Version II)

The BQDC Protocol 6 presented above has a drawback. Unlike dense-coding where the information encoded is sent securely from Alice to Bob, here Mukul, a third party, has access to (a_0, a_1, b_0, b_1). Do Alice and Bob trust Mukul and Neha? If not, can Alice and Bob recover some privacy of the information they wish to send? One way to do so would be to use quantum encodings for $(a_0, a_1), (b_0, b_1)$.

Further, let us consider a different shared entanglement resource – a GHZ state shared between Alice, Bob, Neha (and Mukul) for this variation of BQDC. Note that we can start with $[qqq]_{ABN}$ and use the free channel from Alice to Mukul or Bob to Mukul to create $[qqqq]_{ABMN}$. But for symmetry between Alice and Bob in the protocol we start with $[qqqq]_{ABMN}$ and list other resources accordingly.

Protocol 7. Bidirectional Quantum Dense-Coding (Version II)

Goal: Alice wishes to send two c-bits (a_0, a_1) to Bob. Bob wishes to send two c-bits (b_0, b_1) to Alice
Resources: $[qqqq]_{ABMN}$[a], $[q \to q]_{M\to N}, [q \to q]_{A\to M}$, $[q \to q]_{B\to M}$, $2[c \to c]_{N\to A'}$, $2[c \to c]_{N\to B'}$

1. Before the protocol begins, Alice, Bob, Mukul and Neha share a GHZ state $|\beta^{0000}_{ABMN}\rangle$

2. Alice applies $X^{a_0}Z^{a_1}$ on her part of the GHZ state, Bob applies $X^{b_0}Z^{b_1}$ on his part of the GHZ state

$$(X_A^{a_0} \otimes X_B^{b_0})(Z_A^{a_1} \otimes Z_B^{b_1}) \left|\beta^{0000}_{ABMN}\right\rangle \tag{2.47}$$

$$=(|a_0 b_0 00_{ABMN}\rangle + (-1)^{m_1} \left|\bar{a_0}\bar{b_0}11_{ABMN}\right\rangle) \tag{2.48}$$

$$\text{where} \quad m_0 = a_0 \oplus b_0 \quad \text{and} \quad m_1 = a_1 \oplus b_1$$

3. Alice and Bob send their parts of the GHZ state to Mukul

$$|a_0 b_0 00_{M_1M_2M_3N}\rangle + (-1)^{m_1} \left|\bar{a_0}\bar{b_0}11_{M_1M_2M_3N}\right\rangle \tag{2.49}$$

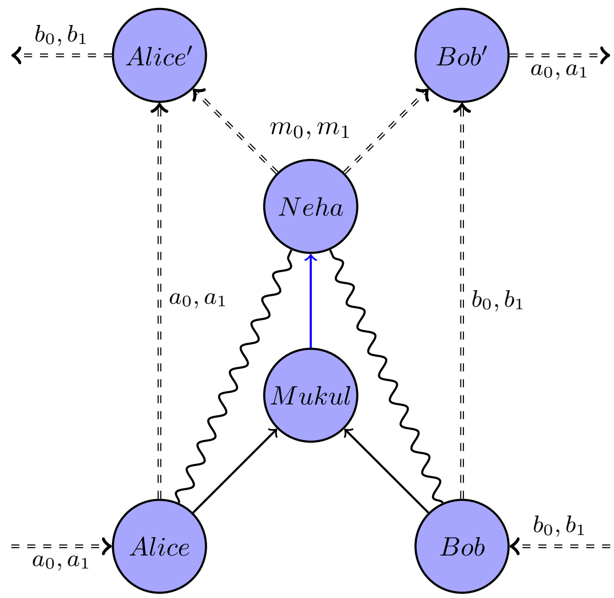


Figure 2.10: Bidirectional Quantum Dense-Coding (Version II)

4. Mukul's aim is to prepare the Bell state $|\beta_{MN}^{m_0m_1}\rangle$ from the GHZ state. Note that m_1 is already encoded in the phase of the GHZ state. To encode m_0 he applies C_{NOT} controlled from M_2 targeted to M_1, followed by a C_{NOT} controlled from M_3 targeted to M_1.

$$|a_0 b_0 00_{M_1M_2M_3N}\rangle + (-1)^{m_1} \left|\bar{a_0}\bar{b_0}11_{M_1M_2M_3N}\right\rangle \tag{2.50}$$

$$\xrightarrow[C_{NOT}^{M_2\rightarrow M_1}]{} |m_0\rangle_{M_1} \left(|b_0 00_{M_2M_3N}\rangle + (-1)^{m_1} \left|\bar{b_0}11_{M_2M_3N}\right\rangle\right) \tag{2.51}$$

$$\xrightarrow[C_{NOT}^{M_3\rightarrow M_1}]{} |m_0 b_0 00_{M_1M_2M_3N}\rangle + (-1)^{m_1} \left|\bar{m_0}\bar{b_0}11_{M_1M_2M_3N}\right\rangle \tag{2.52}$$

5. Now Mukul wishes to discard two qubits to prepare the bell state. He does so by applying C_{NOT} controlled from M_1 targeted to M_3, followed by a C_{NOT} controlled from M_1 targeted to M_2

$$\left|m_0 b_0 00_{M_1 M_2 M_3 N}\right\rangle + (-1)^{m_1}\left|\bar{m_0}\bar{b_0}11_{M_1 M_2 M_3 N}\right\rangle \tag{2.53}$$

$$\xrightarrow[C_{NOT}^{M_1\to M_2}C_{NOT}^{M_1\to M_3}]{} \left|\beta^{m_0 m_1}_{M_1 N}\right\rangle \left|a_0\right\rangle_{M_2} \left|m_0\right\rangle_{M_3} \tag{2.54}$$

6. Mukul discards M_2, M_3 and sends M_1 to Neha. Neha has the state $\left|\beta^{m_0 m_1}_{N_1 N_2}\right\rangle$

7. Neha performs a Bell measurement $\{\left\langle\beta^{m_0 m_1}_{N_1 N_2}\right| | m_0, m_1 \in \{0,1\}\}$ on the two qubits to recover the two bits of information (m_0, m_1)

8. Neha broadcasts (m_0, m_1) to Alice and Bob

9. Alice adds (a_0, a_1) from her memory to (m_0, m_1) to receive (b_0, b_1)

$$b_i = m_i \oplus a_i \quad \forall i \in \{0,1\} \tag{2.55}$$

10. Bob adds (b_0, b_1) from his memory to (m_0, m_1) to receive (a_0, a_1)

$$a_i = m_i \oplus b_i \quad \forall i \in \{0,1\} \tag{2.56}$$

11. Alice receives (b_0, b_1) and Bob receives (a_0, a_1). BQDC complete.

[a] Note that we can start with $[qqq]_{ABN}$ and use the free channel from Alice to Mukul or Bob to Mukul to create $[qqqq]_{ABMN}$. But for symmetry between Alice and Bob we start with $[qqqq]_{ABMN}$

We encoded the information (m_0, m_1) in a quantum manner in this protocol. Nonetheless, the bits m_1 and m_0 are encoded at different levels of security. Here, m_1 is encoded into phase of the GHZ (later to be reduced to Bell state) and therefore makes a_1 and b_1 inaccessible to Mukul or Neha; on the other hand m_0 is quantum-ly encoded by Mukul in Step 4 and 5, and in Equation 2.54 it is evident that Mukul could retrieve the bits a_0 and b_0 if he wishes to. Notice that the protocol hinges on Mukul having the three qubits (Equation 2.50): one with information of a_0, one with information of b_0 and one extra qubit — while this third qubit helps keep track of parity essential for the protocol to succeed, it also makes a_0 and b_0 accessible to Mukul. We will see the third protocol which is completely secure due to a Bell state shared between Alice and Bob.

Nonetheless, we learn that we can perform BQDC in multiple ways with the same advantage (one qubit sent from Mukul to Neha as opposed to two qubits). We also learn a lot from this Protocol 7 specifically since we can implement BQDC using a GHZ state $[qqq]_{ABN}$, while BQT II Protocol 5 necessarily requires $[qq]_{AN}$ and $[qq]_{BN}$. This is interesting in light of the resource inequality:

$$[qq]_{AN} + [qq]_{BN} \geq [qqq]_{ABN} \tag{2.57}$$

This will be relevant for our discussion of the possible duality between BQT and BQDC in the Section 2.8.1.

2.7.3 Masked Encoding and Security

We discuss issues of security in some detail before we present the third version of BQDC. Within the Protocols in the (Classical and Quantum) Butterfly Network, Mukul has the task of encoding information to be sent to Neha, this leaves some information accessible to Mukul. In the case of BQT, the accessibility of the Bell measurement outcomes that Mukul encodes does not pose as a threat to security, since these are not the actual information (qubits being teleported) Alice and Bob wish to share. On the other hand, in the case of CBNC and BQDC, Mukul has access to the very information that Alice and Bob wish to share with each other. This is undesirable from a security point of view. While Alice and Bob require Mukul and Neha to cooperate for the success of the protocols, it would be also be desirable that they cannot extract the information being communicated between Alice and Bob.

We define the term *Masked Encoding* when the information encoded is not accessible to the encoder, in other words, the information is masked from the encoder (here Mukul).

The information accessible to Mukul (the encoder), in the protocols involving the Butterfly Network, is as follows:

- **CBNC Protocol 3**: Mukul has access to a and b
- **BQT I Protocol 4**: Mukul does *not* have access to $|\psi\rangle$ or $|\phi\rangle$
- **BQT II Protocol 5**: Mukul does *not* have access to $|\psi\rangle$ or $|\phi\rangle$
- **BQDC I Protocol 6**: Mukul has access to (a_0, a_1) and (b_0, b_1)
- **BQDC II Protocol 7**: Mukul has access to (a_0,b_0) but *not* (a_1,b_1)
- **BQDC III Protocol 8 (below)**: Mukul does *not* have access to (a_0, a_1) and (b_0, b_1)

2.7.4 Bidirectional Quantum Dense-Coding (Version III)

We have attempted to use quantum encoding in BQDC II Protocol 7 and we were (only) partially successful in masking information from Mukul. In the protocol we present below we build on BQDC I Protocol 6 with an additional Bell pair shared between Alice and Bob, where we achieve masked encoding.

Protocol 8. Bidirectional Quantum Dense-Coding (Version III)

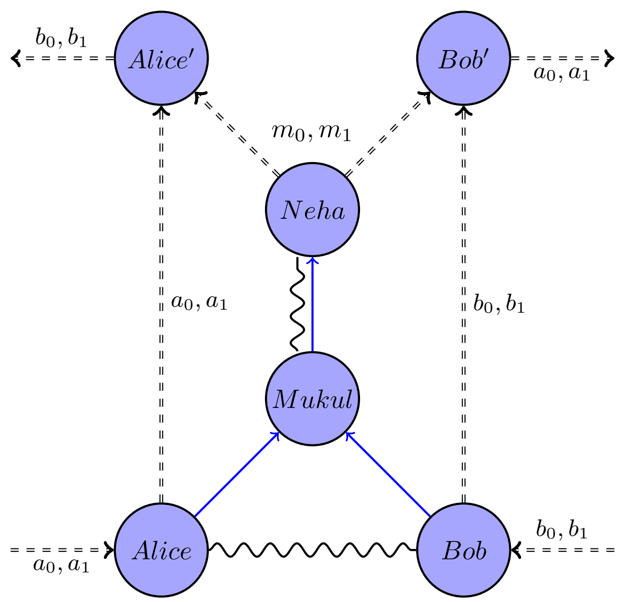


Figure 2.11: Bidirectional Quantum Dense-Coding (Version III)

Goal: Alice wishes to send two c-bits (a_0, a_1) to Bob and Bob wishes to send two c-bits (b_0, b_1) to Alice, through masked encoding

Resources: $[qq]_{AB}, [qq]_{MN}, [q \to q]_{M\to N}, [q \to q]_{A\to M}$, $[q \to q]_{B\to M}$, $2[c \to c]_{N\to A'}$, $2[c \to c]_{N\to B'}$

1. Alice, Bob receive a Bell state $|\beta^{00}_{AB}\rangle$. Mukul and Neha receive Bell state $|\beta^{00}_{MN}\rangle$.

2. Alice applies $X^{a_0}Z^{a_1}$ on her part of the Bell state, Bob applies $X^{b_0}Z^{b_1}$ on his part of the Bell state

$$(X_A^{a_0} \otimes X_B^{b_0})(Z_A^{a_1} \otimes Z_B^{b_1}) \left|\beta_{AB}^{00}\right\rangle = (|a_0 b_{0AB}\rangle + (-1)^{m_1} \left|\bar{a_0}\bar{b}_{0AB}\right\rangle) \equiv |\beta_{AB}^{m_0 m_1}\rangle \tag{2.58}$$

$$\text{where} \quad m_0 = a_0 \oplus b_0 \quad \text{and} \quad m_1 = a_1 \oplus b_1$$

3. Alice and Bob send their parts of the Bell state to Mukul

$$(|a_0 b_{0M_1M_2}\rangle + (-1)^{m_1} \left|\bar{a_0}\bar{b}_{0M_1M_2}\right\rangle) \equiv \left|\beta_{M_1M_2}^{m_0 m_1}\right\rangle \tag{2.59}$$

4. Mukul performs a Bell measurement $\{\left\langle\beta_{M_1M_2}^{m_0m_1}\right| |m_0, m_1 \in \{0,1\}\}$ to get (m_0, m_1)
5. Mukul applies $X^{m_0}Z^{m_1}$ on his part of the Bell state shared with Neha $|\beta_{MN}^{00}\rangle$ and sends his half of the Bell state to Neha
6. Neha performs a Bell measurement $\{\left\langle\beta_{N_1N_2}^{m_0m_1}\right| |m_0, m_1 \in \{0,1\}\}$ on the two qubits to recover the two bits of information (m_0, m_1)
7. Neha broadcasts (m_0, m_1) to Alice and Bob
8. Alice adds (a_0, a_1) from her memory to (m_0, m_1) to receive (b_0, b_1)

$$b_i = m_i \oplus a_i \quad \forall i \in \{0,1\} \tag{2.60}$$

9. Bob adds (b_0, b_1) from his memory to (m_0, m_1) to receive (a_0, a_1)

$$a_i = m_i \oplus b_i \quad \forall i \in \{0,1\} \tag{2.61}$$

10. Alice receives (b_0, b_1) and Bob receives (a_0, a_1). BQDC complete.

Note that the shared entanglement resource in this protocol is $[qq]_{AB}$ and cannot be replaced by $[qq]_{AB'}$ or $[qq]_{A'B}$. We leave it as an open question if one can achive masked encoding starting with a different shared entanglement resource.

Here we share the takeaways for the BQDC protocols, all of which share the communication advantage of requiring Mukul to send a single qubit to Neha (as opposed to two):

Protocol	Non Free Resources	Novelty
BQDC I	$[qq]_{MN}, [q \to q]_{M\to N}$	· Requires only one Bell pair · Requires $[q \to q]_{M\to N}$ not $2[q \to q]_{M\to N}$
BQDC II	$[qqqq]_{ABMN}, [q \to q]_{M\to N}$	· Partially secure with GHZ state · Requires $[q \to q]_{M\to N}$ not $2[q \to q]_{M\to N}$
BQDC III	$[qq]_{AB}, [qq]_{MN}, [q \to q]_{M\to N}$	· Requires one Bell pair $[qq]_{MN}$ · Secure with masked encoding with one Bell pair $[qq]_{AB}$ · Requires $[q \to q]_{M\to N}$ not $2[q \to q]_{M\to N}$

Table 2.3: Takeaway for Bidirectional Quantum Dense-Coding Protocols

2.8 Discussion

In this final Section, we highlight possible impact of this Chapter's work and discuss some interesting directions leading to avenues of future research.

2.8.1 Duality within Bidirectional Teleportation and Dense-Coding

In Subsection 2.4.3 we discussed the resource theory within which teleportation and dense-coding are dual. A natural and interesting question to ask arises:

Is there some version of BQT and BQDC in the butterfly network that are dual?

To fully answer this question one is required to extend the resource theory to multiple parties (including a resource theory of shared entanglement) and systematically define achievable regions. We provide some steps towards solving this open question. We learnt from the duality of teleportation and dense-coding that the coherent versions of protocols are key to showing duality between them. Coherent Dense-Coding is stronger than its classical counterpart and Coherent Teleportation is weaker. Similarly, it might be useful to rework the protocols in this Chapter for BQT and BQDC to their Coherent versions before duality may be analysed. While we leave this question open, to begin with we provide the Coherent version of the Classical Butterfly Network Coding Protocol 3.

Protocol 9. Coherent-Classical Butterfly Network Coding

Goal: Alice wishes to send one bit, $|a\rangle$, coherently to Bob and Bob wishes to send one bit, $|b\rangle$, coherently to Alice
Resources: $[q \to q]_{A \to M}, [q \to q]_{B \to M}, [q \to q]_{M \to N}, [q \to q]_{N \to A'}, [q \to q]_{N \to B'}$

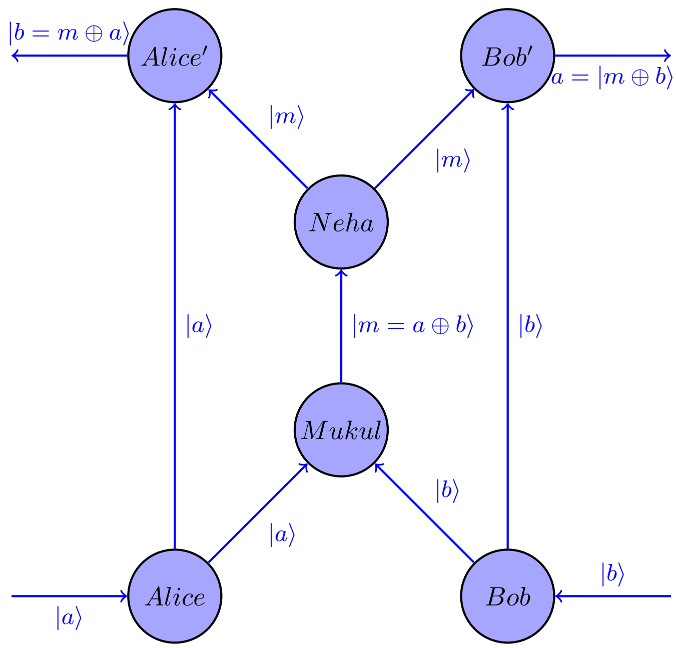


Figure 2.12: Coherent Classical Butterfly Network

1. Alice has $|a\rangle$ and prepares $|a\rangle|a\rangle$ and sends one qubit to Mukul.

2. Bob has $|b\rangle$ and prepares $|b\rangle|b\rangle$ and sends one qubit to Mukul.

3. The state at this step is $|a\rangle_A |a\rangle_{M_1} |b\rangle_{M_2} |b\rangle_B$. Mukul applies C_{NOT} from M_2 to M_1 and discards M_2 to leave the state as $|a\rangle_A |m\rangle_{M_1} |b\rangle_B$ (this is equivalent to the boolean sum step $m = a \oplus b$ in the classical counterpart). Mukul then sends M_1 to Neha.

4. Neha has a qubit in state $|0\rangle$. The state at this step is $|a\rangle_A |m\rangle_{N_1} |0\rangle_{N_2} |b\rangle_B$. She applies C_{NOT} from N_1 to N_2 to give $|a\rangle_A |m\rangle_{N_1} |m\rangle_{N_2} |b\rangle_B$. Neha then sends N_1 to Alice and N_2 to Bob to give $|a\rangle_{A_1} |m\rangle_{A_2} |m\rangle_{B_2} |b\rangle_{B_1}$.

5. Alice applies C_{NOT} from A_1 to A_2.

6. Bob applies C_{NOT} from B_1 to B_2.

7. Final state is $|a\rangle_{A_1} |b\rangle_{A_2} |a\rangle_{B_2} |b\rangle_{B_1}$.

Further, another hurdle in discussing the possible duality between BQDC and BQT is the comparison of shared entanglement. Either one may start with a version of BQT and BQDC which begin with the same shared entanglement resource or one must be able to compare the resources shared by different parties. The closest that we can come to this are BQT II and BQDC II though in the former we have two Bell pairs ($[qq]_{AN}, [qq]_{BN}$) and in the latter we have a three party GHZ state ($[qqq]_{ABN}$). One may attempt to find a masked encoded version of BQDC II that requires $[qq]_{AN}, [qq]_{BN}$ to further help study the possible duality, but it is not clear if such a protocol would be possible.

2.8.2 Quanta-Net and the Boundary Rule

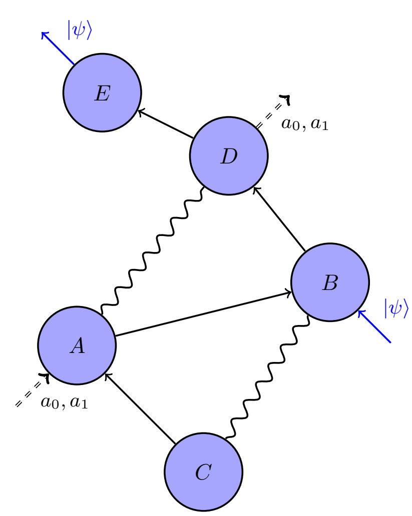


Figure 2.13: An example for a Quantum Network

While designing tasks over a "Quanta-Net" one may be faced with decisions on optimising a set of communication tasks of the form {classical information: for example $(a_0, a_1)_{A\to D}$, quantum information: for example $|\psi\rangle_{B\to E}$} given some set of resources {classical channels, quantum channels, shared entanglement} in a multi-partite setting.

A conceptually interesting yet simple rule of thumb emerges from observing the protocols within this Chapter, that can help prescribe between which nodes shared entanglement would aid in performing certain quantum and classical communication tasks in a large network. First one may draw the networks through their spatial perspective (unlike our distinction of Alice/Alice′..., we now represent them by the same node). Given the nature of Dense-Coding (uses entanglement to encode classical information on quantum channels) and Teleportation (uses entanglement to encode quantum information on classical channels) and the protocols in this chapter the following observation emerges:

> **Boundary Rule:** In the Spatial Perspective of a Network, draw a boundary along nodes that have incoming (outgoing) quantum channels and outgoing (incoming) classical channels. The nodes that this boundary joins require shared entanglement for optimal utilisation of the channel capacity.

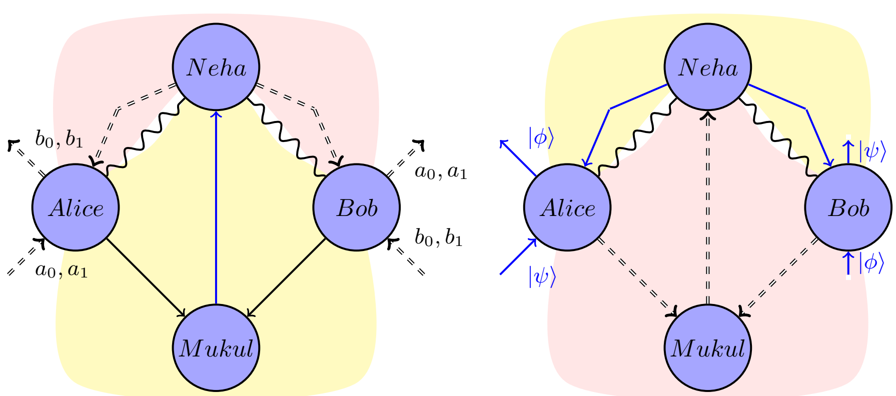


Figure 2.14: Spatial Perspective: BQDC II (left) and BQT II (right)

We show the spatial diagrams (Figure 2.14) for BQT II and BQDC II which serve as interesting examples, since for both the boundary lies along Alice-Neha-Bob (where we shade classical channels as red and quantum channels as blue). Note that the rule does not prescribe the nature of the shared resource (GHZ, Bell, etc), it only provides the nodes that would require shared entanglement. Formalising the Boundary Rule might not be as straightforward, nonetheless it helps us get some insight in tackling difficult network coding tasks.

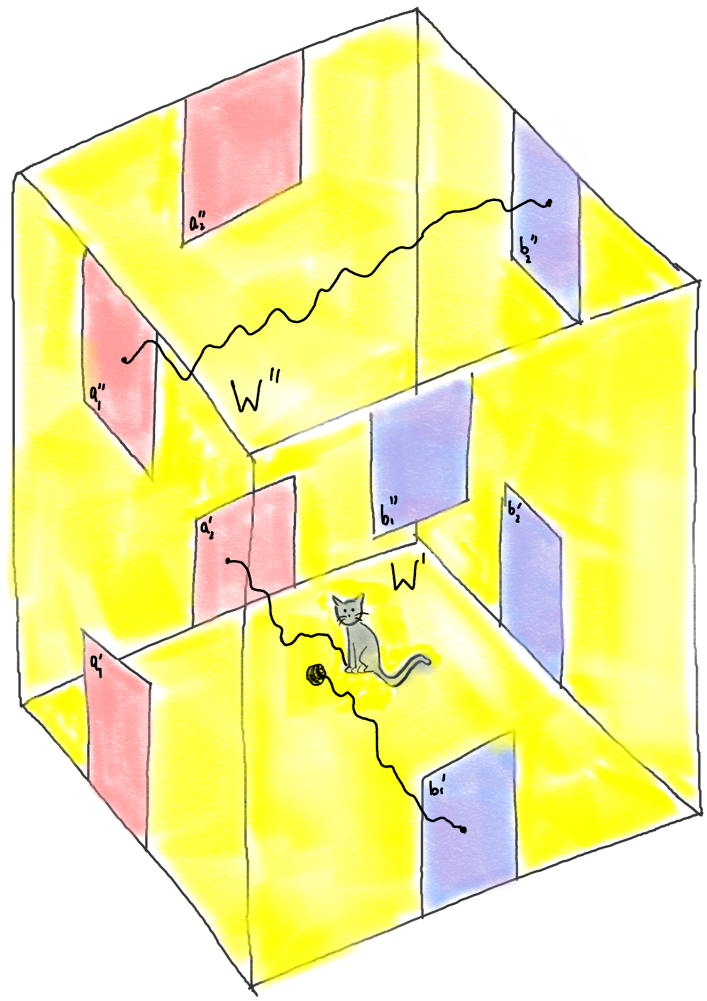
2.8.3 Some final thoughts

Studying multiparty generalisations of Teleportation and Dense-coding such as those in this Chapter may have impact on Quantum Communication tasks over a Quantum Internet. While Classical Communication is cheaper, features such as *Masked Encoding* might be lucrative for security purposes. Generalisations to resource theory of quantum communication to multiple parties may help answer the open question if there are versions of BQDC and BQT that can be shown to be dual. Further, generalisations to post-quantum theories (as was our initial motivation) is one path ahead for this work. Would Teleportation and Dense-Coding analogues be supported in theories with indefinite causality?

In the following Chapters we move to studies within post-quantum theories with indefinite causal structures as promised in the title of the thesis. We will encounter Process Matrices (Chapter 3), the Causaloid Framework (Chapter 4, 5, 6) and the Duotensor Framework (Chapter 6).

Chapter 2: Statement of Contribution

In this Chapter, the main contributions include protocols for Bidirectional Teleportation (Section 2.6) and Bidirectional Dense-Coding (Section 2.7) implemented within the Quantum Butterfly Network of four parties. Namely Protocols 4-8 from Table 2.1 are contributions of this work. I took the lead on this project and the Chapter is written solely by me. It contains unpublished material, presented at various conferences including Quantum Networks Oxford 2017, ICQF India 2017, and YQIS Vienna 2018. The work was done under the supervision of Lucien Hardy. It is a continuation from the my Perimeter Scholar International (PSI) Essay (part of my Master's degree with U Waterloo with PSI certification from Perimeter Institute).



Process in a world
products like yarn unfolding
must never return
lest loop time
and vanish [4]

[4] Chapter 3, Conceptual illustration for the invalid Process Product from Equation 3.4 from Section 3.3

Chapter 3

Tensor Products of Process Matrices

Tensor products over Quantum Channels help quantify Channel Capacities.
Similarly, can tensor products over Process Matrices help quantify Process Capacities?
We present a theorem that qualifies the restriction over tensor products of Processes.

Theories with indefinite causal structure have been studied from both the fundamental perspective of quantum gravity and the practical perspective of information processing as discussed in Section 1.3. Two indefinite causal structure frameworks of interest for the results of this Chapter are the Process Matrix [82] and Quantum Combs [18]. In this Chapter[1], we point out a restriction in forming tensor products of objects with indefinite causal structure within these frameworks: there exist both classical and quantum objects whose tensor products violate the normalisation condition of probabilities, if all local operations are allowed. We obtain a necessary and sufficient condition (Theorem 2) for when such unrestricted tensor products of multi-partite objects are (in)valid. This poses a challenge to extending communication theory to indefinite causal structures, as the tensor product is the fundamental ingredient in the asymptotic setting of communication theory. We discuss a few options to evade this issue. In particular, we show that the sequential asymptotic setting does not suffer the violation of normalisation.

[1] This Chapter is reproduced with modifications from the paper with the same title on work done with Ding Jia [67]

3.1 Towards a Theory of 'Process Communication'..

In the previous Chapter we focused on Quantum Communication tasks such as Teleportation as well as Dense-Coding (motivated from wanting to perform two-way signalling game for theories with indefinite causal structure). Such tasks fall under the purview of Communication theory, at the heart of which lies the physical theory through which communication happens. Going from Classical Communication theory to Quantum Communication required embracing qubits as the unit of information. Going from Quantum Communication (with definite causal structure) to Quantum Communication with indefinite causal structure (let's call this Process Communication) will require understanding the role of (indefinite) causality in communication.

The main motivation of the present work is to pave the way to develop communication theory for frameworks with indefinite causal structure, specifically for the notion of asymptotic capacities. In the usual Shannon asymptotic setting one takes multiple copies of the communication resource such as a channel or a state to define capacity. The result in this paper shows that this setting cannot be extended to quantum theory with indefinite causal structure with straightforward application of tensor products that imposes no restrictions on the allowed local operations. The reason is evidently precisely the indefiniteness of order of events that can lead to causal inconsistencies mathematically seen through the violation of normalisation of probabilities. Further, Guérin et al. [37] considered if maps other than the tensor product could circumvent the restriction provided in this Chapter, through a no-go theorem, to find it would not be possible. Nonetheless, there is a silver lining, we show that if one carefully distinguishes the parallel and sequential asymptotic settings, then the sequential asymptotic setting can be extended to objects with indefinite causal structure salvaging some of the techniques from the Shannon asymptotic setting for developing communication theory for frameworks with indefinite causal structure.

Outline of the Chapter

The major results of this Chapter (Theorem 2, Corollary 2.1, Corollary 2.2) are based on a lemma (Proposition 1) whose content is phrased and proved by Oreshkov and Giarmatzi [83] in the process matrix framework. Note that the present work can equally be carried out in other frameworks such as the Quantum Combs framework [18, 20]. For convenience of directly applying the lemma (Proposition 1) we base the study within the process matrix framework. We provide examples for the certain invalid process products and subsequently discuss the implications of these results.

3.2 Processes

In this Section, we recall the relevant part of the Process Matrix Framework [82, 2, 83] and introduce the nomenclature needed to present our result [66].

Process Matrix Framework

The process matrix framework has two types of objects: closed labs and processes. The closed labs have one input and one output restriction and mathematically can be represented by channels (or Completely Positive Trace-Preserving or CPTP maps). These closed labs are postulated to locally obey quantum theory with definite causal structure. Globally the causal structure can be indefinite modelled through the processes. A process (W) is a linear map from labs (channels) describing local physics to real numbers describing probabilities of observation outcomes. Both channels and states can be seen as special cases of processes.

We use $A, B, C, \cdots$ to denote the parties where local physics takes place. A 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 $\mathcal{H}^{a_2}$. Through the Choi isomorphism [22] processes can be represented as linear operators. A process W associated with parties $A, B, C, \cdots$ is represented as a linear operator $W^{a_1 a_2 b_1 b_2 c_1 c_2 \cdots} \in L(\mathcal{H})$, where $\mathcal{H} := \mathcal{H}^{a_1} \otimes \mathcal{H}^{a_2} \otimes \mathcal{H}^{b_1} \otimes \mathcal{H}^{b_2} \otimes \mathcal{H}^{c_1} \otimes \mathcal{H}^{c_2} \otimes \cdots$. We will sometimes combine the input and output Hilbert spaces and write the process as $W^{abc\cdots}$ for simplicity.

Here is an example of how a Bipartite Process will look like with parties A and B with associated Hilbert spaces for input systems a_1, b_1 and output systems a_2, b_2:

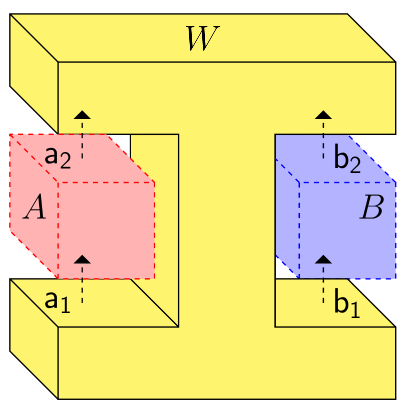


Figure 3.1: Bipartite Process

Physical Conditions on Processes

Similar to states and channels, it is assumed that processes yield positive and normalised outcome probabilities for physical reasons, and can also act on subsystems of local parties. These imply that the process obey the conditions below where the first is a positivity condition and the second is a normalisation of probability condition.

$$W \geq 0 \qquad \text{Positivity} \tag{3.1}$$

$$\operatorname{Tr} W = \dim(\mathcal{H}^{a_2} \otimes \mathcal{H}^{b_2} \otimes \mathcal{H}^{c_2} \otimes \cdots) =: d_O \qquad \text{Normalisation} \tag{3.2}$$

In addition we have:

Proposition 1 (Oreshkov and Giarmatzi [83], rewritten using the language of this chapter)**.** A multi-partite process obeys the normalisation of probability condition if and only if in addition to the identity term it contains at most terms which are type a_1 on some party A.

Types

To understand this proposition we need to introduce the notion of types. A process $W^{ab\cdots}$ can be expanded in the Hilbert-Schmidt basis $\{\sigma_i^x\}_{i=0}^{d_x^2-1}$ of the x subsystem operators $L(\mathcal{H}^x)$ as

$$W^{ab\cdots} = \sum_{i,j,k,l,\cdots} w_{ijkl\cdots} \sigma_i^{a_1} \otimes \sigma_j^{a_2} \otimes \sigma_k^{b_1} \otimes \sigma_l^{b_2} \otimes \cdots\ ,\ w_{ijkl\cdots} \in \mathbb{R}. \tag{3.3}$$

For example, for qubit systems the Pauli basis contains four elements where $\sigma_0 = \mathbb{1}$, and the Pauli operators $\sigma_i, i = 1, 2, 3$. We set the convention to take σ_0^x to be $\mathbb{1}$ for any subsystem x with dimension d. We refer to terms of the form $\sigma_i^x \otimes \mathbb{1}^{\text{rest}}$ for $i \geq 1$ as a type x term, $\sigma_i^x \otimes \sigma_j^y \otimes \mathbb{1}^{\text{rest}}$ for $i, j \geq 1$ as a type xy term etc. The identity term is referred to as a trivial term to be of trivial type.

Restricting attention to some party A, we say that a term is one of the following types on party A regardless of what type it is on the systems of other parties:

Term	Type
$\mathbb{1}$	Trivial
a_1	In-type
a_2	Out-type
a_1 or a_1a_2	Includes In-type
a_2 or a_1a_2	Includes Out-type

Table 3.1: Types of Process Terms

On two parties A and B, terms are of type:

- **Signalling** $A \to B$ if it is a_2b_1 or $a_1a_2b_1$ since A's output is correlated with B's input
- **Signalling** $B \to A$ if it is a_1b_2 or $a_1b_1b_2$ since B's output is correlated with A's input

Coming back to Proposition 1, we can interpret it as allowing for at most signalling channels of the form given above.

3.3 Conditions for forming Process Products

In this section we introduce process products and prove our main results.

Combining Parties

A party $\{A', a'_1, a'_2\}$ and a party $\{A'', a''_1, a''_2\}$ can be combined into a new party $\{A, a_1, a_2\} = \{A'A'', a'_1a''_1, a'_2a''_2\}$ if all channels from a_1 to a_2 can be applied by A. Here $A'A''$ is a shorthand notation for combining parties A' and A'', and xy is a system whose Hilbert space is $\mathcal{H}_x \otimes \mathcal{H}_y$.

Tensor Product on Channels

Such products of parties are implicitly used when one forms tensor products of channels. A channel M is a two-party resource that mediates information between some party A' and some party B'. Given any other channel N mediating information between A'' and B'', the tensor product $M \otimes N$ is a channel associated with $A = A'A''$ and $B = B'B''$, where all channels from a_1 to a_2 can be applied by A and all channels from b_1 to b_2 can be applied by B.

Tensor Product on Processes

Such tensor products are crucial in information theory, as one often studies tasks in the asymptotic setting, where the same resource is used arbitrarily many times. Out of interests, for example in quantum gravity we want to study information communication theory of processes with indefinite causal structure [64]. In order to consider the asymptotic setting for processes we need to define products of processes and check if they are valid processes. Analogous to channel products, for two processes $W^{a'b'\cdots}$ and $Z^{a''b''\cdots}$ with the same number of parties, we tentatively define their product as $P^{ab\cdots} = W^{a'b'\cdots} \otimes Z^{a''b''\cdots}$. It takes in channels of parties $A, B, \cdots$ and outputs probabilities where $A = A'A''$, $B = B'B''$, $\cdots$. The situation for two parties is illustrated in Figure 3.2.

The adaptation of the asymptotic setting in the processes framework needs some qualifications. First, in the context of processes a party represents a localised region of space-time. If multiple copies of a process are used, this introduces multiple copies of the parties. The most natural way to introduce asymptotic setting is to assume that the same parties share multiple copies of the process. Second, as we show below, not all processes allow such sharing.

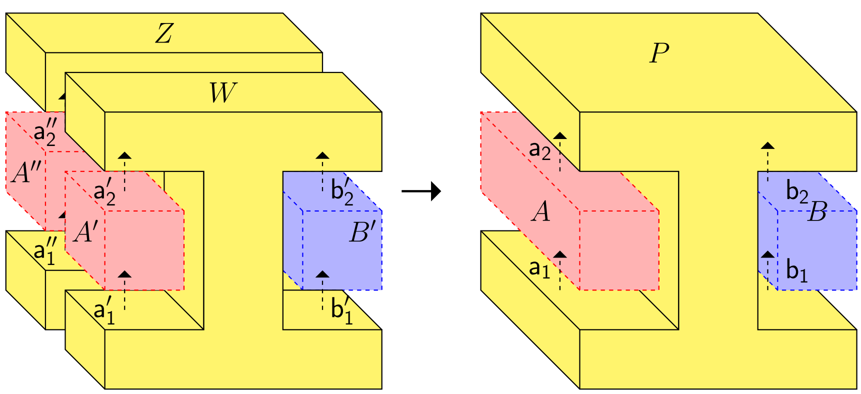


Figure 3.2: Process product for two bipartite processes [2]

Following this construction, the asymptotic setting of a two-party process W^{ab} would require a process $W^{ab\otimes n} = W^{a'b'} \otimes W^{a''b''} \cdots$ that is an n-fold tensor product. The product parties $A = A'A'' \cdots$ and $B = B'B'' \cdots$ each represents a localised region of spacetime where all channels are allowed.

[2]Tikz image made by Ding Jia and reproduced from our work [66]

Example of Invalid Process Product

This asymptotic setting without restriction on the allowed local operations holds without problems for quantum theory with definite causal structure. However, a simple example shows that if arbitrary local operations are allowed, process products are not always valid processes. Consider the process:

$$W^{xy} = \frac{d_O}{2}(\omega^{x_1} \otimes \rho_{\sim}^{x_2 y_1} \otimes \omega^{y_2} + \omega^{x_2} \otimes \rho_{\sim}^{x_1 y_2} \otimes \omega^{y_1}), \tag{3.4}$$

where d_O is the dimension of the outputs, ω is the maximally mixed state and $\rho_\sim := (\mathbb{1} + \sigma_3 \otimes \sigma_3)/4$ is a maximally correlating state that can represent a classical identity channel in the $\{|0\rangle, |1\rangle\}$ basis. This process can be viewed as an equal-weight classical mixture of a channel from x to y and another one from y to x. Suppose A' and B' share a process $W^{a'b'}$ of this form, and A'' and B'' also share a process $W^{a''b''}$ of the same form. The operator $W^{ab} := W^{a'b'} \otimes W^{a''b''}$ for the two parties $A = A'A''$ and $B = B'B''$ is not a valid process. $\rho_\sim^{a_2' b_1'} \otimes \rho_\sim^{a_1'' b_2''}$ includes a term $\sigma_3^{a_2'} \otimes \sigma_3^{a_1''} \otimes \sigma_3^{b_1'} \otimes \sigma_3^{b_2''}$, which leads to a type $a_1 a_2 b_1 b_2$ term and according to Proposition 1 renders the process W^{ab} invalid. Intuitively, $\rho_\sim^{a_2' b_1'} \otimes \rho_\sim^{a_1'' b_2''}$ creates a causal loop and violates the normalisation of probability condition.

Understanding Invalidity

Note that although $W^{a'b'}$ and $W^{a''b''}$ cannot be composed directly, it is possible to have a global process P^{ab} that reduces to the two individual processes upon partial tracing. For example, let A and B have a process of the same form as (3.4). This is a process on the combined parties. The reduced processes $\mathrm{Tr}_{a''b''} W^{ab}$ and $\mathrm{Tr}_{a'b'} W^{ab}$ are exactly $W^{a'b'}$ and $W^{a''b''}$.

The restriction of process products has an analogy with the "non-separability" of entangled states in quantum theory. If ρ^{xy} is entangled, then $\rho^{xy} \neq \rho^x \otimes \rho^y$. Similarly for some processes $W^{ab} \neq W^{a'b'} \otimes W^{a''b''}$. The difference is that for processes tensor products not only may not recover the original process, but may even be invalid.

Note that the processes in the example can be viewed as classical because one can regard it as a classical mixture of classical resources. One can also substitute Choi states of quantum channels for those of classical channels to obtain an example of quantum process that is restricted in forming products. The invalidity of arbitrary products is a feature of quantum as well as classical resources.

We also note that even for channels there exists a similar subtlety in forming products [7, 27, 69, 14]. Given a channel M from A' to B' and another one N from A'' to B'', the product $M \otimes N$ from $A'B''$ and $B'A''$ allows a causal loop and does not preserve probability. Because channels exist in ordinary quantum theory with definite causal structure, one may be tempted to say that the subtlety in forming products is not a new issue brought about by indefinite causal structure theories. It is however debatable whether the above construction is allowed by a theory with definite causal structure. For example, in the general framework of quantum networks [18] this kind of construction is explicitly forbidden by the first quantum combs tensor product rule. The rule requires the preservation of relative ordering of the original systems, and is well-motivated in the context of definite causal structure. In any case the restriction of tensor products is more manifest for processes than for channels. For M and N although one can not combine local parties as $A'B''$ and $B'A''$, one can always combine them as $A'A''$ and $B'B''$ to form a valid tensor product channel. On the other hand, it will be clear from the results below that there are processes W and s for which neither way of combination leads to a valid process.

3.3.1 Restriction on Process Product: a Theorem

The first main result of this Chapter, is the following necessary and sufficient condition characterising when two general multipartite processes cannot (and can) be composed into a valid product process when arbitrary local operations are allowed.

Theorem 2. The product $P = W \otimes Z$ of two processes W and Z is not a valid process if and only if there exists a nontrivial term of W and a nontrivial term of Z that obey two conditions:

1. On any party, where one term is trivial, the other is either trivial or includes the out-type.
2. On any party, where one term is the in-type, the other term includes the out-type.

