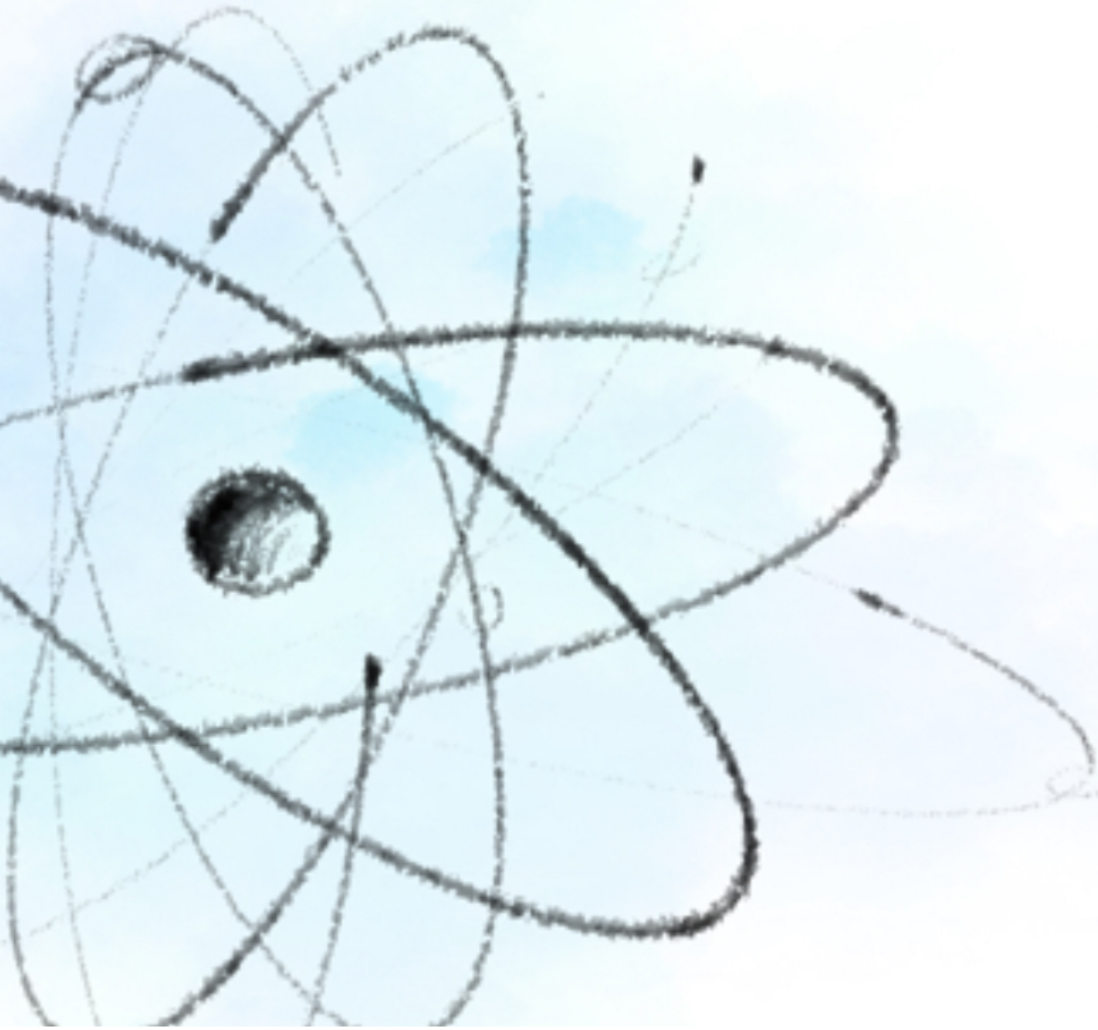
COMPLEMENTARITY, INCOMPATIBILITY, AND IRREVERSIBLE DISTURBANCE

A RESOLUTION TO THE DEBATE BETWEEN BOHR AND HEISENBERG

DAVIDE ROLINO



Università degli Studi di Pavia
Dottorato di Ricerca in Fisica – XXXVIII Ciclo

COMPLEMENTARITY, INCOMPATIBILITY, AND IRREVERSIBLE DISTURBANCE

A Resolution to the Debate Between Bohr and Heisenberg

Davide Rolino

Advisors: Prof. Paolo Perinotti, Dr. Alessandro Tosini

Submitted to the Graduate School of Physics
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN PHYSICS
at the
University of Pavia

Sometimes, to understand what is possible one must describe what is impossible.

Abstract

In 1927, Heisenberg introduced a heuristic argument, based on the famous γ-ray microscope *Gedankenexperiment*, to show that in quantum theory there exist operations that irreversibly disturb the systems on which they act. This argument was intended to show the existence of quantities that cannot be simultaneously measured. However, with a deeper understanding of how information can be manipulated and propagated, it is possible to prove that only the converse relation actually holds. In this thesis, exploiting the framework of Operational Probabilistic Theories (OPTs), we show that the impossibility of performing certain measurements simultaneously implies the existence of operations that irreversibly disturb the systems on which they act. We also present two toy theories, called Minimal Classical Theory (MCT) and Minimal Strongly causal Bilocal Classical Theory (MSBCT), that serve as counterexamples to the converse implication. Even though both theories satisfy full compatibility of observations, they still display irreversibility. Moreover, they are classical, Kochen-Specker, and generalised-non-contextual, yet nonetheless satisfy two quantum no-go theorems: No-Information Without Disturbance (NIWD) and no-broadcasting. This shows that these properties cannot be taken *per se* as signatures of non-classicality.

We also introduce two new classes of operational theories, which we here define and study: Minimal Operational Probabilistic Theories (MOPTs) and Minimal Strongly causal Operational Probabilistic Theories (MSOPTs), of which MCT and MSBCT are representatives, respectively. These theories are characterised by allowing only the minimal possible set of dynamics, with the latter also admitting classical conditioning. We prove that all MOPTs and all MSOPTs whose state spaces contain a spanning set of entangled states necessarily satisfy both NIWD and no-broadcasting.

Finally, we propose an operational definition of Bohr's complementarity, understood as the existence of properties of physical systems that cannot be simultaneously well-defined. We show that complementarity implies incompatibility and, consequently, irreversibility. In the specific case of quantum theory, however, complementarity and incompatibility are proved to coincide.

Contents

Contents **I**

List of figures **VII**

Acronyms **IX**

1 Introduction **1**

1.1 The beginning of quantum theory 2

1.2 Foundations of quantum theory 4

1.3 Uncertainty, complementarity and incompatibility 5

1.3.1 Irreversible disturbance 5

1.3.2 Incompatibility 9

1.3.3 Complementarity 10

1.4 Operational and informational approach 11

1.5 The results of this thesis 12

2 Framework **15**

2.1 A philosophical preamble 15
2.1.1 Under the hood of the framework 16
2.2 The framework of Operational Probabilistic Theories 17
2.2.1 The compositional structure of a theory 17
2.2.2 The probabilistic structure of a theory 34
2.2.3 Coarse-graining: filtering out the uninteresting 38
2.2.4 To quotient or not to quotient 45
2.3 Additional Structure 51
2.3.1 Linear Structure 52
2.3.2 Topological structure 61
2.4 Summary of the assumptions 69

3 Properties of OPTs **71**
3.1 Atomicity, extremality, and more 71
3.1.1 Amygdaloidal OPTs 75
3.2 Properties of the norms 80
3.2.1 Properties of the operational norm 80
3.2.2 Discussion around necessary and sufficient conditions for determinism 84
3.2.3 Properties of the sup norm 87
3.2.4 Properties of Cauchy sequences 87
3.2.5 Cauchy sequences and determinism 91
3.3 Coherently completing a theory 94
3.4 Discriminability 97
3.4.1 Local discriminability 99
3.4.2 Bilocal discriminability 101
3.5 Causal OPTs 105
3.5.1 No-signalling from the future 107
3.5.2 Unique deterministic effect 108
3.5.3 No-signalling without interaction 109
3.5.4 Proportionality 110

3.5.5 Non-causal theories . . . 111
3.6 Classical conditioning and strong causality . . . 114
3.6.1 Properties of strongly causal OPTs . . . 116
3.6.2 Completeness with respect to strong causality . . . 119
3.7 No-restriction hypothesis . . . 122
3.7.1 Relationship with strong causality . . . 124
3.8 Classicality . . . 126
3.8.1 Properties of classical OPTs . . . 126
3.8.2 The relationship with contextuality . . . 128
3.8.3 Classical Theory (CT) . . . 129
3.8.4 What is quantum then? . . . 129
3.8.5 Quantum Theory (QT) . . . 132
3.9 Summary of the assumptions . . . 133

4 Irreversibility, incompatibility & more 135
4.1 Information and disturbance . . . 136
4.1.1 No-information . . . 136
4.1.2 No-disturbance . . . 137
4.1.3 No-Information Without Disturbance (NIWD) . . . 138
4.2 Broadcasting . . . 141
4.3 Compatibility . . . 143
4.3.1 Observation-compatibility . . . 143
4.3.2 Strong-compatibility . . . 144
4.3.3 Weak compatibility . . . 145
4.3.4 Full-compatibility . . . 147
4.3.5 Full-Information Without Disturbance (FIWD) . . . 149
4.3.6 The relationship between full-compatibility and FIWD . . . 149
4.4 Irreversibility . . . 151
4.4.1 The relationship with compatibility . . . 152
4.4.2 The relationship with NIWD . . . 155
4.4.3 The relationship with thermodynamic irreversibility . . . 157