Proof. Suppose W and Z satisfy the conditions and consider the tensor product of the two nontrivial terms. By Proposition 1, to prove that P is invalid we need to show that P contains a nontrivial term that is not type a_1 for any party A. Conditions 1) and 2) guarantee that this is satisfied for the product term we consider.

Conversely, suppose P is not a valid process. By Proposition 1, P contains a nontrivial term that is not type a_1 for any party A. This term must arise out of a tensor product of nontrivial terms over sub-parties A', A''. On any A this term of P is type trivial, a_1a_2 or a_2. We consider each case in turn. For any A where this term is trivial, it must come from a tensor product of terms that are trivial on A. Write this kind of tensor product as $(0,0)$, where 0 denotes the trivial type and $(.,.)$ denote unordered term pairs over A', A''. Next, for any A where this term is type a_1a_2, it must come from a tensor product of the kind $(0,12)$, $(1,12)$, $(1,2)$, $(2,12)$ or $(12,12)$, where 1 and 2 denote in- and out-type respectively. Finally, for any A where this term is type a_2, it must come from a tensor product of the kind $(0,2)$ or $(2,2)$. To sum up, on any A, this term of P comes from a tensor product of the kind $(0,0)$, $(0,12)$, $(0,2)$, $(1,12)$, $(1,2)$, $(2,12)$, $(2,2)$ or $(12,12)$. This implies conditions 1) and 2). □

A useful special case is the condition on two-party processes. Intuitively, the fulfillment of the two conditions in the corollary below give rise to causal loops, which violate the normalised probability condition for processes and hence lead to invalid products.

Corollary 2.1. A product $P^{ab} = W^{a'b'} \otimes Z^{a''b''}$ of two-party processes is not a valid process if and only if: 1) Both W and Z have signalling terms; 2) The Hilbert-Schmidt terms of W and Z put together contain signalling terms of both directions.

Proof. Suppose W and Z obey the two conditions. Then we can pick a signalling term from W of one direction and a signalling term from Z of the other direction. We show that this pair of terms satisfy conditions 1) and 2) in Theorem 2, and hence the product is not a valid process. Neither term is trivial on either of the two parties, so 1) of Theorem 2 is fulfilled. 2) is also fulfilled because the terms signal in different directions.

Conversely, suppose P is not valid. By Theorem 2, there is a nontrivial term from W and a nontrivial term from Z that obey 1) and 2) of Theorem 2. By Proposition 1, both terms are in-type on some party. By 2) of Theorem 2 they must be in-type on different parties, and they include the out-type on the parties where they are not in-type. In other words, they are signalling terms to different directions. This proves conditions 1) and 2) of the statement. □

A product of more than two processes can be constructed iteratively, and the validity of the product process must be checked at each step. If a set of processes cannot form a valid product in one sequence of construction, changing the sequence of construction will not make it valid. This is because the invalid term will always be present.

Corollary 2.1 allows us to straightforwardly identify those two-party processes for which the Shannon asymptotic setting without restriction on local operations is (in)valid.

Corollary 2.2. The n-fold tensor product $W^{ab\otimes n} = W^{a'b'} \otimes W^{a''b''} \cdots$ of a process W^{ab} with itself is not a valid process if and only if it contains signalling terms of both directions.

The above results show that the asymptotic setting without restriction on local operations does not hold for all processes. They suggest two ways to make sense of the asymptotic setting. One can either restrict attention to those processes that have valid products (as characterised by Theorem 2), or try to find a restricted set of local operations for which the products do not violate the normalisation condition of the framework. One option is to only allow non-signalling channels within each product party [27]. We show below that there is another perhaps more justified option, which allows more general local operations and has a clear physical interpretation. This is the sequential asymptotic setting.

Asymptotic Sequential Case

In general, the asymptotic setting can correspond to at least two physical settings. The first is the parallel setting, where two parties share many copies of a resource at the same time (Figure 3.2 depicts this type of tensor product, when time is taken to point upwards, and W and Z are taken to exist at "the same time step"). The second is the sequential setting, where two parties share one copy of a resource at many time steps. In the sequential setting, the local operation a party performs decomposes into operations at different time steps, and these operations follow a definite time sequence. This physical interpretation imposes a natural restriction on the local operations, which can be generalised to processes if different copies of the process appear in a definite temporal order. In this sequential setting, the tensor products of processes obey the normalisation condition.

N-fold Sequential Setting

Suppose n copies of a process W appear in a definite temporal order $W^{a'b'} \prec W^{a''b''} \prec \cdots \prec W^{a^{(n)}b^{(n)}}$. A can apply local operations to systems $a', a'', \cdots, a^{(n)}$, and B to systems $b', b'', \cdots, b^{(n)}$ that obey this temporal order. The local operations that A and B can apply to compose with and close all open systems to obtain probabilities are the so-called n-combs [18]. For A, such an n-comb takes the form $M^{a_2' a_2'' \cdots a_2^{(n)}}_{a_1' a_1'' \cdots a_1^{(n)}}$ and for B it takes the form $N^{b_2' b_2'' \cdots b_2^{(n)}}_{b_1' b_1'' \cdots b_1^{(n)}}$, where the systems obey the temporal order $a_1' \prec a_2' \prec a_1'' \prec a_2'' \prec \cdots \prec a_1^{(n)} \prec a_2^{(n)}$ and $b_1' \prec b_2' \prec b_1'' \prec b_2'' \prec \cdots \prec b_1^{(n)} \prec b_2^{(n)}$.

According to the "universality of quantum memory channels" theorem, the combs M and N can be decomposed into a sequence of memory channels, e.g., $M = M(1)^{e_1 a_2'}_{a_1'} M(2)^{e_2 a_2''}_{e_1 a_1''} \cdots M(n)^{a_2^{(n)}}_{e_{n-1} a_1^{(n)}}$, where $M(i)$ are channels at time steps i with e_i as memory systems that correlate the channels. Similarly, $N = N(1)^{f_1 b_2'}_{b_1'} N(2)^{f_2 b_2''}_{f_1 b_1''} \cdots N(n)^{b_2^{(n)}}_{f_{n-1} b_1^{(n)}}$. Then probability from composing the copies of W with M and N obey the normalisation condition:

$$\begin{aligned} M(\otimes W)N =& [M(1)^{e_1 a_2'}_{a_1'} W^{a_1' b_1'}_{a_2' b_2'} N(1)^{f_1 b_2'}_{b_1'}][M(2)^{e_2 a_2''}_{e_1 a_1''} W^{a_1'' b_1''}_{a_2'' b_2''} N(2)^{f_2 b_2''}_{f_1 b_1''}] \cdots \\ & [M(n)^{a_2^{(n)}}_{e_{n-1} a_1^{(n)}} W^{a_1^{(n)} b_1^{(n)}}_{a_2^{(n)} b_2^{(n)}} N(n)^{b_2^{(n)}}_{f_{n-1} b_1^{(n)}}] = 1 \end{aligned} \tag{3.5}$$

Within each square bracket there is a channel (including states and deterministic effects as special cases) operating on the memory systems, because a process composed with local channels with memory yields a channel [82]. In the end the composition of channels yields the number 1. If one substitutes sub-normalised operators in place of the combs to represent quantum instruments, then it is easy to see that the probabilities must be in the interval $[0, 1]$ and sum to one. Therefore the sequential asymptotic setting generalises to quantum theory with indefinite causal structure. Intuitively, the sequential setting avoids the "causal loop" in (3.4) that violates the normalisation condition by not allowing signalling from a system at a future time step to a system at a past time step.

3.4 Conclusions

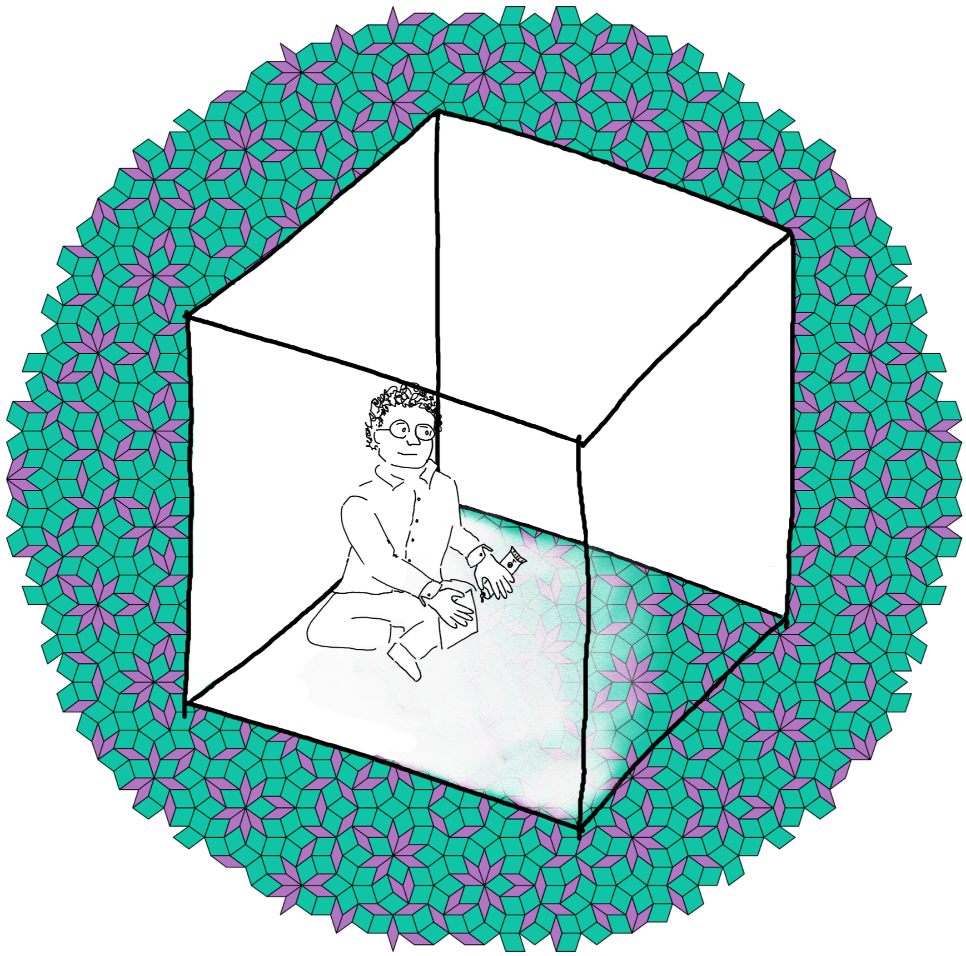
We showed that for processes we cannot take tensor products unrestrictedly, if arbitrary channels are allowed as local operations in a product party. Is this a defect of the process matrix framework itself? One interpretation is that the processes are descriptions of the environment of an entire family of parties [14, 7], and the need to take tensor products of arbitrary processes with indefinite causal structure do not actually arise. Another option is to allow for tensor products of arbitrary parties but restrict the allowed operations in the local parties such that the normalisation condition of processes is preserved [27]. The sequential asymptotic setting we presented above is an example of this kind. A further option is to adopt a more general framework that does not impose a normalisation condition for the matrices/operators that carry the information of the indefinite causal structure. Oreshkov and Cerf's operational quantum theory without predefined time provides an example [81]. We think an important question is to clarify whether these perspectives are compatible with attempts to create processes in the laboratory, because if one can create a process with signalling Hilbert-Schmidt terms in both directions, it is conceivable that one can create more to act jointly on them, and there is no apparent reason why local operations on the joint parties must be restricted.

For communication theory, our results imply that communication tasks defined in the asymptotic limit are not meaningful for processes characterised by Theorem 2 when local operation is unrestricted for the combined parties $A = A'A'' \cdots$, $B = B'B'' \cdots$, $\cdots$. Similarly caution needs to be taken for asymptotic entanglement theory of processes [64]. On the other hand, such issues do not affect one-shot capacities, or asymptotic capacities in the sequential setting. It is possible that some other physically motivated restrictions on local operations yield additional well-defined capacities.

The restriction induces some interesting questions for further research. To what extent does the restriction generalise to indefinite causal structure theories in general [41, 42]? For the particular example we used to demonstrate the restriction, there exists a global process that reduces to the two individual processes. When is this true in general?

Chapter 3: Statement of Contribution

In this Chapter, the main contribution (Section 3.3) includes a theorem for qualifying when tensor product of process matrices are invalid (Theorem 2). This work was co-authored with Ding Jia published in PRA [66, 67]. The project was proposed by Ding Jia and research was done together, Ding largely formalised the theorem. The Chapter has been adapted and re-written from the paper to best fit this thesis.



Person in the box
sorts, composes and then glues
jigsaw data through redundant clues
physical theory emerges
in all its beauty and hues[3]

[3]Chapter 4, illustration for the person "thinking inside the box", on a background of generated Penrose tiling [35] as a dedication to Roger Penrose for his use of a diagrammatic approach in Physics.

Chapter 4

Revisiting the Causaloid Framework

Can a person in a box with all possible operational data uncover the underlying theory?
They compress data through three levels — Tomographic, Compositional and Meta
We revisit the Causaloid framework and present new diagrammatics.

The Causaloid framework introduced by Hardy [41, 42] suggests a research program aimed at finding a theory of Quantum Gravity. On one side General Relativity (GR) while deterministic, features dynamic causal structures; on the other side Quantum Theory (QT) while having fixed causal structures, is probabilistic in nature. It is natural to then expect Quantum Gravity (QG) to house both of the radical aspects of GR and QT, and therefore incorporate *indefinite causal structure*[1]. The Causaloid framework is based on operational methodology – it is based on the assertion that any physical theory, whatever it does, must correlate recorded data. Imagine a person inside a closed space, having access to stacks of cards with recorded data (procedures, outcomes, locations); and the person is tasked with inferring (aspects of) the underlying physical theory that governs the data. The correlation within recorded data due to the physical theory implies that the stacks of cards are filled with (some) redundancy. The person in the box distils away the redundancy by *compressing* the data. We call this *Physical Compression* as it is governed by the nature of the underlying physical theory. In this framework there are three levels of compression: 1) Tomographic Compression, 2) Compositional Compression and 3) Meta Compression.

In this Chapter, we present a new diagrammatic language for physical compression to facilitate exposition of the Causaloid framework, and also provide a review of the Causaloid Framework. In Chapter 5, building upon the work from [41] we will study Meta compression

[1] one can find a broader motivation for the field of *indefinite causal structure* in Section 1.3

and find a hierarchy of theories characterised by Meta Compression for which we will provide a general form. In Chapter 6, we will show that (finite-dimensional) Quantum theory and Classical Probability theory belong to the second rung of this hierarchy, through the application of the Causaloid Framework to the framework of Duotensors [49, 47], following which we discuss ideas around how one may construct theories for the higher rungs of the hierarchy through theories of circuits joined by hyper$_d$wires (hyper-edges that connect d nodes). Finally, we discuss the broad implications of this work for Indefinite Causality.

Outline of the Chapter

Before going further, it is important to note that this Chapter is a descendant to the papers by Hardy that introduced the Causaloid Framework [41, 42], where [42] is a compact version of longer [41]. Therefore, it is inevitable to require exposition of the main aspects of those papers as we present the new diagrammatic representations and before we may discuss the new research contributions that can be found primarily in Chapter 5. While this Chapter can be read by itself, nonetheless we encourage referring to the papers [41, 42] liberally in order for the best comprehension experience. To this end, in Sections of this Chapter one is pointed to the corresponding relevant Sections from the papers [41, 42], in case the reader wishes to supplement their reading.

The Chapter is structured in the following way. In Section 4.1 we will discuss original and newer motivations to take the Causaloid approach. In Section 4.2, we first cover the setup of the Causaloid framework: operational data, the person inside the box, how regions are defined and lastly what we are interested in — predicting certain probabilities; this portion is a condensed version based on Sections 2, 3, 5, 12-15 of [41] and Section 3 of [42]. In the subsequent Section 4.3 we review the fleshed out framework through the three levels of physical compression — namely Tomographic, Compositional and Meta Compression; where we introduce a new diagrammatic representation for physical compression (building upon Sections 16-19 of [41] and Section 4 of [42]). The new diagrammatics will become quite useful as we will see in the following Chapters 5 and 6. We end the Chapter with a succinct synopsis of the three levels of physical compression. The next Chapter will build on the diagrammatics introduced in this Chapter, where we will study Meta Compression through the sufficiency of d-region Compositional compression, that defines a hierarchy of physical theories.

4.1 Why the Causaloid Approach?

In this Section we discuss the original motivations by Hardy behind setting up the Causaloid Framework [41, 42] as well as our motivations for revisiting this approach.

4.1.1 Towards Quantum Gravity...

A theory of Quantum Gravity would reduce to General Relativity on the one hand and Quantum Theory on the other. General Relativity is a theory in which causal structure is dynamical, and the notion of simultaneity is operationally meaningless for space-like separated events. In Quantum Theory, any quantity that is dynamical is subject to quantum indefiniteness (for example, if a particle can go through one slit or another, then it can be indefinite as to which slit it goes through). It follows that we would expect *indefinite causal structure* in a theory of Quantum Gravity.

The question arises how one may find a mathematical basis for such a theory of Quantum Gravity. Hardy points out that an even handed approach is desired given the conceptual novelties brought in by both Quantum Theory (probabilistic nature) and General Relativity (dynamical causality) that are difficult to accommodate in the mathematical structures of the other. Such an even handed approach is possible if one steps outside the confines of the mathematical formulations of either theory. He proposes a general framework to accommodate them: The Causaloid framework ([41, 42]).

> "Hence we adopt the following strategy. We will pick out essential conceptual properties of each theory (QT and GR) and try to find a mathematical framework which can accommodate them." (Section 1, [41])

4.1.2 Letting go of states evolving in time

In order to accommodate theories that may have indefinite causal structure, as a starting point in the Causaloid framework no causal structure is assumed *a priori* between different *regions* (defined subsequently). This is done since it is vital to not rely on the picture of states evolving in time for two reasons. Firstly, the notion of states evolving in time even when possible to recover (such as in General Relativity) or when time is extrinsic to the state being described (such is the case in Quantum Theory) necessarily requires treating space and time on an unequal footing. Hardy conveys this as captured in the following quote:

> "The most obvious issue that arises when attempting to combine QT with GR is that QT has a state on a space-like surface that evolves with respect to an external time whereas in GR time is part of a four dimensional manifold whose behaviour is dynamically determined by Einstein's field equations. We can formulate GR in terms of a state evolving in time - namely a canonical formulation [4, 5]. Such formulations are rather messy (having to deal with the fact that time is not a non-dynamical external variable) and break the elegance of Einstein's manifestly covariant formulation. Given that Einstein's chain of reasoning depended crucially on treating all four space-time coordinates on an equal footing it is likely to be at least difficult to construct QG if we make the move of going from a four dimensional manifold $\mathcal{M}$ to an artificial splitting into a three dimensional spatial manifold Σ and a one dimensional time manifold R..." (Section 11, [41])

But secondly and more importantly, if we try to stick to dynamically determined space-like surfaces that obey some space-time quantum uncertainty, the notion of space-like and time-like separation may itself break down. The quote continues:

> "... But there is a further reason coming from quantum theory that suggests it may be impossible. If the causal structure is dynamically determined then what constitutes a space-like surface must also be dynamically determined. However, in quantum theory we expect any dynamics to be subject to quantum uncertainty. Hence, we would expect the property of whether a surface is space-like or not to be subject to uncertainty. It is not just that we must treat space and time on an equal footing but also that there may not even be a matter-of-fact as to what is space and what is time even after we have solved the equations (see [62, 94]). To this end we will give a framework (which admits a formulation of quantum theory) which does not take as fundamental the notion of an evolving state. The framework will, though, allow us to construct states evolving through a sequence of surfaces. However, these surfaces need not be space-like (indeed, there may not even be a useful notion of space-like)." (Section 11, [41])

Therefore the Causaloid framework without assumptions on causal structure provides quite a general framework. We will see that the (in)definite causal structure is encoded in the second level or Compositional Compression. Note, that if a theory supports states evolving in time it would be possible to recover it, should one desire. Hardy proposes *formalism locality* as an alternate way to formulate a given theory without the evolution of states, as we will see in Chapter 6.

4.1.3 Operational Methodology, Data and Probabilities

If the generality of Causaloid Framework stems from stripping us off *a priori* assumptions of causality and the ontology of space-time is left open, then one may ask what lends us a stable footing that we rely on within this framework? The Causaloid Framework is based on the assertion that *any physical theory, whatever it does, must correlate recorded data.*

This falls under the purview of operational methodology, wherein we find a pragmatic meeting point between differing philosophical positions. Agnostic of a physicist's philosophical tendencies towards (broadly speaking) realism and anti-realism within physics, we expect that a physical theory will *at least* be consistent with experiments that produce certain classical data - such as the setup and observations recorded, say, on a piece of paper. The classical data will have correlations given the underlying physical theory and the goal is to then focus on systematically understanding the correlations within this data. These correlations allow us to compress the recorded data.

Given the difficulties around finding a mathematical framework where Quantum theory and General Relativity meet, let alone finding a coherent ontology between the two, adopting an operational methodology as a basis is a safe route, perhaps even a desperate attempt to revisit the blank canvas before clouding it with assumptions that may not hold. The upside is that whatever we can recover operationally won't be wrong even if incomplete, and hopefully would lead towards a point where we are ready to tack on an ontology, should it be possible and desirable. Hardy discusses this as encapsulated in the following quote:

> "Operationalism played a big role in the discovery of both relativity theory and QT. There are different ways of thinking about operationalism. We can either take it to be fundamental and assert that physical theories are about the behaviour of instruments and nothing more. Or we can take it to be a methodology aimed at finding a theory in which the fundamental entities are beyond the operational realm. In the latter case operationalism helps us put in place a scaffolding from which we can attempt to construct the fundamental theory. Once the scaffolding has served its purpose it can be removed leaving the fundamental theory partially or fully constructed. The physicist operates best as a philosophical opportunist (and indeed as a mathematical opportunist). For this reason we will not commit to either point of view for the time being noting only that the methodology of operationalism serves our purposes. Indeed, operationalism is an important weapon in our armory when we are faced with trying to reconcile apparently irreconcilable theories. A likely reason for

any such apparent irreconcilability is that we are making some unwarranted assumptions beyond the operational realm... The operational methodology is a way of not making wrong statements. If we are lucky we can use it to make progress." (Section 3, [41])

What a physical theory does among other things is help us predict quantities, in particular outcome probabilities, and the Causaloid shares this objective. Therefore, the mathematics probability theory obeys will inform how compression works. Therefore the Causaloid Framework can be regarded as a "general probability theory for theories with indefinite causal structure"; it may also be regarded as a "theory to study correlated data".

4.1.4 Diagrammatics: Old Framework, New Clothes

This brings us to the point where we address why we decide to revisit this framework. The Causaloid Framework, by providing key insights and motivations, has been an important contribution guiding the study of *indefinite causal structures*, often cited for such reasons; and yet, there has been only a handful of papers (for example [74]) specifically building upon the Causaloid framework. This is due to two main reasons.

Firstly, since the introduction of the Causaloid framework, other operational frameworks have come up that can be used to study indefinite causality, such as Quantum Combs [18], Process Matrices [82] and Causal Boxes [86] to name a few. These frameworks are constructed with some assumptions that aid in their direct applicability where they can be seen as generalisations of quantum information processes. This feature leads to the possibility of more specific results that is harder to achieve in the more general Causaloid Framework. Though, depending on the goal in mind a bug can become a feature, and the generality of the Causaloid serves the purpose of being able to accommodate any theory that studies correlated data. To this end, in this work we study further the third level of compression — Meta Compression, to find any broad statements that can be made for physical theories, and in fact it will provide us with a way to categorise theories into a hierarchy.

Secondly, the Causaloid Framework is quite abstract requiring introduction of its own terminology. Perhaps, it explains the sociological phenomenon where researchers studying indefinite causal structures have heard of it, but often do not know much about it. We present a diagrammatic representation for the framework that will perhaps make it somewhat easier for the interested person to be able to understand and work with it. The diagrammatics will prove to be quite powerful, we will see how the *duotensors* [47] relate to the Causaloid framework, primarily through diagrams.

4.2 Causaloid Framework

In this Section we are now ready to cover the setup of the Causaloid framework: recording operational data, the person inside the box, how regions are defined and lastly what we are interested in — predicting probabilities. Note that this portion is based on material from Sections 2, 3, 5, 12-15 of [41] and Section 3 of [42].

4.2.1 Thinking inside the box

As discussed in the previous Section, the Causaloid Framework relies on operational methodology. The way the Causaloid framework is set up is prescribed as follows. We will assume that the *data* obtained from experiments - procedures, settings, outcomes and any other descriptions such as location - is recorded onto cards. Of course since this is classical data it could be instead written on a computer or stored in other ways, nonetheless we may use the idea of cards to setup this framework for the purpose of exposition. A concern around the physicality of classical data — its storage, transportation and processing, may arise. To this end an assumption of *low key* physical devices is made as follows:

> "**The indifference to data principle**: It is always possible to find physical devices capable of storing, transporting, and processing data such that (to within some arbitrarily small error) the probabilities obtained in an experiment do not depend on the detailed configuration of these devices. Such physical devices will be called *low key*." (Section 3, [41])

Now, imagine a person inside a closed space, having access to stacks of cards with recorded data (procedures, outcomes, locations) corresponding to all possible runs of the experiment; and the person is tasked with inferring (aspects of) the underlying physical theory that governs the data. The correlation within recorded data due to the physical theory implies that the stacks of cards are filled with (some) redundancy. The person in the box distils away the redundancy by *compressing* the data.

The person is not able to look outside the box for extra information and must therefore define all concepts using only the information provided through the cards. In Hardy's words this enforces "a particular kind of honesty", and the person is forced to stick to an operational methodology.

Note that there may not be an actual person in a box, rather this *gedanken* setup is to serve the purpose of helping us emulate operational methodology.

4.2.2 Organising Stacks of Cards

Each card records a small amount of proximate data. We may think of the cards representing something analogous to space-time events. One piece of data recorded on any given card will be something we will regard as representing or being analogous to space-time location.

We will consider examples where the data recorded on each card is sorted in the form — (x - *Location*, a - *Actions*, s - *Observation*) (or in short (x,a,s)), where each of the pieces of information are categorised as follows:

1. **Location:** The first piece of data, x, is an observation and represents location. It could be some real physical reference frame such as a GPS system or it could be some other data that we are simply going to regard as representing location.

2. **Actions:** The second piece of data, a, represents some actions. For example it might correspond to the configuration of some knobs we have freedom in setting.

3. **Observations:** The third piece of data, s, represents some further local observations, such as the outcome that is obtained.

Given this form of recorded data, we can organise the cards into the following sets.

- **The stack**, denoted by Y, is the set of cards from a single run of an experiment.

- **The procedure**, denoted by F, corresponds to all cards that are consistent with given function of the settings, where we rewrite actions a as a function $F(x)$ depending on location x (thus the data is of the form (x, $F(x)$, s)). F here stands for the procedure set of cards, the Function $F(x)$ as well as the procedure given by the description of actions, and it will be clear from context which meaning is implied.

- **The full pack**, denoted by V, is the set of all logically possible cards when all possible procedures are taken into account. It is possible that some cards never actually occur in any stack because of the nature of the physical theory but they are included anyway.

The procedure F can be thought of as "what was done". The stacks Y (particular run of procedure giving some outcome) can be thought of as "what was seen".

Given the definitions of the sets we have the following relations between them:

$$Y \quad \textit{(Outcome} \,\big|\, \textit{Procedure)} \quad \subseteq \quad F \quad \textit{(Procedure)} \quad \subseteq \quad V \quad \textit{(Full pack)} \tag{4.1}$$

4.2.3 Operational Regions

V is the complete set of logical possible cards (the stacks). We talked about subsets of V with respect to a - Actions (given by subset Procedure F), and with respect to s - Observations (given by subset Outcome Y); now we look at subsets with respect to x - Locations which we call *regions*. The region R_O is specified by the set of cards from V having $x \in O$. We define R_x to be an elementary region consisting only of the cards having x on them.

Within regions there can be many possible sets of actions (associated with procedures F, F',...) . In a region we consider having an independent choice of which action (procedure) to implement. This captures the notion of "space-time"[2] regions as places where we have local choices. When we have a particular run of the experiment given some procedure we end up with a stack Y of data that can be sorted by regions. Then we find a picture of what happened, laid out in a kind of "space-time".

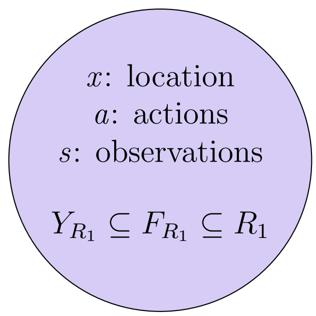


Figure 4.1: Operational Region R_1

For a region R_1 (shorthand for R_{O_1}) we define Y_{R_1} as the cards from the stack Y that are attributed to region R_1:

$$Y_{R_1} = Y \cap R_1 \tag{4.2}$$

Similarly, we define procedure in R_1 as F_{R_1}- the cards from the set F that are attributed to the region R_1

$$F_{R_1} = F \cap R_1 \tag{4.3}$$

Clearly $Y_{R_1} \subseteq F_{R_1} \subseteq R_1$. Given ($Y_{R_1}$, F_{R_1}) we know "what was done" (F_{R_1}) and "what was seen" (Y_{R_1}) in region R_1.

[2]Here, we use the phrase space-time in a non-specific manner. It may be seen as some general notion of spatio-temporal regions that may be distinct from the more specific space-time as known within relativity.

4.2.4 Statement of objective

The objective of the Causaloid Framework is to be able to predict the conditional probabilities of the form

$$\text{prob}(Y_{R_1}|Y_{R_2}, ..., Y_{R_n}, F_{R_1}, F_{R_2}, ..., F_{R_n}) \tag{4.4}$$

without any assumptions of the causal structure, given that R_i, $i \in \{1, 2, ..., n\}$ are n disjoint regions. Operationally, one may interpret the probability as the relative frequencies of the number of cards

$$\frac{N(Y_{R_1}, Y_{R_2}, ..., Y_{R_n}, F_{R_1}, ..., F_{R_n})}{N(Y_{R_2}, ..., Y_{R_n}, F_{R_1}, ..., F_{R_n})} \tag{4.5}$$

when $N(.)$ (the number of cards satisfying the condition within it) becomes large, at least for the denominator.

An important question arises, to begin with are these probabilities even well-defined and if yes, how can we use the Causaloid Framework to efficiently calculate this? Let us address the first part of the question here; the next Section is dedicated to answer the latter. To understand the issue around well-defined probabilities consider three consecutive polarisers (associated with three regions) placed at some angle specified as part of procedure of their respective regions, the outcome notes if a photon passes through or is absorbed. Then the probability of the outcome of the third polariser conditioned on the outcome of the first, as well as the procedures of both of them, given by,

$$\text{prob}(Y_{\text{third polariser}}|Y_{\text{first polariser}}, F_{\text{first polariser}}, F_{\text{third polariser}}) \tag{4.6}$$

is not well defined without consideration of the second polariser and the framework will not be able to predict any value for this conditional probability.

Therefore, to be able to work around such issues we restrict ourselves to a large region R that would contain a large number of (but not all) cards from the full pack V for which all probabilities

$$\text{prob}(Y_R|F_R, C) \tag{4.7}$$

are well-defined given some conditioning C on cards outside R. In the context of the example of the polariser the region R may be considered to be the room in which the experiment takes place and the conditioning C may be all the requirements to avoid "noise" in the experiments. At the level of theory encompassing everything this conditioning may be argued either to take the form of "boundary conditions" (where R is *open* and associated with the "Open Causaloid", see Section 22 of [41]) or not be required (where R is *closed* and associated with the"Universal Causaloid", see Section 30 and 35 of [41])

For the rest of the Chapter we always consider this large region R which is called a "*predictively well-defined region*" and assume the conditioning is taken into account and more concisely we simply write

$$\text{prob}(Y_R|F_R) \tag{4.8}$$

4.3 Three levels of Physical Compression of Data

In this Section, we review how to distil the redundancy in recorded data (cards) through three levels of physical compression to form a "theory of correlated data". By compression we mean the process of organising the data recorded on the cards, and reducing the set of data to the minimum; where the minimum required data can be used to predict any desired probabilities. We call it physical compression, since it is contingent on the physical theory that correlated the data. The levels of compression are as follows:

Pre-Compression: The zeroth level of compression introduces the notion of generalised states and preparations, in absence of assumptions of causal relations between regions, and sets up the stage for the following levels.

Tomographic Compression: The first level of compression pertains to physical compression over single region and here the compression matrices (Λ) and compression sets (Ω) are introduced. Fiducial states are also introduced. This level is closely related to the concept of *tomography* and its relation to Generalised Probability Theories (GPTs) will be discussed.

Compositional Compression: The second level of compression pertains to physical compression over disjoint regions through their *composition*. Here, the notion of *causal adjacency* is introduced. It captures strong causal connections between regions when non-trivial Compositional Compression occurs.

Meta Compression: The third and final level pertains to compression over Tomographic and Compositional Compression, since there may be some universal, region independent *rules* related to the mathematical structure of the theory, which gives the Causaloid $\mathbf{\Lambda}$, seen as a specification of the physical theory itself. Here, the *Causaloid product* $\otimes^{\Lambda}$ that generalises temporal and spatial products is introduced. This is used to take the product over regions and is the culmination of the compression process.

This review of the levels of compression closely follows concepts explained in Sections 16-19 of [41] and Section 4 of [42]; additionally here we present a new diagrammatic representation for the Causaloid framework. The terminology of three levels of compression (Tomographic, Compositional and Meta) is new to this work.

4.3.1 Zeroth level: Pre-Compression

Recall the classical data recorded on cards in an operational manner. A physical theory will capture the correlations in this data and express them through physical concepts. While some of these physical concepts may be theory dependent, the concept of a physical state is central to most (if not all) physical theories. Before we may discuss the three levels of compression, we must focus on how we may represent states in this framework in relation to the data. The definition of states is often either based on a notion of preparation that happens in the past to give the state we have in the present; or based on a notion of measurement that happens in the future of the state that the measurement outcomes tell us something about. Thus these rely on a causal structure that is assumed to be fixed, in so far as the order of events being such that preparation happens in the past and measurement happens in the future. Since we are interested in the absence of assumptions on the underlying causal structure between elementary regions, we need a general notion of physical state, preparations and measurements. Hardy presents the notion of a physical state as follows:

> **The state (in general)** associated with a (generalised) preparation is that thing represented by any mathematical object that can be used to calculate the probability for every outcome of every measurement that may be performed on the system.

Of course the above definition is minimal that centres the operational role of a state while being agnostic of the ontological status of the state. This is ideal since we are building up from operational data. Let us now apply this notion of state to the stacks of cards. Consider a region R_1 inside the predictively well-defined region R (note that R_1 need not be an elementary region). Since $R = R_1 \cup (R - R_1)$ we can write:

$$\begin{aligned} p &= \text{prob}(Y_R | F_R) && (4.9) \\ &= \text{prob}(Y_{R_1} \cup Y_{R-R_1} | F_{R_1} \cup F_{R-R_1}) && (4.10) \end{aligned}$$

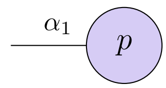
Hardy regards (Y_{R-R_1}, F_{R-R_1}) which happens in $R - R_1$ as a *generalised preparation* for region R_1. The generalised preparation for R_1 is considered to "prepare" a state which is defined soon below. Further, Hardy regards (Y_{R_1}, F_{R_1}) which happens in R_1 as a *measurement* associated with the actions and outcomes seen in R_1. The measurement outcomes for the measurement in R_1 will be labelled by α_1, where α_1 belongs to the (possibly very big) set Γ_{R_1}. That is we have:

$$(Y_{R-R_1}, F_{R-R_1}) \Longleftrightarrow \textit{generalised preparation for } R_1 \tag{4.11}$$

$$(Y^{\alpha_1}_{R_1}, F^{\alpha_1}_{R_1}) \Longleftrightarrow \textit{measurement in } R_1 \tag{4.12}$$

Then we can also label the probabilities with α_1 denoting the probabilities associated with measurement outcomes labelled α_1 as follows:

$$p_{\alpha_1} = \text{prob}(Y^{\alpha_1}_{R_1} \cup Y_{R-R_1} | F^{\alpha_1}_{R_1} \cup F_{R-R_1}) =$$



(4.13)

We will introduce the diagrammatic representation alongside the mathematical definition, similar to Equation 4.13 throughout this Section.

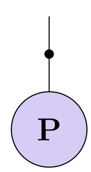
Note that in Equation 4.13 above, the choice of joint probabilities between outcomes is intentional as compared to the conditional probability $\text{prob}(Y_{R_1} | Y_{R-R_1}, F_{R_1} \cup F_{R-R_1})$ which we desire though it introduces more problems. Due to the normalisation in Bayes formula working with conditional probabilities will require dealing with nonlinearities which Hardy points out will "represent an insurmountable problem when we have dynamic causal structure". One can always use the Bayes formula towards the end to calculate the conditional probabilities when required.

Now let us see the definition of state by Hardy in the context of the Causaloid framework:

> **The state (in Causaloid Framework)** for R_1 associated with a generalised preparation in $R - R_1$ is defined to be the thing represented by any mathematical object which can be used to predict p_{α_1} for all measurements in R_1 labelled by the index α_1.

We could simply write the state $\mathbf{P}(R_1)$ as a vector that lists all the probabilities p_{α_1} since it satisfies the above definition:

$$\mathbf{P}(R_1) = \begin{pmatrix} \vdots \\ p_{\alpha_1} \\ \vdots \end{pmatrix} =$$

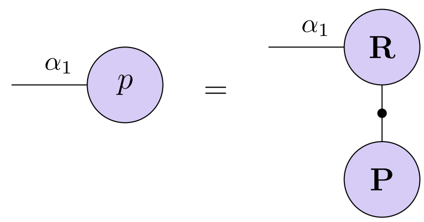


$$\text{where} \quad \alpha_1 \in \Gamma_{R_1} \tag{4.14}$$

where α_1 belongs to the set Γ_{R_1} (or simply Γ_1) which labels all actions and outcomes and we expect the size of the set Γ_1 to be quite big (and possibly infinite).

The objective is to be able to recover the probabilities p_{α_1} and to do so we define another vector $\mathbf{R}_{\alpha_1}$ which represents the measurement $(Y_{R_1}^{\alpha_1}, F_{R_1}^{\alpha_1})$ in region R_1 such that

$$p_{\alpha_1} = \mathbf{R}_{\alpha_1}(R_1) \cdot \mathbf{P}(R_1) \tag{4.15}$$



(4.16)

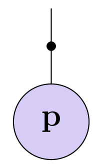
where the vector $\mathbf{R}_{\alpha_1}(R_1)$ has 1 in position α_1 and 0's in all other positions. The inner product of the two vectors gives us p_{α_1} and the above equation may be thought of being some generalisation of the Born rule, under the generalised notions of states $\mathbf{P}(R_1)$ and measurements $\mathbf{R}(R_1)$ given above. Note that Equation 4.15-4.16 is linear in $\mathbf{P}(R_1)$ and $\mathbf{R}(R_1)$ and that in Equation 4.16 the line with a filled black dot represents the inner product.

4.3.2 First level: Tomographic Compression

Now we are prepared to tackle the first level of compression. Here we continue considering a single region R_1. While through Equation 4.14 we were able to calculate (well defined) probabilities p_α in terms of the generalised preparation of state for R_1 and measurements in R_1, notice that the vector $\mathbf{P}(R_1)$ can be quite long and the set Γ_{R_1} can be quite big. A physical theory will (we expect) have some structure for state space that allow for the specification for fewer probabilities (or in other words fewer procedures and outcomes) to be able to predict any p_{α_1}. For example, we encounter this in Quantum Theory through the process of *Tomography*.

Therefore we may replace $\mathbf{P}(R_1)$ with a list of the smallest subset of probabilities from p_{α_1} that allows us to predict any p_{α_1} by means of linear relations. To do so we pick a set of *fiducial* measurements $(Y_{R_1}^{l_1}, F_{R_1}^{l_1})$ labelled by $l_1 \in \Omega_{R_1}$ (or simply Ω_1) where $\Omega_{R_1} \subseteq \Gamma_{R_1}$ and the (new representation of the) state (compared to Equation 4.14) becomes:

$$\mathbf{p}(R_1) = \begin{pmatrix} \vdots \\ p_{l_1} \\ \vdots \end{pmatrix} =$$

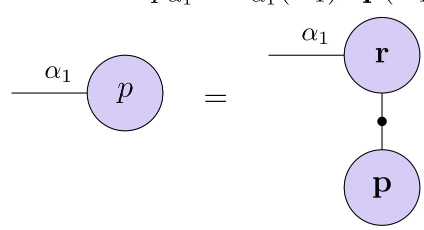


$$\text{where} \quad l_1 \in \Omega_{R_1} \tag{4.17}$$

Note that the choice of the set of fiducial measurements need not be unique, and one may pick any one such set but the size of the set Ω_{R_1} (or the length of $\mathbf{p}(R_1)$) will be the same over any such choice which is the minimum possible size such that $\mathbf{p}(R_1)$ continues to satisfy the definition of a state. The sets Ω_{R_O} (for some region R_O) play an important role in the Causaloid Framework and its size will be informed by the system and underlying physical theory. Of course if no such subset can be found we can always resort to having $\mathbf{p}(R_1) = \mathbf{P}(R_1)$ and $\Omega_{R_1} = \Gamma_{R_1}$ (which will give no compression).

Now that we have the compressed state $\mathbf{p}(R_1)$ we need a way to calculate any general p_α from it, and require a *linear* formula analogous to Equation 4.15. Let us express this as:

$$p_{\alpha_1} = \mathbf{r}_{\alpha_1}(R_1) \cdot \mathbf{p}(R_1) \tag{4.18}$$



(4.19)

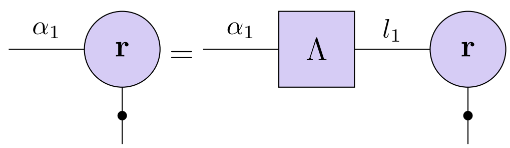
The exact form of $\mathbf{r}_{\alpha_1}(R_1)$ can be found from the recorded data and will contain real numbers, nonetheless we can provide the form for some of these. Given the fiducial measurements $(Y_{R_1}^{l_1}, F_{R_1}^{l_1})$ in R_1, $\mathbf{r}_{\alpha_1}(R_1)$ for $\alpha_1 = l_1$ would simply be

$$\mathbf{r}_{l_1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{for all} \quad l_1 \in \Omega_{R_1} \tag{4.20}$$

where 1 is in the ${l_1}^{th}$ position and all other entries are 0's since this is the only way to ensure $p_{l_1} = \mathbf{r}_{l_1}(R_1) \cdot \mathbf{p}(R_1)$. The fiducial measurements thus form a basis.