4.5 Summary . . . 158

5 Minimal OPTs **161**

5.1 Preliminary results . . . 162

5.1.1 Elementary systems and minimal decomposition . . . 162

5.1.2 Braiding . . . 164

5.2 The definition . . . 166

5.3 Characterisation of instruments and transformations . . . 167

5.3.1 Non-limit instruments and transformations . . . 167

5.3.2 The case of symmetric MOPTs . . . 172

5.3.3 Limits of Cauchy sequences . . . 173

5.3.4 Limits of deterministic transformations . . . 176

5.4 Properties of MOPTs . . . 179

5.4.1 Atomicity of the identity . . . 179

5.4.2 Irreversibility & More . . . 181

5.5 Minimal Classical Theory (MCT) . . . 186

5.5.1 Properties of MCT . . . 188

6 Minimal Strongly causal OPTs **189**

6.1 The definition . . . 190

6.2 Characterisation of instruments and transformations . . . 192

6.3 Properties of MSOPTs . . . 194

6.3.1 Atomicity of the identity . . . 194

6.3.2 Irreversibility & more . . . 205

6.4 Minimal Strongly causal Bilocal Classical Theory (MSBCT) . . . 207

6.4.1 Postulates of the theory . . . 207

6.4.2 Properties of MSBCT . . . 211

6.4.3 Minimal Bilocal Classical Theory (MBCT) . . . 213

6.5 Relationship between an OPT and the corresponding MSOPT . . . 214

6.6 Local discriminability, strong causality and simpliciality . . . 216

7 Complementarity **217**

7.1 Operational definition 217
7.2 Different degrees of complementarity 221
7.3 Complementarity implies incompatibility 223
7.4 The case of quantum theory 228

8 Conclusion **231**

A Category Theory **241**
A.1 Monoidal categories 242
A.2 Braided monoidal categories 246
A.3 Graphical calculus for category theory 248

B Cones **249**
B.1 Proof of Theorem 4 249
B.2 Proof of Theorem 5 251

C The remaining postulates of Quantum Theory (QT) **255**

D Observation-compatibility in CT **259**

E Proof of Lemma 25 **261**

F Details of the proof of Theorem 41 **265**

Acknowledgements **275**

List of publications **277**

Bibliography **279**

Symbols **301**

Index **303**

List of Figures

1.1 Schematic representation of the γ-ray microscope of Heisenberg's *Gedankenexperiment*. 7

2.1 Schematic representation of the Stern-Gerlach experiment. 18

3.1 Example of a set of effects for a system A of size $D_{\mathrm{A}} = 2$, i.e., a bit. 74

3.2 Example of a set of states for a system A of size $D_{\mathrm{A}} = 2$, i.e., a bit. 76

3.3 Example of a noisy set of states for a system A of size $D_{\mathrm{A}} = 2$, i.e., a noisy bit. 77

3.4 Example of an amygdaloidal set of states for a system A of size $D_{\mathrm{A}} = 2$, i.e., an amygdaloidal bit. 78

3.5 Representation of a quantum switch. 112

8.1 Venn diagram summarising the relationships between complementarity, observation-incompatibility, strong-incompatibility, and irreversibility. 237

8.2 Venn diagram summarising the relationship between strong-causality, causality ,and the no-restriction hypothesis. 238

8.3 Venn diagram summarising the relationships between strong-causality, causality, local discriminability, and simpliciality. 239

Acronyms

BCT Bilocal Classical Theory

CA Cellular Automaton

CPTNI Completely Positive Trace-Non-Increasing

CT Classical Theory

DAG Directed Acyclic Graph

DCT Dual Classical Theory

FIWD Full-Information Without Disturbance

FQT Fermionic Quantum Theory

GPT Generalised Probabilistic Theory

LQT Latent Quantum Theory

MBCT Minimal Bilocal Classical Theory

MCT Minimal Classical Theory

MOPT Minimal Operational Probabilistic Theory

MQT Minimal Quantum Theory

MSBCT Minimal Strongly causal Bilocal Classical Theory

MSCT Minimal Strongly causal Classical Theory

MSOPT Minimal Strongly causal Operational Probabilistic Theory

MSQT Minimal Strongly causal Quantum Theory

NIWD No-Information Without Disturbance

OPT Operational Probabilistic Theory

POVM Positive Operator-Valued Measure

QT Quantum Theory

RQT Real Quantum Theory

1 Introduction

1925, Helgoland.

In search of relief from a severe bout of hay fever, Werner Heisenberg retreats to this pollen-free island in the North Sea. There, in isolation, he devotes himself entirely to the emerging field of quantum mechanics, a pursuit that would soon revolutionize our way of doing physics. This is where quantum theory was born.

By the late 19th and early 20th centuries, it had become increasingly clear that classical physics could not fully account for the behaviour of the natural world. Experimental results, such as those related to *black body radiation* and the *photoelectric effect*, resisted explanation within the classical framework. The resolution came in a radically new form: the proposal that energy is not continuous, but quantized. Planck first introduced this idea to explain the spectral distribution of black body radiation [1, 2][1], followed by Einstein, who applied the concept of quantization to light in his explanation of the photoelectric effect [4].

Exploiting this intuition, several other quantum–classical anomalies were successfully addressed:

[1]Interestingly, contrary to what is often suggested in physics textbooks, Planck was already working toward a formula that matched the observed spectrum of black body radiation before the so-called *ultraviolet catastrophe* was formally identified [3]. The name "ultraviolet catastrophe" was used for the first time in 1911 by Erhenfest at the first Solvay Congress.

- *Bohr's atomic model* was introduced to explain the discrete spectral lines of the hydrogen atom through quantised electron orbits. It proposed that electrons can only occupy specific energy levels [5].
- *Compton* demonstrated that X-rays *scatter* off electrons with a shift in wavelength that depends on the scattering angle; evidence consistent with the particle-like behaviour of photons and in contradiction with classical wave theory [6].
- The *Stern–Gerlach experiment* provided evidence for the quantisation of angular momentum. In the experiment, particles with a non-zero magnetic moment were passed through a spatially varying magnetic field, resulting in their deflection into discrete spots on a detection screen [7].

However, despite the significant advances in understanding beyond-classical phenomena, a complete and coherent formulation of a new physical theory was still missing. This changed with Heisenberg's sojourn in Helgoland. During this period, he developed the ideas that led to the publication of his seminal paper "Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen" [8] in September 1925. This work laid the foundation for the *matrix formulation* of quantum mechanics, which was later developed by and in collaboration with Born and Jordan [9, 10]. With this, a comprehensive and systematic exploration of the new quantum theory could finally begin.

To be fair, although the birth of quantum theory is conventionally marked as 1925, celebrating the first coherent formulation in the form of matrix mechanics, Schrödinger was not far behind. He postulated what is now known as the Schrödinger equation in the same year, though its publication came slightly later, in 1926 [11]. The Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x,t) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x,t) \right] \Psi(x,t)$$

describes the time evolution of the wave function $\Psi(x,t)$, the central mathematical object that represents the state of an isolated quantum system in this formulation. Schrödinger's approach offered a different, but ultimately equivalent, perspective on quantum theory, complementing the matrix formulation developed by Heisenberg, Born, and Jordan. This formulation is often referred to as *wave mechanics*.

1.1 The beginning of quantum theory

One might imagine that, once the formalism of quantum theory had been coherently established, physicists could have proceeded directly to apply it, thereby deepening their understanding of the physical world. Yet matters proved far less straightforward. Despite the powerful mathematical tools at their disposal, the new theory remained conceptually elusive: its formalism pointed to behaviours that defied classical intuition and, at times, seemed almost incomprehensible.

Probably the most prominent open question in quantum theory is what is commonly referred to as the *measurement problem*. Although this issue is widely discussed, there is no consensus, at the time of writing this thesis, on whether it should actually be considered a problem at all. The measurement problem brings to light a fundamental tension in quantum theory: the existence of two distinct types of physical evolution. The first is the continuous, deterministic evolution of an isolated quantum system, described by unitary dynamics. The second is the discontinuous, probabilistic evolution associated with measurements, where a quantum system collapses instantaneously into a definite outcome. This apparent duality in the theory's treatment of dynamics has sparked extensive debate and has led to the development of various interpretations of quantum theory, each one proposing a different perspective on the nature of quantum states and the role of measurements in order to address, or circumvent, the measurement problem.

One approach is the *Copenhagen interpretation*, historically associated with Bohr and Heisenberg. It embraces the dual dynamics by positing a fundamental divide between the quantum system and the classical measuring apparatus. Measurements cause a non-unitary, probabilistic collapse of the wave-function, but this process is not described by the theory itself. Instead, it is taken as a primitive postulate [12, 13].

In contrast, *many-worlds interpretations*, initiated by Everett [14, 15], reject the collapse altogether. According to this views, all possible outcomes of a quantum measurement are realized in separate, branching worlds. The evolution of the wave-function is always unitary and deterministic, and the appearance of randomness is an illusion arising from the observer's entanglement with the system.

Objective collapse models, such as the Ghirardi–Rimini–Weber theory [16], propose a modification of quantum dynamics by introducing spontaneous, stochastic collapses of the wave-function. These collapses occur rarely for microscopic systems but frequently enough at the macroscopic scale to explain the emergence of definite outcomes during measurements.

Finally, *hidden variable theories*, most notably Bohmian mechanics [17], maintain that quantum systems possess definite properties at all times. In this framework, the wave-function does not represent probabilities but acts as a guiding field that defines the deterministic evolution of hidden variables, such as the precise positions of particles. Measurement outcomes merely reveal these pre-existing values. In principle, if we had access to the hidden variables, the system would exhibit classical deterministic behaviour. However, since these variables are inaccessible in practice, quantum phenomena appear probabilistic and sometimes paradoxical. Bohmian mechanics thus provides an ontological interpretation of quantum theory that restores determinism at the fundamental level.

In addition to the discussion surrounding the measurement problem, several other topics remain the subject of ongoing debate. Beyond those examined in this thesis—namely, irreversibility, incompatibility, and complementarity—further open questions include for example:

- *Wigner's friend-type paradoxes.* These paradoxes originate from a thought experiment proposed by Eugene Wigner, in which he considers the consequences of one observer being observed by another. Specifically, Wigner imagines observing his friend while the friend performs a quantum measurement [18]. These paradoxes have since evolved to encompass scenarios involving more than two agents, known as *extended Wigner's friend arguments* [19–26]. A problem that is closely related to the measurement problem.
- The *precise definition of classicality.* While it might seem straightforward from an intuitive perspective to distinguish classical from quantum behaviours, defining a rigorous and universally applicable boundary between the two remains a significant challenge. To reliably identify a phenomenon as genuinely quantum, one must be able to exclude the possibility of explaining it through any classical model.

1.2 Foundations of quantum theory

As we have seen in the previous section, despite the rigorous formulations of quantum mechanics, many aspects of the theory remained conceptually unclear. This led to the emergence of a new field of physics: *foundations of quantum theory*. This was a natural consequence of the peculiar nature of quantum theory itself, which drew together like-minded individuals, not only interested in using the theory for practical calculations and technological development, but profoundly intrigued by its inner workings and philosophical implications.

Probably the most recognised result in the field of quantum foundations is *Bell's theorem* [27–29]. This argument was developed by Bell to address the purely foundational question of whether or not quantum theory can be *completed* [12, 30]—that is, extended through the introduction of *hidden variables* that would allow for deterministic predictions. Yet, this result lies at the heart of the development of quantum technologies. Who could have predicted that, decades after its first formulation in 1964, Bell's theorem would become central to a technological revolution? Bell-type theorems underpin modern protocols that certify the presence of entanglement; an essential resource for achieving quantum advantage in information-processing tasks [31–39].

For the sake of completeness, we briefly observe that what Bell showed is that there exists a fundamental trade-off between a theory admitting *hidden variables* and it satisfying the principle of *no-signalling*—the requirement that information cannot travel faster than the speed of light. No theory that simultaneously maintains hidden variables and respects no-signalling can reproduce all the statistical predictions of quantum or post-quantum theories [40–54].

1.3 Heisenberg's *Gedankenexperiment* and Bohr's complementarity with a touch of incompatibility

Among the many phenomena that have been the subject of investigation in quantum foundations, a special place is reserved for *irreversible disturbance*, *complementarity*, and *incompatibility*.

1.3.1 Irreversible disturbance

The notion of *irreversible disturbance* was first introduced in another seminal work by Heisenberg, where he proposed and discussed the famous *Gedankenexperiment* of the γ-ray microscope [55, 56].

Following his contributions to the development of the matrix formulation of quantum theory, Heisenberg turned his attention to one of its most distinctive features: unlike in classical physics, it is impossible to perform a truly non-invasive measurement on a quantum system. Every act of observation necessarily disturbs the system:

> [. . .] in classical physical theories it has always been assumed either that this interaction is negligibly small, or else that its effect can be eliminated from the result by calculations based on 'control' experiments. This assumption is not permissible in atomic physics; the interaction between observer and object causes uncontrollable and large changes in the system being observed, because of the discontinuous changes characteristic of atomic processes [. . .] [56].

In particular, Heisenberg had in mind a relationship expressing the trade-off between the precision of position and momentum measurements for a particle:

$$\Delta x \, \Delta p_x \approx h,$$

and introduced the γ-ray microscope as a thought experiment supporting the validity of this relation. However, today we understand that the matter is more subtle than how Heisenberg presented it. Today, Heisenberg's argument is recognised as a fundamental, yet heuristic, insight in the development of quantum theory, one that captures several distinct features of the formalism, but without providing a rigorous theoretical foundation for any of them. Those were established only through later formalisations.

The *Gedankenexperiment*

The goal of the *Gedankenexperiment* is to determine the position of an electron using a microscope. The electron is assumed to behave like a classical particle, moving along the x-axis, orthogonally to the optical axis of the microscope. The setup is schematically illustrated in figure 1.1.

The functioning of the microscope relies on the scattering of light by the object under observation. In this case, a photon is scattered by the electron and then enters the microscope's lens.

If the wavelength of the incident photons is λ, and ε is the half-angle of the cone defined by the aperture of the lens, then according to classical optics, the resolution limit of the microscope is given by

$$\Delta x = \frac{\lambda}{\sin(\varepsilon)}.$$

However, the interaction between the photon and the electron also imparts momentum to the latter. Due to the Compton effect, the electron receives a recoil proportional to $\frac{h}{\lambda}$, where h is Planck's constant. Because the scattering direction is uncertain, the momentum transfer along the x-axis is also uncertain, and its indeterminacy is approximately

$$\Delta p_x \approx \frac{h}{\lambda} \sin(\varepsilon).$$

This quantifies the unavoidable disturbance introduced by the measurement.

Multiplying these uncertainties yields:

$$\Delta x \Delta p_x \approx \frac{\lambda}{\sin(\varepsilon)} \cdot \frac{h}{\lambda} \sin(\varepsilon) = h,$$

thus recovering Heisenberg's uncertainty relation:

$$\Delta x \Delta p_x \approx h,$$

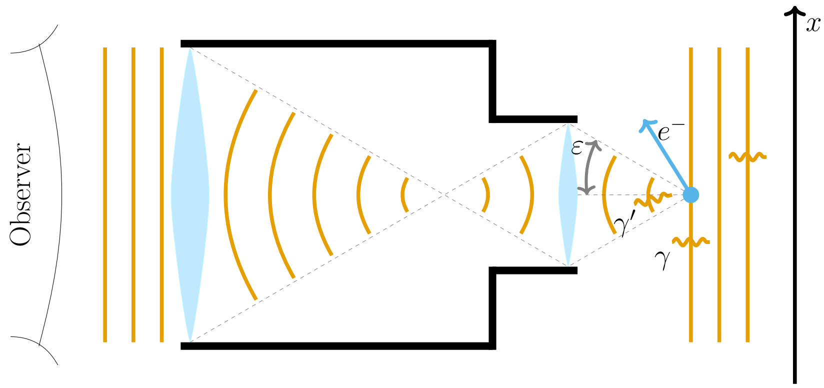


Figure 1.1: Schematic representation of the γ-ray microscope of Heisenberg's *Gedankenexperiment*. An electron, treated as a classical particle, moves along the x-axis, orthogonal to the optical axis of the microscope. The microscope operates by detecting a photon scattered by the electron: after the interaction, the photon (denoted γ') enters the lens, while the electron undergoes a change of trajectory due to the transferred momentum. The parameter ε denotes the half-angle of the cone defined by the aperture of the lens.

which expresses a fundamental limit to how precisely nature allows position and momentum to be known simultaneously. Crucially, this limit arises not from imperfections in the measuring apparatus, but from the very nature of quantum systems and the disturbance caused by observation.

Interpreting the experiment

As one can see now that we have presented the *Gedankenexperiment*, while it is clear that Heisenberg's objective was to demonstrate the validity of his uncertainty relation,

$$\Delta x \, \Delta p_x \approx h,$$

a closer scrutiny reveals certain limitations in his argument. The first concerns what the argument is actually intended to support. It is indeed true that, by following Heisenberg's reasoning, one recovers the uncertainty relation. However, it remains unclear what this relation is precisely referring to: does it characterise a property of the physical system itself, or rather a limitation imposed by the measurement procedure? What is the relationship between the uncertainty relation and the disturbance action? Furthermore, although the thought experiment was devised to highlight a distinctive quantum behaviour, the reasoning behind it remains essentially semi-classical.

This indefiniteness in the argument brought controversies from the very beginning. For example, Bohr, upon reading a draft of Heisenberg's paper, urged him to reconsider the analysis in greater detail, as he was neglecting other physical phenomena that had to be taken into account in addition to Compton scattering. These concerns were acknowledged by Heisenberg and incorporated as a remark in the final version of the article [55]. While Bohr's comment is often interpreted today as evidence that he did not entirely grasp Heisenberg's reasoning—indeed, his observations were correct, but missed the main point of the discussion—it nonetheless highlights that the argument was not being fully understood and required further clarification.

This is reflected in the extensive revisitations of Heisenberg's argument that physicists carried out in the decades that followed, aiming to refine its interpretation and better understand its consequences [17, 57–71].

In recent years, for example, the precise meaning of the quantities Δx and Δp_x has been the subject of active debate, particularly through the contributions of Ozawa [72–74] and Busch [75–77], who offered competing interpretations of the operational content of uncertainty and proposed refined formulations aimed at distinguishing between intrinsic quantum indeterminacy and measurement-induced disturbance.

The modern perspective

Nearly a century after its original formulation, researchers in the field of quantum foundations appear to have reached a consensus regarding the correct interpretation of Heisenberg's *Gedankenexperiment* and the phenomena it encompasses.

First of all, there exist two clearly distinct aspects that can be captured by Heisenberg's relation: the *uncertainty relations* and the *uncertainty principle*. The former concerns the statistics of repeated measurements on an ensemble of equally prepared, identical quantum systems; the latter, on the contrary, concerns a sequence of measurements performed on the same quantum system [63, 70].

Regarding uncertainty relations, subsequent studies provided a rigorous derivation of them starting from the commutation relations, leading to the Robertson-type inequalities [78–81]. Robertson's uncertainty relations have a clear statistical interpretation: they express the impossibility of preparing an ensemble of systems for which both the position and momentum distributions are arbitrarily sharp. In other words, the product of the standard deviations of the two observables, computed over many identically prepared systems, cannot fall below a fixed bound determined by their non-commutativity.

Further studies were also carried out by exploiting the definition of the Shannon entropy, which led to the so-called *entropic uncertainty relations* [82, 83], providing an information-theoretic formulation of the same limitations.

The second aspect, instead, concerns the action of measurement itself, and in particular the phenomenon of *irreversible disturbance*, namely the fact that any measurement on a physical system irreversibly alters its state. Despite its centrality in Heisenberg's original argument, this phenomenon has not yet been fully comprehended within the field of quantum foundations. In particular, an operational definition has long been lacking, a question that will be central in this thesis[2].

This second aspect has also been closely connected to the development of a related, though distinct[3], concept: *incompatibility*, that is, the problem of the joint measurability of different physical quantities.

1.3.2 Incompatibility

As mentioned above, one of the common misinterpretations of Heisenberg's *Gedanken-experiment* is that it concerns the impossibility of jointly measuring certain physical quantities. Regardless of this somewhat mistaken origin, the property of measurement *incompatibility* has since emerged as one of the most distinctive features of the quantum world and a central topic in the study of quantum foundations [84–99].