Now we can see that the general vector $\mathbf{r}_{\alpha_1}(R_1)$ can be expressed as a linear combination over the basis of measurements spanned by $\mathbf{r}_{l_1}(R_1)$. This linear relation can be defined by introducing the compression matrix $\Lambda_{\alpha_1}^{l_1}$ as follows:

$$\mathbf{r}_{\alpha_1} = \sum_{l_1 \in \Omega_{R_1}} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{l_1} \tag{4.21}$$

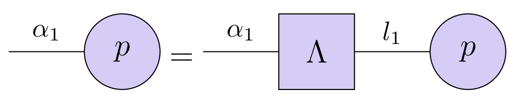


(4.22)

We can take the inner product of Equation 4.21 with some general $\mathbf{p}$ to find a similar relation for probabilities:

$$\mathbf{r}_{\alpha_1} \cdot \mathbf{p} = \sum_{l_1 \in \Omega_{R_1}} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{l_1} \cdot \mathbf{p}$$

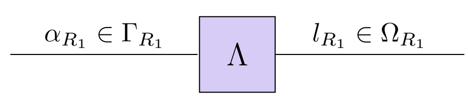
$$\Rightarrow \quad p_{\alpha_1} = \sum_{l_1 \in \Omega_{R_1}} \Lambda_{\alpha_1}^{l_1} p_{l_1} \tag{4.23}$$



(4.24)

It is evident from the definition of the compression matrix $\Lambda_{\alpha_1}^{l_1}$ in Equation 4.21 that its components can be fixed using the entries of $\mathbf{r}$-vectors:

$$\mathbf{r}_{\alpha_1}|_{l_1} = \Lambda_{\alpha_1}^{l_1} :=$$



(4.25)

where $\mathbf{r}_{\alpha_1}|_{l_1}$ is the l_1^{th} component of $\mathbf{r}_{\alpha_1}$.

Note that $\Lambda_{\alpha_1}^{l_1}$ will (often) be a very rectangular matrix since we expect that $|\Gamma_{R_1}| \geq |\Omega_{R_1}|$ where $|.|$ gives you the size of the sets. From the definition for the Λ matrix we also have:

$$\Lambda_{l_1'}^{l_1} = \delta_{l_1'}^{l_1} \quad \text{for} \quad l_1', l_1 \in \Omega_{R_1} \tag{4.26}$$

where $\delta_{l_1'}^{l_1}$ equals 1 if the subscript and superscript are equal and is 0 otherwise.

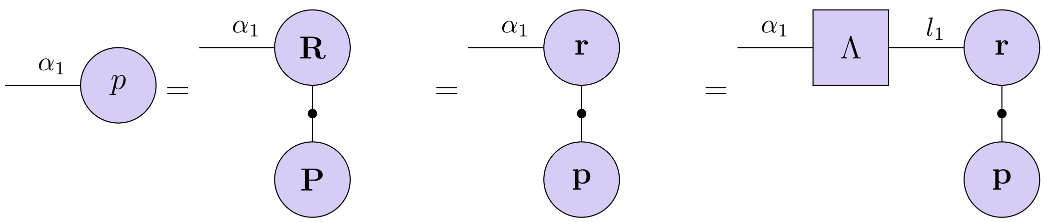
This concludes the details for the first level compression. Let us recapitulate for assimilation. The person in the box in pre-compression phase, first organises the recorded data by regions R, procedures F and outcomes Y and finds the state $\mathbf{P}$ associated with regions. By the structure of the underlying physical theory the state $\mathbf{P}$ can be *compressed* to $\mathbf{p}$ and the measurements $\mathbf{R}_{\alpha_1}$ can be *compressed* to $\mathbf{r}_{\alpha_1}$ where the relation of $\mathbf{r}_{\alpha_1}$ to $\mathbf{r}_{l_1}$ is captured by the Λ matrix:

$$p_{\alpha_1} = \mathbf{R}_{\alpha_1}(R_1) \cdot \mathbf{P}(R_1) \tag{4.27}$$

$$= \mathbf{r}_{\alpha_1}(R_1) \cdot \mathbf{p}(R_1) \tag{4.28}$$

$$= \sum_{l_1 \in \Omega_{R_1}} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{l_1}(R_1) \cdot \mathbf{p}(R_1) \tag{4.29}$$

or diagrammatically we have,



(4.30)

While this might look like some heavy handed machinery, the utility of fiducial measurements and states (and thus, the first level compression) will become evident when we start considering more regions.

Tomographic Compression and Generalised Probability Theory

The first level compression is deeply related to *Generalised Probability Theories (GPTs)*. In the GPT framework one can characterise physical theories by finding the relation between the number of distinguishable states for a system - N and the number of measurement outcomes required - K, to fully characterise the states of the system (when we do not include the normalisation condition), through the function $K(N) = N^r$ where r is theory dependent [40]. For finite-dimensional Classical Probability Theory $r = 1$ while for finite-dimensional Quantum Theory $r = 2$. If the underlying theory qualifies as a GPT the first level compression may be thought of as channel tomography and then the size of the fiducial measurement set is related to $K(N)$ as follows (where the channel input is denoted by $K_i(N_i)$, and channel output is denoted by $K_o(N_o)$):

$$|\Omega_{R_1}| = K_i K_o = N_i^r N_o^r \text{ where} \begin{cases} r = 1 & \text{for Classical Probability Theory} \\ r = 2 & \text{for Quantum Theory} \end{cases} \tag{4.31}$$

Therefore, we will call the first level compression as *Tomographic Compression*. We will also synonymously call the Omega sets Ω_{R_1} for a region as Tomographic Sets and the Lambda matrices for a region Λ as Tomographic Matrices. We will see more of Omega sets and Lambda matrices in the coming sections and they play a central role to the framework.

Note that we have focused on linear relations. One may ask why non-linear relations for physical compression are not considered. The answer may be tied to the nature of Probability theory. Hardy explains:

> "We have just employed linear physical compression here. It is possible that if we employed more general mathematical physical compression (allowing non-linear functions) we could do better. This does not really matter since we are free to choose linear physical compression as the preferred form of physical compression. In fact, it can easily be proven that if we are able to form mixtures of states (as we can in quantum theory) then we cannot do better than linear physical compression (this is not surprising since probabilities combine in a linear way when we form mixtures)."

4.3.3 Second level: Compositional Compression

We have worked with Tomographic Compression that pertains to a single region, and the natural next question to ask is how do physical theories correlate data between multiple disjoint regions. For purposes of exposition let us explain using two regions and extending the following to multiple regions will be straight-forward (discussed at the end of this subsection). Let us consider a composite region consisting of two disjoint regions R_1 and R_2 such that $R_1 \cap R_2 = \phi$ and $R_1, R_2 \subset R$ where R is a larger predictively well-defined region. Since we are considering two regions the generalised preparations of states of one region will depend on the labels of the other region. Let α_1 be the label for measurement outcomes in R_1 and α_2 be the label for measurement outcomes in R_2. Then for R_1 we have

$$(Y_{R_2}^{\alpha_2}, Y_{R-R_1-R_2}, F_{R_2}^{\alpha_2}, F_{R-R_1-R_2}) \Longleftrightarrow \textit{generalised preparation for } R_1 \tag{4.32}$$

$$(Y_{R_1}^{\alpha_1}, F_{R_1}^{\alpha_1}) \Longleftrightarrow \textit{measurement in } R_1 \tag{4.33}$$

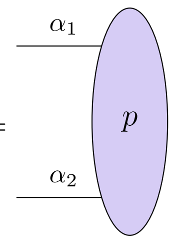
and similarly for R_2 we have

$$(Y_{R_1}^{\alpha_1}, Y_{R-R_1-R_2}, F_{R_1}^{\alpha_1}, F_{R-R_1-R_2}) \Longleftrightarrow \textit{generalised preparation for } R_2 \tag{4.34}$$

$$(Y_{R_2}^{\alpha_2}, F_{R_2}^{\alpha_2}) \Longleftrightarrow \textit{measurement in } R_2 \tag{4.35}$$

For the composite region $R_1 \cup R_2$, the joint probabilities of interest are given by:

$$p_{\alpha_1\alpha_2} = \text{Prob}(Y_{R_1}^{\alpha_1} \cup Y_{R_2}^{\alpha_2} \cup Y_{R-R_1-R_2} | F_{R_1}^{\alpha_1} \cup F_{R_2}^{\alpha_2} \cup F_{R-R_1-R_2}) = \tag{4.36}$$



Given the only compression in our arsenal yet is the first, consider that we can either apply Tomographic compression to region R_1 and to region R_2, or apply Tomographic compression to the composite region $R_1 \cup R_2$. Will these two scenarios be equivalent? We find out.

When we apply Tomographic compression to the region R_1 followed by Tomographic compression on R_2 we get:

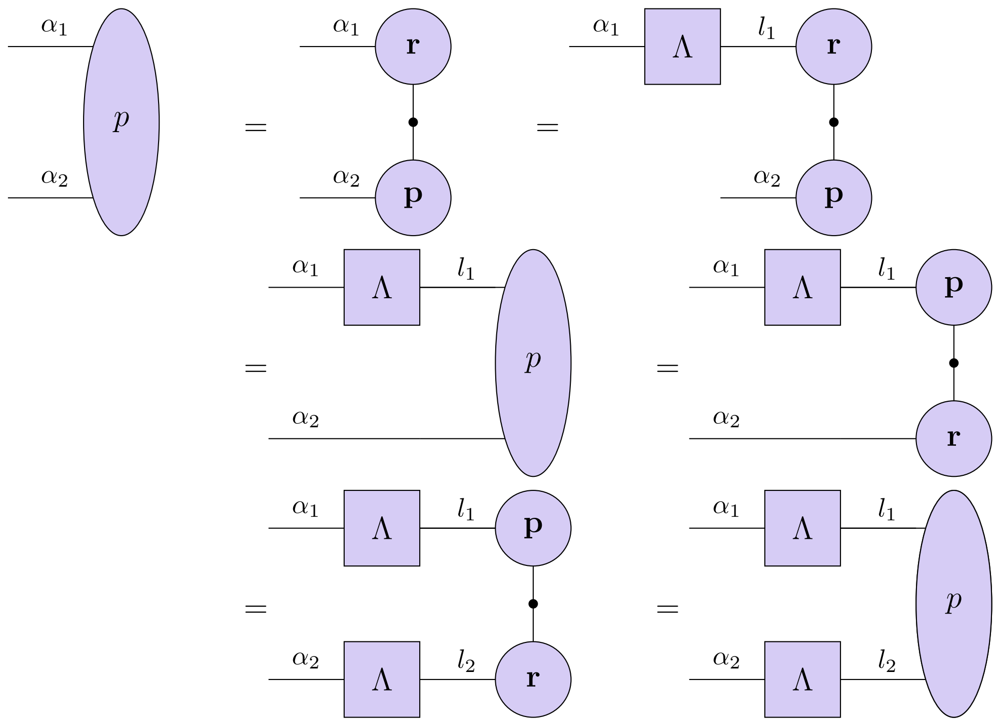
$$\begin{aligned}
p_{\alpha_1\alpha_2} &= \mathbf{r}_{\alpha_1}(R_1) \cdot \mathbf{p}_{\alpha_2}(R_1) && (4.37)\\
&= \sum_{l_1 \in \Omega_1} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{l_1}(R_1) \cdot \mathbf{p}_{\alpha_2}(R_1) && (4.38)\\
&= \sum_{l_1 \in \Omega_1} \Lambda_{\alpha_1}^{l_1} p_{l_1\alpha_2} && (4.39)\\
&= \sum_{l_1 \in \Omega_1} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{\alpha_2}(R_2) \cdot \mathbf{p}_{l_1}(R_2) && (4.40)\\
&= \sum_{l_1 \in \Omega_1} \sum_{l_2 \in \Omega_2} \Lambda_{\alpha_1}^{l_1} \Lambda_{\alpha_2}^{l_2} \mathbf{r}_{l_2}(R_2) \cdot \mathbf{p}_{l_1}(R_2) && (4.41)\\
&= \sum_{l_1 l_2 \in \Omega_1 \times \Omega_2} \Lambda_{\alpha_1}^{l_1} \Lambda_{\alpha_2}^{l_2} \mathbf{r}_{l_2}(R_2) \cdot \mathbf{p}_{l_1}(R_2) && (4.42)\\
&= \sum_{l_1 l_2 \in \Omega_1 \times \Omega_2} \Lambda_{\alpha_1}^{l_1} \Lambda_{\alpha_2}^{l_2} p_{l_1 l_2} && (4.43)
\end{aligned}$$

where we have repeated use of Equation 4.21 for tomographic compression (from 4.37 to 4.38 and from 4.40 to 4.41). Pay close attention to the regions in brackets for the $\mathbf{r}$ and $\mathbf{p}$ vectors! $\mathbf{p}_{\alpha_2}(R_1)$ is the state for region R_1 labelled by generalised preparation label α_2 in region R_2 and similarly $\mathbf{p}_{l_1}(R_2)$ is the state for region R_2 labelled by fiducial preparation label l_1 in region R_1, post Tomographic Compression. Finally notice that we use:

$$l_1 l_2 \in \Omega_1 \times \Omega_2 \tag{4.44}$$

here $\times$ denotes the Cartesian product between the elements of the two sets Ω_1 and Ω_2, for example $\{1, 4, 9\} \times \{1, 2\} = \{11, 12, 41, 42, 91, 92\}$.

The diagrammatic form for Equations 4.37-4.43 are given by:



(4.45)

(4.46)

(4.47)

If instead one chooses to tomographically compress R_2 first and R_1 later then the analogous calculation to Equations 4.37-4.43 will include the following step in place of 4.42:

$$p_{\alpha_1\alpha_2} = \sum_{l_1 l_2 \in \Omega_1 \times \Omega_2} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} \mathbf{r}_{l_1}(R_1) \cdot \mathbf{p}_{l_2}(R_1) \tag{4.48}$$

but the order of first level compression to disjoint regions does not matter and comparing Equation 4.42 and Equation 4.48 we have:

$$\begin{aligned} p_{l_1 l_2} &= \mathbf{r}_{l_1}(R_1) \cdot \mathbf{p}_{l_2}(R_1) = \mathbf{r}_{l_2}(R_2) \cdot \mathbf{p}_{l_1}(R_2) \\ \text{such that} \quad p_{\alpha_1\alpha_2} &= \sum_{l_1 l_2 \in \Omega_1 \times \Omega_2} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} p_{l_1 l_2} \end{aligned} \tag{4.49}$$

Let us now consider the second option, directly applying tomographic compression to the composite region

$$R_{1,2} = R_1 \cup R_2 \tag{4.50}$$

The state for $R_{1,2}$ is any mathematical object that can be used to calculate all $p_{\alpha_1\alpha_2}$ by means of linear relations. The state for $R_{1,2}$ after performing Tomographic Compression on the composite region is defined as as:

$$\mathbf{p}(R_{1,2}) = \begin{pmatrix} \vdots \\ p_{k_1k_2} \\ \vdots \end{pmatrix} \text{, where } \; k_1k_2 \in \Omega_{1,2} \tag{4.51}$$

where $\Omega_{1,2}$ is tomographic set for $R_{1,2}$. For some choice, the fiducial measurements are given by:

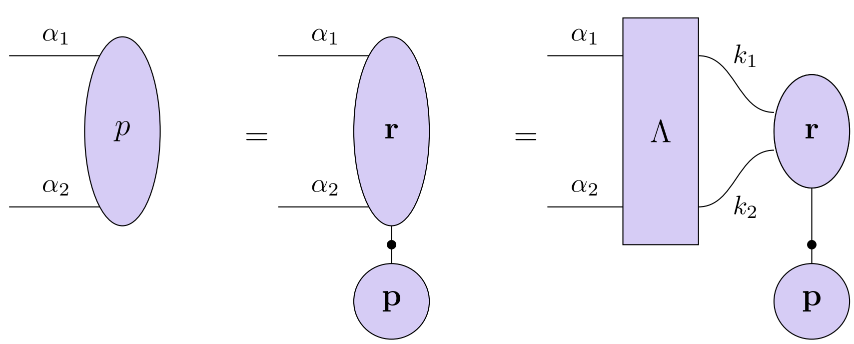
$$\mathbf{r}_{k_1k_2} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{for all} \quad k_1k_2 \in \Omega_{1,2} \tag{4.52}$$

where, the $k_1k_2^{th}$ element is 1 and the rest are 0. The state and fiducial measurements can then be used to calculate probabilities using the generalised Born rule:

$$p_{\alpha_1\alpha_2} = \mathbf{r}_{\alpha_1\alpha_2}(R_{1,2}) \cdot \mathbf{p}(R_{1,2}) \tag{4.53}$$

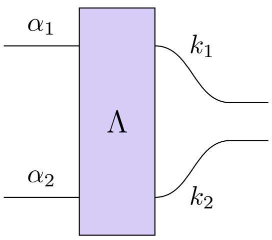
$$= \sum_{k_1k_2 \in \Omega_{1,2}} \Lambda_{\alpha_1\alpha_2}^{k_1k_2} \mathbf{r}_{k_1k_2}(R_{1,2}) \cdot \mathbf{p}(R_{1,2}) \tag{4.54}$$

Equivalently and diagrammatically we have

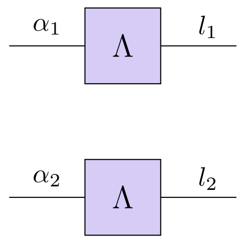


(4.55)

We are ready to compare how the compression is quantified in both cases, by comparing the Ω sets and Λ matrices. We have $\Lambda_{\alpha_1}^{l_1}\Lambda_{\alpha_2}^{l_2}$ where $l_1l_2 \in \Omega_1 \times \Omega_2$ when tomographically compressing each region separately. And we have $\Lambda_{\alpha_1\alpha_2}^{k_1k_2}$ where $k_1k_2 \in \Omega_{1,2}$ when tomographically compressing the composite region.



where $k_1 k_2 \in \Omega_{1,2}$,



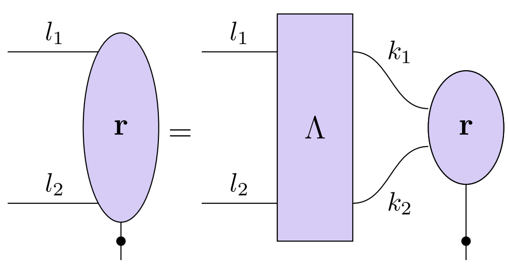
where $l_1 l_2 \in \Omega_1 \times \Omega_2$

It is evident that any compression achieved by $\Lambda_{\alpha_1}^{l_1}\Lambda_{\alpha_2}^{l_2}$ will at least be achieved by $\Lambda_{\alpha_1\alpha_2}^{k_1k_2}$, since it can take the form $\Lambda_{\alpha_1\alpha_2}^{k_1k_2} = \Lambda_{\alpha_1}^{k_1}\Lambda_{\alpha_2}^{k_2}$. In fact, $\Lambda_{\alpha_1\alpha_2}^{k_1k_2}$ will sometimes provide further compression that cannot be found by tomographic compression of region R_1 and tomographic compression of region R_2. Therefore we have the following result which is central to this framework:

$$\Omega_{1,2} \subseteq \Omega_1 \times \Omega_2 \tag{4.56}$$

We may finally define second level or Compositional Compression for composite regions as the compression that is found over and above first level or Tomographic Compression of constituent regions. We define the Compositional Lambda matrix as follows:

$$\mathbf{r}_{l_1 l_2} = \sum_{k_1, k_2 \in \Omega_{1,2}} \Lambda_{l_1 l_2}^{k_1 k_2} \mathbf{r}_{k_1 k_2} \tag{4.57}$$



(4.58)

where $\Lambda_{l_1 l_2}^{k_1 k_2}$ encodes Compositional Compression. Here, $l_1, l_2 \in \Omega_1 \times \Omega_2$ are the labels after tomographic compression of each constituent region, conventionally we will use l's for labelling tomographic compression. Compositional Compression is nontrivial when $\Omega_{1,2}$ is a proper subset of $\Omega_1 \times \Omega_2$ and conventionally we will use k's for compositional compression.

Compositional Compression for multiple disjoint regions can be implemented through an extension of the examples discussed above. Multi-region compression will go through the following label changes

$$l_1 l_2 \dots l_n \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \longrightarrow k_1 k_2 \dots k_n \in \Omega_{1,2\dots n} \tag{4.59}$$

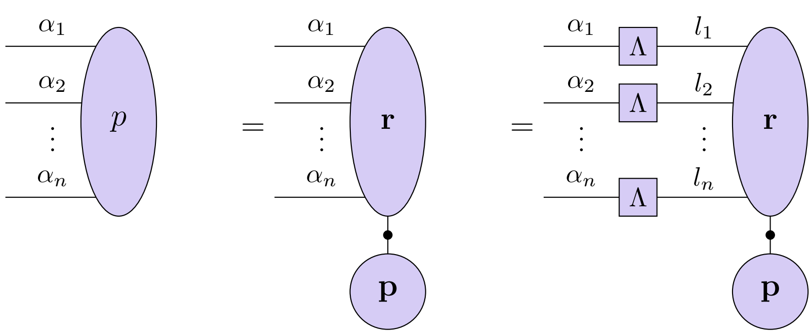
and the associated Lambda matrix, which encodes Compositional Compression takes the form

$$\Lambda^{k_1 k_2 \ldots k_n}_{l_1 l_2 \ldots l_n} \tag{4.60}$$

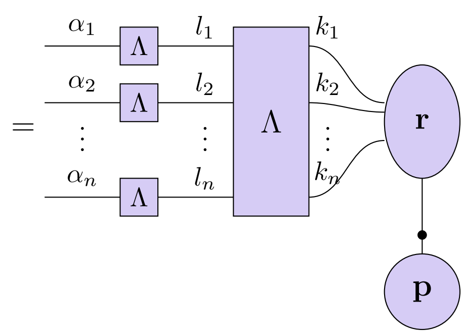
This concludes the details for the second level compression. Let us recapitulate the steps in applying Compositional Compression. The person in the box is interested in n-regions, each associated with the label α_i where i specifies the region. Therefore the person is interested in probabilities of the form $p_{\alpha_1 \ldots \alpha_n}$. They perform tomographic compression on each region. Then they perform compositional compression on the composite region. Mathematically:

$$\begin{aligned}
p_{\alpha_1 \alpha_2 \ldots \alpha_n} &= \mathbf{r}_{\alpha_1 \alpha_2 \ldots \alpha_n} \cdot \mathbf{p} && (4.61)\\
&= \sum_{l_1 l_2 \ldots l_n \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} \ldots \Lambda^{l_n}_{\alpha_n} \mathbf{r}_{l_1 l_2 \ldots l_n} \cdot \mathbf{p} && (4.62)\\
&= \sum_{l_1 l_2 \ldots l_n \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} \ldots \Lambda^{l_n}_{\alpha_n} \sum_{k_1 k_2 \ldots k_n \in \Omega_{1,2,\ldots,n}} \Lambda^{k_1 k_2 \ldots k_n}_{l_1 l_2 \ldots l_n} \mathbf{r}_{k_1 k_2 \ldots k_n} \cdot \mathbf{p} && (4.63)
\end{aligned}$$

and diagrammatically:



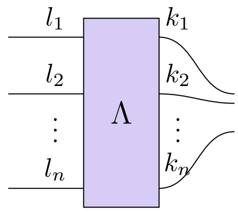
(4.64)



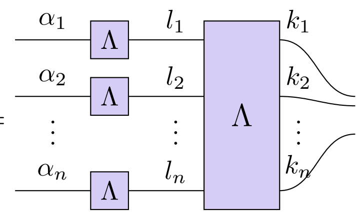
(4.65)

Similar to $\mathbf{r}_{\alpha_1}|_{l_1}$, we can show that the components of the Lambda matrices can be specified through entries of the $\mathbf{r}$-vectors.

$$\mathbf{r}_{l_1 l_2 \ldots l_n}|_{k_1 k_2 \ldots k_n} := \Lambda^{k_1 k_2 \ldots k_n}_{l_1 l_2 \ldots l_n} = \tag{4.66}$$



$$\mathbf{r}_{\alpha_1 \alpha_2 \ldots \alpha_n}|_{k_1 k_2 \ldots k_n} := \sum_{\substack{l_1 l_2 \ldots l_n \\ \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n}} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} \ldots \Lambda^{l_n}_{\alpha_n} \Lambda^{k_1 k_2 \ldots k_n}_{l_1 l_2 \ldots l_n} = \tag{4.67}$$



Causal Adjacency and $\Omega_{1,2}$

Let us take a moment to discuss the implications of Equation 4.56. In what physical situations do we expect to (not) see nontrivial Compositional Compression? If two regions are spatially separated and causally disconnected, then we cannot reduce the number of parameters required to describe them, such as the tensor product in quantum theory, or spatially separated events in relativity - in this case we have $\Omega_{1,2} = \Omega_1 \times \Omega_2$. Even if a causal connection is possible but is mediated through other regions then these other regions can break the correlations between the two regions of interest. Consider for example three quantum gates applied sequentially (like the polariser example from before), then the correlations between the first and third gate will depend on the nature of the second gate (if the second polariser is replaced by a cardboard and new light is sent out then correlations between the first and third polarisers will be broken). In this case as well we have $\Omega_{1,2} = \Omega_1 \times \Omega_2$.

On the other hand if we have strong causal connections between two regions that are not contingent on other regions (for example two consecutive quantum gates, which we can replace by a new composite quantum gate), then we will see that $\Omega_{1,2}$ will be a proper subset of $\Omega_1 \times \Omega_2$. We will then say that the two regions are *causally adjacent* (which we will denote by $R_1 \bowtie R_2$).

When many regions are of interest we can use Compositional Compression between different composite regions to map out causal adjacency. Therefore, Compositional Compression provides us the "mathematical signature for causal structure". We will see some more involved examples of Compositional Compression and a visualisation tool for Omega sets in the following Chapter 5 where we set up the hierarchy and consider multiple regions, and again later in Chapter 6 we see some explicit examples pertaining to the Duotensor Framework (and in turn finite dimensional Quantum Theory).

4.3.4 Third level: Meta Compression

The previous two levels of compression pertain to the compression of single and multiple regions (which lie within a larger predictively well-defined region R), that are encoded combinatorially through the Omega sets and quantitatively through the Lambda matrices. But are the Lambda matrices for different regions independent of each other? A physical theory will in fact often have some structure that will correlate the Lambda matrices itself. The third level of compression will capture this compression of the Lambda matrices, through what is called *the Causaloid*, one of the central mathematical objects in the Causaloid approach, owing the framework its name. It is defined by Hardy as -

> **The causaloid** for a predictively well-defined region R made up of elementary regions R_x is defined to be that thing represented by any mathematical object which can be used to obtain $\mathbf{r}_{\alpha_{\mathcal{O}}}(R_{\mathcal{O}})$ for all measurements $\alpha_{\mathcal{O}}$ in region $R_{\mathcal{O}}$ for all $R_{\mathcal{O}} \subseteq R$.

To begin, let us find a mathematical object that specifies the Causaloid for a predictively well-defined region R, utilising the two levels of compression. The smallest component region of R consists of elementary regions R_x, such that $\cup_x R_x = R$. R can be expressed as consisting non-elementary regions $R_{\mathcal{O}}$ which itself consist of elementary regions. The region $R_{\mathcal{O}}$ can be specified in terms of elementary regions in the following way:

$$R_{\mathcal{O}} = \bigcup_{i \in \mathcal{O}} R_i \quad \text{where} \quad \mathcal{O} = \{x, x', ..., x''\} \tag{4.68}$$

Of course $R_{\mathcal{O}}$ can reduce to a elementary region if $\mathcal{O} = \{x\}$. Now let us discuss the general region $R_{\mathcal{O}}$ through tomographic and compositional compression. In the pre-compression level the general measurement in region $R_{\mathcal{O}}$ is labelled by $\alpha_{\mathcal{O}}$, which decomposes into local measurements at each elementary region part of $\mathcal{O}$:

$$\alpha_{\mathcal{O}} = \alpha_x \alpha_{x'} \cdots \alpha_{x''} \tag{4.69}$$

Each region $R_{\mathcal{O}}$ is associated with tomographic compression. Each composite region $R_{\mathcal{O}}$ (say for $R_{\mathcal{O}} = R_{\mathcal{O}'} \cup R_{\mathcal{O}''} \cup ...$) is associated with compositional compression. But remember that we discussed that tomographic compression of a composite region is equivalent to tomographic compression of its constituents along with compositional compression on the composite region. In light of this it is in fact sufficient to consider tomographic compression of all elementary regions R_x and compositional compression over all composite regions $R_{\mathcal{O}}$. The tomographic compression of elementary region is given by:

$$\Lambda^{l_x}_{\alpha_x}(x, \Omega_x) \tag{4.70}$$

where the Lambda matrix is a function of (x, Ω_x) – the region and the Omega set given that it may not be unique. Similarly, the compositional compression over non-elementary regions $R_{\mathcal{O}}$ is given by

$$\Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}}) \tag{4.71}$$

where, similar to the label $\alpha_{\mathcal{O}}$, the tomographic Omega set label $l_{\mathcal{O}}$ for the fiducial measurements of the composite region will also decompose into the fiducial measurements of the constituent elementary regions, while the compositional Omega set label $k_{\mathcal{O}}$ will not. That is:

$$l_{\mathcal{O}} \equiv l_x l_{x'} \cdots l_{x''} \in \Omega_x \times \Omega_{x'} \times \cdots \times \Omega_{x''} \quad \text{and} \quad k_{\mathcal{O}} \equiv k_x k_{x'} \cdots k_{x''} \in \Omega_{\mathcal{O}} \tag{4.72}$$

Note that if $\mathcal{O} = \{x\}$, then there is no compositional compression, since we are considering a single tomographically compressed region, and thus only a single Omega set Ω_x is relevant:

$$\Lambda^{k_x}_{l_x} = \delta^{k_x}_{l_x} \quad \text{since} \quad l_x, k_x \in \Omega_x \tag{4.73}$$

It is worth noting that fully specifying the Lambda matrix $\Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}})$ would include specifying the region and a choice of Omega set, but it is easy to transform the Lambda matrix from one set of Omega choice to another using the following relation:

$$\Lambda^{k'_{\mathcal{O}}}_{l_{\mathcal{O}}}(\mathcal{O}, \Omega'_{\mathcal{O}}) = \sum_{k_{\mathcal{O}}} [\Lambda^{k_{\mathcal{O}}}_{k'_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}})]^{-1} \Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}}) \tag{4.74}$$

where

$$\Lambda^{k_{\mathcal{O}}}_{k'_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}}) \quad k'_{\mathcal{O}} \in \Omega'_{\mathcal{O}} \tag{4.75}$$

is a square matrix whose inverse exists, given two choices of Omega sets $\Omega_{\mathcal{O}}, \Omega'_{\mathcal{O}}$. Note that the matrix is square since $|\Omega_{\mathcal{O}}| = |\Omega'_{\mathcal{O}}|$. Also note that $k'_{\mathcal{O}}$ and $k_{\mathcal{O}}$ both belong to (different) subsets of $\Omega_x \times \Omega_{x'} \times \cdots \times \Omega_{x''}$ since there is always a $l_{\mathcal{O}}$ for every $k'_{\mathcal{O}}$ and $k_{\mathcal{O}}$. Similar argument gives the relation for transformation of tomographic Lambda matrices $\Lambda^{l_x}_{\alpha_x}(x, \Omega_x)$ pertaining to elementary regions. The transformation rule of Lambdas tells us that considering any one choice of Omega set for each region $\mathcal{O}$ is sufficient for the specification of the Causaloid.

Given these details we can write down one mathematical object that surely specifies the causaloid (denoted by $\mathbf{\Lambda}$):

$$\mathbf{\Lambda} = \left[\begin{array}{ll} \Lambda^{l_x}_{\alpha_x}(x, \Omega_x) & : \text{for a } \ \Omega_x \ \text{ for each elementary } \ R_x \\ \Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}}(\mathcal{O}, \Omega_{\mathcal{O}}) & : \text{for a } \ \Omega_{\mathcal{O}} \text{ for each non-elementary } R_{\mathcal{O}} \subseteq R \end{array} \right] \tag{4.76}$$

Clearly this satisfies the definition of the Causaloid since the lambda matrices can be used to calculate any $\mathbf{r}$-vector using the results stated in previous subsections. From Equations 4.25 and 4.66 we have

$$\mathbf{r}_{\alpha_x}|_{l_x} = \Lambda^{l_x}_{\alpha_x} \quad \text{ and } \quad \mathbf{r}_{\alpha_{\mathcal{O}}}|_{l_{\mathcal{O}}} = \sum_{l_{\mathcal{O}}} \Lambda^{l_x}_{\alpha_x} \Lambda^{l_{x'}}_{\alpha_{x'}} \cdots \Lambda^{l_{x''}}_{\alpha_{x''}} \Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}} \tag{4.77}$$

This specification of the Causaloid is independent of any physical theory. But when considering a particular physical theory the specification of the Causaloid can be compressed given some structure. In such a scenario we expect to be able to calculate some Λ matrices from others and thus, we can take some subset of the Λ matrices. Let us label such a subset by i, then we have

$$\mathbf{\Lambda} = [\Lambda(i) : i = 1 \text{ to } I \mid \textit{rules} \,] \tag{4.78}$$

where *rules* prescribe us the rules for deducing all Λ's from the given subset of $\Lambda(i)$'s. The subset may not be unique, but it will tell us some key aspects of the physical theory. The role of the Causaloid is captured in the following words by Hardy:

> "We will call $\mathbf{\Lambda}$ the *causaloid* (because it contains information about the propensities for different causal structures). This is the central mathematical object in this paper. For any particular physical theory the causaloid is fixed... In fact, once we know the causaloid we can perform any calculation possible in the physical theory. Consequently, the causaloid can be regarded as a specification of a physical theory itself." (Section 4.4, [42])

We may regard going from Equation 4.76 to 4.78 as a kind of physical compression. We name this third level of compression as *Meta Compression* since it does not look like the first two levels and it acts on sets of Lambda's, which is called the Causaloid $\mathbf{\Lambda}$. In the coming Chapter 5 we will study Meta Compression in further detail, providing a class of identities for Compositional Lambdas that will provide a way to classify physical theories.

The Causaloid Product

Let us discuss products. Within Quantum theory there are three basic ways of putting two operators together in quantum theory: $\hat{A}\hat{B}$ (sequential, temporal and causally adjacent), $\hat{A}?\hat{B}$ (temporal but not causally adjacent), and $\hat{A} \otimes \hat{B}$ (spatially separate). One of the goals of the Causaloid Framework, with interest in studying indefinite causality, is to be able to treat space and time on an equal footing and thus to treat these different kinds of products on an an equal footing. To this end the *causaloid product* is defined. While this is in a general framework, the causaloid product can be shown to unify the different types of products within quantum theory.

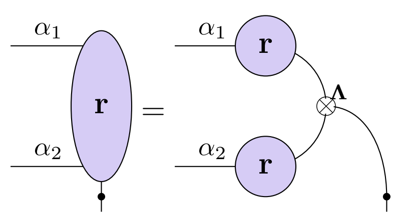
Given $\mathbf{r}_{\alpha_1}$ measurement vector in $R_{\mathcal{O}_1}$ (or simply R_1) and $\mathbf{r}_{\alpha_2}$ measurement vector in $R_{\mathcal{O}_2}$ (or simply R_2) such that $R_{\mathcal{O}_1} \cap R_{\mathcal{O}_2} = \phi$. The causaloid product $\otimes^{\Lambda}$ is defined as:

$$\mathbf{r}_{\alpha_1\alpha_2}(\mathcal{O}_1 \cup \mathcal{O}_2) = \mathbf{r}_{\alpha_1}(\mathcal{O}_1) \otimes^{\Lambda} \mathbf{r}_{\alpha_2}(\mathcal{O}_2) \tag{4.79}$$

or simply

$$\mathbf{r}_{\alpha_1\alpha_2} = \mathbf{r}_{\alpha_1} \otimes^{\Lambda} \mathbf{r}_{\alpha_2} \tag{4.80}$$

where the labels α implicitly specify the regions. Diagrammatically we represent the causaloid product with the symbol $\otimes^{\Lambda}$:

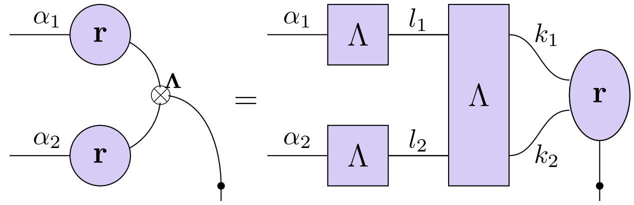


(4.81)

The way to expand the causaloid product is given by Lambda matrices which in turn are given by the causaloid. The component form for Equations 4.79-4.81 using Equation 4.66 is given by:

$$\mathbf{r}_{\alpha_1\alpha_2}|_{k_1k_2} = \sum_{l_1l_2 \in \Omega_1 \times \Omega_2} (\mathbf{r}_{\alpha_1}|_{l_1})(\mathbf{r}_{\alpha_2}|_{l_2})\Lambda^{k_1k_2}_{l_1l_2} \tag{4.82}$$

Diagrammatically we can see that the causaloid product may be expanded as follows:



(4.83)

This can be massaged into a more illuminating form for the purposes of identifying the kinds of different products the unified causaloid product can reduce to:

$$\mathbf{r}_{\alpha_1\alpha_2}|_{k_1k_2} = \sum_{l_1l_2\in\Omega_1\times\Omega_2} (\mathbf{r}_{\alpha_1}|_{l_1})(\mathbf{r}_{\alpha_2}|_{l_2})\Lambda^{k_1k_2}_{l_1l_2} \tag{4.84}$$

$$= \sum_{l_1l_2\in\Omega_{1,2}} (\mathbf{r}_{\alpha_1}|_{l_1})(\mathbf{r}_{\alpha_2}|_{l_2})\delta^{k_1k_2}_{l_1l_2} + \sum_{l_1l_2\in\Omega_1\times\Omega_2-\Omega_{1,2}} (\mathbf{r}_{\alpha_1}|_{l_1})(\mathbf{r}_{\alpha_2}|_{l_2})\Lambda^{k_1k_2}_{l_1l_2} \tag{4.85}$$

$$= (\mathbf{r}_{\alpha_1}|_{k_1})(\mathbf{r}_{\alpha_2}|_{k_2}) + \sum_{l_1l_2\in\Omega_1\times\Omega_2-\Omega_{1,2}} (\mathbf{r}_{\alpha_1}|_{l_1})(\mathbf{r}_{\alpha_2}|_{l_2})\Lambda^{k_1k_2}_{l_1l_2} \tag{4.86}$$

where we first expand the sum into two terms over different sets since $\Omega_{1,2} + (\Omega_1 \times \Omega_2 - \Omega_{1,2}) = \Omega_1 \times \Omega_2$, and in the second step we use the relation

$$\Lambda^{k_1k_2}_{l_1l_2} = \delta^{k_1k_2}_{l_1l_2} \quad \text{for} \quad l_1, l_2 \in \Omega_{1,2} \tag{4.87}$$

There are two special cases for Equation 4.84 given Equation 4.56.

1. *No Compositional Compression* $\Omega_{1,2} = \Omega_1 \times \Omega_2$ such that $|\Omega_{1,2}| = |\Omega_1||\Omega_2|$.

2. *Non-trivial Compositional Compression* $\Omega_{1,2} \subset \Omega_1 \times \Omega_2$ such that $|\Omega_{1,2}| < |\Omega_1||\Omega_2|$.

It follows that

$$\text{if} \quad \Omega_{1,2} = \Omega_1 \times \Omega_2 \quad \text{then} \quad \mathbf{r}_{\alpha_1\alpha_2} = \mathbf{r}_{\alpha_1} \otimes \mathbf{r}_{\alpha_2} \tag{4.88}$$

where $\otimes$ stands for the ordinary tensor product. Hence the ordinary tensor product is a special case of the causaloid product.

We will see in the Chapter 6 through the duotensor formalism, that in (finite-dimensional) quantum theory, typically $\Omega_{1,2} = \Omega_1 \times \Omega_2$ corresponds to the products $\hat{A} \otimes \hat{B}$ as well as $\hat{A}?\hat{B}$. Since the total number of real parameters after taking the product is equal to the product of the number from each operator we have $|\Omega_{1,2}| = |\Omega_1||\Omega_2|$.

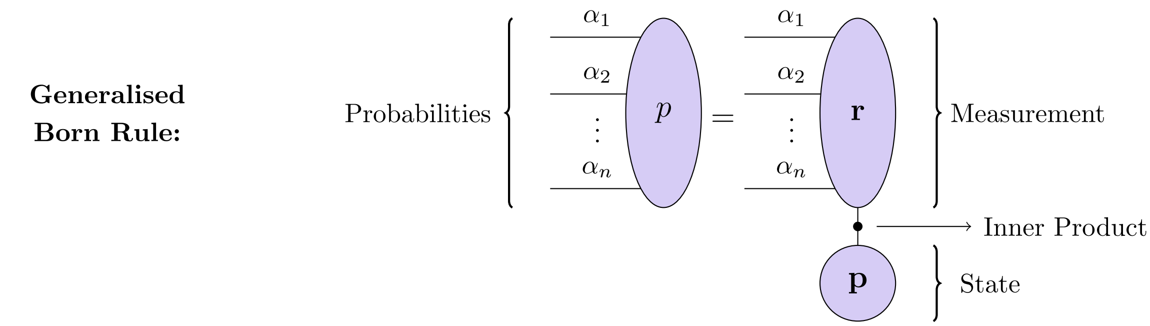
Non trivial Compositional compression occurs ($|\Omega_{1,2}| < |\Omega_1||\Omega_2|$) when two regions are causally adjacent, such that what happens in one region depends, at least partially, on what is done in the other region in a way that cannot be tampered by other regions in R. In quantum theory when we take the product $\hat{A}\hat{B}$, the total number of real parameters in the product is equal to the number in $\hat{A}$ (or $\hat{B}$), giving the special case $|\Omega_{1,2}| = |\Omega_1| = |\Omega_2|$.

Therefore, the causaloid product is expected to bring together all types of different products a physical theory will have to offer, which can be calculated once the causaloid is specified. This concludes the discussion of the three levels of compression.

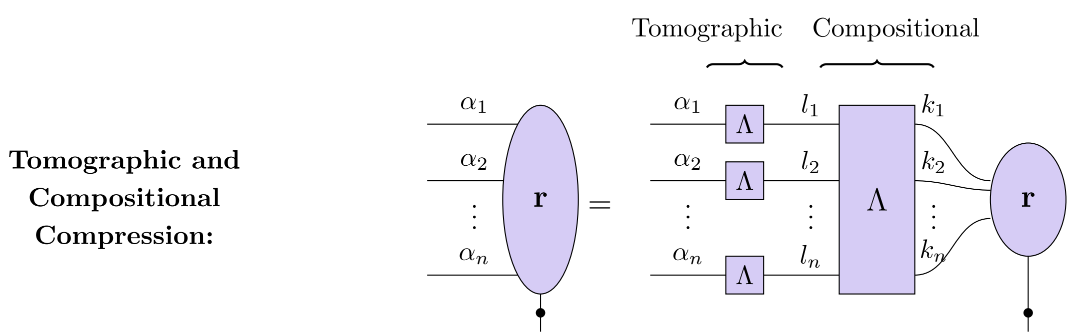
4.3.5 Synopsis: Three Levels of Physical Compression

We presented new diagrammatics for the Causaloid Framework's three levels of physical compression. Here we provide a short and handy synopsis for quick reference.

We considered a large region R which is predictively well-defined. We focus on disjoint regions $R_1, R_2, ...$ within R. We expect the framework to help us predict probabilities. Using the generalised Born rule given linearly in terms of states and measurements gives probabilities.



The **r** measurement vector can be compressed. Tomographic Compression concerns a single region and Compositional Compression concerns multiple disjoint regions.

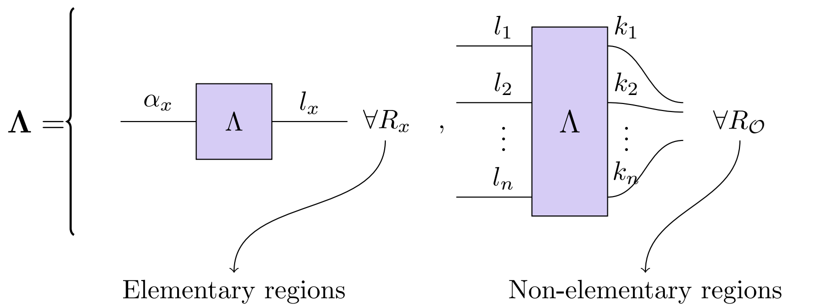


The label change between these levels of compression is given by:

$$
\begin{aligned}
& \alpha_1, \alpha_2, \ldots \alpha_n \in \Gamma_1 \times \Gamma_2 \times \ldots \Gamma_n \\
\longrightarrow \quad & l_1, l_2, \ldots l_n \in \Omega_1 \times \Omega_2 \times \ldots \Omega_n \\
\longrightarrow \quad & k_1, k_2, \ldots k_n \in \Omega_{1,2,\ldots n}
\end{aligned}
$$

Given the first two levels of Compression, we get the set of Lambdas for any region in R which helps us specify *the Causaloid* $\mathbf{\Lambda}$, which may be considered as "a specification of a physical theory itself".

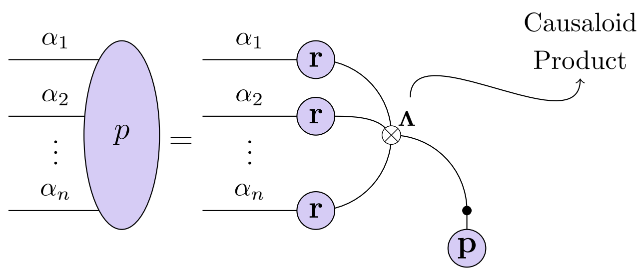
The Causaloid:



(4.89)

The third level or Meta Compression provides us with *rules* to be able to calculate any Lambda Matrix given a reduced set which gives a shorter specification of the Causaloid. The Causaloid in turn helps us define the Causaloid Product $\otimes^{\Lambda}$, that unifies different kinds of spatio-temporal products within a theory, and gives us any general $\mathbf{r}$-vectors.

Causaloid Product:



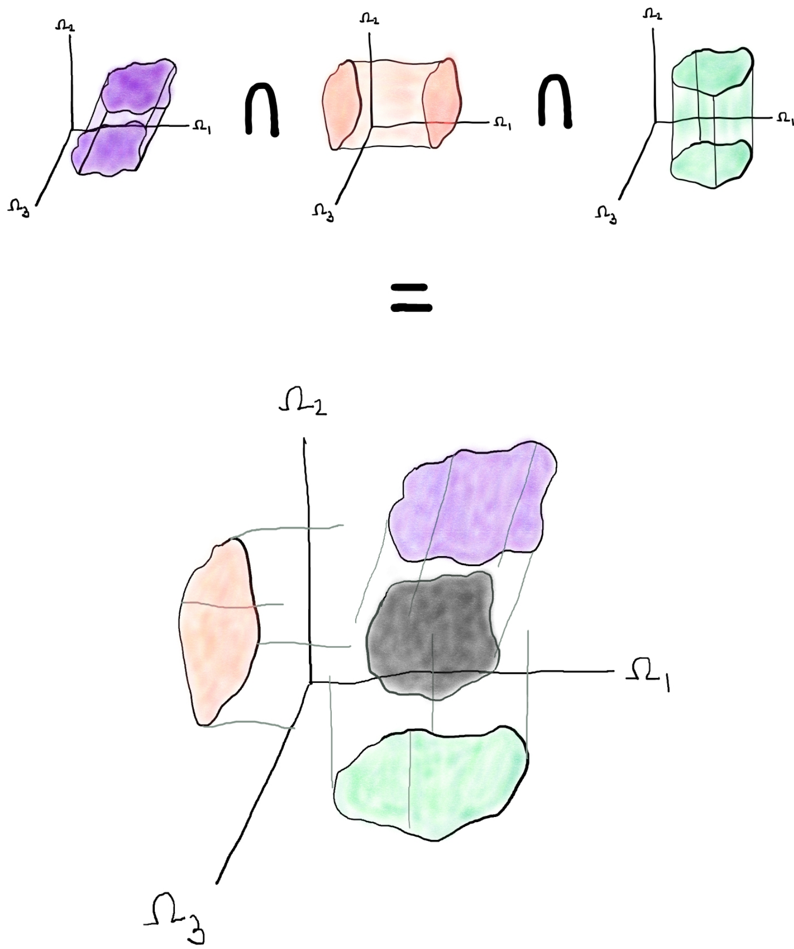
This brings us full circle. The framework prescribes a way to organise recorded data using the structure of physical theories, by distilling down to a compressed version and studying correlations.

The work presented in this Chapter was born out of a motivation to revisit the Causaloid Framework, a seminal work that paved the way for many works on indefinite causal structures. To this end, we presented a diagrammatic representation for the three levels of physical compression, along with a review of the framework. We bear the hope that this diagrammatic review will make this framework more accessible to those interested, who are engaged studying indefinite causal structures.

In the coming Chapters, we will utilise the diagrammatics presented here, to study Meta-Compression further giving us a Hierarchy (Chapter 5); and to diagrammatically apply the Causaloid Framework to the Duotensor Framework, placing it into the second rung of the Hierarchy (Chapter 6).

Chapter 4: Statement of Contribution

In this Chapter, I provide an accessible review of the Causaloid Framework by Lucien Hardy [41][42] for anyone interested in becoming familiar with this framework. The main contribution of this Chapter are new diagrammatics defined in Section 4.3, as well as updated nomenclature around the three levels of compression. This was largely worked on by myself and the next two Chapters build upon the new diagrammatics. The Chapter is solely written by me, based on mentioned sections the papers [41][42] by Lucien Hardy. This Chapter contains unpublished work that was presented at various conferences (QISS HK 2020, Q-Turn 2020 [90], APS 2021 [91], Quantizing Time 2021 [89], QPL 2021) as the work developed.



"Why so meta?",
said the sets
to the matrices
at the intersection of
Cartesian products
"Well, we ran out of data to compress!" [3]

[3]Chapter 5, illustration for Equation 5.63 using visualisation tools from Section 5.3

Chapter 5

A Hierarchy through Meta-Compression

Physical Compression gives the specification of a theory: the Causaloid $\mathbf{\Lambda}$
What can we learn about the causal structure through Meta-Compression of $\mathbf{\Lambda}$?
We study Meta-Compression and characterise a hierarchy of theories

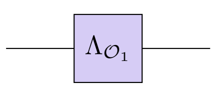
Meta Compression, the third level of physical compression from the Causaloid Framework (CF), gives us a reduced set of Lambda matrices required to specify the Causaloid $\mathbf{\Lambda}$ and some *rules* that tell us how to calculate all Lambda matrices from this reduced set. We propose a way to categorise families of physical theories into a *Hierarchy*, by studying Meta Compression and characterising the Causaloid **Lambda**. The rungs of the Hierarchy are given by what we call $\Lambda_{\mathcal{O}_d}$*-sufficiency*. We will utilise CF's diagrammatic representation from Chapter 4 to do so. A known Hierarchy for Generalised Probability Theories (GPTs) [99, 40, 93] , that differentiates Classical Probability Theory from Quantum theory, is contingent on Tomographic Compression. The Hierarchy, presented here, is different, it is contingent on Compositional Compression that captures the causal structure and places Quantum theory as well as Classical Probability theory on the same rung in this respect, as we see in Chapter 6.

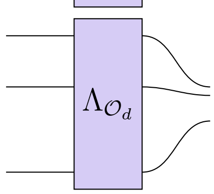
Outline of the Chapter

In Section 5.1 we define the Hierarchy through characterising the Causaloid **Lambda** through $\Lambda_{\mathcal{O}_d}$-sufficiency. In Section 5.2 we show an example of Lambda matrix relations by Hardy from [41]. We will fully characterise such relations. In order to do so, we present a Toolkit to work with Cartesian products of Omega sets and subsets thereof, that will be crucial to formalise our results. Then we begin giving the form for Lambda matrices one rung of the Hierarchy at a time, in Sections 5.4-5.6. The Section 5.4 is the trivial rung where no Compositional Compression occurs. Section 5.5, the second rung, will consider a tripartite region and consists of detailed causal relations and Lambda matrix as well as Omega set results. In Section 5.6 we see how to continue the characterisation for the higher rungs, and will provide a general form for the Omega sets. In Section 5.7 we discuss the road ahead: what we have learnt by revising the Causaloid Framework and the directions in which the current work may be applied.

5.1 Defining the Hierarchy

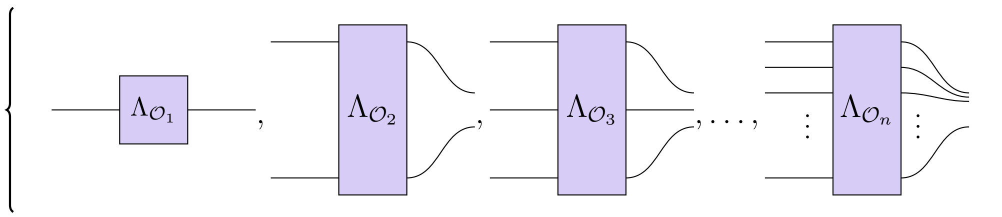
As before, we consider the predictively well-defined region R made of (disjoint) elementary regions R_x, and a general region $R_{\mathcal{O}}$ where the region can be composed of elementary regions $\mathcal{O} = \{x, x', ...x''\}$ such that $R_{\mathcal{O}} = R_x \cup R_{x'} \cup ...R_{x''}$. We will say that the *size* $|.|$ of a region labelled by $\mathcal{O}$ is the number of elementary regions it consists of, for example $|\mathcal{O} = \{x, x', x'', x'''\}| = 4$. The size of R will be the number of elementary regions it consists of, which we will say is n. We will now introduce a shorthand for Lambda matrices particularly useful for the hierarchy we will define. Recall that the specification of the causaloid from Equation 4.76 includes Tomographic Compression matrices for single elementary regions $\Lambda^{l_x}_{\alpha_x}$ and Compositional Compression matrices for non-elementary regions $\Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}}$. We can refer to the Compression matrices through a shorthand using the size of their associated region while leaving out the information of indices:

$\Lambda_{\mathcal{O}_1} =$  $:= \text{any} \quad \Lambda^{l_x}_{\alpha_x} = \Lambda^{k_x}_{\alpha_x} \quad \text{where} \quad |\mathcal{O} = \{x\}| = 1$ (5.1)

$\Lambda_{\mathcal{O}_d} =$  $:= \text{any} \quad \Lambda^{k_{\mathcal{O}}}_{l_{\mathcal{O}}} \text{ or } \Lambda^{k_{\mathcal{O}}}_{\alpha_{\mathcal{O}}} \quad \text{where} \quad |\mathcal{O} = \{x, x', ...x''\}| = d > 1$ (5.2)

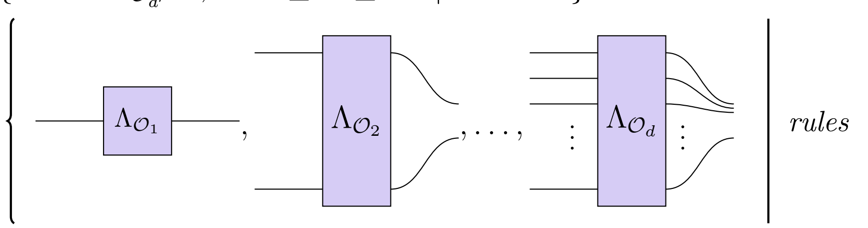
Then we can re-specify the causaloid:

$$\mathbf{\Lambda} = \{ \quad \text{all} \quad \Lambda_{\mathcal{O}_d} \quad , \quad \forall 1 \leq d \leq n \} \tag{5.3}$$

$=$  (5.4)

Now we are ready to define the Hierarchy. We ask when and under what conditions is d-(elementary)-region compositional compression sufficient to provide all possible m-region compositional compression matrices for $n \geq m > d$. That is under what conditions does Meta compression lead to the specification of the causaloid to become:

$$\mathbf{\Lambda} \equiv \mathbf{\Lambda}_{\mathcal{O}_d} = \{ \quad \text{all} \quad \Lambda_{\mathcal{O}_{d'}} \quad , \quad \forall 1 \leq d' \leq d \quad | \quad \textit{rules} \quad \} \quad \text{where} \quad d < n$$

$=$  (5.5)

If d-(elementary)-region compositional compression is sufficient to calculate all higher region compositional compression, for the *smallest* possible d, then we say that the physical theory is $\Lambda_{\mathcal{O}_d}$-sufficient[1] and thus, belongs to the d^{th} rung of the Hierarchy. The Causaloid after Meta Compression for a $\Lambda_{\mathcal{O}_d}$-sufficient physical theory is denoted by $\mathbf{\Lambda}_{\mathcal{O}_\mathbf{d}}$ (in so far that the region R is large enough to well represent a physical theory). In such a case it would be possible to find functions of the form:

$$\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}, \ldots, \Lambda_{\mathcal{O}_d}) \quad \text{and} \quad \Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2}, \ldots, \Omega_{\mathcal{O}_d}) \quad \text{where} \quad d < m \leq n \tag{5.6}$$

We will provide concrete steps to calculate other Lambda matrices (the *rules*). We provide tools to work with Omega sets and present a formula for how the n-region Omega sets are determined in terms of d-region Omega sets when we have d-(elementary)-region compositional compression sufficiency.

[1]Usage: $\Lambda_{\mathcal{O}_d}$-sufficiency (*noun*), $\Lambda_{\mathcal{O}_d}$-sufficient (*adjective)*

5.2 Proof of concept *a la Hardy*

To begin understanding the kind of conditions one can find to calculate a Lambda matrix from others, we will first go over an example provided by Hardy in Section 23, [41], as a proof of concept, before we are ready to dive into the deep end of the pool.

Consider three regions $R_1, R_2, R_3 \in R$. We begin after taking Tomographic Compression into account. Now, we start with $\mathbf{r}_{l_1 l_2 l_3} \cdot \mathbf{p}$ upon which compositional compression is first applied to region $R_2 \cup R_3$ and then applied to region $R_1 \cup R_2$.

$$\begin{aligned}
& \mathbf{r}_{l_1 l_2 l_3}(R_1 \cup R_2 \cup R_3) \cdot \mathbf{p}(R_1 \cup R_2 \cup R_3) && (5.7)\\
= \; & \mathbf{r}_{l_2 l_3}(R_2 \cup R_3) \cdot \mathbf{p}_{l_1}(R_2 \cup R_3) && (5.8)\\
= \; & \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{k_2' k_3}(R_2 \cup R_3) \cdot \mathbf{p}_{l_1}(R_2 \cup R_3) && (5.9)\\
= \; & \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{l_1 k_2' k_3}(R_1 \cup R_2 \cup R_3) \cdot \mathbf{p}(R_1 \cup R_2 \cup R_3) && (5.10)\\
= \; & \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{l_1 k_2'}(R_1 \cup R_2) \cdot \mathbf{p}_{k_3}(R_1 \cup R_2) && (5.11)\\
= \; & \sum_{k_1 k_2 \in \Omega_{1,2}} \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_1 k_2'}^{k_1 k_2} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{k_1 k_2}(R_1 \cup R_2) \cdot \mathbf{p}_{k_3}(R_1 \cup R_2) && (5.12)\\
= \; & \sum_{k_1 k_2 \in \Omega_{1,2}} \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_1 k_2'}^{k_1 k_2} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{k_1 k_2 k_3}(R_1 \cup R_2 \cup R_3) \cdot \mathbf{p}(R_1 \cup R_2 \cup R_3) && (5.13)
\end{aligned}$$

where $\mathbf{p}_{l_1}$ is the tomographically compressed state for region $R_2 \cup R_3$, and $\mathbf{p}_{k_3}$ is the state for region $R_1 \cup R_2$ after compositional compression was done on $R_2 \cup R_3$. This calculation is reminiscent of a calculation we saw earlier (Equation 4.37-4.43) since arguments of regions are shifted around in a similar way. One can also perform these calculations diagrammatically but we skip the same for now. Effectively we have:

$$\mathbf{r}_{l_1 l_2 l_3} \cdot \mathbf{p} = \sum_{k_1 k_2 \in \Omega_{1,2}} \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_1 k_2'}^{k_1 k_2} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{k_1 k_2 k_3} \cdot \mathbf{p} \tag{5.14}$$

Since this is true for any $\mathbf{p}$ we have

$$\mathbf{r}_{l_1 l_2 l_3} = \sum_{k_1 k_2 \in \Omega_{1,2}} \sum_{k_2' k_3 \in \Omega_{2,3}} \Lambda_{l_1 k_2'}^{k_1 k_2} \Lambda_{l_2 l_3}^{k_2' k_3} \mathbf{r}_{k_1 k_2 k_3} \tag{5.15}$$

We also have through the definition of three region compositional compression on $R_1 \cup R_2 \cup R_3$

$$\mathbf{r}_{l_1 l_2 l_3} = \sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \Lambda^{j_1 j_2 j_3}_{l_1 l_2 l_3} \mathbf{r}_{j_1 j_2 j_3} \tag{5.16}$$

where we use j as indices instead of k. We will stick to the convention of using $l, k, j, i, ..$ for $\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}, \Lambda_{\mathcal{O}_3}, \Lambda_{\mathcal{O}_4}...$ to be able to differentiate the compositional compression applied to different number of regions as relevant. We will also continue to solely use k indices to refer to general compositional compression if differentiating by the size of the region is not relevant. Comparing the Equations 5.15 and 5.16 gives us the following Lambda relations conditional on some Omega relations.

$$\text{If} \quad \Omega_{1,2,3} = \Omega_{1,2} \times \Omega_{\not{2},3} \quad \text{and} \quad \Omega_{2,3} = \Omega_{2,\not{3}} \times \Omega_{\not{2},3} \quad \text{then} \quad \Lambda^{k_1 k_2 k_3}_{l_1 l_2 l_3} = \sum_{k'_2 \in \Omega_{2\not{3}}} \Lambda^{k_1 k_2}_{l_1 k'_2} \Lambda^{k'_2 k_3}_{l_2 l_3} \tag{5.17}$$

where the notation $\Omega_{\not{2},3}$ means that we form the set of all k_3 for which there exists an element in $k_2 k_3 \in \Omega_{2,3}$ with matching k_3 (for example if $\Omega_{2,3} = \{12, 13, 24, 34\}$ then $\Omega_{\not{2},3} = \{2, 3, 4\}$). We will discuss this notation further in Section 5.3. This gives us an example of how Lambda matrices can be related to each other given some conditions on the Omega sets are met. By no means are these the most general conditions on the Omega sets to find relations between the Lambda matrices and thus we ask can we do better? Hardy leaves us with a research program:

> "It should be possible to characterise all possible relationships between lambda matrices so we know how much freedom we have in specifying the causaloid. These constraints are likely to give us deep insight into the possible nature of physical theories." (Section 23, [41])

We attempt to characterise precisely these relationships between Lambda matrices and Omega sets through studying the Hierarchy and share the insights these provide us with in the remainder of this Chapter.

5.3 Toolkit for Working with Omega Sets

Now we will provide some simple mathematical tools based on set theory applied to product sets. These tools will help us work with Omega sets, and will be used in the following Subsections to help study the Hierarchy.

5.3.1 Projections

We will now formally define the projection map which we used in Equation 5.17. Consider the indices $k_1 k_2 \dots k_n \in \Omega_{1,2,\dots,n}$ where $\Omega_{1,2,\dots,n} = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$. There may be a situation where we only want to discuss the indices $k_1, k_2, \dots, k_{n-1}$, what Omega set would these indices belong to? The projection map $\mathcal{P}_m(.)$ goes from some number of regions to a smaller set, where the m^{th} region is removed, notated as follows:

$$\mathcal{P}_m : \Omega_{1,2,\dots,m,\dots,n} \to \Omega_{1,2,\dots,\not{m},\dots,n} \tag{5.18}$$

In the case when we apply the projection map to a set which can be factorised into a Cartesian product of sets, the projection map will yield a new set that can also be factorised:

$$\mathcal{P}_i : \Omega_1 \times \dots \times \Omega_i \dots \times \Omega_n \to \Omega_1 \times \dots \times \Omega_{i-1} \times \Omega_{i+1} \dots \times \Omega_n \tag{5.19}$$

Here are some examples of application of the projection map:

$$\begin{aligned}
\mathcal{P}_1 &: \{(1,1),(1,2),(2,1),(2,3),(4,2)\} \to \{(\not{1},1),(\not{1},2),(\not{2},1),(\not{2},3),(\not{4},2)\} = \{1,2,3\} \\
\mathcal{P}_2 &: \{(1,1),(1,2),(2,1),(2,3),(4,2)\} \to \{(1,\not{1}),(1,\not{2}),(2,\not{1}),(2,\not{3}),(4,\not{2})\} = \{1,2,4\}
\end{aligned}$$

It must be pointed out that the projection map causes loss of information. Consider some set $\Omega_{1,2} \subseteq \Omega_1 \times \Omega_2$. If we apply the projection map $\mathcal{P}_1$ we have $\Omega_{\not{1},2}$ and if we apply $\mathcal{P}_2$, we have $\Omega_{1,\not{2}}$. Observe that generally $\Omega_{1,2} \neq \Omega_{1,\not{2}} \times \Omega_{\not{1},2}$. In fact, we find that they are strictly related in the following way (a result which will become important later, see 5.5.1):

$$\Omega_{1,2} \;\subseteq\; \Omega_{1,\not{2}} \times \Omega_{\not{1},2} \;\subseteq\; \Omega_1 \times \Omega_2 \tag{5.20}$$

5.3.2 Filling

The filling map $\mathcal{F}_{\mathcal{O}}(.)$ is in some sense the opposite of the projection map, it takes some Omega set $\Omega_{i,j}$ and pads it with Tomographic Omega sets for additional regions specified to make it a valid Omega set for the composite region labelled by $\mathcal{O}$, where $i, j \in \mathcal{O}$:

$$\mathcal{F}_{\mathcal{O}} : \Omega_{i,j} \to \Omega_1 \times \dots \times \Omega_{i,j} \times \dots \times \Omega_n \quad \text{where} \quad i, j \in \mathcal{O} = \{1, 2, \dots, n\} \tag{5.21}$$

For example:

$$\mathcal{F}_{\{1,2,4,5,7\}} : \Omega_{2,4} \to \Omega_1 \times \Omega_{2,4} \times \Omega_5 \times \Omega_7 \tag{5.22}$$

The filling map will be useful when trying to take intersections of sets that do not have matching regions such as $\Omega_{1,2}$ and $\Omega_{2,3}$. In this case we will fill both sets to $\mathcal{O} = \{1,2,3\}$ and then take the intersection $F_{\mathcal{O}}(\Omega_{1,2}) \cap F_{\mathcal{O}}(\Omega_{2,3})$.

5.3.3 Permutations and Order

When considering some predictively well-defined region R consisting of elementary regions R_x, the label x is an integer and sets an arbitrary convention for the order of the regions. The order of the regions and this convention do not have any physical bearing, nonetheless it is useful to represent calculations, and one would like to stick to a certain convention when performing any calculations. The standard convention we define is the order of labels in increasing x such that for some Omega set $\Omega_{i,j,\ldots,n}$, we have $i < j < \cdots < n$.

It may so happen that upon applying projection and filling maps on Omega sets, one may end up with a set where the standard convention cannot be followed, for example a term such as $\Omega_{1,4,\not{5}} \times \Omega_2$, but it may be desirable to be able to bring back the set to the standard order for the purpose of representing the calculation. In such a case the ordering map $\mathcal{O}(.)$ (not to be confused with the label for a composite region $\mathcal{O}$ which shows up as a subscript for regions) will become useful.

The ordering map $\mathcal{O}(.)$ is a permutation map that permutes back the regions of an Omega set back to the standard ordering of $R_1, R_2, ..., R_n$. If the regions are some permutation σ of $\{1, 2, \ldots, n\}$ then the ordering map is the application of its inverse, namely σ^{-1}:

$$\mathcal{O} : \Omega_{\sigma(1,2,\ldots,n)} \rightarrow \Omega_{1,2,\ldots,n} \tag{5.23}$$

From our example if $\Omega_{1,4,\not{5}} = \{(1, 2), (1, 4), (5, 6)\}$ and $\Omega_2 = \{3, 7\}$ then:

$$\mathcal{O}(\Omega_{1,4,\not{5}} \times \Omega_2) = \{(1, 3, 2), (1, 3, 4), (5, 3, 6), (1, 7, 2), (1, 7, 4), (5, 7, 6)\} \tag{5.24}$$

where one can see the elements $3, 7$ of Ω_2 in the second position between elements of Ω_1 and Ω_4. Note that $\mathcal{O}(\Omega_{1,4,\not{5}} \times \Omega_2) = \mathcal{O}(\Omega_2 \times \Omega_{1,4,\not{5}})$ since the elements of a set are not ordered.

5.3.4 Visualising Omega Sets

While sets themselves are not ordered and do not have much structure, the Causaloid Framework works extensively with Cartesian products of sets and their subsets, the relations between these sets are crucial. It would thus be beneficial to provide a way to visualise Omega sets and action of the maps $(\mathcal{P}_m(.), \mathcal{F}_{\mathcal{O}}(.), \mathcal{O}(.))$ discussed until now. This will aid with grasping the main results of this work in a better manner.

Any Omega set of an elementary region will have some elements, we use integers as elements but in principle they can be any label. We can place the elements visually on a line.

Consider the example $\Omega_1 = \{1,2,3,4,5,6,7,8,9\}$, Visually we have:

$$\Omega_1 \equiv \quad (1)^{\bullet} \quad (2)^{\bullet} \quad (3)^{\bullet} \quad (4)^{\bullet} \quad (5)^{\bullet} \quad (6)^{\bullet} \quad (7)^{\bullet} \quad (8)^{\bullet} \quad (9)^{\bullet} \tag{5.25}$$

but clearly the order of placing the elements does not matter. Now consider if we use the filling map $\mathcal{F}_{\{1,2\}}(\Omega_1) = \Omega_1 \times \Omega_2$ where $\Omega_2 = \{1,2,3\}$. We can visually represent products of elementary Omega sets on a grid, which will always look like a rectangular block:

$$\Omega_1 \times \Omega_2 \equiv \begin{array}{ccccccccc} (1,3)^{\bullet} & (2,3)^{\bullet} & (3,3)^{\bullet} & (4,3)^{\bullet} & (5,3)^{\bullet} & (6,3)^{\bullet} & (7,3)^{\bullet} & (8,3)^{\bullet} & (9,3)^{\bullet} \\ (1,2)^{\bullet} & (2,2)^{\bullet} & (3,2)^{\bullet} & (4,2)^{\bullet} & (5,2)^{\bullet} & (6,2)^{\bullet} & (7,2)^{\bullet} & (8,2)^{\bullet} & (9,2)^{\bullet} \\ (1,1)^{\bullet} & (2,1)^{\bullet} & (3,1)^{\bullet} & (4,1)^{\bullet} & (5,1)^{\bullet} & (6,1)^{\bullet} & (7,1)^{\bullet} & (8,1)^{\bullet} & (9,1)^{\bullet} \end{array} \tag{5.26}$$

the ordering map will change the axis of the grid to match the standard convention but does not add any insight visually. Moving to projections, the possible projection maps on two regions visually simply reduce a grid back to a line $\mathcal{P}_1 = \Omega_2$ and $\mathcal{P}_2 = \Omega_1$, thus justifying the name. Further sets of the form $\Omega_{1,\not{2}} \times \Omega_{\not{1},2}$ are products and will always visually take the form of a block (under some shuffling of rows and columns if need be, which is a valid operation for sets). We now revisit Equation 5.20 and the three sets $\Omega_{1,2}$, $\Omega_{1,\not{2}} \times \Omega_{\not{1},2}$ and $\Omega_1 \times \Omega_2$. Let's say $\Omega_{1,2} = \{(4,2),(5,1),(5,2),(5,3),(6,2)\}$, then a visual kind of proof for Equation 5.20 looks like:

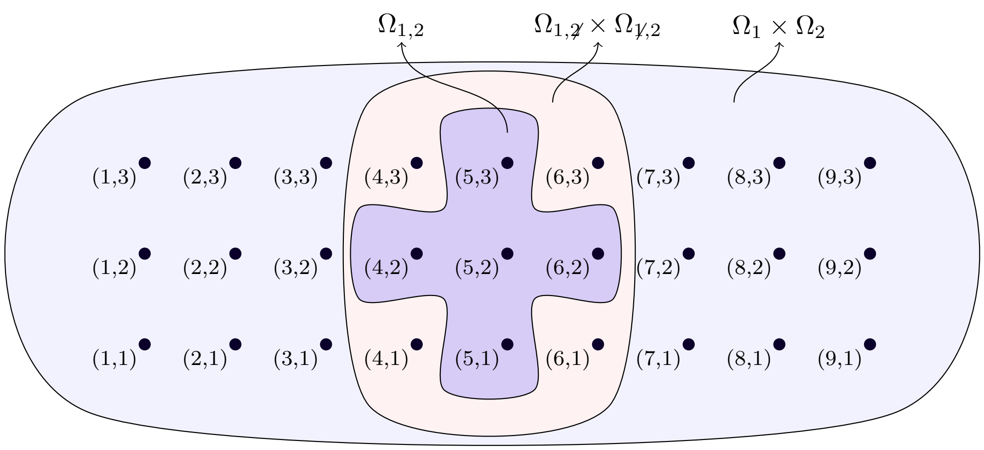


Figure 5.1: Visual form of $\Omega_{1,2} \subseteq \Omega_{1,\not{2}} \times \Omega_{\not{1},2} \subseteq \Omega_1 \times \Omega_2$

This concludes the guide to our toolkit, let us begin applying it!

5.4 $\Lambda_{\mathcal{O}_1}$-Sufficiency

Physical theories belonging to the lowest rung of the Hierarchy ($d = 1$) will have Tomographic Compression and trivial Compositional Compression. We are interested in finding functions following the form given in Equation 5.6 for $d = 1$:

$$\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}) \quad \text{and} \quad \Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}) \quad \text{for} \quad m > 1 \tag{5.27}$$

In absence of any Compositional Compression we will have $\Omega_{1,2,\ldots,m}(\Omega_{\mathcal{O}_1}) = \Omega_1 \times \Omega_2 \times \ldots \Omega_m$ and the Lambda matrices will take the form:

$$\begin{aligned} \Lambda^{k_1 k_2 \ldots}_{l_1 l_2 \ldots} &= \Lambda^{k_1}_{l_1} \Lambda^{k_2}_{l_2} \cdots = \delta^{k_1}_{l_1} \delta^{k_2}_{l_2} \ldots \\ \Lambda^{k_1 k_2 \ldots}_{\alpha_1 \alpha_2 \ldots} &= \Lambda^{k_1}_{\alpha_1} \Lambda^{k_2}_{\alpha_2} \ldots \end{aligned} \tag{5.28}$$

and so on, where in the latter equation we incorporate Tomographic Compression to give $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1})$. $\Lambda_{\mathcal{O}_1}$-Sufficiency is easy to specify due to the absence of causally adjacent regions, and makes the physical theory a simple one, one where causality does not play a role in correlating regions, perhaps meaning such physical theories do not have a concept of time.

5.5 $\Lambda_{\mathcal{O}_2}$-Sufficiency

Physical theories belonging to the next rung of the Hierarchy ($d = 2$) will have Tomographic Compression and non-trivial Compositional Compression. We are interested in finding functions following the form given in Equation 5.6 for $d = 2$:

$$\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}) \quad \text{and} \quad \Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2}) \quad \text{for} \quad m > 2 \tag{5.29}$$

Let us begin by finding the general form when $m = 3$ for $\Lambda_{\mathcal{O}_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$. Consider a general tripartite composite region of R_1, R_2, R_3 within the predictively well-defined R. Of interest are probabilities that after Tomographic compression are as follows:

$$\begin{aligned} p_{\alpha_1,\alpha_2,\alpha_3} &= \text{Prob}(Y^{\alpha_1}_{R_1} \cup Y^{\alpha_2}_{R_2} \cup Y^{\alpha_3}_{R_3} \cup Y_{R-R_1-R_2-R_3} | F^{\alpha_1}_{R_1} \cup F^{\alpha_2}_{R_2} \cup F^{\alpha_3}_{R_3} \cup F_{R-R_1-R_2-R_3}) \\ &= \sum_{l_1 \in \Omega_{R_1}} \sum_{l_2 \in \Omega_{R_2}} \sum_{l_3 \in \Omega_{R_3}} \Lambda^{l_1}_{\alpha_1} \Lambda^{l_2}_{\alpha_2} \Lambda^{l_3}_{\alpha_3} \mathbf{r}_{l_1,l_2,l_3} \cdot \mathbf{p} \end{aligned} \tag{5.30}$$

Now for three regions there are three $\Lambda_{\mathcal{O}_2}$ for regions $R_1 \cup R_2$, $R_2 \cup R_3$ and $R_1 \cup R_3$ respectively[2] and a single $\Lambda_{\mathcal{O}_3}$ for region $R_1 \cup R_2 \cup R_3$. Let us apply all of these to $\mathbf{r}_{l_1 l_2 l_3} \cdot \mathbf{p}$

[2] In general for m regions we will have $\binom{m}{2} = \frac{n(n-1)}{2}$ number of $\Lambda_{\mathcal{O}_2}$s.

in the order of Compositional Compression to the region $R_1 \cup R_2$, followed by, to the region $R_2 \cup R_3$, followed by, to the region $R_1 \cup R_3$, finally followed by Compositional compression to all three regions $R_1 \cup R_2 \cup R_3$:

$$\begin{aligned}
& \mathbf{r}_{l_1 l_2 l_3}(R_1 \cup R_2 \cup R_3) \cdot \mathbf{p}(R_1 \cup R_2 \cup R_3) && (5.31)\\
= & \mathbf{r}_{l_1 l_2}(R_1 \cup R_2) \cdot \mathbf{p}_{l_3}(R_1 \cup R_2) && (5.32)\\
= & \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k'_1 k'_2}(R_1 \cup R_2) \cdot \mathbf{p}_{l_3}(R_1 \cup R_2) && (5.33)\\
= & \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k'_2 l_3}(R_2 \cup R_3) \cdot \mathbf{p}_{k'_1}(R_2 \cup R_3) && (5.34)\\
= & \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k_2 k'_3}(R_2 \cup R_3) \cdot \mathbf{p}_{k'_1}(R_2 \cup R_3) && (5.35)\\
= & \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k'_1 k'_3}(R_1 \cup R_3) \cdot \mathbf{p}_{k_2}(R_1 \cup R_3) && (5.36)\\
= & \sum_{k_1 k_3 \in \Omega_{1,3}} \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k_1 k_3}_{k'_1 k'_3} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k_1 k_3}(R_1 \cup R_3) \cdot \mathbf{p}_{k_2}(R_1 \cup R_3) && (5.37)\\
= & \sum_{k_1 k_3 \in \Omega_{1,3}} \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{k_1 k_3}_{k'_1 k'_3} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{k_1 k_2 k_3} \cdot \mathbf{p} && (5.38)\\
= & \sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \sum_{k_1 k_3 \in \Omega_{1,3}} \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{j_1 j_2 j_3}_{k_1 k_2 k_3} \Lambda^{k_1 k_3}_{k'_1 k'_3} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{j_1 j_2 j_3} \cdot \mathbf{p} && (5.39)
\end{aligned}$$

where we have used regions arguments for $\mathbf{r}$ and $\mathbf{p}$ carefully, similar to what we have seen before. In short we have:

$$\mathbf{r}_{l_1 l_2 l_3} \cdot \mathbf{p} = \sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \sum_{k_1 k_3 \in \Omega_{1,3}} \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{j_1 j_2 j_3}_{k_1 k_2 k_3} \Lambda^{k_1 k_3}_{k'_1 k'_3} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{j_1 j_2 j_3} \cdot \mathbf{p} \quad (5.40)$$

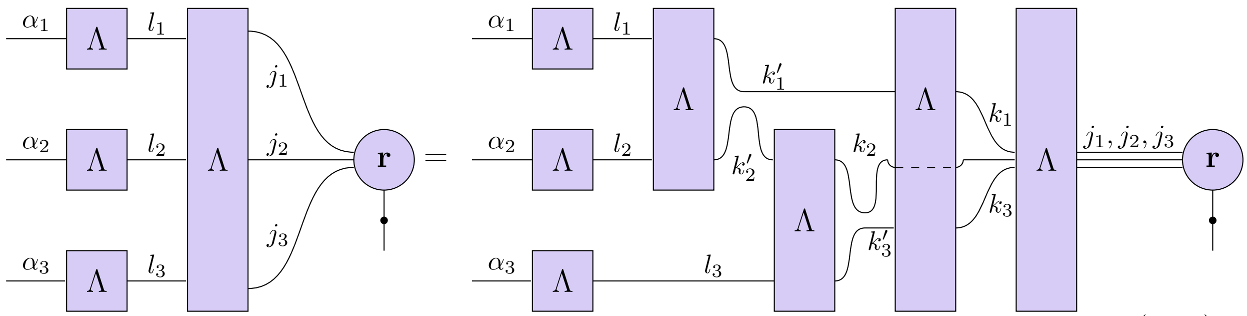
By definition of three region compositional compression we also have

$$\mathbf{r}_{l_1 l_2 l_3} \cdot \mathbf{p} = \sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \Lambda^{j_1 j_2 j_3}_{l_1 l_2 l_3} \mathbf{r}_{j_1 j_2 j_3} \cdot \mathbf{p} \quad (5.41)$$

where we use j indices for $\Lambda_{\mathcal{O}_3}$ matrices. Both Equation 5.40 and Equation 5.41 are true for all $\mathbf{p}$. Comparing them we get:

$$\sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \Lambda^{j_1 j_2 j_3}_{l_1 l_2 l_3} \mathbf{r}_{j_1 j_2 j_3} = \sum_{j_1 j_2 j_3 \in \Omega_{1,2,3}} \sum_{k_1 k_3 \in \Omega_{1,3}} \sum_{k_2 k'_3 \in \Omega_{2,3}} \sum_{k'_1 k'_2 \in \Omega_{1,2}} \Lambda^{j_1 j_2 j_3}_{k_1 k_2 k_3} \Lambda^{k_1 k_3}_{k'_1 k'_3} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \mathbf{r}_{j_1 j_2 j_3} \quad (5.42)$$

Equation 5.42 combined with tomographic compression as seen in Equation 5.30 can be represented diagrammatically as follows (where the dashed line passes over Λ (not through)):



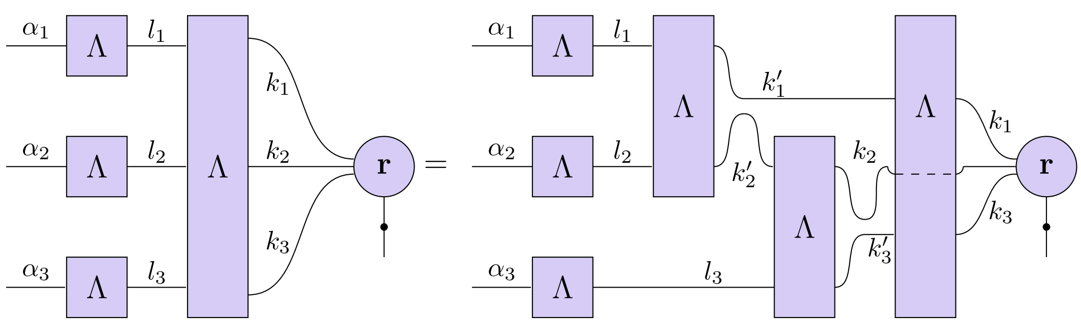
$$(5.43)$$

The above Equations 5.30-5.43 are always true in the Causaloid Framework irrespective of the Hierarchy, but when we have $\Lambda_{\mathcal{O}_2}$-Sufficiency we expect that $\Lambda^{j_1j_2j_3}_{k_1k_2k_3}$ will reduce to $\delta^{j_1}_{k_1}\delta^{j_2}_{k_2}\delta^{j_3}_{k_3}$. This is the case since any $\Lambda_{\mathcal{O}_3}$ would be decomposable into $\Lambda_{\mathcal{O}_2}, \Lambda_{\mathcal{O}_1}$ and since we have applied all possible $\Lambda_{\mathcal{O}_2}, \Lambda_{\mathcal{O}_1}$, no more compression would be possible. Let us incorporate this into Equation 5.42 to get:

$$\sum_{j_1j_2j_3\in\Omega_{1,2,3}} \Lambda^{j_1j_2j_3}_{l_1l_2l_3}\mathbf{r}_{j_1j_2j_3} = \sum_{j_1j_2j_3\in\Omega_{1,2,3}} \sum_{k_1k_3\in\Omega_{1,3}} \sum_{k_2k'_3\in\Omega_{2,3}} \sum_{k'_1k'_2\in\Omega_{1,2}} \delta^{j_1}_{k_1}\delta^{j_2}_{k_2}\delta^{j_3}_{k_3}\Lambda^{k_1k_3}_{k'_1k'_3}\Lambda^{k_2k'_3}_{k'_2l_3}\Lambda^{k'_1k'_2}_{l_1l_2}\mathbf{r}_{j_1j_2j_3}$$

$$\sum_{k_1k_2k_3\in\Omega_{1,2,3}} \Lambda^{k_1k_2k_3}_{l_1l_2l_3}\mathbf{r}_{k_1k_2k_3} = \sum_{k_1k_3\in\Omega_{1,3}} \sum_{k_2k'_3\in\Omega_{2,3}} \sum_{k'_1k'_2\in\Omega_{1,2}} \Lambda^{k_1k_3}_{k'_1k'_3}\Lambda^{k_2k'_3}_{k'_2l_3}\Lambda^{k'_1k'_2}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.44}$$

Therefore with $\Lambda_{\mathcal{O}_2}$-Sufficiency Equation 5.42 combined with tomographic compression as seen in Equation 5.30 diagrammatically reduces to:



$$(5.45)$$

We now explore how Equation 5.42 reduces to simpler forms of $\Lambda^{k_1k_2k_3}_{l_1l_2l_3}$ and $\Omega^{k_1k_2k_3}_{l_1l_2l_3}$ under all possible different causal relations between the regions R_1, R_2 and R_3 if we have $\Lambda_{\mathcal{O}_2}$-Sufficiency. This will fully characterise the tripartite composite region for which Hardy gave an example as we saw in Subsection 5.2.

5.5.1 Tripartite Region: Causal Relations Results

If two regions, say R_1 and R_2 are causally adjacent then we use the notation $R_1 \bowtie R_2$ as a specification of their causal relation. If they aren't causally adjacent we simply write R_1, R_2. Recall when R_1 and R_2 are not causally adjacent then the compositional compression does not provide any additional compression, and we have $\Omega_{1,2} = \Omega_1 \times \Omega_2$ and $\Lambda^{k_1,k_2}_{l_1,l_2} = \delta^{k_1}_{l_1}\delta^{k_2}_{l_2}$ after tomographic compression has already been applied.

Let us provide a table for all possible tripartite causal relations and the corresponding forms for $\Lambda_{\mathcal{O}_2}$ matrices which we will use to simplify Equation 5.42 (and consequently Equation 5.45) to calculate $\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ and $\Omega_{1,2,3}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$ in terms of the causal relations between R_1, R_2, R_3. (While we work with elementary regions, the calculations in this Section will hold for three tomographically compressed composite regions as well.)