In its original and most widely studied formulation, *compatibility* pertains to Positive Operator-Valued Measures (POVMs). Let $[\![\mathrm{E}_x]\!]_{x\in\mathsf{X}}$ and $[\![\mathrm{F}_y]\!]_{y\in\mathsf{Y}}$ be two POVMs on the Hilbert space $\mathcal{H}_\mathrm{A}$. They are said to be compatible if there exists a third POVM $[\![\mathrm{C}_{x,y}]\!]_{x,y\in\mathsf{X}\times\mathsf{Y}}$ on $\mathcal{H}_\mathrm{A}$ such that:

$$\begin{aligned} \mathrm{E}_x &= \sum_{y\in\mathsf{Y}} \mathrm{C}_{x,y} \quad \forall x\in\mathsf{X}, \\ \mathrm{F}_y &= \sum_{x\in\mathsf{X}} \mathrm{C}_{x,y} \quad \forall y\in\mathsf{Y}. \end{aligned} \tag{1.1}$$

In words, two POVMs are compatible if there exists a joint measurement whose outcomes can be classically post-processed to recover the statistics of both original measurements.

If no such joint measurement exists, the two POVMs are said to be *incompatible*.

[2] We remark that in this thesis we are not interested in deriving uncertainty relations connecting different subsequent measurements on the same quantum system, but rather in identifying the conditions under which such a measurement-induced disturbance is present.

[3] For a more detailed discussion, see section 4.4.1.

1.3.3 Complementarity

First introduced by Bohr in his 1928 Como lecture [100] as a more general principle from which Heisenberg's disturbance and uncertainty relation was to be derived, *complementarity* quickly became a cornerstone of quantum theory. However, as with many of Bohr's ideas, it has remained the subject of lively debate—partly owing to his notoriously cryptic style of exposition[4].

In the literature, three principal lines of research on complementarity can be identified. The first consists of the initial discussions by Bohr himself [100, 101] and Pauli [102].

The second line, often referred to as the "which-way" approach, was initiated by Englert [103]. In this framework, complementarity is understood as a quantitative trade-off between which-path information and fringe visibility in interference experiments, such as the double-slit experiment or the Mach-Zehnder interferometer [104–111].

The third, more recent, line of research investigates complementarity in terms of quantum observables, and particularly the incompatibility of POVMs [71, 76, 112–118].

Although each of these approaches provides valuable insight, we argue that they are ultimately unsatisfactory. On the one hand, the which-way paradigm—though inspired by Bohr's discussions on complementarity centred on wave-particle duality—adheres too closely to a specific physical model. It privileges one interpretative framework over others. In line with our operational perspective, which aims to remain agnostic regarding particular models or interpretations, we regard this model-dependent notion of complementarity as limited in scope. On the other hand, more recent approaches tend to conflate complementarity with the notion of incompatibility. While the two concepts are closely related—we show here that they are actually equivalent in standard quantum theory—, we maintain that they should remain conceptually distinct. Complementarity, in our view, deserves an independent operational definition that captures its foundational role without being reduced to incompatibility.

As proposed in Ref. [119]—and here discussed in chapter 7—, we introduce a novel definition of complementarity that preserves the original spirit of Bohr's formulation while retaining full generality. In our framework, complementarity is defined as the impossibility of making simultaneous statements about certain properties of a physical system.

[4] According to an anecdote, it took years before anyone noticed that one of Bohr's works had been published with two pages swapped.

1.4 Operational and informational approach

As we have seen, there are limitations to what can be understood about quantum theory using the standard formalisms of matrix and wave mechanics. To overcome these limitations, researchers in the field of quantum foundations have begun to explore alternative, more abstract frameworks.

A pivotal moment in this direction was the work of Birkhoff and von Neumann, "The Logic of Quantum Mechanics" [120]. In this paper, the authors proposed a new kind of logic—now known as *quantum logic*—based on the lattice structure of closed subspaces of Hilbert spaces. Their aim was "to discover what logical structure one may hope to find in physical theories which, like quantum mechanics, do not conform to classical logic." While their approach was rooted in the mathematical structure of quantum theory, it marked a conceptual shift: rather than taking classical logic as a given, they asked what kind of logic might be appropriate for theories exhibiting quantum-like behaviour.

Importantly, the work of Birkhoff and von Neumann opened the door to a more abstract, theory-agnostic perspective on physical theories. Although their immediate focus was quantum theory, the structural features they examined—such as the lattice of propositions—suggested that it was possible to study the logical architecture of a physical theory independently of the specific formalism it employed. This marked a foundational shift: rather than deriving insights from within quantum theory, one could investigate the broader class of theories that share certain structural or operational characteristics.

Building on this idea, the research line on *quantum logic* emerged [121–123], seeking to characterize and generalize the logical and algebraic underpinnings of quantum theory. These developments gradually gave rise to a more general ambition: to understand quantum theory not in isolation, but in relation to a wider space of conceivable physical theories.

This ambition finds its modern expression in what we call the *operational and informational approach* to quantum foundations [124–127]. Within this framework, quantum theory is viewed as a theory of information processing, and is characterized by operational principles concerning which information-theoretic tasks are possible or impossible. This perspective seeks to identify what distinguishes quantum theory from other conceivable theories by characterizing the tasks that can or cannot be performed within each [124–126, 128–134]

This approach also has roots in the field of *quantum information*, which treats quantum mechanics itself as a resource for processing and transmitting information [38, 135–141]. In this sense, quantum information not only provided new technological applications, but also reshaped our foundational understanding of quantum theory by giving operational and information-theoretic constraints a central role.

The scenario of alternative theories that we consider here is captured by the *framework of Operational Probabilistic Theories (OPTs)* [130, 133, 134, 142, 143]. This framework has been specifically devised to survey and compare general physical theories that share the same compositional structure as quantum theory—that is, the way operations can be combined to build up experiments—without assuming any particular underlying mathematical formalism. In this sense, it allows one to study quantum theory "from the outside." Moreover, it provides a suitable playground for exploring general structural and logical dependencies among the properties that physical theories may exhibit.

The OPT framework also connects with other approaches, such as the less structured *framework of Generalised Probabilistic Theories (GPTs)* [144–153], as well as with the diagrammatic, category-theoretic approach often referred to as *quantum pictorialism* [154–157].

1.5 The results of this thesis

In this thesis, we continue the tradition of quantum foundations by investigating some of the most intriguing phenomena that characterise the quantum world. In particular, we focus on the notions of *irreversible disturbance*, *incompatibility*, *complementarity*, and the relationships between them.

In Theorem 30 we show that *observation-incompatibility* (Definition 37)—the existence of measurements that cannot be performed simultaneously—implies *irreversible disturbance* (Definition 44) à la Heisenberg at the level of operations—i.e., the existence of operations that irreversibly alter the state of the system on which they are acting. However, the converse does not hold. To demonstrate this, we provide two toy theories as counterexamples: *Minimal Classical Theory (MCT)* (section 5.5) [98] and *Minimal Strongly causal Bilocal Classical Theory (MSBCT)* (section 6.4) [158].

- MCT is a toy theory obtained by restricting the sets of allowed operations of Classical Theory (CT)[5], while keeping its sets of states and measurements untouched. Specifically, we maintain only those processes necessary for compatibility with the structure of the framework of OPTs: these include only operations in which, following a measurement, a random state is prepared, as well as the identity and swap operations (and compositions and limits of sequences thereof).

- MSBCT is similarly obtained by restricting the set of allowed operations in

[5]Here we use CT to indicate the operational version of classical theory [133, 159]

> Bilocal Classical Theory (BCT) as above, but now allowing also for classical conditioning—that is, all operations where the action performed is conditioned on the classical outcome of a previous operation.

Both of these theories exhibit irreversible disturbance, yet all their measurements are jointly measurable; thus, they lack observation-incompatibility.

Furthermore, these two theories are classical—in the sense of being simplicial, with all pure states perfectly discriminable—, and they are also Kochen-Specker and generalised-non-contextual. Nonetheless, they still satisfy *No-Information Without Disturbance (NIWD)* and *no-broadcasting*, showing that these features cannot be taken *per se* as signatures of non-classicality.

The way we introduce these two theories is as particular instances of two broader classes of theories: *Minimal Operational Probabilistic Theories (MOPTs)* (chapter 5) and *Minimal Strongly causal Operational Probabilistic Theories (MSOPTs)* (chapter 6). MOPTs are *minimal* versions of OPTs, obtained by allowing only preparations, measurements, permutations of systems, and arbitrary compositions and limits thereof. A MSOPT is then built by reintroducing conditional operations.

We show that all MOPTs and all MSOPTs that have a spanning set of entangled states satisfy both NIWD and no-broadcasting, independently of the nature of their systems. The reason is that these theories do not admit a non-trivial decomposition of the identity transformation (Theorem 38, Theorem 41). A condition that we show implies the aforementioned no-go theorems (Theorem 26, Theorem 27).

In order to coherently build these new classes of theories, we establish a series of new results concerning properties of Cauchy sequences within the framework of OPTs. These results, presented mainly in section 3.2.4 and section 3.2.5, are developed to characterise the full set of operations of MOPTs and MSOPTs, which are required to be Cauchy-complete (Assumption 4). We also show how to coherently perform the Cauchy-completion of a theory (Theorem 8) and how to construct a coherent OPT by adding all conditional operations (Theorem 21), a fundamental result for the definition of MSOPTs.

Having completed the characterisation of the relationship between compatibility and irreversibility, in the final chapter of this thesis we broaden our scope by introducing an operational definition of Bohr's *complementarity* (Definition 62) and characterising its relationship with incompatibility. We show that complementarity implies incompatibility (Corollary 16) and, consequently, irreversibility. We also show that in the specific case of quantum theory complementarity and incompatibility coincide (section 7.4).

2 Framework

2.1 A philosophical preamble

HUMANS have always been fascinated by the natural world that surrounds them and, for the most various reasons—genuine curiosity, survival, wealth, power, religion—have always tried to describe and comprehend it and its inner workings. Evidence of this pursuit can be found in ancient artefacts such as the Mayan and Aztec calendars, as well as in early philosophical traditions. Among these, metaphysics played a foundational role in the development of what we now call *physics*.

Physics is often regarded as distinct from metaphysics and philosophy because of its definiteness: the idea that it rests on clear, objective assumptions and produces precise, testable predictions. To some extent this is true. The mathematical formalism underlying physical theories allows for the construction of coherent models that align with experimental data.