	Causal Relations	$\Lambda^{k_1k_3}_{k'_1k'_3}$	$\Lambda^{k_2k'_3}_{k'_2l_3}$	$\Lambda^{k'_1k'_2}_{l_1l_2}$
1	R_1, R_2, R_3	$\delta^{k_1}_{k'_1}\delta^{k_3}_{k'_3}$	$\delta^{k_2}_{k'_2}\delta^{k'_3}_{l_3}$	$\delta^{k'_1}_{l_1}\delta^{k'_2}_{l_2}$
2	$R_1 \bowtie R_2$	$\delta^{k_1}_{k'_1}\delta^{k_3}_{k'_3}$	$\delta^{k_2}_{k'_2}\delta^{k'_3}_{l_3}$	$\Lambda^{k'_1k'_2}_{l_1l_2}$
3	$R_2 \bowtie R_3$	$\delta^{k_1}_{k'_1}\delta^{k_3}_{k'_3}$	$\Lambda^{k_2k'_3}_{k'_2l_3}$	$\delta^{k'_1}_{l_1}\delta^{k'_2}_{l_2}$
4	$R_1 \bowtie R_3$	$\Lambda^{k_1k_3}_{k'_1k'_3}$	$\delta^{k_2}_{k'_2}\delta^{k'_3}_{l_3}$	$\delta^{k'_1}_{l_1}\delta^{k'_2}_{l_2}$
5	$R_1 \bowtie R_2$ and $R_2 \bowtie R_3$	$\delta^{k_1}_{k'_1}\delta^{k_3}_{k'_3}$	$\Lambda^{k_2k'_3}_{k'_2l_3}$	$\Lambda^{k'_1k'_2}_{l_1l_2}$
6	$R_1 \bowtie R_3$ and $R_2 \bowtie R_3$	$\Lambda^{k_1k_3}_{k'_1k'_3}$	$\Lambda^{k_2k'_3}_{k'_2l_3}$	$\delta^{k'_1}_{l_1}\delta^{k'_2}_{l_2}$
7	$R_1 \bowtie R_2$ and $R_1 \bowtie R_3$	$\Lambda^{k_1k_3}_{k'_1k'_3}$	$\delta^{k_2}_{k'_2}\delta^{k'_3}_{l_3}$	$\Lambda^{k'_1k'_2}_{l_1l_2}$
8	$R_1 \bowtie R_2$ and $R_2 \bowtie R_3$ and $R_1 \bowtie R_3$	$\Lambda^{k_1k_3}_{k'_1k'_3}$	$\Lambda^{k_2k'_3}_{k'_2l_3}$	$\Lambda^{k'_1k'_2}_{l_1l_2}$

Table 5.1: Form for $\Lambda_{\mathcal{O}_2}$ given Causal Relations between R_1, R_2, R_3

1. Causal relation R_1, R_2, R_3:

It is clear from our study of $\Lambda_{\mathcal{O}_1}$-sufficiency that when R_1, R_2 and R_3 are not causally adjacent then no compositional compression occurs and we have,

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \delta^{k_1}_{l_1}\delta^{k_2}_{l_2}\delta^{k_3}_{l_3} \quad \text{where} \quad \Omega_{1,2,3} = \Omega_1 \times \Omega_2 \times \Omega_3 \tag{5.46}$$

2. Causal relation $R_1 \bowtie R_2$:

For $R_1 \bowtie R_2$, and R_3 not causally adjacent to either R_1 or R_2 we calculate the form for Lambda matrix starting from Equation 5.42:

$$\sum_{k_1k_2k_3\in\Omega_{1,2,3}} \Lambda^{k_1k_2k_3}_{l_1l_2l_3}\mathbf{r}_{k_1k_2k_3} = \sum_{k_1k_3\in\Omega_{1,3}} \sum_{k_2k_3'\in\Omega_{2,3}} \sum_{k_1'k_2'\in\Omega_{1,2}} \Lambda^{k_1k_3}_{k_1'k_3'}\Lambda^{k_2k_3'}_{k_2'l_3}\Lambda^{k_1'k_2'}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.47}$$

$$= \sum_{k_1k_3\in\Omega_1\times\Omega_3} \sum_{k_2k_3'\in\Omega_2\times\Omega_3} \sum_{k_1'k_2'\in\Omega_{1,2}} \delta^{k_1}_{k_1'}\delta^{k_3}_{k_3'}\delta^{k_2}_{k_2'}\delta^{k_3'}_{l_3}\Lambda^{k_1'k_2'}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.48}$$

$$= \sum_{k_1k_2\in\Omega_{1,2}} \Lambda^{k_1k_2}_{l_1l_2}\mathbf{r}_{k_1k_2l_3} \tag{5.49}$$

which gives us:

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \Lambda^{k_1k_2}_{l_1l_2}\delta^{k_3}_{l_3} \quad \text{where} \quad \Omega_{1,2,3} = \Omega_{1,2}\times\Omega_3 \tag{5.50}$$

There are two similar cases to $R_1 \bowtie R_2$ that also have two regions causally adjacent and thus have similar calculations to Equations 5.47-5.49. We omit the calculations and provide the results below.

3. Causal relation $R_2 \bowtie R_3$

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \delta^{k_1}_{l_1}\Lambda^{k_2k_3}_{l_2l_3} \quad \text{where} \quad \Omega_{1,2,3} = \Omega_1\times\Omega_{2,3} \tag{5.51}$$

4. Causal relation $R_1 \bowtie R_3$

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \Lambda^{k_1k_3}_{l_1l_3}\delta^{k_2}_{l_2} \quad \text{where} \quad \Omega_{1,2,3} = \mathcal{O}(\Omega_{1,3}\times\Omega_2) \tag{5.52}$$

where above we use the ordering map $\mathcal{O}(.)$, which reorders the indices $k_1k_3k_2$ back to the standard order $k_1k_2k_3$. We have formally defined the ordering map $\mathcal{O}(.)$ in Subsection 5.3.

5. Causal relation $R_1 \bowtie R_2$ and $R_2 \bowtie R_3$

Let us now consider the scenario when two pairs of regions are causally adjacent (we will see an example in Section 6.3.6). We simplify Equation 5.42 to find the following:

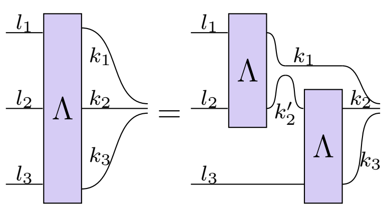
$$\sum_{k_1k_2k_3\in\Omega_{1,2,3}} \Lambda^{k_1k_2k_3}_{l_1l_2l_3}\mathbf{r}_{k_1k_2k_3} = \sum_{k_1k_3\in\Omega_{1,3}} \sum_{k_2k_3'\in\Omega_{2,3}} \sum_{k_1'k_2'\in\Omega_{1,2}} \Lambda^{k_1k_3}_{k_1'k_3'}\Lambda^{k_2k_3'}_{k_2'l_3}\Lambda^{k_1'k_2'}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.53}$$

$$= \sum_{k_1k_3\in\Omega_1\times\Omega_3} \sum_{k_2k_3'\in\Omega_{2,3}} \sum_{k_1'k_2'\in\Omega_{1,2}} \delta^{k_1}_{k_1'}\delta^{k_3}_{k_3'}\Lambda^{k_2k_3'}_{k_2'l_3}\Lambda^{k_1'k_2'}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.54}$$

$$= \sum_{k_2k_3\in\Omega_{2,3}} \sum_{k_1k_2'\in\Omega_{1,2}} \Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2}\mathbf{r}_{k_1k_2k_3} \tag{5.55}$$

There is a subtle point to be made here. When we have a Lambda matrix of the form $\Lambda^{k_1k_2}_{l_1l_2}$, where $l_1l_2 \in \Omega_1 \times \Omega_2$ then k_1k_2 spans $\Omega_{1,2}$ since there is a l_1l_2 for every k_1k_2 given that $\Omega_{1,2} \subseteq \Omega_1 \times \Omega_2$. This is not true for the term $\Lambda^{k_2k_3}_{k_2'l_3}$ that shows up in Equation 5.55. While k_2k_3 do belong to $\Omega_{2,3}$, they do not span it since k_2' belongs to the set $\Omega_{\not{1},2}$ and therefore restricts k_2 which must belong to the set $\Omega_{\not{1},2}$ *as well as* $\Omega_{2,\not{3}}$.

Now a question arises as to what form does $\Omega_{1,2,3}$ take? We may be tempted to say that since $k_1 \in \Omega_{1,\not{2}}$, $k_3 \in \Omega_{\not{2},3}$ and $k_2 \in \Omega_{\not{1},2} \cap \Omega_{2,\not{3}}$ then $\Omega_{1,2,3}$ should be equal to $\Omega_{1,\not{2}} \times (\Omega_{\not{1},2} \cap \Omega_{2,\not{3}}) \times \Omega_{\not{2},3}$, but this is not true. In the same way that for some $\Lambda^{k_1k_2}_{l_1l_2}$, k_1 belongs to $\Omega_{1,\not{2}}$ and k_2 belongs to $\Omega_{\not{1},2}$ but k_1k_2 belongs to $\Omega_{1,2}$ and not $\Omega_{1,\not{2}} \times \Omega_{\not{1},2}$, our temptation would generally not hold true, since Equation 5.20 gives us $\Omega_{1,2} \subseteq \Omega_{1,\not{2}} \times \Omega_{\not{1},2}$. Instead, we must think of $k_1k_2k_3$ indices together and use the filling map. Clearly $k_1k_2k_3$ must belong to $\mathcal{F}_{\{1,2,3\}}(\Omega_{1,2})$ *as well as* $\mathcal{F}_{\{1,2,3\}}(\Omega_{2,3})$, that is $k_1k_2k_3 \in \mathcal{F}_{\{1,2,3\}}(\Omega_{1,2}) \cap \mathcal{F}_{\{1,2,3\}}(\Omega_{2,3})$. We may even show as a consequence of Equation 5.20 that $(\Omega_{1,2} \times \Omega_3) \cap (\Omega_1 \times \Omega_{2,3}) \subseteq \Omega_{1,\not{2}} \times (\Omega_{\not{1},2} \cap \Omega_{2,\not{3}}) \times \Omega_{\not{2},3}$. Therefore we have the following result:



$$\equiv \Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'} \Lambda^{k_2k_3}_{k_2'l_3} \Lambda^{k_1k_2'}_{l_1l_2} \ , \ k_2' \in \Omega_{\not{1},2} \tag{5.56}$$

$$\begin{aligned} \text{where} \quad \Omega_{1,2,3} &= \mathcal{F}_{\{1,2,3\}}(\Omega_{1,2}) \cap \mathcal{F}_{\{1,2,3\}}(\Omega_{2,3}) \\ &= (\Omega_{1,2} \times \Omega_3) \cap (\Omega_1 \times \Omega_{2,3}) \end{aligned} \tag{5.57}$$

The above result may remind us of the proof of concept example by Hardy (Equation 5.17) presented before but what we have here is more general. Instead of choosing some $\Lambda_{\mathcal{O}_2}$s and finding a condition under which it is equivalent to $\Lambda_{\mathcal{O}_3}$ (Equation 5.17), we have started without assuming any causal relations for three regions (Equation 5.42) and worked our way down to give the most general form for $\Omega_{1,2,3}$ for some causal relations.

The following two cases similarly also have two pairs of regions being causally adjacent and thus have similar calculations to Equations 5.53-5.55. We omit the calculations and provide the results below.

6. Causal relation $R_1 \bowtie R_3$ and $R_2 \bowtie R_3$

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_3'} \Lambda^{k_1k_3}_{l_1k_3'} \Lambda^{k_2k_3'}_{l_2l_3} \ , \ k_3' \in \Omega_{\not{2},3} \tag{5.58}$$

$$\text{where} \quad \Omega_{1,2,3} = (\mathcal{O}\,(\Omega_{1,3} \times \Omega_2)) \cap (\Omega_1 \times \Omega_{2,3}) \tag{5.59}$$

Here we have used the ordering map $\mathcal{O}(.)$, which reorders the indices $k_1k_3k_2$ back to the standard order $k_1k_2k_3$.

7. Causal relation $R_1 \bowtie R_2$ and $R_1 \bowtie R_3$

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k'_1} \Lambda^{k_1k_3}_{k'_1l_3}\Lambda^{k'_1k_2}_{l_1l_2} \;,\; k'_1 \in \Omega_{1,2,3} \tag{5.60}$$

$$\text{where} \quad \Omega_{1,2,3} = \left(\mathcal{O}\left(\Omega_{1,3} \times \Omega_2\right)\right) \cap \left(\Omega_{1,2} \times \Omega_3\right) \tag{5.61}$$

Same as above the ordering map $\mathcal{O}(.)$, reorders the indices $k_1k_3k_2$ back to $k_1k_2k_3$.

8. Causal relation $R_1 \bowtie R_2$, $R_2 \bowtie R_3$ and $R_1 \bowtie R_3$

Finally we come to the most general scenario where all three regions are causally adjacent with each other. Equation 5.42 cannot be simplified further and we have:

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k'_1,k'_2,k'_3} \Lambda^{k_1k_3}_{k'_1k'_3}\Lambda^{k_2k'_3}_{k'_2l_3}\Lambda^{k'_1k'_2}_{l_1l_2} \;,\; k'_1,k'_2,k'_3 \in \Omega_{1,2} \times \Omega_{2,3} \tag{5.62}$$

$$\text{where} \quad \Omega_{1,2,3} = \left(\Omega_{1,2} \times \Omega_3\right) \cap \left(\Omega_1 \times \Omega_{2,3}\right) \cap \left(\mathcal{O}\left(\Omega_{1,3} \times \Omega_2\right)\right) \tag{5.63}$$

Having studied all possible causal relations, let us tabulate our results – the forms $\Omega_{1,2,3}$ and $\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$.

	Causal Relations	$\Omega_{1,2,3}$	$\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$
1	R_1, R_2, R_3	$\Omega_1 \times \Omega_2 \times \Omega_3$	$\delta^{k_1}_{l_1}\delta^{k_2}_{l_2}\delta^{k_3}_{l_3}$
2	$R_1 \bowtie R_2$	$\Omega_{1,2} \times \Omega_3$	$\Lambda^{k_1k_2}_{l_1l_2}\delta^{k_3}_{l_3}$
3	$R_2 \bowtie R_3$	$\Omega_1 \times \Omega_{2,3}$	$\delta^{k_1}_{l_1}\Lambda^{k_2k_3}_{l_2l_3}$
4	$R_1 \bowtie R_3$	$\mathcal{O}(\Omega_{1,3} \times \Omega_2)$	$\Lambda^{k_1k_3}_{l_1l_3}\delta^{k_2}_{l_2}$
5	$R_1 \bowtie R_2$ and $R_2 \bowtie R_3$	$(\Omega_{1,2} \times \Omega_3) \cap (\Omega_1 \times \Omega_{2,3})$	$\sum_{k'_2} \Lambda^{k_2k_3}_{k'_2l_3}\Lambda^{k_1k'_2}_{l_1l_2}$
6	$R_1 \bowtie R_3$ and $R_2 \bowtie R_3$	$\mathcal{O}\left(\Omega_{1,3} \times \Omega_2\right) \cap (\Omega_1 \times \Omega_{2,3})$	$\sum_{k'_3} \Lambda^{k_1k_3}_{l_1k'_3}\Lambda^{k_2k'_3}_{l_2l_3}$
7	$R_1 \bowtie R_2$ and $R_1 \bowtie R_3$	$\mathcal{O}\left(\Omega_{1,3} \times \Omega_2\right) \cap (\Omega_{1,2} \times \Omega_3)$	$\sum_{k'_1} \Lambda^{k_1k_3}_{k'_1l_3}\Lambda^{k'_1k_2}_{l_1l_2}$
8	$R_1 \bowtie R_2$ and $R_2 \bowtie R_3$ and $R_1 \bowtie R_3$	$(\Omega_{1,2} \times \Omega_3) \cap (\Omega_1 \times \Omega_{2,3})$ $\cap\, \mathcal{O}\left(\Omega_{1,3} \times \Omega_2\right)$	$\sum_{k'_1k'_2k'_3} \Lambda^{k_1k_3}_{k'_1k'_3}\Lambda^{k_2k'_3}_{k'_2l_3}\Lambda^{k'_1k'_2}_{l_1l_2}$

Table 5.2: $\Omega_{1,2,3}$ and $\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ given Causal Relations between R_1, R_2, R_3

To complete the specification of $\Lambda_{\mathcal{O}_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ we include tomographic compression:

$$\Lambda^{k_1k_2k_3}_{\alpha_1\alpha_2\alpha_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}) = \sum_{l_1l_2l_3 \in \Omega_1 \times \Omega_2 \times \Omega_3} \Lambda^{l_1}_{\alpha_1}\Lambda^{l_2}_{\alpha_2}\Lambda^{l_3}_{\alpha_3}\Lambda^{k_1k_2k_3}_{l_1l_2l_3}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}) \tag{5.64}$$

5.5.2 Going beyond Tripartite Regions: $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$, $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$

We have shown how to characterise $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ and $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$ for $m = 3$. How do we go about doing the same for higher m? We will provide the *strategy* that can be used to do all calculations. To begin with we find the Equation analogous to Equation 5.43 by applying all possible $\Lambda_{\mathcal{O}_2}$ matrices (there would be $\binom{m}{2}$ of such matrices) followed by $\Lambda_{\mathcal{O}_m}$. If we have $\Lambda_{\mathcal{O}_2}$-sufficiency we have $\Lambda^{k_1 k_2 \ldots k_m}_{l_1 l_2 \ldots l_m} = \delta^{k_1}_{l_1} \delta^{k_2}_{l_2} \ldots \delta^{k_m}_{l_m}$ giving us an Equation analogous to Equation 5.45. Depending upon causal relations – absence or presence of causal adjacency between each pair of regions – one may simplify the equation further to get the form for $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ and $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$.

It is time to address an important point that we had deferred until now. We discussed that we must apply all $\binom{m}{2}$ of $\Lambda_{\mathcal{O}_2}$ matrices. Does the order of applying these matrices matter? Did the order matter in the tripartite region case? It turns out that the order only affects the details around summations of the form for $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$, such that some order may yield simpler calculations but one may apply any order of $\Lambda_{\mathcal{O}_2}$ matrices as long as each one of them is applied once. On the other hand the form for $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$ will be independent of the order of application of $\Lambda_{\mathcal{O}_2}$. We will call this property *associativity of compositional compression*:

> **Associativity of Compositional Compression:** When applying more than one compositional compression matrix, the total compression, seen through the effective Lambda matrix and associated Omega set, is independent of the order of application of the compositional matrices.

We are ready to provide an important result of this work:

Since $\Omega_{\mathcal{O}_m}$ is independent of the order of application of $\Lambda_{\mathcal{O}_2}$, we are able to present the general form for $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$ when we have $\Lambda_{\mathcal{O}_2}$-sufficiency:

$$\Omega_{\mathcal{O}_m} = \bigcap_{i \neq j \in \mathcal{O}_m}^{\binom{m}{2}} \mathcal{O}\left(\mathcal{F}_{\mathcal{O}_m}\left(\Omega_{i,j}\right)\right) \quad \text{where} \quad \mathcal{O}_m = \{1, 2, \ldots, m\} \tag{5.65}$$

where we have used the filling and ordering maps associated with a region with label $\mathcal{O}_m = \{1, 2, \ldots, m\}$ with size $|\mathcal{O}_m| = m$

One can then for every pair of regions R_i, R_j which are not causally adjacent replace $\Omega_{i,j}$ with $\Omega_i \times \Omega_j$ to simplify the form.

As an example of the application of the strategy, if we have $m = 4$ for four regions R_1, R_2, R_3 and R_4 such that $R_1 \bowtie R_2$ and $R_2 \bowtie R_3$ and $R_3 \bowtie R_4$ then one can go through the steps to find:

$$\Lambda^{k_1 k_2 k_3 k_4}_{l_1 l_2 l_3 l_4} = \sum_{k'_1 k'_2 k'_3} \Lambda^{k_3 k_4}_{k'_3 l_4} \Lambda^{k_2 k'_3}_{k'_2 l_3} \Lambda^{k'_1 k'_2}_{l_1 l_2} \quad \text{where} \quad k'_1 k'_2 k'_3 \in \Omega_{1,2} \times \Omega_{\not{2},3} \tag{5.66}$$

$$\begin{aligned} \Omega_{1,2,3,4} &= (\Omega_{1,2} \times \Omega_3 \times \Omega_4) \cap (\Omega_1 \times \Omega_{2,3} \times \Omega_4) \cap (\Omega_1 \times \Omega_2 \times \Omega_{3,4}) \\ &= (\Omega_{1,2} \times \Omega_{3,4}) \cap (\Omega_1 \times \Omega_{2,3} \times \Omega_4) \end{aligned} \tag{5.67}$$

5.6 $\Lambda_{\mathcal{O}_d}$-Sufficiency

We have shown how to characterise $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2})$ and $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \Omega_{\mathcal{O}_2})$. How do we go about doing the same for a general rung of the Hierarchy? We augment the *strategy* provided so that we may be able to calculate the forms for $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \ldots, \Lambda_{\mathcal{O}_d})$ and $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \ldots, \Omega_{\mathcal{O}_d})$ for $d \leq m$ when we have $\Lambda_{\mathcal{O}_d}$-sufficiency.

To begin with we find the Equation analogous to Equation 5.43 by applying all possible $\Lambda_{\mathcal{O}_d}$ matrices (there would be $\binom{m}{d}$ of such matrices) followed by $\Lambda_{\mathcal{O}_m}$. If we have $\Lambda_{\mathcal{O}_2}$-sufficiency we have $\Lambda^{k_1 k_2 \ldots k_m}_{l_1 l_2 \ldots l_m} = \delta^{k_1}_{l_1} \delta^{k_2}_{l_2} \ldots \delta^{k_m}_{l_m}$ giving us an Equation analogous to Equation 5.45. Depending upon causal relations – absence or presence of causal adjacency between every d region composite – one may simplify the equation further to get the form for $\Lambda_{\mathcal{O}_m}(\Lambda_{\mathcal{O}_1}, \ldots, \Lambda_{\mathcal{O}_d})$ and $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \ldots, \Omega_{\mathcal{O}_d})$.

Since $\Omega_{\mathcal{O}_m}$ is independent of the order of application of $\Lambda_{\mathcal{O}_d}$ due to *associativity of compositional compression*, we conjecture that the general form for $\Omega_{\mathcal{O}_m}(\Omega_{\mathcal{O}_1}, \ldots \Omega_{\mathcal{O}_d})$ when we have $\Lambda_{\mathcal{O}_d}$-sufficiency is:

$$\Omega_{\mathcal{O}_m} = \bigcap_{\forall\ \mathcal{O}_d \subseteq \mathcal{O}_m}^{\binom{m}{d}} \mathcal{O}\left(\mathcal{F}_{\mathcal{O}_m}\left(\Omega_{\mathcal{O}_d}\right)\right)$$
$$\text{where} \quad \mathcal{O}_d \subseteq \mathcal{O}_m \quad \text{such that} \quad |\mathcal{O}_d| = d \quad \text{and} \quad \mathcal{O}_m = \{1, 2, \ldots, m\} \tag{5.68}$$

where we have used the filling and ordering maps associated with a region with label $\mathcal{O}_m = \{1, 2, \ldots, m\}$ with size $|\mathcal{O} = \{1, 2, \ldots, m\}| = m$ and $\mathcal{O}_d$ is any region such that $\mathcal{O}_d \subseteq \mathcal{O}_m$ and $|\mathcal{O}_d| = d$.

This concludes our work regarding the Hierarchy. We will apply the Causaloid Framework to the Duotensor Framework in the coming Chapter.

5.7 The Road Ahead...

The work presented in the current as well as previous Chapter was born out of a motivation to revisit the Causaloid Framework, a seminal work that motivated many works on indefinite causal structures. To this end, we presented a diagrammatic representation for the three levels of physical compression (Chapter 4), along with a review of the framework. We bear the hope that this diagrammatic review will make this framework more accessible to those interested, who are engaged in studying indefinite causal structures.

> "Although correlations with indefinite causal structure had been defined in more general settings (e.g., Hardy's causaloid and Oreshkov and Giarmatzi's general processes), their quantitative features have not been studied in detail."
> (Jia, [65])

Further, the other crucial contribution of this work is substantiating how to handle correlations when a causal structure is not assumed, through the study of Meta Compression in detail. This led to the presentation of a Hierarchy for physical theories.

What does the Hierarchy signify physically? To answer this let us go back to the space of Generalised Probabilistic Theories (or GPTs). An existing hierarchy through the value of r in $K = N^r$ (Sorkin's Hierarchy [93], and the introduction of this relation by Wootters [99]) can be ascribed to Tomographic Compression (Section 4.3.2) in terms of the Causaloid Framework, though it has been studied earlier in the context of GPTs [40] and can be attributed to tomographic properties of states, transformations etc. The hierarchy presented in this work, pertains to the space of GPTs with indefinite causal structure (of which definite causal structure is a special case) and can be ascribed to Compositional Compression and in turn the causal structure. We will see in the next Chapter that Quantum Theory (QT) and Classical Probability Theory (CPT) will belong to the same rung in terms of this work's hierarchy.

Therefore we may think of the space of GPTs with indefinite causal structure structured by two independent hierarchies. We schematically show this in Figure 5.2.

Through the progress, our work leaves us with many open questions:

- How one may fill the space of GPTs through the hierarchy presented in this work?
- How the techniques from studies of GPTs may translate possibly to this hierarchy, and how this hierarchy can further tell us about the space of GPTs.

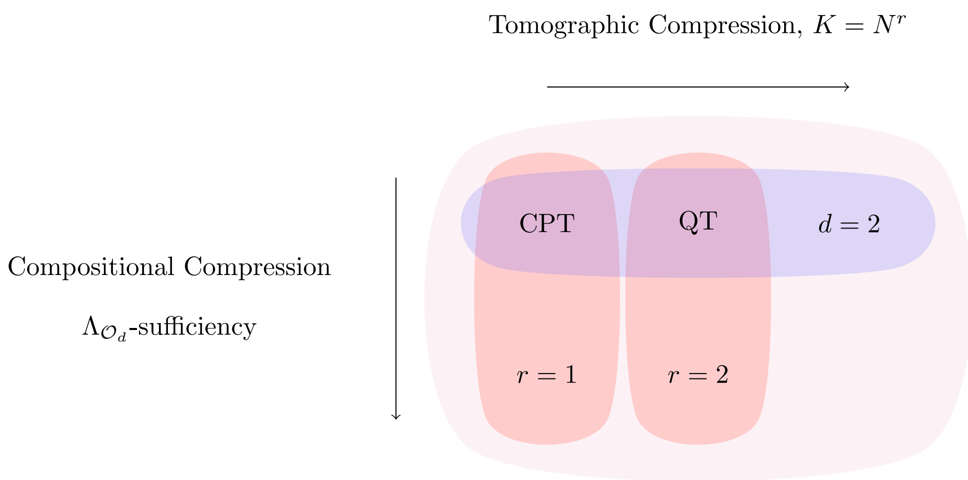


Figure 5.2: The space of Generised Probability Theories with respect to the two Hierarchies

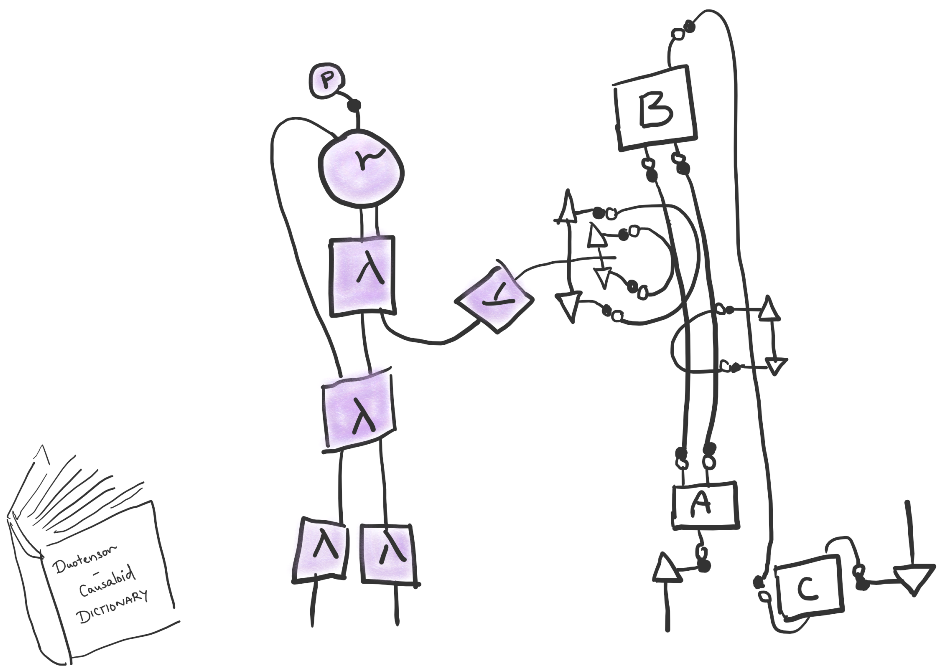
- Do known frameworks for indefinite causal structures fall within this Hierarchy and if so where?
- How does (indefinite) causality manifest in the Hierarchy? How does Meta Compression look like for a theory with dynamic causal structure such as relativity? Since the Causaloid is an operational framework, we would need an operational formalism for relativity (such as [50]) to achieve this.
- We will see that the duotensor framework has Omega sets that satisfy $\Omega_{1,2} = \Omega_{1,\not{2}} \times \Omega_{\not{1},2}$; can we find physical examples of Omega sets that satisfy $\Omega_{1,2} \subset \Omega_{1,\not{2}} \times \Omega_{\not{1},2}$ and if so, when?

Revisiting the Causaloid Framework has led us down a road that leaves us with the large research program of pursuing the characterisation of indefinite causality. We invite the interested to join us in researching the answers to these questions.

The next Chapter focuses on studying the rungs of the Hierarchy through the application of the Causaloid Framework to the Duotensor Framework (which includes finite quantum theory), and in turn will help us learn how the framework can be used practically.

Chapter 5: Statement of Contribution

In this Chapter, the main contribution is towards providing a hierarchy of physical theories with respect to Meta Compression by providing the definition of $\Lambda_{\mathcal{O}_d}$-sufficiency, building upon Chapter 4. The project was initiated by me and the problem statement of characterising Meta Compression was presented by Lucien Hardy. The results around characterising the hierarchy throughout the Chapter are my contribution. The Chapter is solely written by me and contains unpublished work that was presented at various conferences (QISS HK 2020, Q-Turn 2020 [90], APS 2021 [91], Quantizing Time 2021 [89], QPL 2021) as the work developed.



Duotensor met the Causaloid
in the land of diagrams
thus populating
the second rung
of the hierarchy[3]

[3]Chapter 6, illustration for the duotensor circuit and the corresponding Causaloid diagram considered in Section 6.3.6

Chapter 6

The Hierarchy's Second Rung: Duotensors meet the Causaloid

Meta Compression with $\Lambda_{\mathcal{O}_d}$ sufficiency provides us with a Hierarchy for physical theories. Which rungs do Quantum theory (QT), Classical Probability Theory (CPT) belong to? *We show using the Duotensor Framework that QT and CPT are* $\Lambda_{\mathcal{O}_2}$ *sufficient.*

Meta Compression, the third level of physical compression in the Causaloid framework, provides a compressed specification of the Causaloid $\mathbf{\Lambda}$ given some rules (Chapter 4). We characterised Meta Compression through $\Lambda_{\mathcal{O}_d}$ sufficiency which gave us some general rules for the rungs of a Hierarchy for physical theories (Chapter 5). A natural question arises, where do existing theories, such as Quantum theory (QT) and Classical Probability Theory (CPT) live within this Hierarchy? In this Chapter, we apply the Causaloid Framework to operational circuits pertaining to generalised probability theories with directed wires called the *duotensors* [49], which can be used to describe finite-dimensional QT and CPT as special cases. We (diagrammatically) identify the three levels of physical compression, of the Causaloid framework, in the Duotensor Framework and provide the structure of the identified Lambda matrices as well as a more specific manifestation of the Meta Compression rules. This application is possible since the Duotensor Framework is formulated in a manner having *formalism locality* (shortened to f-locality), a concept introduced by Hardy in [46], which allows us to do calculations pertaining to a region by using mathematical objects only pertaining to said region. We show that duotensors are $\Lambda_{\mathcal{O}_2}$ sufficient, thus populating the second rung. In contrast to Tomographic Compression which separates QT and CPT through the value of r in $K = N^r$ (refer to Section 4.3.2 and for more context Hardy's QT from five reasonable axioms paper [40], Sorkin's Hierarchy [93], and the introduction of this relation by Wootters [99]), the Hierarchy is more concerned with Compositional Compression and places QT and CPT together in this respect.

We also briefly discuss how one may generalise the structure of the duotensors circuits to be composed with what we will call *hyper$_d$wires* instead of wires, these are hyper-edges that connect d regions. Such generalisation would be $\Lambda_{\mathcal{O}_d}$ sufficient by construction and can be used to populate the higher rungs of the hierarchy. We speculate upon what we may learn from such generalisations, once constructed and lay out some future directions.

6.1 Applying the Causaloid Framework

The Causaloid framework aims to help us find the structure of a (potentially unknown) theory given sufficient operational data as a means of theory construction. To do so it starts from regions that are on equal footing to avoid making assumptions on causality within the theory. Eventually we would like to use the Causaloid Framework to help us find clues towards the structure of Quantum Gravity. This is undoubtedly a challenging research program. To prepare for this challenge can we learn how to apply the Causaloid Framework to existing theories, such as Quantum Theory and Classical Probability Theory? Indeed one could start from operational data but that would entail relearning a lot of research that has led to these well established theories. Additionally, we would have to start from scratch for each such theory. Instead we propose to start from formulations of these theories that make the application of the Causaloid Framework possible at a suitable intermediate step that circumvents starting from operational data. Further, since the space of Generalised Probability Theories (GPTs) house both finite dimensional QT and CPT, among other theories (such as the PR-Box world), applying the Causaloid Framework to a framework of GPTs with directed wires called the *Duotensor Framework* is ideal for this purpose.

In the past two Chapters we focused on reviewing and providing a diagrammatic language for the Causaloid Framework (Chapter 4) and studying Meta Compression, thereby characterising the hierarchy (Chapter 5). The application of the Causaloid Framework to existing physical theories will help us learn a few things- how to identify the levels of physical compression (and thus the central mathematical object – the Causaloid $\mathbf{\Lambda}$) through the mathematical objects of a theory; how the hierarchy plays into these identified mathematical objects corresponding to Tomographic and Compositional Λ matrices; and finally we learn some insights about the physical theory itself, through its place in the hierarchy and since the Causaloid $\mathbf{\Lambda}$ can be seen as the specification of the physical theory itself. Once we are able to see how the Causaloid Framework applies to existing theories, we discuss some preliminary ideas in Section 6.4 towards a future research project of constructing frameworks for the higher rungs, in order to populate the hierarchy with the application of the Causaloid framework in mind.

6.1.1 Formalism-local and Operational Formulations

To apply the Causaloid Framework to some formulation of a theory, as opposed to starting with operational data, we must meet certain conditions on the formulation of said theory. Let us see why. Recall that in the Causaloid Framework we work with operational data (obtained through some experiments) in the form of piles of cards. In lieu of these cards with data we require formulations that are operational in nature. This means that the Causaloid Framework is compatible with theories formulated in an operational manner. Since physical theories are validated experimentally, having operational formulations of a theory are certainly useful as evident from the field of Quantum Information (which serves as an operational formulation of Quantum theory).

Further, in the Causaloid Framework, piles of cards are organised into elementary region sets R_x, which have cards with the same location x, such that arbitrary regions are on an equal footing prior to compositional compression. Any calculations for probabilities of a region $R_{\mathcal{O}}$ can be done through the use of mathematical objects pertaining to some predictively well-defined region R which contains $R_{\mathcal{O}}$. This was crucial since this allowed for performing calculations pertaining to a region without referring to regions outside it, specially when the causal relations between these regions are not presumed *a priori*. Indeed, the Causaloid framework would not carry full generality when applied to formulations of theories that involve evolution of states in time, that restrict the kind of $R_{\mathcal{O}}$ one can consider. Therefore, we require formulations of existing theories that have, what Hardy calls, the property of *formalism locality (f-locality))*[1] as defined by him in [49] as follows:

> "**Formalism Locality (F-Locality):** A formalism for a physical theory is said to have the property of "formalism locality" if we can do calculations pertaining to any region of space-time employing only mathematical objects associated with that region." (Section 1, [49])

It must be emphasised though that f-locality is not a property of a theory but rather of its formulation. Formulations of theories with evolution of state in time, where the state is stretched over a large (or infinite) spatial surface, such as in standard quantum theory aren't f-local. Similarly, formulations of theories through boundary conditions where the

[1] Usage: formalism locality or f-locality (*noun*), formalism local or f-local (*adjective)*

boundary lies far from regions of which quantities need be calculated (such as scattering amplitudes calculations in studies of Quantum Field Theories) aren't f-local either. One may even argue that non-f-local formulations of theories aren't truly compatible with a strict operational methodology, in so far that experiments can be conducted in arbitrary localised regions of space-time and referring to regions outside where and when an experiment occurs would inevitably require some assumptions on the structure of causality in the theory (assumptions we are not ready to make), which may not be evident from within the localised experiment region itself.

The Causaloid Framework is setup operationally and in a f-local manner. Therefore, to apply it to an existing theory while circumventing starting from data would require us to find or construct a formalism-local operational formulation of said theory. Part of Hardy's long standing research program (summarised in [46]) since the Causaloid papers [41, 42], has been towards formulating existing theories in an operational and f-local manner, be it for Quantum theory ([46, 47]), for Quantum Field theory (this is briefly done in Part V, [50]), or for General Relativity ([50]), so that they may be leveraged to construct an operational, f-local path towards Quantum Gravity. For the purposes of this work we apply the Causaloid Framework to the Duotensor Framework, as mentioned before.

6.1.2 Outline of the Chapter

The remaining Chapter is structured as follows. Section 6.2 introduces the reader to the basics of the Duotensor Framework based on the paper [49]. Section 6.3 will contain the main contribution of this Chapter, where we will identify the three levels of physical compression from the Causaloid Framework in the Duotensor Framework. Since both the Causaloid and Duotensor Framework support diagrammatic representations, we will apply the Causaloid Framework diagrammatically. We make use of the modified version of (Tikz) duotenzor package [45] to produce the many diagrams in this Chapter. Here, we will also see the Lambda matrix relations and Omega set relations that we derived in Chapter 5 when solved further given the duotensorial structure. In Section 6.4 we will discuss how one may generalise the duotensors and in Section 6.4.2 we discuss open questions around how the hierarchy may relate to existing literature around *causal indefiniteness*.

6.2 Duotensor Framework

We are now ready to give a primer for the *Duotensor Framework* [46, 49], set up to treat probabilistic circuits. The duotensor framework consists of operations, which come from the operational description of some experiment. Operations have input and output wires that join them to other operations. Operations with only outputs are called *preparations* and operations with only input are analogous to measurements called *results*. Finally operations with both inputs and outputs may be thought of as transformations. But the framework unifies all of these under operations in the spirit of keeping regions on a causally equal footing. These operations together when joined are referred to as fragments. A fragment with no open input or output wires is called a circuit. Fragments (the operational description) are then associated with duotensors (the mathematical objects), which gives us a way to calculate probabilities. Duotensors are a lot like tensors but additionally have two kinds of inputs and two kinds of outputs (diagrammatically denoted with black or white dots) that led to the name "duo"-tensors.

The duotensors pertain to Generalised Probability Theories that can be set up in circuit form with directed wires. One can "upload" physical theories in to the Duotensor Framework if the physical theory can be expressed though operations and wires and satisfies the two assumptions of the Duotensor Framework (Section 12, [49]). Finite dimensional Quantum theory when "uploaded" to the Duotensor Framework gives the Operator Tensor Formulation of Quantum theory [47, 48].

In this Section we review the operational descriptions within the Duotensor Framework, how probabilities are set up and finally how to calculate them using duotensorial expressions, based on Section 2 to Section 7 of [49]. Alongside, through use of footnotes, we also *identify* how a concept of the Duotensor Framework (DF) will relate to the Causaloid framework (CF), connecting to the work to be presented in Section 6.3. We use footnotes to avoid the mixing of concepts of the Duotensor and Causaloid Frameworks for new readers, while also catering to the curiosity of more eager readers who are (somewhat) familiar to the duotensors. The reader may also safely choose to ignore these footnotes as we will go through them at length in the coming Section.

6.2.1 Operational descriptions

The Duotensor Framework is an operational framework. We begin by describing the experiments – through apparatus, knob settings, system types, outcome sets etc. – which are depicted using operations joined by wires[2].

[2] This will correspond to CF's Pre-Compression stage of Physical Compression

Operations

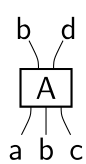
An operation[3] - A (denoted by capital letter in sans serif font), corresponds to one use of some apparatus. It is specified using the following features:

1. *Input and Outputs:* These can be thought of as apertures through which systems enter and exit the apparatus once. These systems passing in and out are operationally specified through system *types* $\mathsf{a}, \mathsf{b}, \ldots$ (note the sans serif font). We will denote inputs by $\mathsf{w}^-(\mathsf{A})$ and outputs by $\mathsf{w}^+(\mathsf{A})$. These types will affect composition of operations.

2. *Settings:* denoted by $\mathsf{s}(\mathsf{A})$ is part of the specification of the operation. It consists of knob settings and other adjustable parameters of the apparatus, that can affect outcomes.

3. *Outcomes:* denoted by $\mathsf{o}(\mathsf{A})$ is a set associated with the operation. All elements of $\mathsf{o}(\mathsf{A})$ are possible meter readings, detector clicks etc. that tell us what "happened" in the operation A. We will be interested in the probabilities of such outcomes.

Therefore, the full specification of operation A, is given by $\{\mathsf{s}(\mathsf{A}), \mathsf{o}(\mathsf{A}), \mathsf{w}^-(\mathsf{A}), \mathsf{w}^+(\mathsf{A})\}$. The operation A is represented diagrammatically by a box with inputs and outputs with system types, and symbolically with subscripts (superscripts) for inputs (outputs). Note that the integer to the system type in the symbolic notation helps identify apart two distinct input/output which may have the same type (for example here $\mathsf{b}_2, \mathsf{b}_4$), while in the diagrammatic notation such distinction is unambiguous from the diagram.



$$\iff \quad \mathsf{A}_{\mathsf{a}_1\mathsf{b}_2\mathsf{c}_3}^{\mathsf{b}_4\mathsf{d}_5} \tag{6.1}$$

Wires

We use wires[4] to join different operations when a system exits from an aperture of some operation and enters the aperture of another. Wires are governed by the following rules:

[3] Operations will correspond to CF's Elementary Regions

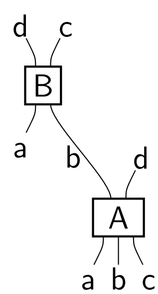
[4] Wires tell us which regions are Causally Adjacent $\bowtie$.

1. *One wire:* Since wires are operational equivalent of the passage of some system therefore a wire can connect to only one input and one output, consequently only one wire may connect an input and output

2. *Type matching:* Since wires represent some system therefore joining an output and input is possible only if their types match, (in the example below wire of type b joins A operation's output to B operation's input).

3. *No closed loops*[5]*:* Since wires are directed from outputs to inputs, it is demanded that when we go from one operation to another via wires that we never reach the same operation again. This rules out closed time-like loops.

Wires are represented diagrammatically by a line and symbolically are seen through repeated indices - when a subscript of some operation and superscript of another are the same. Consider the example of a wire joining two operations A and B:



$$\iff \qquad \mathsf{A}^{\mathsf{b}_4\mathsf{d}_5}_{\mathsf{a}_1\mathsf{b}_2\mathsf{c}_3}\mathsf{B}^{\mathsf{d}_7\mathsf{c}_8}_{\mathsf{a}_6\mathsf{b}_4} \tag{6.2}$$

Note that the symbolic notation contains integer subscripts on the system input and output types, since repeated indices represent the wires, and the type and integer should match. In the diagrammatic notation since the wires are evident within the diagrams, integer subscripts aren't necessary.