However, works in the field of quantum foundations shows that such definiteness is only a first-order approximation. Like any intellectual endeavour, physics is built upon assumptions, many of which are philosophical in nature—for example, that physical law do not change in time or are the same throughout all our universe. The key difference is that physicists are not always consciously aware of them [160]. Heisenberg expressed this point vividly in his book “Physics and Philosophy”:

> What we observe is not nature itself, but nature exposed to our method

of questioning.

At this stage the reader may wonder why such an introduction appears in a chapter devoted to a physical framework. The reason is that any framework capable of describing a wide range of physical theories must rest on certain foundational assumptions. These assumptions inevitably carry philosophical weight, and it is important to make them explicit from the outset.

A bit like Descartes, who embarked on a personal journey to identify something that could not be doubted—culminating in the famous conclusion "I think, therefore I am"—we must first identify the foundational elements of physical theories in order to build a general framework. What are the minimal ingredients that any physical theory must possess? These will serve as the fundamental building blocks upon which the framework is constructed.

2.1.1 Under the hood of the framework

The framework is based on the scientific-philosophical doctrine of *operationalism* proposed by Bridgman, arguing that a physical concept is meaningful only if it is defined by a method of measurement [161]. In the words of the author himself:

> In general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations [161].

In practice, this principle is implemented in the framework by taking as primitives not physical quantities, but the operations that can be performed in a laboratory. It is the operations themselves that define the physical quantities.

Specifically, the first two primitives of the framework are *systems* and *tests*. *Systems* represent the physical entities that are probed in a laboratory—for example, an electron, a molecule, or a radiation field. *Tests* represent the physical processes—occurring between systems—that constitute experiments, such as the single use of a laboratory device.

By combining these two primitives, we obtain a powerful descriptive framework that allows us to mathematically model any conceivable experiment. However, one essential ingredient is still missing: the ability to make predictions. In order for the framework to yield meaningful insights into physical phenomena, it must also allow for quantitative predictions that can be compared with empirical data. This is the motivation for introducing the third primitive of the framework: *probabilities*[1].

[1] Through the probabilistic structure, the framework can also be seen as a non trivial extension of probability calculus. Furthermore, following the reasoning delineated in Refs. [162–165], it can also be treated as an extension of logical reasoning.

From a methodological perspective, the framework adopts a *black-box* approach to physical processes, where tests are characterized by their operational connectivity, and probabilities are assigned through a suitable calculus [143, 166].

2.2 The framework of Operational Probabilistic Theories

We can now formally introduce the *framework of Operational Probabilistic Theories (OPTs)* [129, 130, 133, 134, 142, 143, 167].

2.2.1 The compositional structure of a theory

As previously mentioned, the framework of OPTs aims at describing generic theories of information processing—including Classical Theory (CT) and Quantum Theory (QT)[2] alongside a universe of other alternative theories—starting from their compositional structure, which is based on the notions of systems and tests.

Let Θ be a generic OPT. The *systems* characterising this theory—i.e., the set of physical entities that are subject of study in this theory—are denoted with capital Roman letters $\mathrm{A}, \mathrm{B}, \ldots \in \mathsf{Sys}(\Theta)$. The set of *tests*—i.e., the physical processes that can be implemented within the theory—is denoted by $\mathsf{Test}(\Theta)$. For any pair of systems A and B, the subset $\mathsf{Test}(\mathrm{A} \to \mathrm{B}) \subset \mathsf{Test}(\Theta)$ collects all tests from system A to system B.

In QT, systems correspond to complex Hilbert spaces, while tests are described by *quantum instruments*—that is, collections of Completely Positive Trace-Non-Increasing (CPTNI) maps (*quantum operations*) whose coarse-graining yields a *quantum channel*, namely a Completely Positive Trace-Preserving (CPTP) map.

A particular test from an input system A to an output system B is denoted by $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, where X represents the *outcome space* of the test; a set containing all the possible *outcomes* of the experiment. When clear from the context we will often use the shorthand notation $\mathsf{T} \equiv \mathsf{T}_{\mathsf{X}} \equiv \mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}$ to denote tests.

To each outcome $x \in \mathsf{X}$ is associated an *event* $\mathscr{T}_x \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$ representing the realization of a particular physical occurrence. Therefore, a tests is an *indexed family* of events: $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, with $\mathscr{T}_x \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$ for all $x \in \mathsf{X}$.

For example, the test describing a Stern-Gerlach experiment [7], schematically illustrated in figure 2.1, is $[\![\mathscr{T}_\uparrow, \mathscr{T}_\downarrow]\!]$, where the two events $\mathscr{T}_\uparrow$, $\mathscr{T}_\downarrow$ represent the outcomes corresponding to the particle being measured in a spin-up or spin-down

[2] As for the case of CT, we will adopt the acronym QT to specifically refer to the operational formulation of quantum theory [124–126, 128–132].

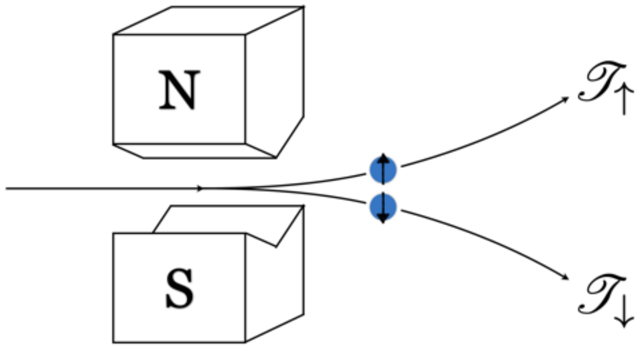


Figure 2.1: Schematic representation of the Stern-Gerlach experiment. The test describing the experiment is $[\![\mathscr{T}_\uparrow, \mathscr{T}_\downarrow]\!]$. The two events $\mathscr{T}_\uparrow$, $\mathscr{T}_\downarrow$ represent the two occurrences of the particle being measured in a spin-up or spin-down state, respectively.

state, respectively.

The idea behind the notion of a test is to generalise the concept quantum instrument to generic theories of information-processing.

Among all tests defined in a theory, there is a subset of particular importance: the tests whose outcome space consists of a single element—that is, a *singleton set*—denoted by $\star := \{*\}$. These tests are called *singleton tests*. The single events that compose them are called *deterministic* and operationally model processes that do not provide information. In QT, deterministic events are *quantum channels*.

Remark 1

A key point is that *tests are the primitive concept of the framework*, not events. To define an OPT Θ, one specifies the set of tests $\mathsf{Test}(\Theta)$, from which the full set of events $\mathsf{Event}(\Theta)$ is determined.
Consequently, it may happen that two OPTs Θ_1 and Θ_2 share the same event set, $\mathsf{Event}(\Theta_1) \equiv \mathsf{Event}(\Theta_2)$, while differing in their test sets, $\mathsf{Test}(\Theta_1) \neq \mathsf{Test}(\Theta_2)$. In such a case, they represent distinct OPTs.

For the remaining of our discussion we will make the assumption that all outcome spaces are finite dimensional sets.

Assumption 1: Finite outcome spaces

For any test $\mathsf{T}_\mathsf{X} \in \mathsf{Test}(\Theta)$ in an OPT Θ, the cardinality of the outcome space X is assumed to be finite, i.e., $|\mathsf{X}| < \infty$.

However, in general, the framework does not exclude the possibility of also allowing infinite, possibly continuous, outcome spaces [134, 143].

In this thesis, we also assume that the order of the events in a test is significant.

Assumption 2: The order matters

If two tests consist of the same events in a different order—e.g., $\mathsf{T}_\mathsf{X} \equiv [\![\mathscr{T}_1, \mathscr{T}_2]\!]$ and $\mathsf{T}'_\mathsf{X} \equiv [\![\mathscr{T}_2, \mathscr{T}_1]\!]$—they are *not* equal. The only exception is when all the events of the tests coincide.

Remark 2

We highlight that, although Assumption 2 has been commonly assumed in previous works—for instance, in [158]—its rationality has been recently questioned [168]. Indeed, assuming that the order of the elements of two instruments makes them distinct lacks a proper operational justification, and might even introduce contextual dependence (see the discussion in the last part of section 2.2.4).

We also make the following remark regarding the notation we adopt.

Remark 3

Throughout this thesis we will employ two distinct notations when dealing with ensembles of events. The first notation, $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$, is used to explicitly denote tests. The alternative notation, $\{\mathscr{T}_x\}_{x\in\mathsf{X}}$, is instead employed to denote generic ensembles of events, which may not correspond to proper tests of an OPT. This distinction will be particularly useful when discussing generalised instruments (section 2.3.1) and limits thereof.

Remark 4

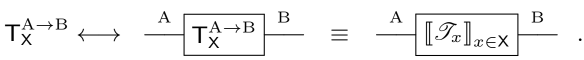
A subtle but important point that has to be clarified is related to the nature of systems. In the theoretical description of an experiment, it is often necessary to consider multiple systems that are operationally

indistinguishable—i.e., identical up to relabelling. For instance, one may study two photonic modes of the same kind, each admitting the same set of allowed operations. Although these are distinct instances of physical systems, they share the same operational behaviour.
To avoid interpretative ambiguities when such identical systems appear multiple times in a single experiment, it is convenient to regard every system as a representative of an equivalence class, called a *system type*. This abstraction separates a system's *identity* (a particular instance) from its *type* (the class of systems with the same operational structure). In this way, reasoning about experiments with repeated appearances of the same kind of system becomes unambiguous.
From now on, references to system types will be left implicit.

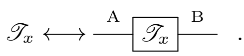
Diagrammatic representation

A particularly elegant feature of the framework is that it supports a fully diagrammatic representation of computations. In fact, every element of an OPT admits a graphical depiction. Systems are represented as wires, while tests and events are represented as boxes connected by these wires.

For example, a test $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{B})$ is represented as:



Likewise, a single event $\mathscr{T}_x \in \mathsf{Event}(\mathrm{A}\to\mathrm{B})$ within the test is represented by:



As a convention, the input–output direction in the diagrams is taken to go from left to right. However, this does not imply a preferred direction for the flow of information[3].

Now that we have defined the basic elements of our framework, we can discuss how they can be composed. This is an extremely important step in the construction of the framework since, as mentioned in the introduction, one of its key objectives is to describe physical theories that exhibit a compositional structure analogous to that of QT.

[3] In the subclass of *causal OPTs* (section 3.5), a preferred direction for the flow of information is instead fixed and conventionally chosen to be from left to right.

Sequential composition

The first requirement is that the framework must support *sequential composition* of tests. This operation models the performance of multiple operations (experiments) on the same physical system.

Formally, for any triple of systems $\mathrm{A}, \mathrm{B}, \mathrm{C} \in \mathsf{Sys}(\Theta)$, there exists a map

$$\Box : \mathsf{Test}(\mathrm{A} \to \mathrm{B}) \times \mathsf{Test}(\mathrm{B} \to \mathrm{C}) \to \mathsf{Test}(\mathrm{A} \to \mathrm{C})$$

called *sequential composition* such that for any $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{G}_{\mathsf{Y}}^{\mathrm{B}\to\mathrm{C}} \in \mathsf{Test}(\mathrm{B} \to \mathrm{C})$, the composite test is defined as

$$\mathsf{G}_{\mathsf{Y}}^{\mathrm{B}\to\mathrm{C}} \Box \mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \coloneqq (\mathsf{G} \Box \mathsf{T})_{\mathsf{X}\times\mathsf{Y}}^{\mathrm{A}\to\mathrm{C}} = (\mathsf{GT})_{\mathsf{X}\times\mathsf{Y}}^{\mathrm{A}\to\mathrm{C}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{C})\,, \tag{2.1}$$

We will often omit the composition symbol $\Box$ in favour of the more compact notation $(\mathsf{GT})_{\mathsf{X}\times\mathsf{Y}}^{\mathrm{A}\to\mathrm{C}}$.

As the definition suggests, we assume that the outcome space of the composite test is the Cartesian product of the outcome spaces of the component tests. That is, each outcome of the composite test is uniquely identified by an ordered pair (x, y) of outcomes from the component tests. This provides a natural one-to-one correspondence between the events of the composed test and pairs of events from the original tests.

In formulae, if $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$ and $\mathsf{G}_{\mathsf{Y}}^{\mathrm{B}\to\mathrm{C}} \equiv [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}}$, then:

$$[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \Box [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \coloneqq [\![\mathscr{G}_y \Box \mathscr{T}_x]\!]_{(x,y)\in\mathsf{X}\times\mathsf{Y}} = \left[\!\!\left[(\mathscr{G}\mathscr{T})_{(x,y)}\right]\!\!\right]_{(x,y)\in\mathsf{X}\times\mathsf{Y}}. \tag{2.2}$$

This induces the map for sequential composition at the level of events:

$$\Box : \mathsf{Event}(\mathrm{A} \to \mathrm{B}) \times \mathsf{Event}(\mathrm{B} \to \mathrm{C}) \to \mathsf{Event}(\mathrm{A} \to \mathrm{C})\,,$$

explicitly defined by

$$\mathscr{G}_y \Box \mathscr{T}_x \coloneqq \mathscr{G}\mathscr{T}_{(x,y)}. \tag{2.3}$$

The two maps, for tests and for events, are related via the identity:

$$(\mathsf{GT})_{\mathsf{X}\times\mathsf{Y}}^{\mathrm{A}\to\mathrm{B}} \coloneqq [\![\mathscr{G}_y \Box \mathscr{T}_x]\!]_{(x,y)\in\mathsf{X}\times\mathsf{Y}}\,. \tag{2.4}$$

One final assumption is needed to make this operation well-defined. What happens when one of the composed tests has a singleton outcome space? We assume that for any outcome space X,

$$\star \times \mathsf{X} = \mathsf{X} = \mathsf{X} \times \star, \tag{2.5}$$

and at the level of outcomes,

$$(*, x) = x = (x, *)\,, \quad \forall x \in \mathsf{X}. \tag{2.6}$$

In words, we identify outcome pairs involving the singleton element with the non-trivial outcome alone. This reflects the fact that the singleton outcome occurs with certainty and does not contribute additional information to the composite outcome.

Continuing with the example of the Stern–Gerlach experiment, composing two such experiments—possibly with different orientations of the magnetic field—results in a test with four outcomes. For instance, if the magnetic fields are orthogonal, the corresponding instrument is described by

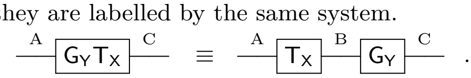
$$[\![\mathscr{T}_{\rightarrow}\mathscr{T}_{\uparrow},\ \mathscr{T}_{\leftarrow}\mathscr{T}_{\uparrow},\ \mathscr{T}_{\rightarrow}\mathscr{T}_{\downarrow},\ \mathscr{T}_{\leftarrow}\mathscr{T}_{\downarrow}]\!]\,.$$

Similarly, a photonic mode undergoing a unitary evolution $\mathscr{U}$, followed by a beam splitter and a photodetector—described by a Positive Operator-Valued Measure (POVM) $[\![\mathrm{a}_1, \mathrm{a}_2]\!]$—yields a new POVM with two possible outcomes:

$$[\![\mathrm{a}_1\mathscr{U},\ \mathrm{a}_2\mathscr{U}]\!]\,.$$

The application of the unitary evolution does not change the number of outcomes of the experiment.

Following the convention that the input-output direction is diagrammatically represented as going from left to right, the operation of sequential composition is depicted by connecting the output wire of the first process to the input wire of the second, provided they are labelled by the same system.

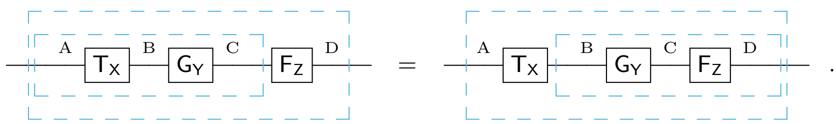


We highlight that, following our convention, the diagrammatic order of composition appears reversed compared to the traditional notation used in formulas, where the information flow is from right to left.

Within the framework, the operation of sequential composition is also required to satisfy the *associativity* property. This means that when composing multiple operations—specifically, three or more—the outcome is independent of how the operations are grouped. For any three tests $\mathsf{T_X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, $\mathsf{G_Y} \in \mathsf{Test}(\mathrm{B} \to \mathrm{C})$, and $\mathsf{F_Z} \in \mathsf{Test}(\mathrm{C} \to \mathrm{D})$, it holds that:

$$\mathsf{F_Z}\,(\mathsf{G_Y}\mathsf{T_X}) = (\mathsf{F_Z}\mathsf{G_Y})\,\mathsf{T_X},$$

and diagrammatically:



Associativity captures the idea that the way we group operations is merely a mathematical requirement, without any operational or physical meaning.

Note that for any triple of outcome spaces X, Y, $\mathsf{Z} \in \mathsf{Out}(\Theta)$, there is a natural one-to-one correspondence between the compound outcome spaces $(\mathsf{X} \times \mathsf{Y}) \times \mathsf{Z}$ and $\mathsf{X} \times (\mathsf{Y} \times \mathsf{Z})$, given by the identification:

$$((x, y), z) \leftrightarrow (x, y, z) \leftrightarrow (x, (y, z)), \quad \forall x \in \mathsf{X}, \forall y \in \mathsf{Y}, \forall z \in \mathsf{Z}. \tag{2.7}$$

Associativity for sequential composition of tests (2.1) induces associativity also for the sequential composition of events. For any three events $\mathscr{T}_x \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$, $\mathscr{G}_y \in \mathsf{Event}(\mathrm{B} \to \mathrm{C})$, and $\mathscr{F}_z \in \mathsf{Event}(\mathrm{C} \to \mathrm{D})$—where we are supposing that these events belong to appropriate tests of the theory—, it holds that

$$\mathscr{F}_z (\mathscr{G}_y \mathscr{T}_x) = (\mathscr{F}_z \mathscr{G}_y) \mathscr{T}_x. \tag{2.8}$$

The framework also requires the *presence of an identity element* for the operation of sequential composition. This is realised by the presence of a family of tests I_Θ, called *identities*. For any system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, there exists a test $\mathsf{I}^{\mathrm{A}\to\mathrm{A}} \in \mathsf{I}_\Theta \subset \mathsf{Test}(\mathrm{A} \to \mathrm{A})$, called *identity for system* A, such that, for any system $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and for any test $\mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$:

$$\mathsf{I}^{\mathrm{B}\to\mathrm{B}} \mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}} = \mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}} = \mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}} \mathsf{I}^{\mathrm{A}\to\mathrm{A}}. \tag{2.9}$$

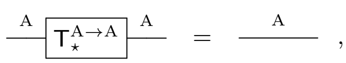
The identity represents the operation of doing nothing to a physical system. Operationally, it models situations in which the system undergoes no interaction, such as propagation through a non-interactive medium, or more generally the mere persistence of the system without any process affecting its state.

Using (2.9), one can show that the identity test for any system is unique. Moreover, by exploiting the relationship (2.4), it follows that identity tests are singleton: $\mathsf{I}_\star^{\mathrm{A}\to\mathrm{A}} := [\![\mathscr{I}_\mathrm{A}]\!]$, for any system A. By the correspondence between maps and events (via (2.6)), the identity condition for events becomes:

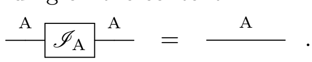
$$\mathscr{I}_\mathrm{B} \mathscr{T}_x = \mathscr{T}_x = \mathscr{T}_x \mathscr{I}_\mathrm{A}, \quad \forall \mathscr{T}_x \in \mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}}, \quad \forall \mathsf{T}_\mathsf{X}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B}). \tag{2.10}$$

By the uniqueness of the identity test, the identity event is also unique.

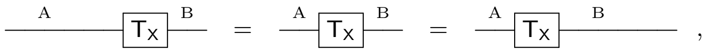
Diagrammatically the identity tests and events will not be explicitly represented in wired boxes, adopting the contracted notation in which identities tests and events will be simply represented as wires. The wire will represent the test:



or the identity event depending on the context:

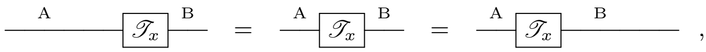


The uniqueness of identity tests and events implies a one-to-one correspondence with the systems of the theory. As a result, operations can "slide" freely along identity wires:



(2.11)

for any test $\mathsf{T}_\mathsf{X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, and similarly



for any event $\mathscr{T}_x \in \mathsf{T}_\mathsf{X}$.