Fragments

Fragments[6] are formed by wiring together some operations while obeying all wiring rules, and are quite versatile. They are denoted by capital alphabets in sans serif font, such

[5]In principle one may relax this as long as the probabilities are well-defined. We will utilise such an exception, with no physical repercussions, in order to apply CF to DF, to be introduced in Equation 6.63.

[6]A fragment made of more than one operation will correspond to CF's composite regions

as fragment F. For example, the composition of A and B is a fragment, say $\mathsf{F} = \mathsf{AB}$ (here we are suppressing the subscripts and superscripts of the fragments for brevity). Fragments in general will have some open outputs and inputs after wiring together the operations and will be associated with settings $\mathsf{s}(\mathsf{F})$, outcomes $\mathsf{o}(\mathsf{F})$, similar to operations. The inputs and outputs of the operations that make it, will either remain open $\mathsf{w}^-(\mathsf{F})$, $\mathsf{w}^+(\mathsf{F})$ or become associated with internal wiring $\mathsf{w}(\mathsf{F})$ of the fragment. For example for the fragment $\mathsf{F} = \mathsf{AB}$, the inputs and outputs of A, B, $\mathsf{w}^-(\mathsf{A})$, $\mathsf{w}^+(\mathsf{A})$, $\mathsf{w}^-(\mathsf{B})$, $\mathsf{w}^+(\mathsf{B})$ will give the inputs and outputs for fragment F: $\mathsf{w}^-(\mathsf{F})$, $\mathsf{w}^+(\mathsf{F})$ and additionally some internal wiring $\mathsf{w}(\mathsf{F})$ (in this case it is the wire with system type b). We will denote $\mathsf{w}^-(\mathsf{F})$, $\mathsf{w}^+(\mathsf{F})$, $\mathsf{w}(\mathsf{F})$ together as $\mathsf{w}^{\mathsf{all}}(\mathsf{F})$. The setting and outcome sets need not be Cartesian products of the operations the fragment is made of[7]. A fragment may be a part of a bigger fragment. Fragments may therefore be regarded as the circuit equivalent of arbitrary regions of space-time. A fragment may be made up of disjoint fragments. An operation is a special case of a fragment. Another special case of a fragment is a circuit, explained below.

Circuits

Circuits[8] are fragments with no open inputs or outputs left after wiring its operations. Thus, for a circuit C, we have $\mathsf{w}^-(\mathsf{C}), \mathsf{w}^+(\mathsf{C})$ as empty specifications and $\mathsf{w}^{\mathsf{all}}(\mathsf{C}) \equiv \mathsf{w}(\mathsf{C})$. Similar to fragments a circuit may consist of disjoint circuits. Circuits play an important role in calculating probabilities.

6.2.2 Probabilities

A fragment is something that "happens". The probability associated with this fragment happening is defined as the probability some outcome happens, given settings and wiring.

$$\mathsf{Prob}(\mathsf{A}) = \mathsf{Prob}(\mathsf{x_A} \in \mathsf{o}(\mathsf{A})|\mathsf{s}(\mathsf{A}), \mathsf{w}^{\mathsf{all}}(\mathsf{A})) \tag{6.3}$$

One may similarly define conditional probabilities for two disjoint (or non-overlapping) Fragments A, B by the following expression

$$\mathsf{Prob}(\mathsf{A}|\mathsf{B}) = \mathsf{Prob}(\mathsf{x_A} \in \mathsf{o}(\mathsf{A})|\mathsf{s}(\mathsf{A}), \mathsf{w}^{\mathsf{all}}(\mathsf{A}), \mathsf{x_b} \in \mathsf{o}(\mathsf{B}), \mathsf{s}(\mathsf{B}), \mathsf{w}^{\mathsf{all}}(\mathsf{B})) \tag{6.4}$$

Immediately, we must ask when are these probabilities well conditioned?

[7]This is due to CF's Compositional Compression Structure
[8]Circuits will correspond to CF's predictively well-defined regions

Well Conditioned Probabilities:

$$\begin{aligned} &\text{The probability } \mathsf{Prob}(\mathsf{A}|\mathsf{B}) \text{is well conditioned} \\ &\text{if} \quad \mathsf{Prob}(\mathsf{A}|\mathsf{BC}) = \mathsf{Prob}(\mathsf{A}|\mathsf{BD}) \text{ for all } \mathsf{C}, \mathsf{D} \end{aligned} \tag{6.5}$$

In such a case, we can see that $\mathsf{Prob}(\mathsf{A}|\mathsf{B})$ is fully determined by the fragments A, B. While probabilities are not always well conditioned, the framework provides us with a condition that guarantees well-conditioned probabilities.

> **Assumption 1**[9] The probability, Prob(A), for any circuit, A is well conditioned. The probability is determined by $\mathsf{s}(\mathsf{A}), \mathsf{w}^{\mathsf{all}}(\mathsf{A}), \mathsf{o}(\mathsf{A})$ and is independent of operational descriptions belonging outside the circuit.

Using Bayes rule and the above Assumption one can show (detailed steps available on page 11, [49]) that for two circuits A, B we have

$$\mathsf{Prob}(\mathsf{AB}) = \mathsf{Prob}(\mathsf{A})\mathsf{Prob}(\mathsf{B}) \tag{6.6}$$

While the Assumption applies to circuits we do want to also be able to work with fragments. To this effect the notion of *equivalence relations* is introduced. We say, roughly speaking, that two objects are equivalent if they share the same probabilistic properties. To define equivalence relations the $p(.)$ function is introduced as follows:

$$p(\alpha\mathsf{A} + \beta\mathsf{B} + ...) = \alpha\mathsf{Prob}(\mathsf{A}) + \beta\mathsf{Prob}(\mathsf{B}) + ... \tag{6.7}$$

where α, β are real numbers and A, B are circuits. Clearly this implies that $p(\mathsf{A}) = \mathsf{Prob}(\mathsf{A})$. Similarly we have $p(\mathsf{AB}) = p(\mathsf{A})p(\mathsf{B})$ from Equation 6.6. We are now ready to define what is meant by *Equivalence* denoted by the symbol $\equiv$.

> **Equivalence**: We say two expressions are equivalent
>
> $$\begin{aligned} \text{expression}_1 \;&\equiv\; \text{expression}_2 \\ \text{if} \quad p(\text{expression}_1\mathsf{F}) &= p(\text{expression}_2\mathsf{F}) \end{aligned} \tag{6.8}$$
>
> for any fragment F such that $\text{expression}_1\mathsf{F}$ and $\text{expression}_2\mathsf{F}$ are some linear sum of circuits as required by the function $p(.)$.

[9] Therefore circuits correspond to predictively well-defined regions.

Equality implies equivalence but not the other way around, making equivalence a weaker concept. Equivalence will be very useful (as we shall see later) since it allows us to compare objects even if they are not associated with well conditioned probabilities as long as they combine with some fragment to give well conditioned probabilities. One important equivalence we will make of is $\mathsf{Prob}(\mathsf{A}) \equiv \mathsf{A}$ for any circuit A, since the fragment F that can make a circuit into a circuit is another circuit and using 6.6 we can write $p(\mathsf{AF}) = p(\mathsf{A})p(\mathsf{F}) = p(\mathsf{Prob}(\mathsf{A})\mathsf{F})$.

6.2.3 Duotensors

In this part, we show how to go from operations to duotensors, which may be considered as going from the operational descriptions - the physics, to the corresponding mathematics. We begin with an operation A with only one output (called a *preparation*) and an operation B with only one input (called an *result*), each with system type 'a' which can be joined together with a wire to give the simplest circuit. Then we give the duotensor associated with this circuit. To do so, the concepts of *fiducial results* and *fiducial preparations* are introduced, side by side due to the similarity of their structures (up to direction of time).

Fiducial Results:
An operation that has only inputs is called an *result*. Consider the result $\mathsf{B}_{\mathsf{a}_1}$ with input of system type a.
$\mathsf{X}^{a_1}_{\mathsf{a}_1}$, where $a_1 = 1$ to K_a, is a *fiducial set* of results for system type a if we can write

$$\mathsf{B}_{\mathsf{a}_1} \equiv B_{a_1} \mathsf{X}^{a_1}_{\mathsf{a}_1} \tag{6.9}$$

for any result $\mathsf{B}_{\mathsf{a}_1}$.
B_{a_1} is a duotensor that supplies the coefficients, which has K_a real (possibly negative) entries.

Fiducial Preparations:
An operation that has only outputs is called a *preparation*. Consider the preparation $\mathsf{A}^{\mathsf{a}_2}$ with output of type a.
${}_{a_2}\mathsf{X}^{\mathsf{a}_2}$, where $a_2 = 1$ to K_a, is a *fiducial set* of preparations for system type a if we can write

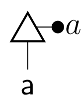
$$\mathsf{A}^{\mathsf{a}_2} \equiv {}^{a_2}A \ {}_{a_2}\mathsf{X}^{\mathsf{a}_2} \tag{6.10}$$

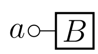
for any preparation $\mathsf{A}^{\mathsf{a}_2}$.
${}^{a_2}A$ is a duotensor that supplies the coefficients, which has K_a real (possibly negative) entries.

Let us unpack Equations 6.9 and 6.10. These are hybrid equations - they contain operations A, B, denoted with sans serif font; as well as duotensors A, B denoted with the usual math font. The sans serif indices of operations signify system type, while math font indices of duotensors run over the K_a entries in this example. The usual convention of summing over repeated indices (of the same font) is used.

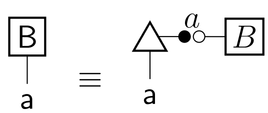
The fiducial set of results (or preparations) for a given system type (here a) is minimal, such that no set of fewer than K_{a} results (or preparations) can have this property. Such a fiducial set can always be found since in the worst case there would be one fiducial result (or preparation) for each possible result (or preparation). In general though the fiducial set of results (or preparations) will be much smaller than the set of results (or preparations) [10]. The fiducial set of results or preparations are the bridging elements of the hybrid equations which map the physics (operations) to the mathematics (duotensors).

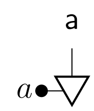
Further, we can write each element of Equations 6.9 and 6.10 in diagrammatic form:

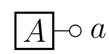
$\mathsf{X}^{a_1}_{\mathsf{a}_1} \iff$  (6.11)

$B_{a_1} \iff$  (6.12)

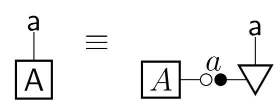
Then the diagrammatic form for Equation 6.9 would be:

 (6.13)

${}_{a_2}\mathsf{X}^{\mathsf{a}_2} \iff$  (6.14)

${}^{a_2}A \iff$  (6.15)

Then the diagrammatic form for Equation 6.10 would be:

 (6.16)

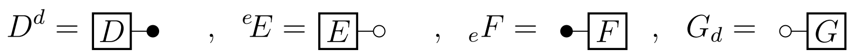
Equations 6.13 and 6.16 are hybrid diagrams corresponding to the hybrid equations 6.9 and 6.10. Hybrid diagrams have wires running vertically for the operational description and have links running horizontally for the duotensors. Horizontal links represent the summation over the corresponding index (here a).

Black and White Dots

Notice the black and white dots used in the diagrammatic representation – these play an important role. A duotensor, generally, can have four kinds of indices – subscripts and pre-subscripts corresponding to results and, superscripts and pre-superscripts corresponding to preparations. The reason for four indices is that there are two independently chosen basis sets associated with every index - a fiducial set of effects and a fiducial set of preparations. Diagrammatically the two basis sets are depicted using black or white dots, the convention

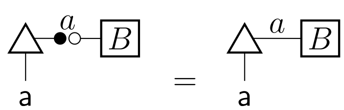
[10] Fiducial results and preparations will relate to CF's tomographic compression.

is depicted below:

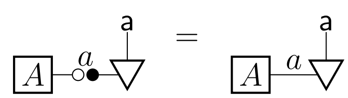


(6.17)

A *rule* is imposed on composing duotensors. Only white and black dots may be matched together as seen implicitly before in Equations 6.13 and 6.16. Symbolically repeated indices between (pre-)subscript and (pre-)superscript can be summed over as long as they lie on the same side of the duotensors (along with matching types and fonts), this is the symbolic equivalent of the black and white dots matching rule. This additional structure with the black and white dots (or the four indices) that distinguishes these from tensors lends the duotensors their name. In the above example, D^d can be matched with G_d, while ^eE can be matched with $_eF$. The horizontal link interrupted by a black and a white dot indicate that we sum over the index, and can be replaced with a line :



(6.18)

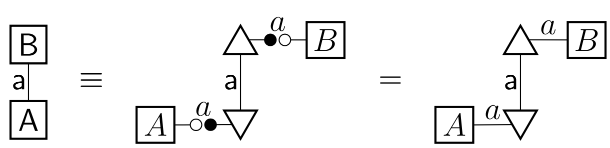


(6.19)

We will soon discuss the available conventions for using white and black dots as well as how one may flip colours (or symbolically hop them over to the other side) using the *hopping metric*. For the purpose of the work to be presented in the Section 6.3, we always use black dots for fiducial results (such as $\mathsf{X}^{a_1}_{\mathsf{a}_1}$) and fiducial preparations (such as $_{a_2}\mathsf{X}^{\mathsf{a}_2}$), thus unambiguously leaving the duotensors with white dots.

The simple circuit

We can combine operation A and operation B giving the simplest possible circuit. We use the Equations 6.13 and 6.16 to expand the circuit (notice equivalence symbol ≡):



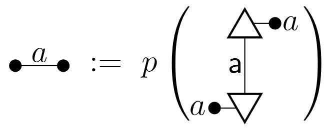
(6.20)

the above diagrams symbolically give:

$$\mathsf{A}^{\mathsf{a}_1}\mathsf{B}_{\mathsf{a}_1} \equiv {}^{a_1'}A \;\; {}_{a_1'}\mathsf{X}^{\mathsf{a}_1} \;\; \mathsf{X}^{a_1}_{\mathsf{a}_1} \;\; B_{a_1} \tag{6.21}$$

We are now ready to define the *hopping metric*, which essentially tells us how the two independently chosen basis sets associated with every index (lending duotensors their structure) talk to each other.

Hopping metric: denoted by ●–●[11], is a duotensor, defined using the $p(\cdot)$ function



$$\Leftrightarrow \qquad {}_{a'_1}g^{a_1} := p\left({}_{a'_1}\mathsf{X}^{\mathsf{a}_1}\mathsf{X}^{a_1}_{\mathsf{a}_1}\right) \tag{6.22}$$

and its inverse is represented by ○–○.

$$\circ\!-\!\circ \;\Leftrightarrow\; {}^{a'_1}g_{a_1} \tag{6.23}$$

While ●–● has real non-negative entries, ○–○ has possibly negative real entries.

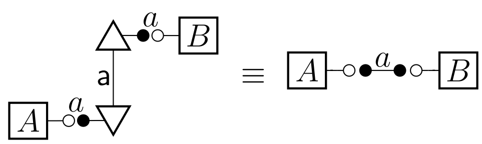
Identity: One can also introduce ●–○ and ○–●. Since ○–○ is the inverse of ●–● we have:

$$\left(\bullet\!-\!\bullet\!\circ\!-\!\circ = \bullet\!-\!\circ \;\Leftrightarrow\; {}_{a_1}g^{a''_1}\,{}^{a'_1}g_{a''_1} := {}^{a'_1}_{a_1}\delta\right) \text{ and } \left(\circ\!-\!\circ\!\bullet\!-\!\bullet = \circ\!-\!\bullet \;\Leftrightarrow\; {}^{a''_1}g_{a'_1}\,{}_{a''_1}g^{a_1} := \delta^{a_1}_{a'_1}\right) \tag{6.24}$$

This implies that both ●–○ and ○–● are equal to the identity. Therefore one can introduce pairs of black and white dots, or delete them:

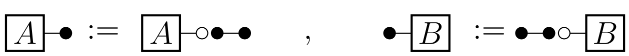
$$-\!\circ\!\bullet\!- \;=\; \text{——} \;=\; -\!\bullet\!\circ\!- \tag{6.25}$$

Back to the simple circuit, using the definition of the hopping metric we can reduce Diagram 6.20 to:



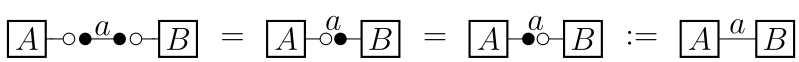
(6.26)

The hopping metric can be used to flip the colour of dots, as shown in :



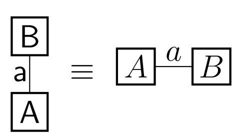
(6.27)

Relations 6.27 applied to 6.26 gives:



(6.28)

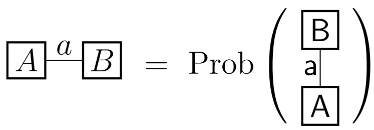
Using 6.20, 6.26, 6.28 we get the equivalence between the operational side (in this case the simple circuit, but can be generalised to fragments) on the left hand side with the associated duotensor on the right hand side:



(6.29)

[11] The hopping metric will be important for CF's compositional compression.

We can go further to give an equality relation. Recall that $\mathsf{Prob}(\mathsf{C}) \equiv \mathsf{C}$ for any circuit C. The duotensor is in fact equal to the probability associated with the circuit:



$$, \quad \mathsf{Prob}(\mathsf{A}^{\mathsf{a}_1}\mathsf{B}_{\mathsf{a}_1}) = {}^{a_1}A \; {}_{a_1}B = A^{a_1} \; B_{a_1} \tag{6.30}$$

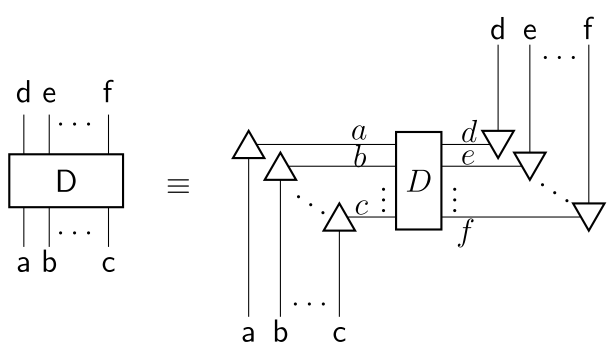
Note the similarity of structure of the duotensor and the circuit, which will hold true for general circuits as well, once we have full decomposibility. Hardy points out:

> "It is striking that the probability for a circuit is given by a duotensor calculation that looks the same as the circuit itself... In the diagrammatic case we need only rotate the diagram through 90° and change the font from sans serif to normal maths font. In the symbolic case we need only change the font." (Section 6.9, [47])

Full Decomposability

Operations will not always have single inputs and outputs, how do we map multiple inputs and outputs to fiducial sets of results and preparations? We require some extra structure to find the corresponding duotensor calculations for a general operation. The following assumption addresses this issue:

> **Assumption 2: Operations are fully decomposable:** Any operation is assumed to be equivalent to a linear combination of operations each of which consists of a fiducial set of results for each input and a fiducial set of preparations for each output. Hence, this assumption is equivalent to the statement that any operation, $\mathsf{D}^{\mathsf{d}_4\mathsf{e}_5\ldots\mathsf{f}_6}_{\mathsf{a}_1\mathsf{b}_2\ldots\mathsf{c}_3}$, can be written, in diagrammatic notation, as



(6.31)

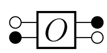
and correspondingly in symbolic notation as,

$$\mathsf{D}^{\mathsf{d}_4\mathsf{e}_5\dots\mathsf{f}_6}_{\mathsf{a}_1\mathsf{b}_2\dots\mathsf{c}_3} \equiv \mathsf{X}^{a_1}_{\mathsf{a}_1}\mathsf{X}^{b_2}_{\mathsf{b}_2}\cdots\mathsf{X}^{c_3}_{\mathsf{c}_3}\ {}^{d_4e_5\dots f_6}D_{a_1b_2\dots c_3}\ {}_{d_4}\mathsf{X}^{\mathsf{d}_4}\,{}_{e_5}\mathsf{X}^{\mathsf{e}_5}\cdots{}_{f_6}\mathsf{X}^{\mathsf{f}_6} \tag{6.32}$$

Physically, this assumption has to do with recognising that inputs and outputs are associated with *physical systems* entering and exiting the operation and thus is a statement on the statistics of a composite system. It is shown in [47, 48] that for the space of generalised probabilistic theories, full decomposibility is equivalent to tomographic locality[12], which states that the state of a composite system can be determined from the statistics collected by making measurements on its components.

What is a duotensor?

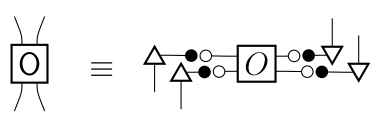
Upon discussing full decomposability we are ready to discuss general duotensors with multiple horizontal wires corresponding to preparations and/or results. As mentioned earlier, duotensors can have four kinds of indices - (pre-)subscripts (on the left side of the duotensor box) and (pre-)superscripts (on the right side of the duotensor box). We also saw how to use the hopping metric to flip the colours (or hop the indices to the other side). With multiple wires from a duotensor in general one may have any combination of coloured dots, consider the example of an operation O:



(6.33)

This can become messy. Therefore we share three possible conventions to represent duotensors, each with its own significance. Let's consider the operation O (from above) and how to represent its associated duotensor in the different conventions:

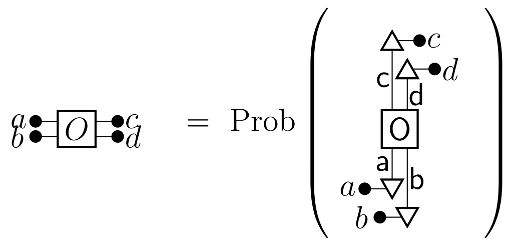
1. *All white dots* A duotensor with all white dots provides the weights in the sum over fiducial elements.



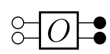
(6.34)

2. *All black dots* A duotensor with all black dots corresponds to the fiducial probabilities when fiducial preparations (results) are placed on each input (output) of an operation. Capping the inputs and outputs of O with fiducial elements gives:

[12]The assumption of full decomposibility will provide a specific form for CF's tomographic lambda matrices and omega sets.



3. *Standard form* is when we have all the indices on the right hand side so we have only superscripts and subscripts. The standard form only invokes the use of fiducial results (but not fiducial preparations). Diagrammatically this corresponds to having all white dots on the left and all black dots on the right

   (6.35)

   One can connect duotensors up in standard form without using the hopping metric since black dots will always be put next to white dots.

These conventions tell us what do the black and white dots signify. It is clear that a duotensor with all black dots corresponds to the probability of the particular fiducial result(s) and/or preparation(s) to have happened, and thus must be non-negative (and real). On the other hand a duotensor with all white dots corresponds to coefficients that weigh the probabilities to give probabilities for any result(s) and/or preparation(s) to have happened, and are real but may be negative, depending on the theory. While classical probability theory gives non-negative entries, quantum theory in general requires negative entries.

Similar to the steps shown in the simple circuit going from operations to duotensors one may use all the tools provided to work with any fragment or circuit. Examples can be found in [49, 47]. Further, Quantum theory can be formulated through the duotensors in a f-local manner by mapping operations with operators that are associated with the space of Hermitian operators on complex Hilbert spaces, that give further structure to the duotensors, seen in detail in [47].

We now move on to the application of the Causaloid framework to the Duotensor Frameowrk (Section 6.3) where we will stick to the convention of all white dots for duotensors of operations and all black dots for fiducial elements. We will also make use of the hopping metric and identity duotensors as required.

6.3 Causaloid meets Duotensors

We are now ready to apply the Causaloid Framework's three levels of physical compression to the Duotensor Framework. We will progressively identify descriptions and objects of the Duotensor Framework with those from the Causaloid Framework, and in the process use everything we have learnt so far from Chapter 4 until now. Since we will refer to both frameworks we will often make use of the shorthand DF (for the Duotensor Framework) and CF (for the Causaloid Framework). We consider a circuit R to serve as exposition of the identification process and then provide results that are true for general circuits. The identification process has the following six steps:

1. **Operational Data:** We begin by identifying how operational descriptions (operations, wires, settings, outcomes) from DF would translate to the organisation of data in CF, including identification of elementary regions.

2. **Probabilities:** We identify well-conditioned probabilities for circuits (Assumption 1) with the notion of predictively well-defined regions in CF, as well as use the $p(.)$ function for probabilities associated with regions.

3. **Zeroth level:** Then we consider a general circuit for exposition and identify probabilities p_{α_1} for some region R_1 as well as $\mathbf{R}(R_1)$ and $\mathbf{P}(R_1)$.

4. **First level:** Here, we identify fiducial measurements that give the Tomographic Lambda matrices and Omega sets. These have a specific form due to DF's full decomposability (Assumption 2).

5. **Second level:** We will consider two regions of the circuit for exposition and identify the Compositional Lambda matrix and Omega set, which have a specific form in terms of the hopping metric and identity maps, due to the structure of wires.

6. **Third level:** Finally we consider three regions. As an example we consider the causal relation $R_1 \bowtie R_2$ and $R_2 \bowtie R_3$ (that corresponds to Case 5 from Section 5.5). We will see that the sums in the expression for $\Lambda_{\mathcal{O}_3}(\Lambda_{\mathcal{O}_2}, \Lambda_{\mathcal{O}_1})$ will vanish. We will also see specification for the Causaloid for the Duotensor Framework which we show to be $\Lambda_{\mathcal{O}_2}$-sufficient.

The identification process will primarily be done using diagrammatics of both frameworks - those introduced in Chapter 4 for CF and those of DF from [49], reviewed in Section 6.2. The identifications for the circuit R will then be used to provide the general form for Lambda matrices and Omega sets for the Duotensor Framework, thus showing us that it is $\Lambda_{\mathcal{O}_2}$ sufficient.

6.3.1 Operational Data

In CF, recall from Section 4.2.2 and 4.2.3 we have operational data recorded on cards in the form (x - *Location*, a - *Actions*, s - *Observation*) according to which cards are organised into sets of cards *regions* (through locations), *procedures* (through actions) and *outcomes* (through observations). We are circumventing the step of starting from operational data, therefore we identify how the operational descriptions from DF (Section 6.2.1) can be parsed to fit the data format of CF.

1. **Location:** In DF the notion of location is not explicit but captured through operations which happen in a localised region of space-time. Note that the same operation can be used multiple times, in such a case one may distinguish them with an integer subscript or using a different letter. We will consider each instance of an operation to represent an elementary region (representing cards attributed to a single location). In terms of notations, each operation will be assigned with a position label (an integer), such that operations $\{\mathsf{A}, \mathsf{B}, \mathsf{C}, \dots\}$ are associated with regions $(\{R_1, R_2, R_3, ...\})$ (where we use the position of the letter in the alphabet as the region label)

$$\text{Operations } \ \mathsf{A}, \mathsf{B}, \mathsf{C}, \dots \iff \textit{Elementary regions } \ R_1, R_2, R_3, \dots \tag{6.36}$$

   It follows that fragments that are composed of more than one operation will represent composite regions and thus a fragment say $\mathsf{F} = \mathsf{AB}$ will be associated with the region $R_1 \cup R_2$. Further, for CF's predictively well-defined region R, we use the symbol R (which as we see will correspond to a circuit).

2. **Actions:** CF's notion of actions for some region, say R_1, corresponds to DF's settings $\mathrm{s}(\mathsf{A})$ and input and output wire types $\mathrm{w}(\mathsf{A})$, of the operation A. These will map to the *procedure* set F_{R_1} which may also be denoted by F_{A}.

$$\mathsf{s}(\mathsf{A}), \mathsf{w}^-(\mathsf{A}), \mathsf{w}^+(\mathsf{A}) \iff \textit{Procedure } \ F_{R_1} = F \cap R_1 \tag{6.37}$$

   An issue arises for the above identification for fragments. While CF's procedure is local on the elementary regions, fragments in DF have internal wiring that live on pairs of operations, which do not seem locally describable. Hardy discusses this issue in [43] through the notion of connections as part of "Object Oriented Operationalism". Nonetheless, this does not hinder the application of CF to DF done through elementary regions for which the identification holds.

3. **Observations:** CF's notion of observations for some region, say R_1, given the actions, corresponds to DF's outcomes $\mathsf{o}(\mathsf{A})$ of the operation A, given the setting, inputs, outputs (and wiring). This will provide the *outcome* sets Y_{R_1} or Y_{A}.

$$\mathsf{o}(\mathsf{A}) \iff \textit{Outcome}\ \ Y_{R_1} = Y \cap R_1 \tag{6.38}$$

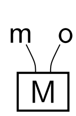
   The above identification can also be used for fragments.

It is important to note that while in DF's operational description one is provided with settings and outcomes, if one were to obtain operational data as required by CF, one would have to repeat the operation/fragment/circuit multiple times until all settings are "chosen" and all outcomes "happen".

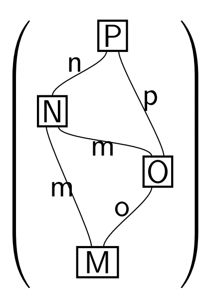
6.3.2 Probabilities

In the Causaloid framework, as well as the Duotensor framework, one important objective is to be able to calculate probabilities (an objective shared by operational theories and in particular generalised probability theories). It is important to be able to tell when calculations give well-defined probabilities. Recall from CF's Section 4.2.4 that for a composite region $R = \bigcup_x R_x$, one either considers conditions C such that $\mathrm{prob}(Y_R|F_R, C)$ is well-defined or carefully choose the composite region R such that $\mathrm{prob}(Y_R|F_R)$ is predictively well-defined without any conditioning. While fragments do not have *well-conditioned probabilities*, DF's Assumption 1 guarantees that circuits will. Therefore, if one is interested in some operations (elementary regions) one may consider circuit(s), which we denote with R, that contain these operations as the predictively well-defined region R. Since probabilities of circuits are multiplicative one may simply focus also on the circuit(s) that contain the operations (regions) of interest.

For example, if one is interested in the probabilities for some operation M (region R_{13}), one can start with $\mathrm{prob}(Y_R|F_R)$ where the predictively well-defined region R is given by the circuit $\mathsf{R} = \mathsf{MNOP}$ which contains the operation M:



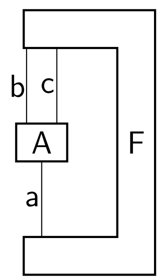
$, \quad \mathrm{prob}(Y_R|F_R) = \mathrm{prob}$



(6.39)

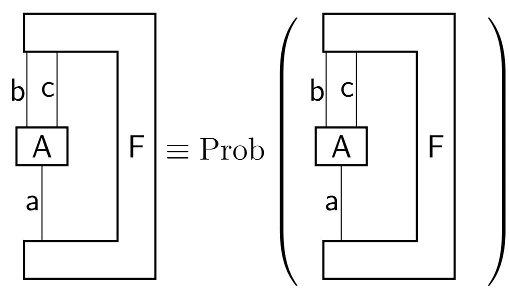
6.3.3 Level Zero: Pre-Compression

We are ready to identify the levels of physical compression (analogous to the steps from the Section 4.3). For exposition we will consider an example circuit R (which serves as the predictively well-defined region) here, throughout until the end of Meta Compression. Consider a region R_1 associated with an operation A that is within the circuit R (we have arbitrarily chosen an operation A with one input and two outputs for the purposes of exposition but one may easily consider instead an operation with different number of input(s) and output(s)). The operation A of interest and the rest of the circuit, denoted by the fragment F[13], compose to give the circuit $\mathsf{R} = \mathsf{AF}$:



(6.40)

Since we are interested in probabilities associated with this circuit recall that from DF we have that a circuit is *equivalent* to its associated probability, that is:



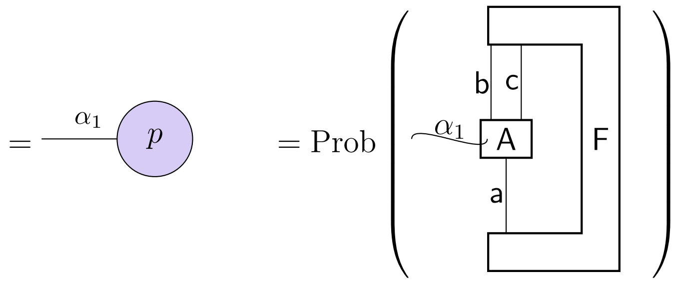
(6.41)

This distinction between equality and equivalence is important to note going forward. Through the prescription of CF we are interested in the probabilities:

$$p_{\alpha_1} = \text{Prob}(Y_{R_1}^{\alpha_1} \cup Y_{R-R_1} | F_{R_1}^{\alpha_1} \cup F_{R-R_1}) \tag{6.42}$$

$$= \mathsf{Prob}(\mathsf{x_A} \in \mathsf{o(A)}, \mathsf{x_f} \in \mathsf{o(F)} | \mathsf{s(A)}, \mathsf{w^-(A)}, \mathsf{w^+(A)}, \mathsf{s(F)}, \mathsf{w^{all}(F)}) \tag{6.43}$$

[13]It is not a coincidence that the fragment F resembles a Quantum Comb [18], since they are indeed closely related concepts. Unlike general Process Matrices that allow causal indefiniteness, DF's fragments have a definite causal structure

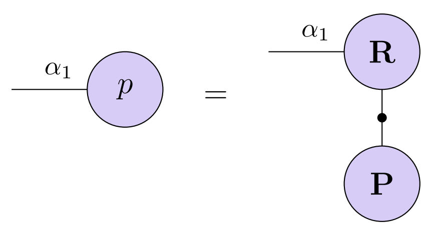


(6.44)

where we identify DF's equivalent object. Notice the α_1 label in Equation 6.44, we explain this in a moment. Let us recall (Section 4.3.1) CF's notion of a state:

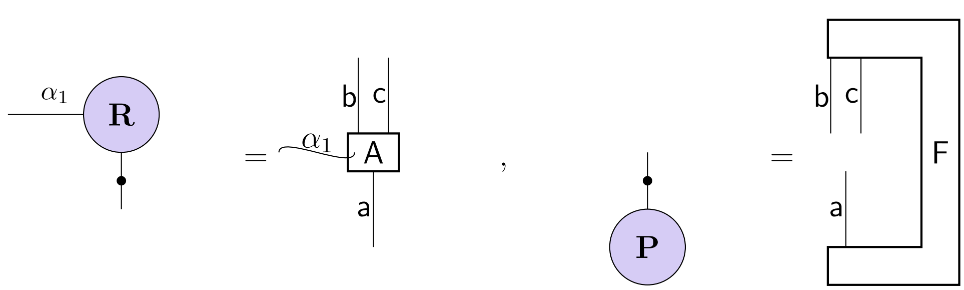
> **The state** for R_1 associated with a generalised preparation in $R - R_1$ is defined to be the thing represented by any mathematical object which can be used to predict p_{α_1} for all measurements in R_1 labelled by the index α_1.

where (Y_{R-R_1}, F_{R-R_1}) is *generalised preparation* for region R_1, and (Y_{R_1}, F_{R_1}) is a *measurement* associated with the actions and outcomes seen in R_1. The measurement outcomes for the measurement in R_1 will be labelled by α_1 which gave us CF's generalised Born rule:



(6.45)

Note that when we refer to CF we will use *generalised preparation* (where causal structure is not assumed) which should be disambiguated with the DF's concept of preparation. DF comes equipped with a definite causal structure where results and preparations are distinguished. α_1 introduced on the right hand side of Equation 6.44 labels the outcome and setting for the operation A, which is not explicitly drawn in DF, but we will draw it explicitly to aid identification with CF. We can now identify the components (measurement vector $\mathbf{R}(R_1)$, state $\mathbf{P}(R_1)$) from CF's generalised Born rule 6.45 with the elements in Equation 6.44.



(6.46)

Mapping Vectors to Fragments and Duotensors

Let us try to understand this identification. Clearly an operation (elementary region) is a particular kind of fragment. We explain how we can identify both $\mathbf{R}(R_1)$ and $\mathbf{P}(R_1)$ with fragments. We have worked with tomographic compression of two separate regions (Equation 4.49) where we can show $p_{\alpha_1\alpha_2} = \mathbf{r}_{\alpha_1}(R_1)\cdot\mathbf{p}_{\alpha_2}(R_1) = \mathbf{r}_{\alpha_2}(R_2)\cdot\mathbf{p}_{\alpha_1}(R_2)$ and in the case where two regions make a predictively well-defined region such as we have here for A and F then the object that represents the state given by the generalised preparation of one region can be seen as the measurement vector for the other region! Therefore it is in fact expected that fragments/operations can be identified as either (generalised preparation or measurement) as required by the context of which region we are writing in terms of, and the inner product corresponds to joining of wire(s) (which is why we do not need to provide the internal structure of the fragment F). In fact the role of state vector is only for the purposes of scaffolding and we later only focus on $\mathbf{r}$-vectors to do calculations, as captured by Hardy as follows:

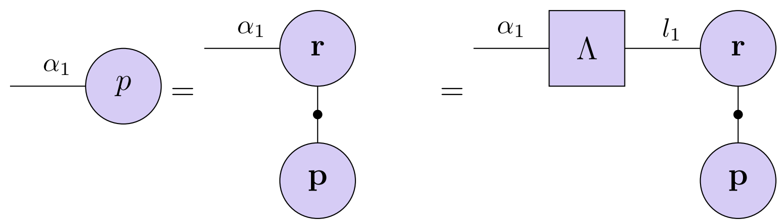
> "It is these r-type vectors that we use in real calculations. The state vector, p, is akin to scaffolding - it can be dispensed with once the structure of the r-type vectors is in place..." (Section 4.2, [46])

Further, let us address something implicit in the above equations. Recall that $\mathbf{R}$ and $\mathbf{P}$ (and subsequently $\mathbf{r}$ and $\mathbf{p}$) are prescribed to be vectors. On the other hand fragments (and subsequently duotensors) have more tensorial of a structure. How can we equate them, and more importantly why are they represented with different kinds of objects? To equate them one can "flatten" a tensor (or duotensor) into a vector (at the cost of losing some structure). By "flattening" we mean taking the components of the tensor and arranging them in a column as the components of a vector. In this example $\mathsf{A}_{\mathsf{a}}^{\mathsf{bc}}$ flattens into $\mathbf{R}$ and $\mathsf{F}_{\mathsf{bc}}^{\mathsf{a}}$ flattens into $\mathbf{P}$. It is natural that the same object is represented by different objects in the two Frameworks. CF deals with operational data that is organised and studied to find the correlations a physical theory provides; and thus it is expected to *a priori* not have the full structure in its description of the state and measurement of a region. The structure is recovered through the levels of compression. DF on the other hand has a more specific structure inbuilt even at the level of operational description, through the notion of systems that enter and exit. One may thus think of the Causaloid Framework working with a number (the size of the vector) while the Duotensor Framework comes equipped with a physically meaningful factorisation (through system types) of the same number.

This concludes the Pre-Compression phase.

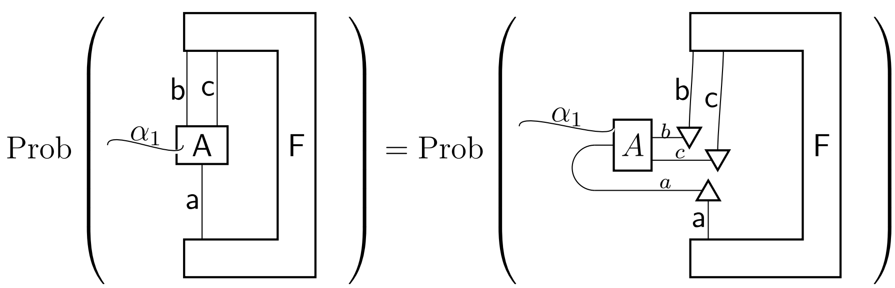
6.3.4 Level One: Tomographic Compression

CF's Tomographic compression (defined in Section 4.3.2), through a fiducial set of measurements on the region R_1, replaces the uncompressed state and measurement vector $\mathbf{P}$ and $\mathbf{R}$ with the vectors $\mathbf{p}$ and $\mathbf{r}$ of size $|\Omega_{R_1}|$ (corresponding to the size of the fiducial measurement set). Fiducial measurement vectors $\mathbf{r}_{l_1}$ can be linearly combined using the tomographic compression matrix $\Lambda^{l_1}_{\alpha_1}$ where $\alpha_1 \in \Gamma_{R_1}$ and $l_1 \in \Omega_{R_1}$, such that $\Omega_{R_1} \subseteq \Gamma_{R_1}$, to give $\mathbf{r}_{\alpha_1}$ and thus p_{α_1}. Recall the following equation:

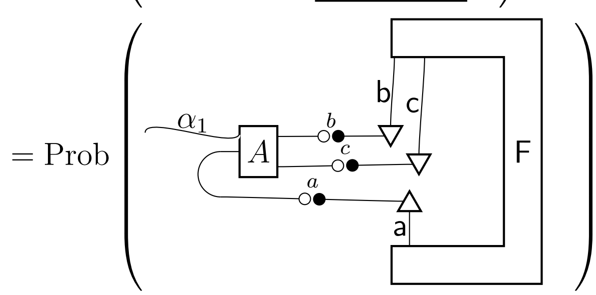


(6.47)

DF's fiducial results ($\mathsf{X}^{a}_{\mathsf{a}}$) and fiducial preparations ($_{b_2}\mathsf{X}^{\mathsf{b}_2}$, $_{c_2}\mathsf{X}^{\mathsf{c}_2}$) together play the role of CF's fiducial measurements. Let us see how. We expand the circuit R using hybrid diagrams that involve fiducial results and preparations on operation A:

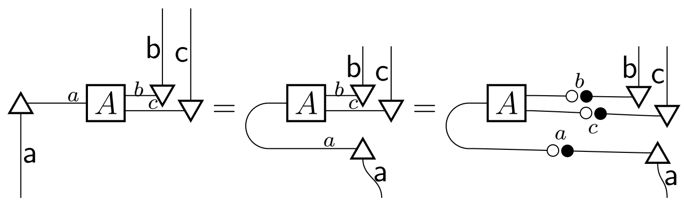


(6.48)



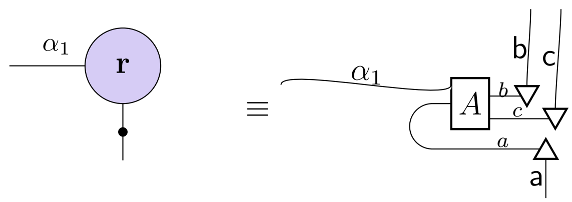
(6.49)

We slightly tweak the DF diagrammatics and bend the fiducial results wires coming out on the left back to the right side. Therefore we will have duotensors with wires on the right for both preparations and results. This is purely towards aiding the identification with the CF diagrammatics. In Equation 6.48 we have used the following duotensor manipulations where we introduced black and white dots since —○●— = —— = —●○—:



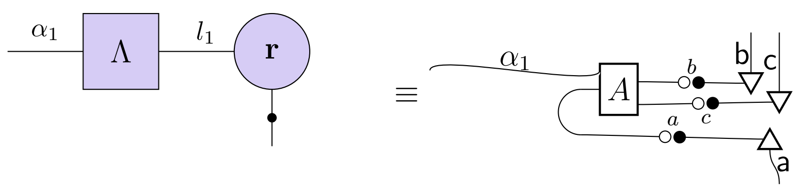
(6.50)

We compare Equation 6.47 (CF) with Equation 6.48 (DF), thus identifying $\mathbf{r}_{\alpha_1}$:



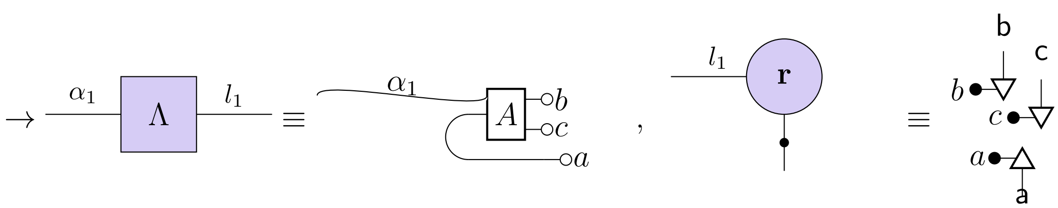
(6.51)

We continue by expanding $\mathbf{r}_{\alpha_1} = \sum_{\Omega_1} \Lambda_{\alpha_1}^{l_1} \mathbf{r}_{l_1}$ and comparing it to Equation 6.49,



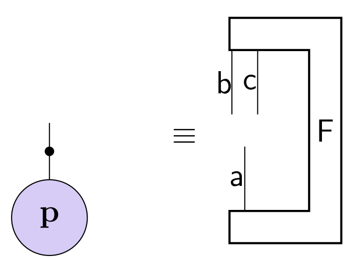
(6.52)

to identify $\Lambda_{\alpha_1}^{l_1}$ with a duotensor with *white dots*, that provides coefficients, as well as $\mathbf{r}_{l_1}$ with *black dots* corresponding to fiducial preparations and results:



(6.53)

The state $\mathbf{p}$ that lists the probabilities for fiducial measurements p_{l_1} is associated with the fragment F (same as for $\mathbf{P}$) since it continues to satisfy the definition of the state and one can always truncate $\mathbf{P}$ to contain only the entries where $\alpha_1 = l_1$ to get $\mathbf{p}$:



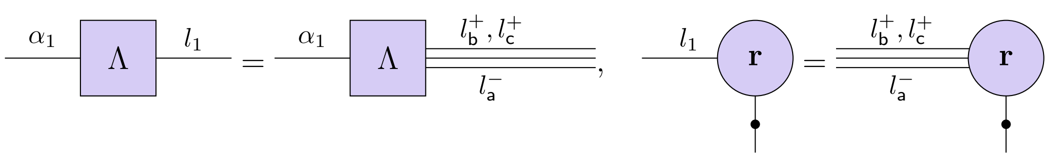
(6.54)

Let us discuss the role of DF's Assumption 2: *full decomposability* for CF's first level of physical compression. Tomographic Compression captures the number of measurements required to do Tomography of the region, here R_1. Full decomposability goes further, it tells us the number of fiducial preparations required for tomography of each of the outputs of an operation and number of fiducial results required for tomography of each of the inputs of an operation, thus giving additional structure. Full decomposability translates to *tomographic locality* in the space of GPTs.

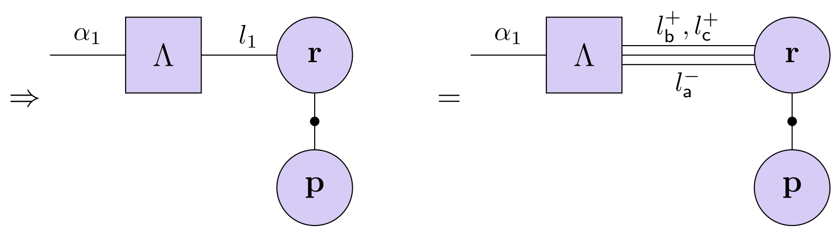
Given full decomposibility, Omega sets, here Ω_{R_1}, can be factorised in terms of DF's inputs and output. For A from our example we see that.

$$l_1 \in \Omega_1 \;\longrightarrow\; l_{\mathsf{a}}^-, l_{\mathsf{b}}^+, l_{\mathsf{c}}^+ \in \Omega_{\mathsf{a}}^- \times \Omega_{\mathsf{b}}^+ \times \Omega_{\mathsf{c}}^+$$
$$\text{where } l_{\mathsf{a}}^- \in \Omega_{\mathsf{a}}^- = \{1, \ldots, K_{\mathsf{a}}\}\ ,\ l_{\mathsf{b}}^+ \in \Omega_{\mathsf{b}}^+ = \{1, \ldots, K_{\mathsf{b}}\}\ ,\ l_{\mathsf{c}}^+ \in \Omega_{\mathsf{c}}^+ = \{1, \ldots, K_{\mathsf{c}}\} \quad (6.55)$$

where we use + for Omega sets of preparations and − for Omega sets of results as a way to distinguish them which will become useful to discuss Compositional Compression in DF, though strictly from the standpoint of CF we may not be able to discern between them from the data of a single region. Let us incorporate the form for the Omega set to update the Causaloid Diagrams of our example of the Duotensor Framework. We see that:



(6.56)



(6.57)

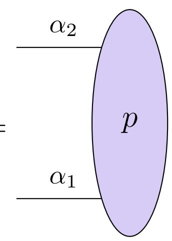
For a general operation O (Region R_{15}) (or fragment) we will have the Omega set factorise:

$$\Omega_{R_{15}} = \left(\bigtimes_{\mathsf{i} \in \mathsf{w}^-(\mathsf{O})} \Omega_{\mathsf{i}}^- \right) \times \left(\bigtimes_{\mathsf{o} \in \mathsf{w}^+(\mathsf{O})} \Omega_{\mathsf{o}}^+ \right) \quad (6.58)$$

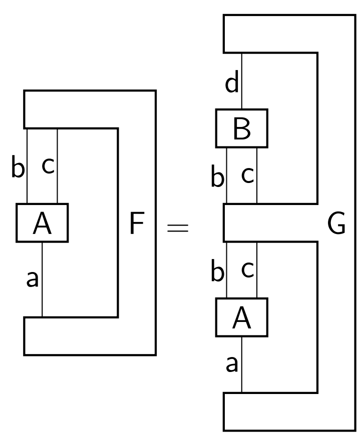
6.3.5 Level Two: Compositional Compression

Compositional Compression is physical compression for composite regions over and above Tomographic Compression of the constituent regions (Section 4.3.3). To see Compositional Compression we will consider two elementary regions, operation A (region R_1) and operation B (region R_2), corresponding to a $\Lambda_{\mathcal{O}_2}$ matrix. We are interested in probabilities p_{α_1,α_2} for the composite region $R_1 \cup R_2$ (Equation 4.36, reproduced below). We will discuss $\Lambda_{\mathcal{O}_d}$ in the Meta Compression section where we will use concepts we learnt in Chapter 5.

$$p_{\alpha_1\alpha_2} = \text{Prob}(Y_{R_1}^{\alpha_1} \cup Y_{R_2}^{\alpha_2} \cup Y_{R-R_1-R_2} | F_{R_1}^{\alpha_1} \cup F_{R_2}^{\alpha_2} \cup F_{R-R_1-R_2}) = \tag{6.59}$$



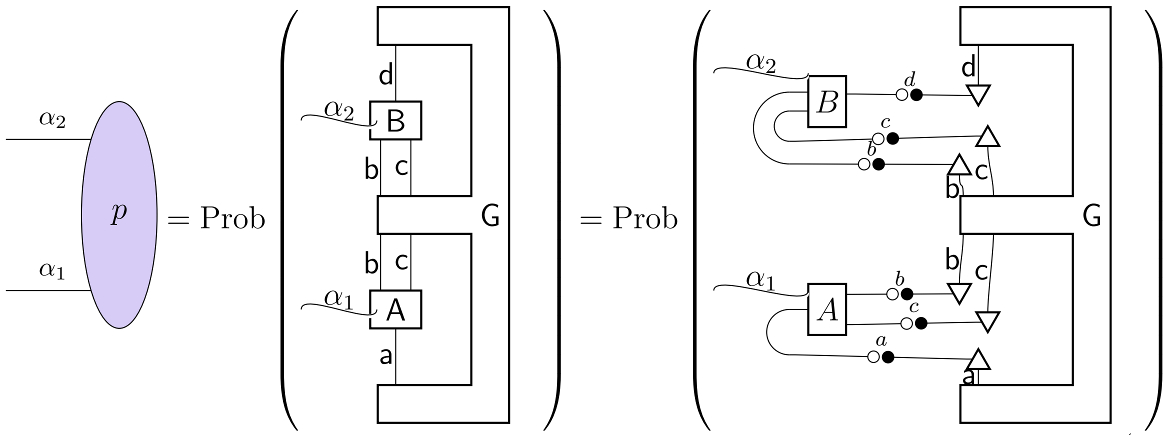
The operation A (region R_1) and the operation B (region R_2) live in the circuit R from before. We have simply expanded the fragment F in terms of the operation B and some fragment G (for rest of the circuit):



(6.60)

Comparing both sides it is clear that the composition of the operation B and the fragment G gives the fragment F or $\mathsf{F} = \mathsf{GB}$, such that $\mathsf{R} = \mathsf{FA} = \mathsf{GAB}$.

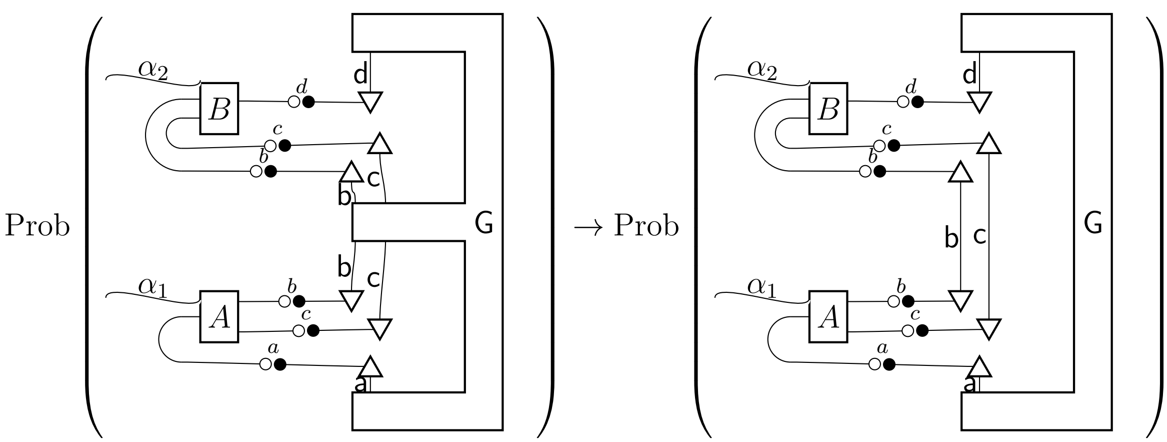
We can identify p_{α_1,α_2} with the predictively well-defined probabilities of the circuit R, indexed with α_1 and α_2 (analogous to Equation 6.44). Further, we can expand A and B in terms of fiducial preparations and results, analogous to Equations 6.48-6.49, which capture Tomographic Compression. Therefore we have:



(6.61)

The Equation above will help us identify $\Lambda^{l_1}_{\alpha_1}$, $\Lambda^{l_2}_{\alpha_2}$, $\mathbf{r}_{l_1}$ and $\mathbf{r}_{l_2}$ (analogous to Equation 6.51-6.53). Next, we would like to identify Compositional Compression through $\Lambda^{k_1,k_2}_{l_1,l_2}$. To do so we need to further specify the fragment G, to tell us if A and B are causally adjacent. In the case that A and B are not causally adjacent, there would be *no* Compositional Compression, such that $\Lambda^{k_1,k_2}_{l_1,l_2} = \delta^{k_1}_{l_1}\delta^{k_2}_{l_2}$ and thus $\mathbf{r}_{k_1k_2} = \mathbf{r}_{l_1l_2} = \mathbf{r}_{l_1} \otimes \mathbf{r}_{l_2}$.

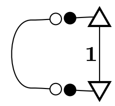
But we would like to see the case where we do have some non-trivial Compositional Compression. Therefore, let us consider that A and B are causally adjacent ($R_1 \bowtie R_2$), and redraw the circuit connecting the relevant wires between A and B:

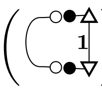


(6.62)

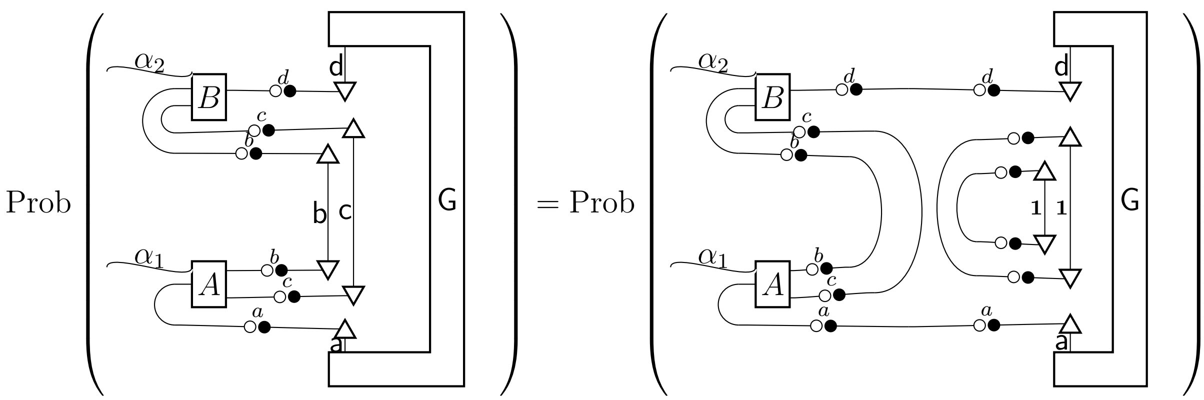
In order to identify the Lambda matrix $\Lambda^{k_1,k_2}_{l_1,l_2}$, we need to manipulate the expression in 6.62 through the use of what we call the *identity circuit*, which we define below.

Identity Circuit: denoted by I (capital I in sans serif font), is a special circuit, consisting of a single operation with one input and one output of the same type (**1**), associated with a single fiducial result or measurement ($K_{\mathbf{1}} = 1$), where the output connects back to the input.

Identity Circuit:  (6.63)

Such a circuit with an arbitrary system type would clearly violate one of DF's rules of wiring: the *no closed loop* condition, primarily since it may lead to inconsistencies, thereby not allowing probabilities to be defined. For the identity circuit the system type, denoted by **1** instead of a letter, has only a single fiducial result or preparation, thus giving $K_{\mathbf{1}} = 1$, thus the type **1** can be seen as the smallest non-null subset for any system type **s**. This renders the setting and outcome set of the operation empty since there will be no degrees of freedom. Therefore, we do not draw an actual operation. The system type **1** does not correspond to any real system rather it just *is*, a vestige of something that happened. The probability of the identity circuit is also well-defined: Prob() = 1. Therefore, we make the exception despite the *no closed loop* rule.

Let us now modify the expression in 6.62 by inserting an identity circuit, I, for *each* wire that connects the two regions R_1 and R_2, in order to help us identify $\Lambda_{\mathcal{O}_2}$ and $\mathbf{r}_{k_1,k_2}$. The sole purpose of the insertion of the identity circuit(s), I, is to make sure that the elements of the Omega set $\Omega_{1,2}$ can be compared to $\Omega_1 \times \Omega_2$.

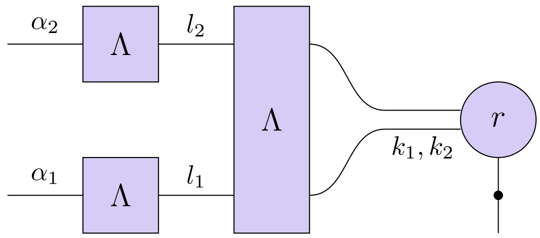


(6.64)

In DF's Equation 6.6 we saw that for circuits Q, R, ... we have Prob(QR...) = Prob(Q)Prob(R)....
Therefore, we justify Equation 6.64 as follows:

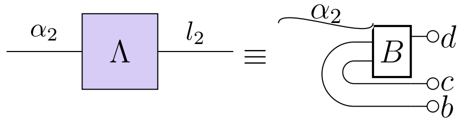
$$\mathsf{Prob(RII)} = \mathsf{Prob(R)Prob(I)Prob(I)} = \mathsf{Prob(R)} \quad \text{since} \quad \mathsf{Prob(I)} = 1 \tag{6.65}$$

Recall from Section 4.3.3, that the diagram from the Causaloid Framework with the two levels of compression for regions R_1 and R_2 is as follows:

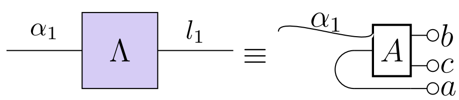


(6.66)

We are now ready to proceed with the identification process. Comparing the right hand side expression of 6.64 (DF) to CF's expression 6.66 we identify the Tomographic compression matrices (same as in Section 6.3.4),

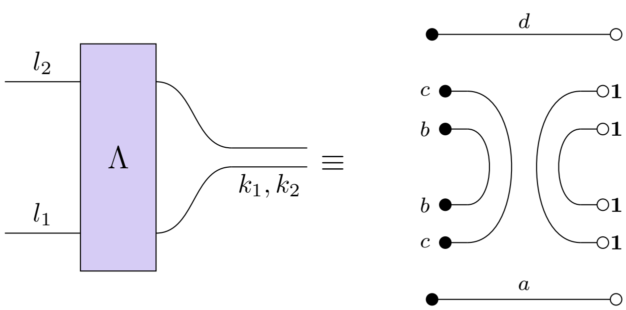


$$= \left({}^{d_2}B_{b_2 c_2} \right)_{\alpha_2} \tag{6.67}$$



$$= \left({}^{b_1 c_1}A_{a_1} \right)_{\alpha_1} \tag{6.68}$$

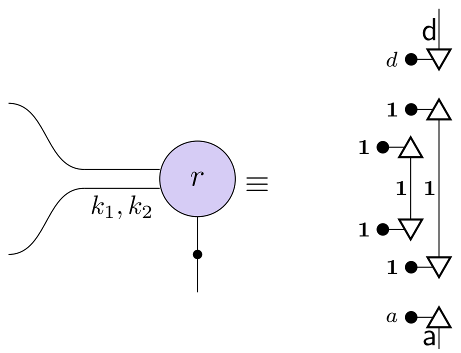
which are duotensors with white dots, and we identify the Compositional compression matrix as follows,



$$= {}^{d_2'}_{d_2}\delta \; {}_{b_1}g^{b_2} \; {}_{c_1}g^{c_2} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{a_1'}_{a_1}\delta \tag{6.69}$$

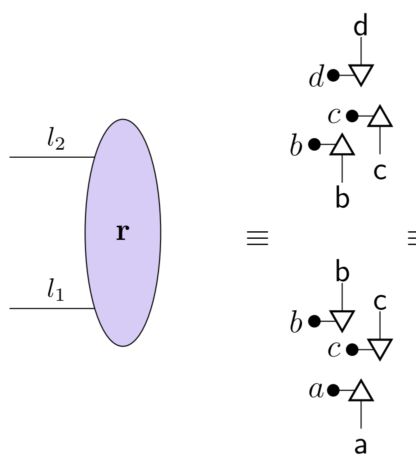
where we have also written the symbolic equivalents as well. The diagrammatics are definitely easier to help understand the identifications. The Compositional compression matrix $\Lambda^{k_1,k_2}_{l_1,l_2}$ is composed of the hopping metric duotensor •—• of the type(s) being compressed and the inverse ○—○ of type **1**, as well as identity duotensor •—○ for uncompressed types (defined in Equations 6.22, 6.23 and 6.24). The map from the tomographic indices l_1, l_2 to the compositional indices k_1, k_2 correspond to going from black to white dots in the DF diagram. This will be true for Compositional Lambda matrices in the Duotensor Framework.

We, lastly, identify the compositionally compressed measurement vector $\mathbf{r}_{k_1,k_2}$,



$$\equiv {}_{d_2'}\mathsf{X}^{\mathsf{d}_2'}\ {}_{\mathbf{1}}g^{\mathbf{1}}\ {}_{\mathbf{1}}g^{\mathbf{1}}\ \mathsf{X}^{\mathsf{a}_1'}_{a_1'} \tag{6.70}$$

which is composed of fiducial preparations and fiducial results as well as the hopping metric •—• with system type $\mathbf{1}$. Compare this to the tomographically compressed measurement vector $\mathbf{r}_{l_1,l_2}$ which we identify from Equation 6.61 (prior to Compositional Compression):



$$\equiv {}_{d_2}\mathsf{X}^{\mathsf{d}_2}\ \mathsf{X}^{\mathsf{c}_2}_{c_2}\ \mathsf{X}^{\mathsf{b}_2}_{b_2}\ {}_{b_1}\mathsf{X}^{\mathsf{b}_1}\ {}_{c_1}\mathsf{X}^{\mathsf{c}_1}\ \mathsf{X}^{\mathsf{a}_1}_{a_1} \tag{6.71}$$

which consists solely of fiducial results and fiducial preparations. Thus, we see that the hopping metric, as seen in Equation 6.69, is vital to Compositional Compression of adjacent regions and thus plays a role in determining the causal structure. In this respect, it is similar to the metric in General Relativity. This completes the identification process.

Let us now see the form of the Compositional Omega set $\Omega_{1,2}$ (for region $R_1 \cup R_2$) and compare it to the Tomographic Omega set $\Omega_1 \times \Omega_2$ (for regions R_1 and R_2) for our example circuit R. We also see how the factorisability of the Tomographic Omega set due to full decomposibility (Equation 6.55) factors into the Compositional Omega set.

Tomographic Omega sets for the regions R_1 and R_2 using full decomposability is given by

$$l_1 \in \Omega_1 \;\longrightarrow\; l^-_{\mathsf{a}}, l^+_{\mathsf{b}}, l^+_{\mathsf{c}} \in \Omega^-_{\mathsf{a}} \times \Omega^+_{\mathsf{b}} \times \Omega^+_{\mathsf{c}} \tag{6.72}$$

$$l_2 \in \Omega_2 \;\longrightarrow\; l^-_{\mathsf{b}}, l^-_{\mathsf{c}}, l^+_{\mathsf{d}} \in \Omega^-_{\mathsf{b}} \times \Omega^-_{\mathsf{c}} \times \Omega^+_{\mathsf{d}} \tag{6.73}$$

where $l^-_{\mathsf{a}} \in \Omega^-_{\mathsf{a}} = \{1, \ldots, K_{\mathsf{a}}\}$ and so on. The $-$ and $+$ denote if the set corresponds to inputs or outputs respectively. The Cartesian product of these two Omega sets gives:

$$l_1, l_2 \in \Omega_1 \times \Omega_2 \;\longrightarrow\; l^-_{\mathsf{a}}, l^+_{\mathsf{b}}, l^+_{\mathsf{c}}, l^-_{\mathsf{b}}, l^-_{\mathsf{c}}, l^+_{\mathsf{d}} \in \Omega^-_{\mathsf{a}} \times \Omega^+_{\mathsf{b}} \times \Omega^+_{\mathsf{c}} \times \Omega^-_{\mathsf{b}} \times \Omega^-_{\mathsf{c}} \times \Omega^+_{\mathsf{d}} \tag{6.74}$$

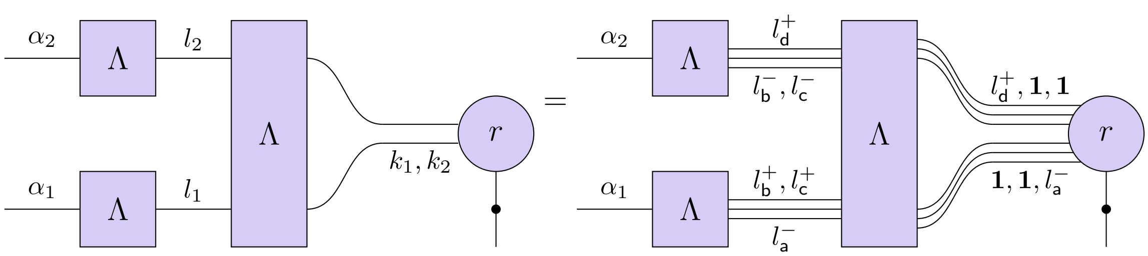
The Composition Omega set for the region $R_1 \cup R_2$ is given by,

$$k_1, k_2 \in \Omega_{1,2} \;\longrightarrow\; l^-_{\mathsf{a}}, \mathbf{1}^+, \mathbf{1}^+, \mathbf{1}^-, \mathbf{1}^-, l^+_{\mathsf{d}} \in \Omega^-_{\mathsf{a}} \times \{1\} \times \{1\} \times \{1\} \times \{1\} \times \Omega^+_{\mathsf{d}} \tag{6.75}$$

where $\{1\}$ is a set with single element corresponding to the system type $\mathbf{1}$ (from the identity circuit), allowing for the sets $\Omega_{1,2}, \Omega_1 \times \Omega_2$ to be comparable. Notice, that we continue to use the tomographic compression set indices for the inputs/outputs given that the system types are not affected by Compositional compression. More importantly notice that $\Omega_{1,2}$ also supports a factorised form. This clearly gives us the size of the Omega sets,

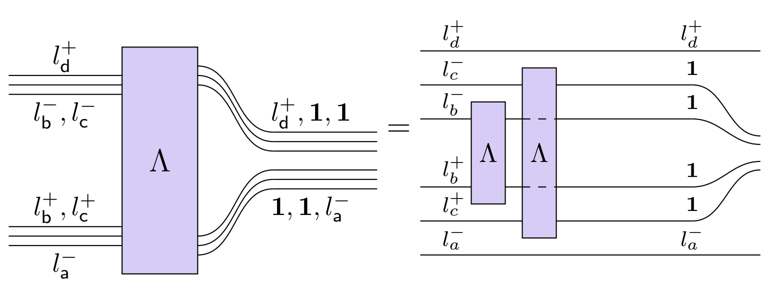
$$|\Omega_1 \times \Omega_2| = K_{\mathsf{a}} K^2_{\mathsf{b}} K^2_{\mathsf{c}} K_{\mathsf{d}} \quad , \quad |\Omega_{1,2}| = K_{\mathsf{a}} K_{\mathsf{d}} \tag{6.76}$$

The Causaloid diagram 6.66 can be redrawn to incorporate the factorised Omega sets



(6.77)

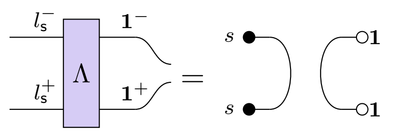
Further, the Lambda matrix $\Lambda^{k_1,k_2}_{l_1,l_2}$ itself can be factorised, diagrammatically we have,



(6.78)

symbolically we have $\quad \Lambda^{k_1,k_2}_{l_1,l_2} = \Lambda^{l^-_{\mathsf{a}}, \mathbf{1}^+, \mathbf{1}^+, \mathbf{1}^-, \mathbf{1}^-, l^+_{\mathsf{d}}}_{l^-_{\mathsf{a}}, l^+_{\mathsf{b}}, l^+_{\mathsf{c}}, l^-_{\mathsf{b}}, l^-_{\mathsf{c}}, l^+_{\mathsf{d}}} = \delta^{l^-_a}_{l^-_a} \Lambda^{\mathbf{1}^+, \mathbf{1}^-}_{l^+_{\mathsf{b}}, l^-_{\mathsf{b}}} \Lambda^{\mathbf{1}^+, \mathbf{1}^-}_{l^+_{\mathsf{c}}, l^-_{\mathsf{c}}} \delta^{l^+_d}_{l^+_d}$ (6.79) where the Lambda matrices of the form $\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l^+_{\mathsf{s}},l^-_{\mathsf{s}}}$ correspond to a wire of system type s, shared between the two regions R_1 (operation A) and R_2 (operation B), or equivalently correspond to the internal wiring of the fragment for the composite region $R_1 \cup R_2$ (fragment $\mathsf{D} = \mathsf{AB}$). The diagrammatic form for $\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l^+_{\mathsf{s}},l^-_{\mathsf{s}}}$ for the wire type s shared between any two regions is:

$$\left(\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l^+_{\mathsf{s}},l^-_{\mathsf{s}}} = \right.$$



$$\left. \right) , \quad \mathbf{1}^+, \mathbf{1}^- \in \{1, 1\} \ , \ l^+_{\mathsf{s}}, l^-_{\mathsf{s}} \in \Omega^+_{\mathsf{s}} \times \Omega^-_{\mathsf{s}} \tag{6.80}$$

We now generalise the form of $\Lambda_{\mathcal{O}_2}$ for any two elementary regions:

For a fragment $\mathsf{J} = \mathsf{KL}$ made of two general operations K and L (region $R_8 \cup R_9$) we will have the Compositional Compression matrix take the form:

$$\Lambda^{k_8,k_9}_{l_8,l_9} = \left(\prod_{\mathsf{i} \in \mathsf{w}^-(\mathsf{J})} \delta^{l^-_{\mathsf{i}}}_{l^-_{\mathsf{i}}} \right) \left(\prod_{\mathsf{o} \in \mathsf{w}^+(\mathsf{J})} \delta^{l^+_{\mathsf{o}}}_{l^+_{\mathsf{o}}} \right) \left(\prod_{\mathsf{s} \in \mathsf{w}(\mathsf{J})} \Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l^+_{\mathsf{s}},l^-_{\mathsf{s}}} \right) \tag{6.81}$$

The Omega sets take the form:

$$\Omega_1 \times \Omega_2 = \mathcal{O}\left(\left(\bigtimes_{\mathsf{i} \in \mathsf{w}^-(\mathsf{J})} \Omega^-_{\mathsf{i}} \right) \times \left(\bigtimes_{\mathsf{o} \in \mathsf{w}^+(\mathsf{J})} \Omega^+_{\mathsf{o}} \right) \times \left(\bigtimes_{\mathsf{s} \in \mathsf{w}(\mathsf{J})} \Omega^-_{\mathsf{s}} \times \Omega^+_{\mathsf{s}} \right) \right) \tag{6.82}$$

$$|\Omega_1 \times \Omega_2| = \left(\prod_{\mathsf{i} \in \mathsf{w}^-(\mathsf{J})} K_{\mathsf{i}} \right) \left(\prod_{\mathsf{o} \in \mathsf{w}^+(\mathsf{J})} K_{\mathsf{o}} \right) \left(\prod_{\mathsf{s} \in \mathsf{w}(\mathsf{J})} K^2_{\mathsf{s}} \right) \tag{6.83}$$

$$\Omega_{1,2} = \mathcal{O}\left(\left(\bigtimes_{\mathsf{i} \in \mathsf{w}^-(\mathsf{J})} \Omega^-_{\mathsf{i}} \right) \times \left(\bigtimes_{\mathsf{o} \in \mathsf{w}^+(\mathsf{J})} \Omega^+_{\mathsf{o}} \right) \times \left(\bigtimes_{\mathsf{s} \in \mathsf{w}(\mathsf{J})} \{\mathbf{1}^-\} \times \{\mathbf{1}^+\} \right) \right) \tag{6.84}$$

$$|\Omega_{1,2}| = \left(\prod_{\mathsf{i} \in \mathsf{w}^-(\mathsf{J})} K_{\mathsf{i}} \right) \left(\prod_{\mathsf{o} \in \mathsf{w}^+(\mathsf{J})} K_{\mathsf{o}} \right) \tag{6.85}$$

where in Equation 6.84, for every $\mathsf{s} \in \mathsf{w}(\mathsf{J})$ instead of its associated Omega set we put the set consisting of a single element $\{1\}$ (coming from the identification with the Identity Circuit).

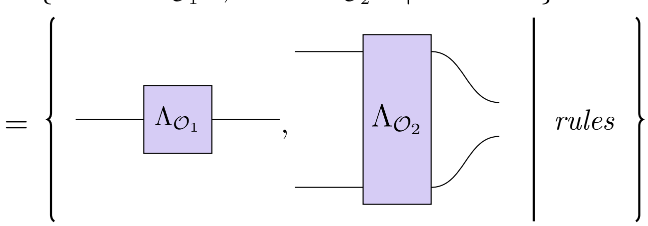
6.3.6 Level Three: Meta Compression

To complete the application of Causaloid Framework to the Duotensor Framework, we have to identify the last level of physical compression, Meta Compression, which specifies the Causaloid $\mathbf{\Lambda}$, the central mathematical object of the Causaloid Framework that can be seen as a specification of the physical theory itself.

Further, in Chapter 5, we introduced a hierarchy based on Meta Compression where $\Lambda_{\mathcal{O}_d}$-sufficiency provided the rungs of the hierarchy. In this Section we wish to strengthen our main claim of this Chapter, that the Duotensor Framework falls on the second rung of the Hierarchy and is $\Lambda_{\mathcal{O}_2}$-sufficient. We provide the outline of how we will go about showing this claim. This section is the culmination of the many concepts introduced in Chapters 4, 5 and the current Chapter.

Recall from Equation 5.5 that for $\Lambda_{\mathcal{O}_2}$-sufficiency the Causaloid $\mathbf{\Lambda}$ is given by

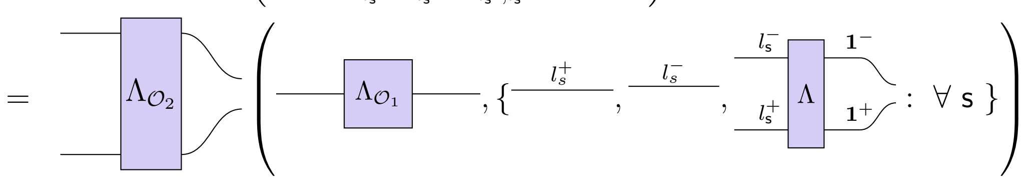
$$\mathbf{\Lambda} \equiv \mathbf{\Lambda}_{\mathcal{O}_d} = \{ \quad \text{all} \quad \Lambda_{\mathcal{O}_1} \quad , \quad \text{all} \quad \Lambda_{\mathcal{O}_2} \quad | \quad \textit{rules} \quad \} \tag{6.86}$$



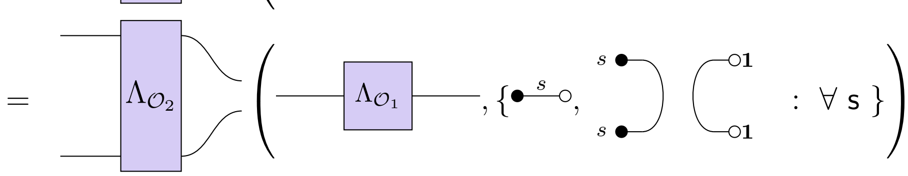
(6.87)

In the previous Section 6.3.5, we characterised the general form of $\Lambda_{\mathcal{O}_2}$ which may be written (symbolically and diagrammatically) as,

$$\Lambda_{\mathcal{O}_2}\left(\Lambda_{\mathcal{O}_1}, \{\delta^{l_{\mathsf{s}}^-}_{l_{\mathsf{s}}^-}, \delta^{l_{\mathsf{s}}^+}_{l_{\mathsf{s}}^+}, \Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{s}}^+,l_{\mathsf{s}}^-} \;:\; \forall\, \mathsf{s}\,\}\right) \text{ where } \mathsf{s} \text{ is any possible type} \tag{6.88}$$



(6.89)



(6.90)

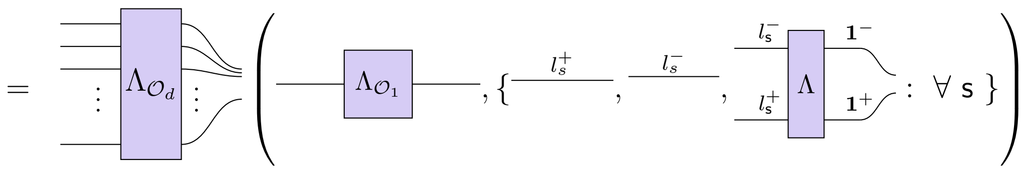
in terms of the identity duotensors and hopping metric (and inverse) duotensors (along with $\Lambda_{\mathcal{O}_1}$ in the case where Compositional Compression is done without prior Tomographic Compression).

Further, recall that $\Lambda_{\mathcal{O}_2}$-sufficiency is guaranteed if we can have the form of $\Lambda_{\mathcal{O}_d}$ be,

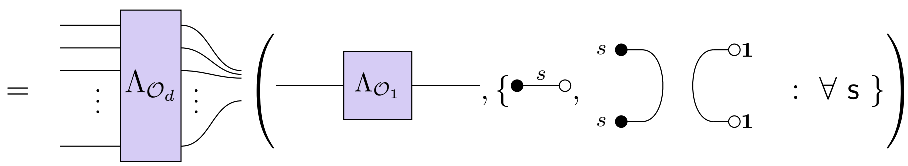
$$\Lambda_{\mathcal{O}_d}\left(\Lambda_{\mathcal{O}_1}, \Lambda_{\mathcal{O}_2}\right) \quad \text{for} \quad d > 2 \tag{6.91}$$

Therefore to show $\Lambda_{\mathcal{O}_2}$-sufficiency of the Duotensor Framework, combining the form of $\Lambda_{\mathcal{O}_2}$ (Equations 6.88, 6.89, 6.90) with the form of $\Lambda_{\mathcal{O}_d}$ from Equation 6.91, we need to show that the following is true,

$$\Lambda_{\mathcal{O}_d}\left(\Lambda_{\mathcal{O}_1}, \{\delta^{l_{\mathsf{s}}^-}_{l_{\mathsf{s}}^-}, \delta^{l_{\mathsf{s}}^+}_{l_{\mathsf{s}}^+}, \Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{s}}^+,l_{\mathsf{s}}^-} \; : \; \forall \; \mathsf{s} \; \}\right) \text{ where } \mathsf{s} \text{ is any possible type} \tag{6.92}$$



(6.93)

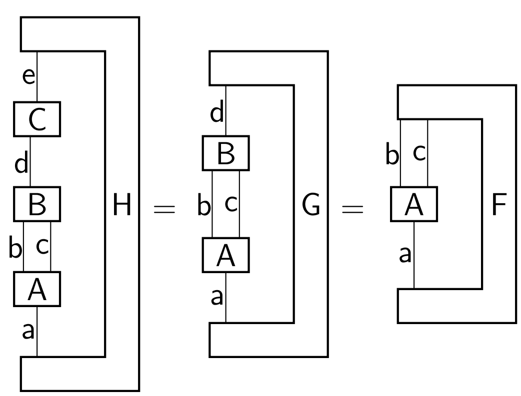


(6.94)

thus showing that any $\Lambda_{\mathcal{O}_d}$ can be written down in terms of the identity duotensors and hopping metric (and inverse) duotensors (along with $\Lambda_{\mathcal{O}_1}$).

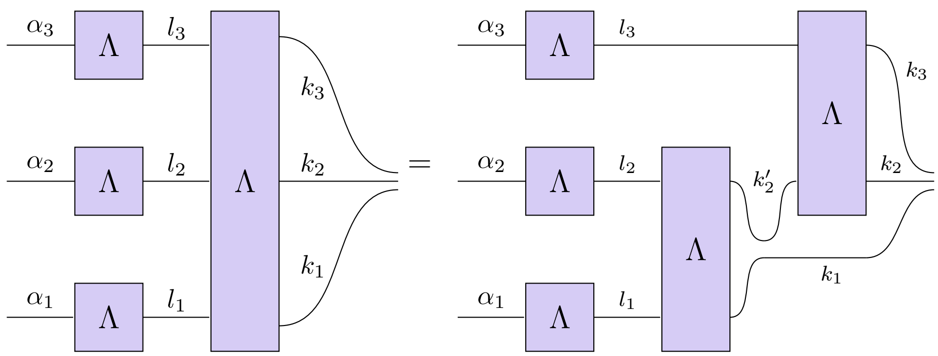
We will explicitly show that Equations 6.92, 6.93, 6.94 are true for $d = 3$ by considering a tripartite region with given causal structure in DF, going through the identification process of Lambda matrices (as shown in Sections 6.3.4 and 6.3.5) and using the results from Section 5.5. We then show that it will hold true for a general value of d by providing some simple arguments. Now that we have shared our goal and the means by which we will show this, let us get right to it.

Consider three operations A, B and C corresponding to three elementary regions within the circuit R in DF. This is the same circuit R as before, where $\mathsf{R} = \mathsf{ABCH} = \mathsf{ABG} = \mathsf{AF}$.