Parallel composition

In addition to describing how experiments can be composed in sequence, the framework must also account for the possibility of performing experiments *simultaneously* on different physical systems. This models, for example, the situation in which two physicists carry out independent experiments in separate laboratories at the same time.

To formalize this scenario, the framework first requires the ability to describe *composite systems*. For any two systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, their composite is denoted by $\mathrm{AB} \in \mathsf{Sys}(\Theta)$. This composition is captured by a map

$$\boxtimes : \mathsf{Sys}(\Theta) \times \mathsf{Sys}(\Theta) \to \mathsf{Sys}(\Theta)\,.$$

However, we will always omit the operator $\boxtimes$, favouring the compact notation AB for composite systems.

Furthermore, we require the map to be *associative*

$$(AB)\,C = A\,(BC) \tag{2.12}$$

and the existence of a special system that serves as the *identity element* for parallel composition. This system is denoted by I and is called the *trivial system*. By definition, it satisfies the condition:

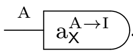
$$\mathrm{IA} = \mathrm{A} = \mathrm{AI}, \quad \forall \mathrm{A} \in \mathsf{Sys}(\Theta)\,. \tag{2.13}$$

Using the latter property (2.13), it is immediate to show that the trivial system I is unique.

From the point of view of interpretation, the trivial system I represents "nothing the theory cares to describe" [142]. This does not necessarily mean that it encodes inaccessible or hidden information, but rather that it corresponds to degrees of freedom in which the physicist has no interest. A typical example of a physical system described by the trivial system is one that is discarded after a measurement.

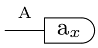
Consider a particle hitting a detector. We might model the measurement process as an evolution $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}$ from a system A—the particle—to a classical system B representing the measurement apparatus. However, in many such scenarios, we are not interested in keeping track of the post-measurement state of the apparatus: we care only about the measurement outcome $x \in \mathsf{X}$. In such cases, it is more appropriate to describe the process as a test from A to the trivial system I.

Tests of this kind are called *observation-tests*: $\mathrm{a}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{I}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{I})$. To distinguish them from generic tests, observation-tests are denoted using lowercase Roman letters. Since the trivial system is not represented diagrammatically, observation-tests are depicted with a rounded edge in place of the output wire:



.

The events composing an observation-test are called *observations*:

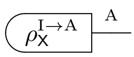


$$, \quad \forall \mathrm{a}_x \in \mathrm{a}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{I}} \equiv [\![\mathrm{a}_x]\!]_{x\in\mathsf{X}} \,.$$

Similarly, there are tests whose input system is the trivial one:

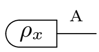
$$\rho_{\mathsf{X}}^{\mathrm{I}\to\mathrm{A}} \in \mathsf{Test}(\mathrm{I} \to \mathrm{A}) \,.$$

These are called *preparation-tests*, and they model the preparation of a system in a given ensemble of states. Diagrammatically, preparation-tests are represented with a rounded edge in place of the input wire:



.

The events composing a preparation-test are called *preparations*:



$$, \quad \forall \rho_x \in \rho_{\mathsf{X}}^{\mathrm{I}\to\mathrm{A}} \equiv [\![\rho_x]\!]_{x\in\mathsf{X}} \,.$$

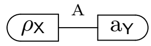
To distinguish preparation-tests from general tests, we denote them using Greek letters.

Preparations and observation-tests generalise the concepts of density matrices and POVMs, respectively, from QT.

To write the equations not in diagrammatic form, we introduce the following notation: the *round ket* $|\cdot)$ denotes a preparation, while the *round bra* $(\cdot|$ denotes an observation.

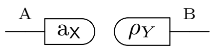
Remark 5

In any OPT, it is always assumed that for every system A, the sets of preparation-tests $\mathsf{Test}(\mathrm{I} \to \mathrm{A})$ and observation-tests $\mathsf{Test}(\mathrm{A} \to \mathrm{I})$ are non-empty. This assumption ensures that the framework can always describe at least *prepare-and-measure scenarios*:



.

Furthermore, this also guarantees within an OPT the existence of tests with non-trivial input and output systems:



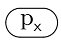
$\in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$,

for any A, B of the theory. Sometimes, this test will be referred to as *measure-and-re-prepare*.

Last but not least, we have *scalar-tests*, whose events are called *scalars*. These are tests whose input and output systems are both the trivial system:

$$\mathrm{p}_x \in \mathsf{Event}(\mathrm{I} \to \mathrm{I}) , \quad \forall \mathrm{p}_x \in \mathrm{p}_{\mathsf{X}}^{\mathrm{I} \to \mathrm{I}} \equiv [\![\mathrm{p}_x]\!]_{x \in \mathsf{X}} \in \mathsf{Test}(\mathrm{I} \to \mathrm{I}) .$$

Diagrammatically, scalar-tests are depicted as rounded boxes without input or output wires. When convenient, they may also be represented simply by their label, without any graphical structure:



, p_{x}.

Now that we have a well-defined notion of parallel composition of systems, it is natural to ask how one can act *independently* on different systems within a composite one. In particular, we are interested in describing *local operations* that act on subsystems independently. This leads to the notion of *parallel composition for tests*.

Given any two arbitrary tests $\mathsf{T}_{\mathsf{X}}^{\mathrm{A} \to \mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{G}_{\mathsf{Y}}^{\mathrm{C} \to \mathrm{D}} \in \mathsf{Test}(\mathrm{C} \to \mathrm{D})$, their parallel composition is defined as

$$\mathsf{T}_{\mathsf{X}}^{\mathrm{A} \to \mathrm{B}} \boxtimes \mathsf{G}_{\mathsf{Y}}^{\mathrm{C} \to \mathrm{D}} := (\mathsf{T} \boxtimes \mathsf{G})_{\mathsf{X} \times \mathsf{Y}}^{\mathrm{AC} \to \mathrm{BD}} .$$

This composite test belongs to $\mathsf{Test}(\mathrm{AC} \to \mathrm{BD})$ and describes the simultaneous execution of two tests on independent systems, i.e., experimenters performing independent experiments on different physical systems. We refer to such tests as *compound-local-tests*—an analogous terminology is used for their events.

Formally, the *parallel composition map* is defined as

$$\boxtimes : \mathsf{Test}(\mathrm{A} \to \mathrm{B}) \times \mathsf{Test}(\mathrm{C} \to \mathrm{D}) \to \mathsf{Test}(\mathrm{AC} \to \mathrm{BD}) ,$$

which exists for any quadruple of systems A, B, C, D $\in \mathsf{Sys}(\Theta)$.

As in the case of sequential composition, we assume that the outcome set of the composite test is the Cartesian product of the outcome sets of the individual tests—that is, each pair of outcomes labels a distinct composite event. Furthermore, we require adherence to (2.5) and (2.6).

The parallel composition of tests naturally induces a corresponding operation on events. For any quadruple of systems A, B, C, D $\in \mathsf{Sys}(\Theta)$, the map

$$\boxtimes : \mathsf{Event}(\mathrm{A} \to \mathrm{B}) \times \mathsf{Event}(\mathrm{C} \to \mathrm{D}) \to \mathsf{Event}(\mathrm{AC} \to \mathrm{BD})$$

is defined by

$$\mathscr{T}_x \boxtimes \mathscr{G}_y := (\mathscr{T} \boxtimes \mathscr{G})_{(x,y)} .$$

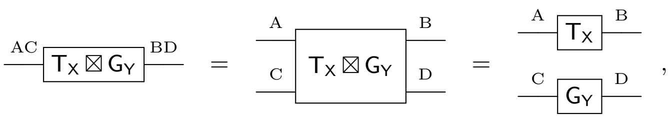
The relationship between the composite test and its events is then given by:

$$(\mathsf{T} \boxtimes \mathsf{G})^{\mathrm{AC}\to\mathrm{BD}}_{\mathsf{X}\times\mathsf{Y}} := \left[\!\!\left[(\mathscr{T} \boxtimes \mathscr{G})_{(x,y)} \right]\!\!\right]_{(x,y)\in\mathsf{X}\times\mathsf{Y}} .$$

An example of such a test is the simultaneous application of two Stern–Gerlach experiments to two independent beams of particles. This results in a four-outcome compound experiment, corresponding to the four possible joint outcomes:

$$[\![\mathscr{T}_\uparrow \mathscr{G}_\uparrow, \mathscr{T}_\uparrow \mathscr{G}_\downarrow, \mathscr{T}_\downarrow \mathscr{G}_\uparrow, \mathscr{T}_\downarrow \mathscr{G}_\downarrow]\!] .$$

Diagrammatically, the parallel composition of tests is depicted as:

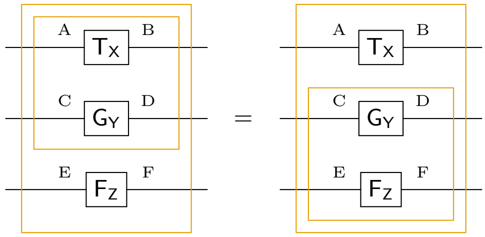


for any tests $\mathsf{T}_\mathsf{X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{G}_\mathsf{Y} \in \mathsf{Test}(\mathrm{C} \to \mathrm{D})$. The same diagrammatic notation is adopted for events.

By convention, tests and events that are composed in parallel are diagrammatically represented by translating the corresponding inline equation, as read from left to right, into a diagram drawn from top to bottom. We will see, however, that it is always possible to exchange this order.

As with sequential composition, the parallel composition must satisfy *associativity*. That is, for any three tests $\mathsf{T}_\mathsf{X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, $\mathsf{G}_\mathsf{Y} \in \mathsf{Test}(\mathrm{C} \to \mathrm{D})$, and

$\mathsf{F_Z} \in \mathsf{Test}(\mathrm{E} \to \mathrm{F})$, the following holds:



, (2.14)

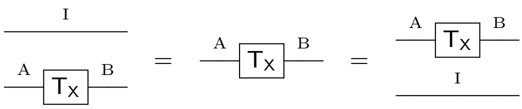
and the analogous identity holds at the level of events:

$$(\mathscr{T}_x \boxtimes \mathscr{G}_y) \boxtimes \mathscr{F}_z = \mathscr{T}_x \boxtimes (\mathscr{G}_y \boxtimes \mathscr{F}_z)\,,$$

for any $\mathscr{T}_x \in \mathsf{T_X}$, $\mathscr{G}_y \in \mathsf{G_Y}$, and $\mathscr{F}_z \in \mathsf{F_Z}$.