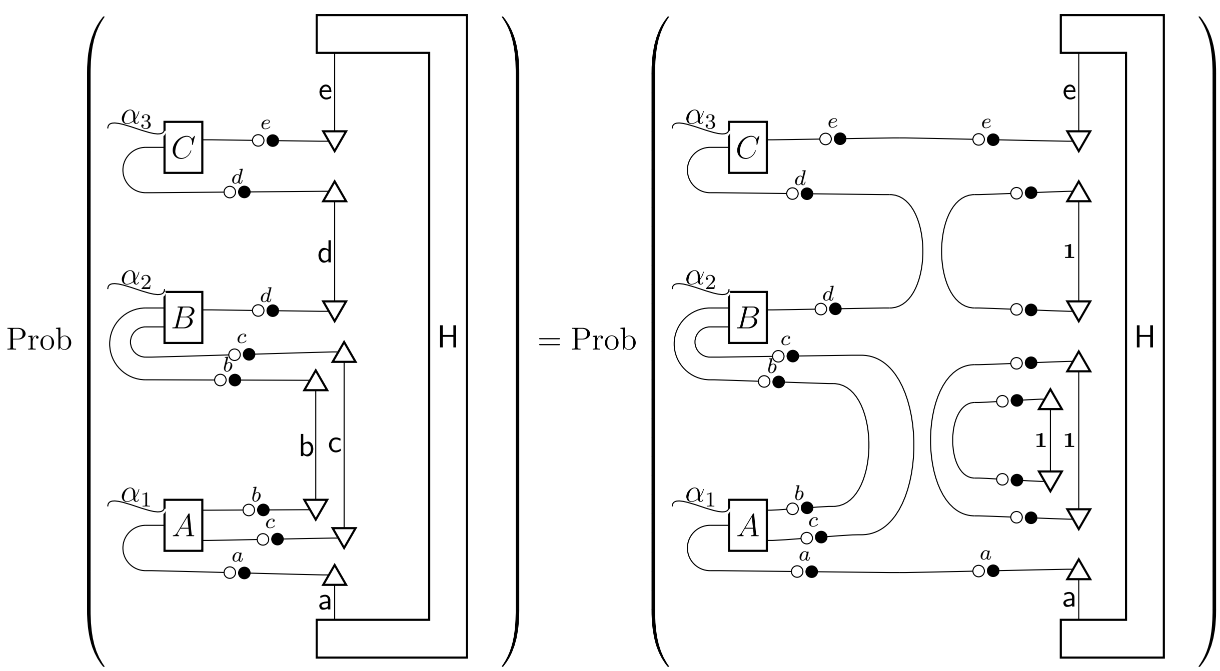
(6.95)

Here the causal relations between the operations (elementary regions) are: $\mathsf{A} \bowtie \mathsf{B}$ (or $R_1 \bowtie R_2$) and $\mathsf{B} \bowtie \mathsf{C}$ (or $R_2 \bowtie R_3$) which corresponds to Case 5 from Section 5.5.1, where we showed that if we have $\Lambda_{\mathcal{O}_2}$-sufficiency and the causal relations stated before then the following Causaloid equation for Lambda matrices holds true:



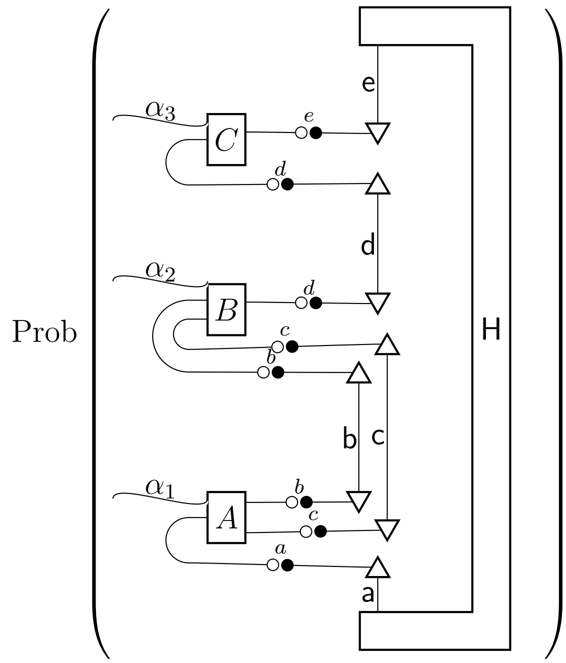
(6.96)

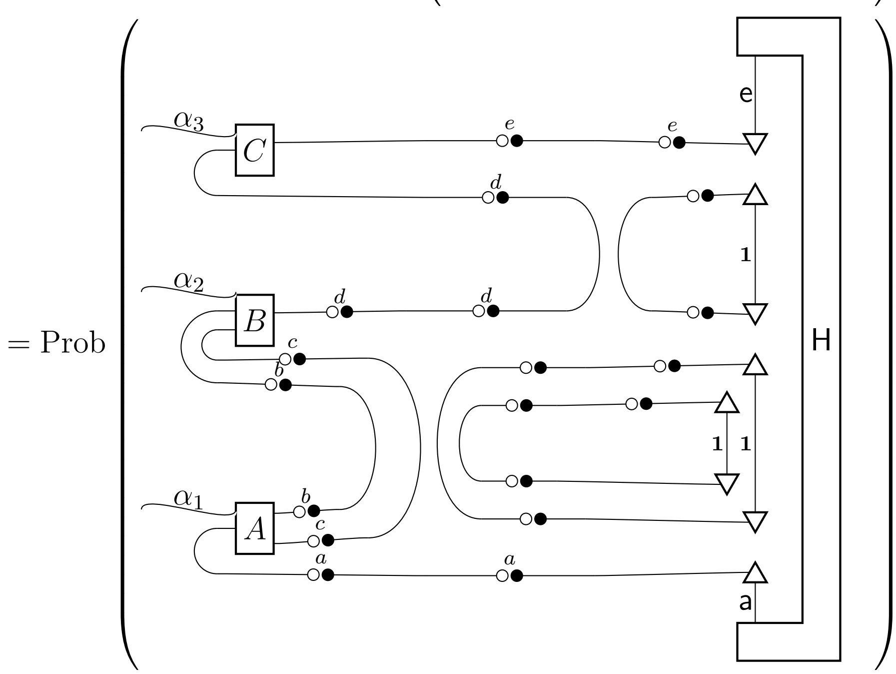
We will identify the Lambda matrices $\Lambda^{k_1,k_2,k_3}_{l_1,l_2,l_3}$, $\Lambda^{k_1,k_2'}_{l_1,l_2}$ and $\Lambda^{k_2,k_3}_{k_2',l_3}$ to check if Equation 6.96 (or equivalently $\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'} \Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2}$) holds for our DF example. To identify $\Lambda^{k_1,k_2,k_3}_{l_1,l_2,l_3}$, we do expansions of fiducial preparations and fiducial results of the circuit R and insertions of identity circuits, hopping metric (and its inverse) and identity duotensors (analogous to those seen in Section 6.3.5):



(6.97)

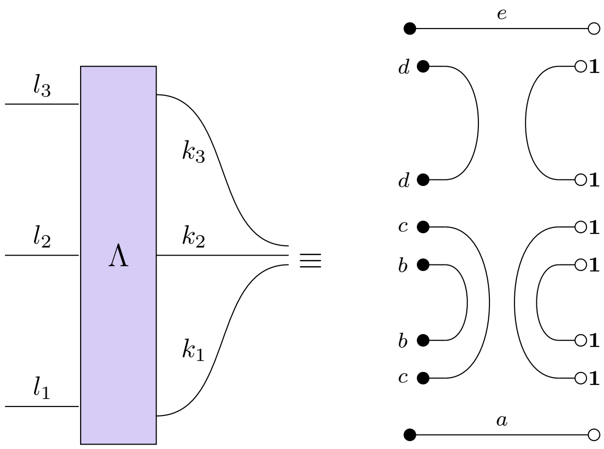
To identify $\Lambda^{k_1,k_2'}_{l_1,l_2}$ and $\Lambda^{k_2,k_3}_{k_2',l_3}$, we do expansions of fiducial preparations and fiducial results and insertions of identity circuits, hopping metric (and its inverse) as before with some additional identity duotensors:





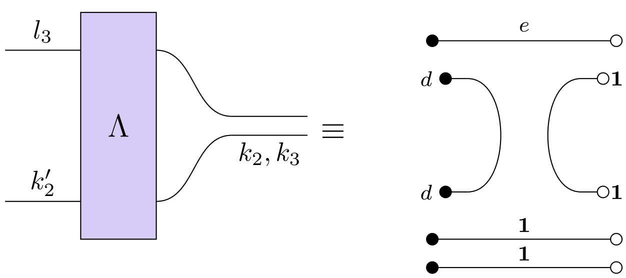
(6.98)

We can identify $\Lambda^{k_1,k_2,k_3}_{l_1,l_2,l_3}$ from the right hand side expression of 6.97,



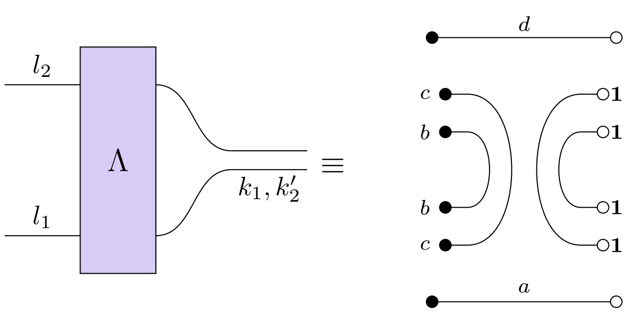
$$= {}^{e_3'}_{e_3}\delta \; {}_{d_2}g^{d_3} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}_{b_1}g^{b_2} \; {}_{c_1}g^{c_2} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{a_1'}_{a_1}\delta \quad (6.99)$$

We can identify $\Lambda^{k_2,k_3}_{k_2',l_3}$ from the right hand side expression of 6.98,



$$= {}^{e_3'}_{e_3}\delta \; {}_{d_2'}g^{d_3} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}_{\mathbf{1}}\delta \; {}^{\mathbf{1}}_{\mathbf{1}}\delta \quad (6.100)$$

Similarly we can identify $\Lambda^{k_1,k_2'}_{l_1,l_2}$ from the right hand side expression of 6.98,



$$= {}^{d_2'}_{d_2}\delta \; {}_{b_1}g^{b_2} \; {}_{c_1}g^{c_2} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{a_1'}_{a_1}\delta \quad (6.101)$$

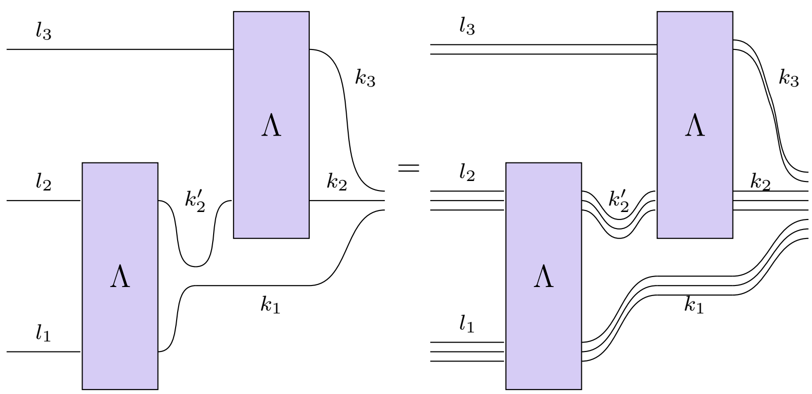
From the DF structure of these three Lambda matrices, it is clear diagrammatically that $\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'} \Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2}$, or symbollically as well,

$${}^{e_3'}_{e_3}\delta \; {}_{d_2}g^{d_3} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}_{b_1}g^{b_2} \; {}_{c_1}g^{c_2} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{a_1'}_{a_1}\delta = ({}^{d_2'}_{d_2}\delta \; {}_{b_1}g^{b_2} \; {}_{c_1}g^{c_2} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{a_1'}_{a_1}\delta)({}^{e_3'}_{e_3}\delta \; {}_{d_2'}g^{d_3} \; {}^{\mathbf{1}}g_{\mathbf{1}} \; {}^{\mathbf{1}}_{\mathbf{1}}\delta \; {}^{\mathbf{1}}_{\mathbf{1}}\delta) \quad (6.102)$$

where we have used the symbolic form provided in Equations 6.99,6.100 and 6.101. Further the Tomographic Compression matrices $\Lambda^{l_1}_{\alpha_1}, \Lambda^{l_2}_{\alpha_2}$ and $\Lambda^{l_3}_{\alpha_3}$ will be the same in both these cases. Therefore Equation 6.96 is satisfied.

Thus, we have shown that for this example the Lambda matrix for three regions can be written in terms of Lambda matrices for two regions. Now we will show something stronger, that the Lambda matrix for three regions can be written through hopping metric, its inverse and identity duotensors. Let us go through the following elaborate steps, towards this goal.

We have seen that the Tomographic as well as Compositional Compression Omega sets factorise, such as in Equation 6.77. The Causaloid diagram 6.96 can be redrawn to incorporate the factorised Omega sets,



(6.103)

where the Omega sets are of the form,

$$\begin{aligned} & l_1, l_2, l_3 \in \Omega_1 \times \Omega_2 \times \Omega_3 \\ =& l^-_{\mathsf{a}}, l^+_{\mathsf{b}}, l^+_{\mathsf{c}}, l^-_{\mathsf{b}}, l^-_{\mathsf{c}}, l^+_{\mathsf{d}}, l^-_{\mathsf{d}}, l^+_{\mathsf{e}} \in \Omega^-_{\mathsf{a}} \times \Omega^+_{\mathsf{b}} \times \Omega^+_{\mathsf{c}} \times \Omega^-_{\mathsf{b}} \times \Omega^-_{\mathsf{c}} \times \Omega^+_{\mathsf{d}} \times \Omega^-_{\mathsf{d}} \times \Omega^+_{\mathsf{e}} \end{aligned} \tag{6.104}$$

$$\begin{aligned} & k_1, k_2' \in \Omega_{1,2} \\ =& l^-_{\mathsf{a}}, \mathbf{1}^+, \mathbf{1}^+, \mathbf{1}^-, \mathbf{1}^-, l^+_{\mathsf{d}} \in \Omega^-_{\mathsf{a}} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \Omega^+_{\mathsf{d}} \end{aligned} \tag{6.105}$$

$$\begin{aligned} & k_2, k_3 \in (\Omega_{\not{1},2} \times \Omega_3) \cap \Omega_{2,3} \\ =& \mathbf{1}^-, \mathbf{1}^-, \mathbf{1}^+, \mathbf{1}^-, l^+_{\mathsf{e}} \in \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \Omega^+_{\mathsf{e}} \end{aligned} \tag{6.106}$$

$$\begin{aligned} & k_1, k_2, k_3 \in \Omega_{1,2,3} \\ =& l^-_{\mathsf{a}}, \mathbf{1}^+, \mathbf{1}^+, \mathbf{1}^-, \mathbf{1}^-, \mathbf{1}^+, \mathbf{1}^-, l^+_{\mathsf{e}} \in \Omega^-_{\mathsf{a}} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \{\mathbf{1}\} \times \Omega^+_{\mathsf{e}} \end{aligned} \tag{6.107}$$

Here we have used Equation 5.57 in Equation 6.106. We can solve Equation 6.103 further, by substituting the factorised forms for the Lambda matrices:

$$\Lambda^{k_1,k_2'}_{l_1,l_2} = \Lambda^{l_{\mathsf{a}}^-,\mathbf{1}^+,\mathbf{1}^+,\mathbf{1}^-,\mathbf{1}^-,l_{\mathsf{d}}^+}_{l_{\mathsf{a}}^-,l_{\mathsf{b}}^+,l_{\mathsf{c}}^+,l_{\mathsf{b}}^-,l_{\mathsf{c}}^-,l_{\mathsf{d}}^+} = \delta^{l_a^-}_{l_a^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{b}}^+,l_{\mathsf{b}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{c}}^+,l_{\mathsf{c}}^-}\delta^{l_d^+}_{l_d^+} \tag{6.108}$$

$$\Lambda^{k_2,k_3}_{k_2',l_3} = \Lambda^{\mathbf{1}^-,\mathbf{1}^-,\mathbf{1}^+,\mathbf{1}^-,l_{\mathsf{e}}^+}_{\mathbf{1}^-,\mathbf{1}^-,l_{\mathsf{d}}^+,l_{\mathsf{d}}^-,l_{\mathsf{e}}^+} = \delta^{\mathbf{1}^-}_{\mathbf{1}^-}\delta^{\mathbf{1}^-}_{\mathbf{1}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{d}}^+,l_{\mathsf{d}}^-}\delta^{l_e^+}_{l_e^+} \tag{6.109}$$

Substituting 6.108 and 6.109 into Equation 6.103, and upon simplifying we get:

$$\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'}\Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2} \tag{6.110}$$

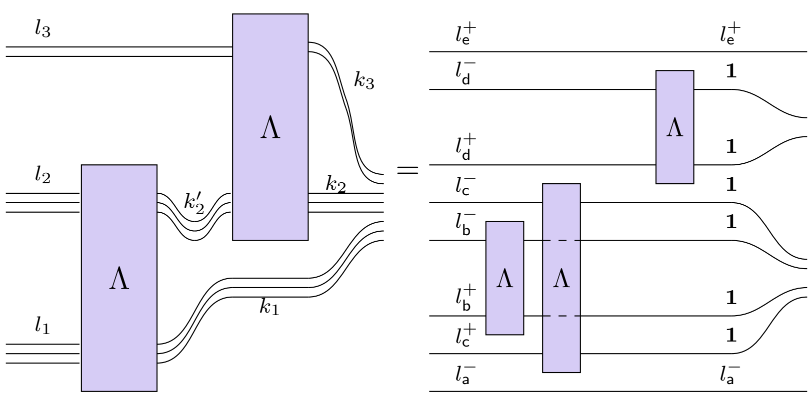
$$= \sum_{k_2'\in\{1\}\times\{1\}\times\Omega_{\mathsf{d}}^+}\Lambda^{\mathbf{1}^-,\mathbf{1}^-,\mathbf{1}^+,\mathbf{1}^-,l_{\mathsf{e}}^+}_{\mathbf{1}^-,\mathbf{1}^-,l_{\mathsf{d}}^+,l_{\mathsf{d}}^-,l_{\mathsf{e}}^+}\Lambda^{l_{\mathsf{a}}^-,\mathbf{1}^+,\mathbf{1}^+,\mathbf{1}^-,\mathbf{1}^-,l_{\mathsf{d}}^+}_{l_{\mathsf{a}}^-,l_{\mathsf{b}}^+,l_{\mathsf{c}}^+,l_{\mathsf{b}}^-,l_{\mathsf{c}}^-,l_{\mathsf{d}}^+} \tag{6.111}$$

$$= \sum_{k_2'\in\{1\}\times\{1\}\times\Omega_{\mathsf{d}}^+}(\delta^{\mathbf{1}^-}_{\mathbf{1}^-}\delta^{\mathbf{1}^-}_{\mathbf{1}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{d}}^+,l_{\mathsf{d}}^-}\delta^{l_e^+}_{l_e^+})(\delta^{l_a^-}_{l_a^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{b}}^+,l_{\mathsf{b}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{c}}^+,l_{\mathsf{c}}^-}\delta^{l_d^+}_{l_d^+}) \tag{6.112}$$

$$= \delta^{l_{\mathsf{a}}^-}_{l_{\mathsf{a}}^-}\left(\sum_{\{1\}\times\{1\}}\delta^{\mathbf{1}^-}_{\mathbf{1}^-}\delta^{\mathbf{1}^-}_{\mathbf{1}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{b}}^+,l_{\mathsf{b}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{c}}^+,l_{\mathsf{c}}^-}\right)\left(\sum_{\Omega_{\mathsf{d}}^+}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{d}}^+,l_{\mathsf{d}}^-}\delta^{l_{\mathsf{d}}^+}_{l_{\mathsf{d}}^+}\right)\delta^{l_{\mathsf{e}}^+}_{l_{\mathsf{e}}^+} \tag{6.113}$$

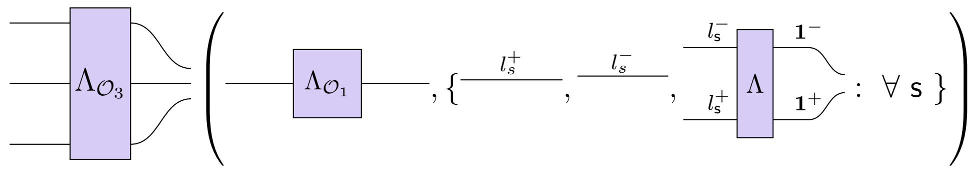
$$\Rightarrow \Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \delta^{l_{\mathsf{a}}^-}_{l_{\mathsf{a}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{b}}^+,l_{\mathsf{b}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{c}}^+,l_{\mathsf{c}}^-}\Lambda^{\mathbf{1}^+,\mathbf{1}^-}_{l_{\mathsf{d}}^+,l_{\mathsf{d}}^-}\delta^{l_{\mathsf{e}}^+}_{l_{\mathsf{e}}^+} \tag{6.114}$$

This can be shown diagrammatically as follows,



(6.115)

where dotted wires signify that the wire goes over (and not through) the Lambda matrix. Notice that the sum in the equation $\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'}\Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2}$ which holds true in CF for the given causal relations, vanishes in the simplified form of DF. In fact this is true irrespective of the causal relations, in DF the sum will always vanish. This is an important result, because we are able to show that $\Lambda^{k_1,k_2,k_3}_{l_1,l_2,l_3}$ satisfies the form:

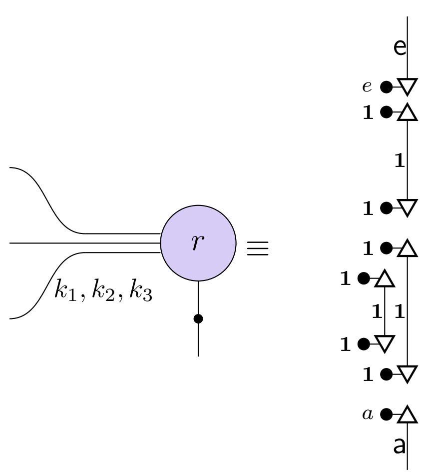


(6.116)

(6.117)

An Observation At this point, after the heavy lifting we have done, a simple, almost trivial sounding observation, needs to be made. A wire, by definition of the operational description of the Duotensor Framework, can only connect at most two disjoint regions. The wires are also at the heart of Compositional Compression (as seen in this Section). Therefore, any $\Lambda_{\mathcal{O}_d}$ for $d > 2$ will be expressible using the two region DF elements (the hopping metric, its inverse and identity duotensors). Thus, the Duotensor Framework is $\Lambda_{\mathcal{O}_2}$-sufficient, belonging to the second rung of the hierarchy (as defined in Chapter 5), and thereby finite-dimensional Quantum theory and Classical Probability theory are $\Lambda_{\mathcal{O}_2}$-sufficient as well.

To complete the identification process we identify $\mathbf{r}_{k_1,k_2,k_3}$. Since $\Lambda^{k_1k_2k_3}_{l_1l_2l_3} = \sum_{k_2'} \Lambda^{k_2k_3}_{k_2'l_3}\Lambda^{k_1k_2'}_{l_1l_2}$ we have $(j_1, j_2, j_3 \equiv k_1, k_2, k_3) \in \Omega_{1,2,3}$ where strictly speaking, indices j_1, j_2, j_3 belong to the $\Lambda_{\mathcal{O}_3}$ and indices k_1, k_2, k_3 belong to $\Lambda_{\mathcal{O}_2}$. Therefore, for both the diagrams (the one with $\Lambda_{\mathcal{O}_3}$ and the one with $\Lambda_{\mathcal{O}_2}$) the same identification of measurement vector $\mathbf{r}_{k_1,k_2,k_3}$ holds, given below,



$$\equiv {}_{e_3'}\mathsf{X}^{\mathsf{e}_3'}\ {}_{\mathbf{1}}g^{\mathbf{1}}\ {}_{\mathbf{1}}g^{\mathbf{1}}\ {}_{\mathbf{1}}g^{\mathbf{1}}\ \mathsf{X}^{\mathsf{a}_1'}_{a_1'} \tag{6.118}$$

6.4 Populating Higher Rungs

The Causaloid Framework contains the space of generalised probabilistic theories including those with indefinite causal structure. As a means to characterise this space we introduced the Hierarchy in Chapter 5 which places theories on different rungs based on Meta-Compression. The Duotensor Framework lives on the second rung of the hierarchy. We would like to explore the space of generalised probabilistic theories that live on higher rungs, in order to explore causal indefiniteness. In this Section, we propose the research program of constructing theories formulated in an operational and f-local manner that would populate the higher rungs, and what we may learn from them. We provide some ideas of how one would go about such a construction.

6.4.1 Hyper$_d$wired Circuits

The operations in the Duotensor Framework can be represented through nodes of *Directed Acyclic Graphs*, or DAGs in short, and the wires would serve as directed edges connecting the nodes. DAGs are particularly useful in studies of *causality* as they can be used to represent the direction of time when closed time loops are not allowed (thus the Acyclicness). This connects back to the condition on wires in DF such that if one starts from an operation and follows the wires, they do not come back to the same operation. DAGs have been used to draw graphs for causal connections between regions, such as in the gedanken EPR paradox experiment. But it also means that Causal DAG pictures come with specific assumptions on notions of time and causality which is restrictive for the purposes of our pursuit to study indefinite causal structures – assumptions that the Causaloid framework does not make. This observation is captured by Jia:

> "...although many causal frameworks are based on DAGs, there are reasonable frameworks that are naturally associated with hypergraphs rather than graphs, such as Hardy's causaloid framework of indefinite causal structure" (Jia, [65])

The Hierarchy comes into play here as it distinguishes causal frameworks based on *fixed* DAGs (which fall on a single rung) from other causal frameworks, such as those based on hyper-edges. At the heart of DAGs and the Duotensor Framework are edges or wires. Therefore, we would like to construct theories with operations composed with hyper-edges. We propose that one may attempt to do so by generalising the Duotensor Framework, which is formulated operationally and in a f-local manner. We would keep the assumptions of well conditioned probabilities and full decomposability intact but modify the structure of operations and wires. Instead of operations with inputs and outputs, we propose operations to have d kinds of connections (d-puts). Operations could then be connected to other operations through hyper$_d$wires (or hyper-edges with d nodes). These generalisations would be constructed such that they continue to be operational and formalism local, and thus be amenable to the application of the Causaloid Framework. The modified structure for operations and hyper$_d$wires will be captured through Compositional Compression and in turn Meta Compression affecting their place in the hierarchy, such that theories with hyper$_d$wires for some d would be constructed to be $\Lambda_{\mathcal{O}_d}$-sufficient.

We do not expect the entire space of such constructions to be physical, belonging to the world. One may ask what would the construction of them then achieve? The study of Generalised Probability Theories led to reconstruction of quantum theory from few information theoretic postulates, that could discern between quantum theory and classical probabilistic theory, which shed new insights on our understanding of quantum theory. Similarly, we expect the study of theories constructed with hyper-edges would shed some light on indefinite causality which may become relevant for the regime of quantum gravity.

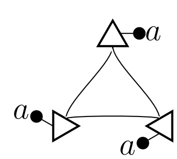
We present some ideas for the construction for $d = 3$ which we call the *triotensors*, followed by some ideas for the case $d = 4$ which we will call *quadrotensors*.

Triotensors

The triotensors would correspond to having hyper-edges that connect three operations. Instead of outputs and inputs, there would be three kinds of "puts", and instead of fiducial preparations and results, there would be three kinds of fiducial sets, each associated with two kinds of basis (corresponding to the black and white dots), therefore they would have six indices. Thus the name triotensors.

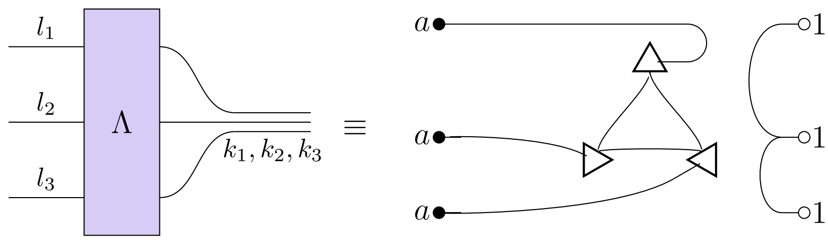
Since inputs and outputs corresponded to apertures through which systems came in and out, a sense of causality was established through them. With three kinds of "puts" the notion of causality will have to be interpreted differently, and there may be a few possibilities: one may conservatively attribute the three "puts" to a past and two future directions (or two past and one future direction), or allow for a completely different kind of causality.

One important step would be to define the generalisation of the hopping metric, here we show a possibility of representing the hopping metric with all black dots



(6.119)

The Compositional Compression for three regions would provide compression through the hopping metric, similar to what we saw in the duotensors, while two region Compositional Compression would essentially yield no compression. The Lambda matrix would be constructed as follows



(6.120)

where the associated Lambda matrix and Omega set would be,

$$\Lambda^{k_1,k_2,k_3}_{l_1,l_2,l_3} = \Lambda^{\mathbf{1},\mathbf{1},\mathbf{1}}_{l_1,l_2,l_3} \tag{6.121}$$

$$\Omega_{1,2,3} = \{\mathbf{1}\},\{\mathbf{1}\},\{\mathbf{1}\}\ ,\ \Omega_1 = \Omega_2 = \Omega_3 = \Omega_a \tag{6.122}$$

When the Duotensor Framework is adapted to Quantum theory, an operation is replaced by an operator, where each set of fiducial preparations and results is associated with the space of Hermitian operators on the complex Hilbert space (whose dimension is dependent on the system type). It seems that a quantum version of the triotensors may run into some mathematical issues, since if we associate some vector space with each of the three fiducial sets, there may be difficulty finding a map that takes three vector spaces to give a real positive number, though such a map would be required since we would like to calculate probabilities. However, a generalisation of classical probability theory through the triotensors would not be affected by this issue and would thus be possible.

Quadrotensor

The quadrotensors, would correspond to having hyper-edges that connect four operations. Instead of outputs and inputs, there would be four kinds of "puts", and instead of fiducial preparations and results, there would be four kinds of fiducial sets, each associated with two kinds of basis (corresponding to the black and white dots), therefore they would have eight indices. Thus the name quadrotensors.

With four kinds of "puts" the quadrotensors could possibly be interpreted/constructed to have two past and two future directions (that is two independent directions of time). This could be quite interesting to study as one could model a definite causal structure on two arrows of time that would emerge as indefinite causality on a single arrow of time.

The mathematical obstacles towards a quantum generalisation for the triotensors would go away when attempting a quantum generalisation of qudrotensors, since one may be able to choose vectors spaces to associate with the four fiducial sets, such that a map on four vector spaces yields a positive real number (since we are interested in probabilities). A natural choice for such vector spaces would be the quaternions, that extend from complex numbers.

6.4.2 Beyond the Hierarchy

There may indeed be theories that do not fit into the hierarchy, here we discuss open questions that we would like to address or tackle in the future.

In this Chapter we discussed applying the Causaloid Framework to existing theories, such as Quantum theory, though we did not talk about the other obvious candidate – General Relativity (GR). Since GR exhibits dynamic causal structures, it would an interesting project to attempt application of the Causaloid Framework to GR. For this an operational and f-local formulation of GR is required. Hardy presented such a formulation called Operational General Relativity in [50]. Thus, a possible future direction would be to apply the Causaloid Framework to Operational General Relativity, and study GR's Compositional Compression and Meta Compression. Given that Compositional Compression in the Duotensor Framework was expressed in terms of the hopping metric and inverse, one wonders if we would see some similar appearance of the metric tensor while studying Compositional Compression in GR.

The other obvious direction to take would be to characterise existing frameworks, such as the Process Matrix Framework, Quantum Combs, Processes (the category theory kind), that study causal indefiniteness and ask if the hierarchy is a meaningful characterisation for them and if not, then what would be a good way to study how they relate to the space of GPTs with indefinite causality.

Chapter 6: Statement of Contribution

In this Chapter, the main contribution (Section 6.3,6.4) includes applying the Causaloid Framework (Chapter 4) to other operational and formalism-local frameworks such as the Duotensor Framework by Lucien Hardy [46]. We find that the Duotensor Framework populates the second rung of the Hierarchy presented in Chapter 5. The project was initiated by Lucien Hardy and the work was done together. The Chapter is solely written by me and contains unpublished work that was discussed in its nascent stages at various conferences (QISS HK 2020, Q-Turn 2020 [90], APS 2021 [91], Quantizing Time 2021 [89], QPL 2021) as the work developed.

Chapter 7

Conclusion

In this thesis, we began by considering bidirectional communication tasks and provided protocols for bidirectional quantum teleportation and bidirectional dense-coding. An open problem is to extend the quantum communication resource theory to multi-party scenarios and to find if Bidirectional Quantum Teleportation and Bidirectional Dense-Coding can be dual through such a resource theory. Further, we considered tensor products of processes to provide a theorem for when such tensor products are (in)valid.

In the second half, we considered an operational approach to study indefinite causality in the space of GPTs, by revisiting the Causaloid Framework and providing a diagrammatic representation. Further, we studied Meta-Compression, the third level of physical compression in the Causaloid Framework, and were able to characterise a hierarchy of theories pertaining to Causal Structure. Further, we applied the Causaloid Framework to the Duotensor Framework, thus showing the applicability of the Causaloid Framework, and we showed that the Duotensor Framework belongs to the second rung of our Hierarchy. Where would General Relativity fall into this hierarchy, as well as other frameworks for indefinite causality? We wish to study this in future work.

Many open questions emerge upon the introduction of this hierarchy, and we hope that the indefinite causal structure community may join us in revisiting the Causaloid Framework to explore open problems. We invite the interested to help us work towards these problems.

I would like to thank the reader for their time, their interest and for making it to the end of this thesis. We hope you may take with you a few insights and ideas.

References

[1] Rudolf Ahlswede, Ning Cai, S-YR Li, and Raymond W Yeung. Network information flow. *IEEE Transactions on information theory*, 46(4):1204–1216, 2000.

[2] Mateus Araújo, Cyril Branciard, Fabio Costa, Adrien Feix, Christina Giarmatzi, and Časlav Brukner. Witnessing causal nonseparability. *New Journal of Physics*, 17(10):102001, 2015.

[3] Mateus Araújo, Fabio Costa, and Časlav Brukner. Computational advantage from quantum-controlled ordering of gates. *Physical Review Letters*, 113(25):250402, 2014.

[4] Richard Arnowitt, Stanley Deser, and Charles W Misner. Gravitation: an introduction to current research. 1962.

[5] Abhay Ashtekar. New variables for classical and quantum gravity. *Physical review letters*, 57(18):2244, 1986.

[6] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce. Teleportation in general probabilistic theories. In *Proceedings of Symposia in Applied Mathematics*, volume 71, pages 25–48. American Mathematical Society Providence, RI, 2012.

[7] Ämin Baumeler and Stefan Wolf. Device-independent test of causal order and relations to fixed-points. *New Journal of Physics*, 18(3):035014, 2016.

[8] Jessica Bavaresco, Mateus Araújo, Časlav Brukner, and Marco Túlio Quintino. Semi-device-independent certification of indefinite causal order. *Quantum*, 3:176, 2019.

[9] John S Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[10] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical review letters*, 70(13):1895, 1993.

[11] Charles H Bennett and Stephen J Wiesner. Communication via one-and two-particle operators on einstein-podolsky-rosen states. *Physical review letters*, 69(20):2881, 1992.

[12] Gilles Brassard. Teleportation as a quantum computation. *arXiv preprint quant-ph/9605035*, 1996.

[13] Časlav Brukner. Quantum causality. *Nature Physics*, 10(4):259–263, 2014.

[14] Časlav Brukner. Private communication, 2017.

[15] Marcello Caleffi and Angela Sara Cacciapuoti. Quantum switch for the quantum internet: Noiseless communications through noisy channels. *IEEE Journal on Selected Areas in Communications*, 38(3):575–588, 2020.

[16] Carlos Cardoso-Isidoro and Francisco Delgado. Featuring causal order in teleportation with two quantum teleportation channels. In *Journal of Physics: Conference Series*, volume 1540, page 012024. IOP Publishing, 2020.

[17] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Physical Review A*, 86(4):040301, 2012.

[18] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009.

[19] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Informational derivation of quantum theory. *Physical Review A*, 84(1):012311, 2011.

[20] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2):022318, 2013.

[21] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2):022318, August 2013.

[22] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear algebra and its applications*, 10(3):285–290, 1975.

[23] Boris S Cirel'son. Quantum generalizations of bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.

[24] John F Clauser and Abner Shimony. Bell's theorem. experimental tests and implications. *Reports on Progress in Physics*, 41(12):1881, 1978.

[25] Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201, 1997.

[26] Richard Cleve, Peter Hoyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 236–249. IEEE, 2004.

[27] Fabio Costa. Private communication, 2017.

[28] Borivoje Dakić and Časlav Brukner. *Quantum Theory and Beyond: Is Entanglement Special?*, page 365–392. Cambridge University Press, 2011.

[29] Rafael de la Madrid Modino. *Quantum Mechanics in rigged Hilbert space language.* PhD thesis, Ph. D. thesis, Universidad de Valladolid, 2001.

[30] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777, 1935.

[31] Adrien Feix, Mateus Araújo, and Časlav Brukner. Quantum superposition of the order of parties as a communication resource. *Physical Review A*, 92(5):052326, 2015.

[32] Richard Phillips Feynman. Space-time approach to non-relativistic quantum mechanics. In *Feynman's Thesis—A New Approach To Quantum Theory*, pages 71–109. World Scientific, 2005.

[33] Hong-Zi Fu, Xiu-Lao Tian, and Yang Hu. A general method of selecting quantum channel for bidirectional quantum teleportation. *International Journal of Theoretical Physics*, 53(6):1840–1847, 2014.

[34] Christopher A Fuchs. Quantum mechanics as quantum information (and only a little more). *arXiv preprint quant-ph/0205039*, 2002.

[35] Penrose Tiling Generator. https://misc.0o0o.org/penrose/.

[36] Philippe Allard Guérin, Adrien Feix, Mateus Araújo, and Časlav Brukner. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Physical review letters*, 117(10):100502, 2016.

[37] Philippe Allard Guérin, Marius Krumm, Costantino Budroni, and Časlav Brukner. Composition rules for quantum processes: a no-go theorem. *New Journal of Physics*, 21(1):012001, 2019.

[38] Gus Gutoski and John Watrous. Toward a general theory of quantum games. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 565–574. ACM, 2007.

[39] Jiu-Cang Hao, Chuan-Feng Li, and Guang-Can Guo. Controlled dense coding using the greenberger-horne-zeilinger state. *Physical Review A*, 63(5):054301, 2001.

[40] Lucien Hardy. Quantum theory from five reasonable axioms. *arXiv preprint quant-ph/0101012*, 2001.

[41] Lucien Hardy. Probability theories with dynamic causal structure: a new framework for quantum gravity. *arXiv preprint gr-qc/0509120*, 2005.

[42] 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.

[43] Lucien Hardy. Operational structures as a foundation for probabilistic theories, June 2009. PIRSA:09060015 see, https://pirsa.org.

[44] Lucien Hardy. Quantum gravity computers: On the theory of computation with indefinite causal structure. In *Quantum reality, relativistic causality, and closing the epistemic circle*, pages 379–401. Springer, 2009.

[45] Lucien Hardy. The duotenzor drawing package. *See http://tug. ctan. org/tex-archive/graphics/duotenzor*, 2010.

[46] Lucien Hardy. Formalism locality in quantum theory and quantum gravity. *Philosophy of quantum information and entanglement*, pages 44–62, 2010.

[47] Lucien Hardy. Reformulating and reconstructing quantum theory. *arXiv preprint arXiv:1104.2066*, 2011.

[48] Lucien Hardy. The operator tensor formulation of quantum theory. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1971):3385–3417, 2012.

[49] Lucien Hardy. A formalism-local framework for general probabilistic theories, including quantum theory. *Mathematical Structures in Computer Science*, 23(2):399–440, 2013.

[50] Lucien Hardy. Operational general relativity: Possibilistic, probabilistic, and quantum. *arXiv:1608.06940*, 2016.

[51] Lucien Hardy. The construction interpretation: Conceptual roads to quantum gravity. *arXiv preprint arXiv:1807.10980*, 2018.

[52] Lucien Hardy and Robert Spekkens. Why physics needs quantum foundations. *arXiv preprint arXiv:1003.5008*, 2010.

[53] Aram Harrow. Coherent communication of classical messages. *Physical review letters*, 92(9):097902, 2004.

[54] Shima Hassanpour and Monireh Houshmand. Bidirectional quantum teleportation and secure direct communication via entanglement swapping. *arXiv preprint arXiv:1411.0206*, 2014.

[55] Shima Hassanpour and Monireh Houshmand. Bidirectional quantum controlled teleportation by using epr states and entanglement swapping. *arXiv preprint arXiv:1502.03551*, 2015.

[56] Shima Hassanpour and Monireh Houshmand. Bidirectional quantum teleportation via entanglement swapping. In *2015 23rd Iranian conference on electrical engineering*, pages 501–503. IEEE, 2015.

[57] Shima Hassanpour and Monireh Houshmand. Bidirectional teleportation of a pure epr state by using ghz states. *Quantum Information Processing*, 15(2):905–912, 2016.

[58] Masahito Hayashi, Kazuo Iwama, Harumichi Nishimura, Rudy Raymond, and Shigeru Yamashita. Quantum network coding. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 610–621. Springer, 2007.

[59] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[60] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Unified approach to quantum capacities: towards quantum noisy coding theorem. *Physical Review Letters*, 85(2):433, 2000.

[61] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81(2):865, 2009.

[62] Chris J Isham. Canonical quantum gravity and the problem of time. In *Integrable systems, quantum groups, and quantum field theories*, pages 157–287. Springer, 1993.

[63] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[64] Ding Jia. Quantum indefinite spacetime. *UWSpace*, 2017. http://hdl.handle.net/10012/11998.

[65] Ding Jia. Quantifying causality in quantum and general models. *arXiv preprint arXiv:1801.06293*, 2018.

[66] Ding Jia and Nitica Sakharwade. Process products need qualification. *arXiv:1706.05532*, 2017.

[67] Ding Jia, Nitica Sakharwade, et al. Tensor products of process matrices with indefinite causal structure. *Physical Review A*, 97(3):032110, 2018.

[68] Aleks Kissinger and Sander Uijlen. Picturing indefinite causal structure. *arXiv preprint arXiv:1701.00659*, 2017.

[69] Marius Krumm. Private communication, 2017.

[70] Matthew S Leifer and Robert W Spekkens. Towards a formulation of quantum theory as a causally neutral theory of bayesian inference. *Physical Review A*, 88(5):052130, 2013.

[71] Debbie Leung, Jonathan Oppenheim, and Andreas Winter. Quantum network communication—the butterfly and beyond. *IEEE Transactions on Information Theory*, 56(7):3478–3490, 2010.

[72] XS Liu, GL Long, DM Tong, and Feng Li. General scheme for superdense coding between multiparties. *Physical Review A*, 65(2):022304, 2002.

[73] GW Mackey. The mathematical foundations of quantum mechanics, a lecture-note volume by wa benjamin. *Inc., New York-Amsterdam*, 1963.

[74] Sonia Markes and Lucien Hardy. Entropy for theories with indefinite causal structure. *Journal of Physics: Conference Series*, 306(1):012043, 2011.

[75] Lluís Masanes and Markus P Müller. A derivation of quantum theory from physical requirements. *New Journal of Physics*, 13(6):063001, 2011.

[76] Chiranjib Mukhopadhyay and Arun Kumar Pati. Superposition of causal order enables perfect quantum teleportation with very noisy singlets. *arXiv preprint arXiv:1901.07626*, 2019.

[77] Sreraman Muralidharan and Prasanta K Panigrahi. Perfect teleportation, quantum-state sharing, and superdense coding through a genuinely entangled five-qubit state. *Physical Review A*, 77(3):032321, 2008.

[78] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[79] Ognyan Oreshkov. Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics. *Quantum*, 3:206, 2019.

[80] Ognyan Oreshkov and Nicolas J Cerf. Operational formulation of time reversal in quantum theory. *Nature Physics*, 11:853, 2015.

[81] Ognyan Oreshkov and Nicolas J Cerf. Operational quantum theory without predefined time. *New Journal of Physics*, 18(7):073037, 2016.

[82] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092, 2012.

[83] Ognyan Oreshkov and Christina Giarmatzi. Causal and causally separable processes. *New Journal of Physics*, 18(9):093020, 2016.

[84] Daniele Oriti. *Approaches to quantum gravity: Toward a new understanding of space, time and matter*. Cambridge University Press, 2009.

[85] Paolo Perinotti. Causal structures and the classification of higher order quantum computations. In *Time in physics*, pages 103–127. Springer, 2017.

[86] Christopher Portmann, Christian Matt, Ueli Maurer, Renato Renner, and Björn Tackmann. Causal boxes: quantum information-processing systems closed under composition. *IEEE Transactions on Information Theory*, 63(5):3277–3305, 2017.

[87] Helmut Reeh and Siegfried Schlieder. Comments on the unit ”a r ä equivalence of lorentzin variant fields. *Il Nuovo Cimento (1955-1965)*, 22(5):1051–1068, 1961.

[88] Carlo Rovelli. *Quantum Gravity*. Cambridge University Press, 2004.

[89] Nitica Sakharwade. Hierarchy of theories with indefinite causal structures: A second look at the causaloid framework, June 2021. PIRSA:21060122 see, https://pirsa.org.

[90] Nitica Sakharwade and Lucien Hardy. Hierarchy of theories with indefinite causal structure. 2020. Q-Turn, https://www.q-turn.org/wp-content/uploads/2020/11/Q-Turn_2020_paper_185.pdf.

[91] Nitica Sakharwade and Lucien Hardy. Hierarchy of theories with indefinite causal structures: A second look at the causaloid framework. *Bulletin of the American Physical Society*, 2021.

[92] Chitra Shukla, Anindita Banerjee, and Anirban Pathak. Bidirectional controlled teleportation by using 5-qubit states: a generalized view. *International Journal of Theoretical Physics*, 52(10):3790–3796, 2013.

[93] Rafael D Sorkin. Quantum mechanics as quantum measure theory. *Modern Physics Letters A*, 9(33):3119–3127, 1994.

[94] William G Unruh and Robert M Wald. Time and the interpretation of canonical quantum gravity. *Physical Review D*, 40(8):2598, 1989.

[95] Robert M Wald. *General Relativity*. University of Chicago press, 1984.

[96] John Watrous. Theory of quantum information (notes from fall 2011). In *John Watrous's Lecture Notes*, page Lec 05. 2011.

[97] John Archibald Wheeler. Information, physics, quantum: The search for links. *1990*, pages 3–29, 1990.

[98] Mark M Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[99] William K Wootters. Local accessibility of quantum states. *Complexity, entropy and the physics of information*, 8:39–46, 1990.

[100] Xin-Wei Zha, Zhi-Chun Zou, Jian-Xia Qi, and Hai-Yang Song. Bidirectional quantum controlled teleportation via five-qubit cluster state. *International Journal of Theoretical Physics*, 52(6):1740–1744, 2013.