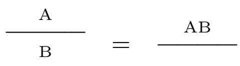
We also assume that the identification of outcome spaces under parallel composition follows the rule given in (2.7).

The identity element for parallel composition is given by the identity test on the trivial system: $\mathscr{I}_{\mathrm{I}}$. That is, for any systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and any test $\mathsf{T_X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, one has:



. (2.15)

Moreover, identities must compose in the following way:



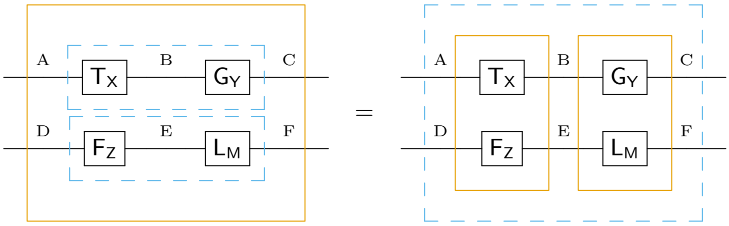
,

expressing that doing nothing separately to systems A and B amounts to doing nothing to their composite AB.

As far as reconstructing the compositional structure of QT, we are doing well: we have defined two operations—sequential and parallel composition—and required that the set of tests (and consequently that of events) is closed under both. Furthermore, the parallel composition map also applies to systems, and we assume that $\mathsf{Sys}(\Theta)$ is closed under this operation.

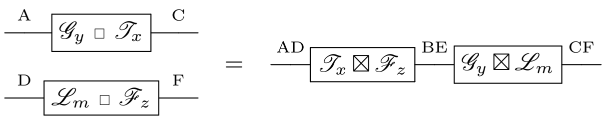
In the Hilbert space formulation of QT, sequential composition corresponds to the usual composition of functions, while parallel composition—both for systems (Hilbert spaces) and operations (CPTNI maps)—is implemented via the (algebraic) tensor product.

To complete the compositional structure, two final requirements are needed. The first one is that the sequential and parallel composition operations commute. For any systems $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F} \in \mathsf{Sys}(\Theta)$ and tests $\mathsf{T}_\mathsf{X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, $\mathsf{G}_\mathsf{Y} \in \mathsf{Test}(\mathrm{B} \to \mathrm{C})$, $\mathsf{F}_\mathsf{Z} \in \mathsf{Test}(\mathrm{D} \to \mathrm{E})$, and $\mathsf{L}_\mathsf{M} \in \mathsf{Test}(\mathrm{E} \to \mathrm{F})$, we require:



, (2.16)

or, equivalently



,

where here we expressed the relationship in the case of events. In words, this expresses that one can either first compose in parallel and then in sequence, or vice versa, and obtain the same result.

We highlight that, as in the case of (2.6) and (2.7), the commutativity requirement (2.16) also induces a corresponding identification of the outcome spaces:

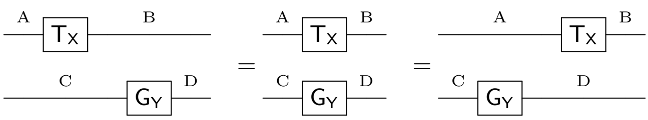
$$(\mathsf{X} \times \mathsf{Y}) \times (\mathsf{Z} \times \mathsf{M}) \longleftrightarrow (\mathsf{X} \times \mathsf{Z}) \times (\mathsf{Y} \times \mathsf{M}) \,,$$

which, at the level of individual outcomes, reads:

$$\begin{pmatrix} (x, y) \\ (z, m) \end{pmatrix} \longleftrightarrow \left(\begin{pmatrix} x \\ z \end{pmatrix}, \begin{pmatrix} y \\ m \end{pmatrix} \right) \longleftrightarrow (x, y, z, m) \longleftrightarrow (x, z, y, m) \,,$$

for any $x \in \mathsf{X}$, $y \in \mathsf{Y}$, $z \in \mathsf{Z}$, and $m \in \mathsf{M}$.

Thanks to (2.16), we can now further enhance our ability to "slide" tests—and consequently events—along the wires, extending the principle from (2.11):



. (2.17)

The identities in (2.17) demonstrate that the positioning of boxes on different wires carries no operational significance. The only relevant information is the *input–output connectivity*—that is, which wires are linked by which tests. There is no underlying *spatiotemporal* structure assumed in these diagrams. In other words, we do not interpret the diagram as embedded in a space-time: the only meaningful relations between tests are those induced by their connections in the network.

Furthermore, another consequence of (2.17) is that any compound-local-test can always be decomposed as the sequential composition of two tests of the form $\mathsf{T}_\mathsf{X} \boxtimes \mathsf{I}_\star \in \mathsf{Test}(\mathrm{AC} \to \mathrm{BC})$ and $\mathsf{I}_\star \boxtimes \mathsf{G}_\mathsf{Y} \in \mathsf{Test}(\mathrm{AC} \to \mathrm{AD})$. These will be referred to as the *local-test of system* A, and the *local-test of system* C, *from system* AC, respectively:

$$(\mathsf{T}_\mathsf{X} \boxtimes \mathsf{I}_\star) \mathbin{\square} (\mathsf{I}_\star \boxtimes \mathsf{G}_\mathsf{Y}) = \mathsf{T}_\mathsf{X} \boxtimes \mathsf{G}_\mathsf{Y} = (\mathsf{I}_\star \boxtimes \mathsf{G}_\mathsf{Y}) \mathbin{\square} (\mathsf{T}_\mathsf{X} \boxtimes \mathsf{I}_\star) \,.$$

Reversible tests and events

To conclude our discussion on the compositional structure, we need to introduce a particular class of tests and events: the family of *reversible tests* and *reversible events*.

Definition 1 (Reversible tests and events)

A test $\mathsf{R}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ is said to be *reversible* if there exists another test $\left(\mathsf{R}^{-1}\right)^{\mathrm{B}\to\mathrm{A}}$, called the *inverse*, such that:

$$\left(\mathsf{R}^{-1}\right)^{\mathrm{B}\to\mathrm{A}} \mathsf{R}^{\mathrm{A}\to\mathrm{B}} = \mathsf{I}_\star^{\mathrm{A}\to\mathrm{A}},$$

and

$$\mathsf{R}^{\mathrm{A}\to\mathrm{B}} \left(\mathsf{R}^{-1}\right)^{\mathrm{B}\to\mathrm{A}} = \mathsf{I}_\star^{\mathrm{B}\to\mathrm{B}}.$$

Analogously, an event $\mathscr{R} \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$ is said to be *reversible* if there exists an event $\mathscr{R}^{-1} \in \mathsf{Event}(\mathrm{B} \to \mathrm{A})$, also called the *inverse*, such that:

$$\mathscr{R}^{-1}\mathscr{R} = \mathscr{I}_\mathrm{A}, \qquad \mathscr{R}\mathscr{R}^{-1} = \mathscr{I}_\mathrm{B}.$$

From this definition, the following properties of reversible tests and events immediately follow:

Lemma 1

Every reversible test is singleton, and every reversible event is deterministic.

Lemma 2

Reversibility is preserved under both sequential and parallel composition.

Lemma 3

For any system A in an OPT Θ, the set of reversible tests from A to itself forms a group. The same holds for reversible events.

In our discussion thus far, we have already encountered a reversible test: the identity.

The introduction of reversible test also allows to define the notion of *operational equivalence* between systems. Two systems A and B are said to be operationally equivalent

$$\mathrm{A} \cong \mathrm{B} \tag{2.18}$$

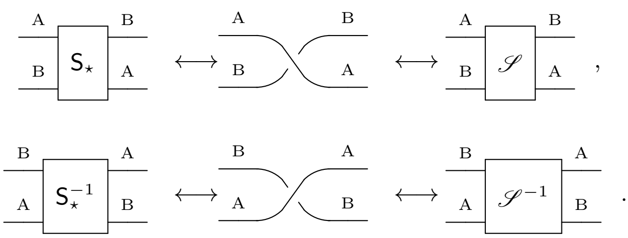
if there exists a reversible test $\mathsf{R} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, i.e., whose input and output systems are A and B or vice versa.

Braiding

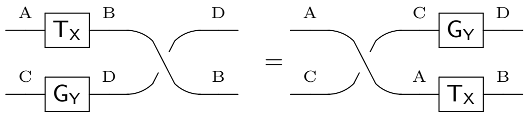
We are now in a position to state the final requirement that completes the compositional structure of OPTs, making it structurally analogous to that of QT. We require the existence of a family of reversible tests, called *braidings*, which allow two agents to exchange systems. More precisely, given any two systems A, $\mathrm{B} \in \mathsf{Sys}(\Theta)$, there exist two singleton reversible tests:

- the braiding $\mathsf{S}_\star^{\mathrm{AB}\to\mathrm{BA}} = \{\mathscr{S}_{\mathrm{A,B}}\}$, and
- its inverse $\left(\mathsf{S}^{-1}\right)_\star^{\mathrm{BA}\to\mathrm{AB}} = \left\{\mathscr{S}^{-1}_{\mathrm{A,B}}\right\}$.

They are represented diagrammatically as:



These tests must satisfy the *naturality* condition[4]:



, (2.19)

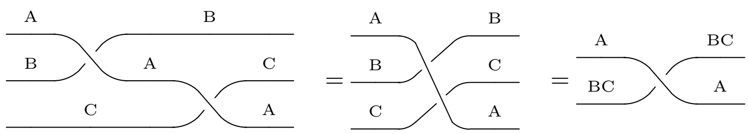
meaning that tests and events can slide along the wires even when those wires are intertwined by a braiding operation. In other words, equation (2.11) must continue to hold in the presence of braiding.

As in all analogous situations encountered so far, equation (2.19) also induces a natural one-to-one correspondence between the outcome spaces:

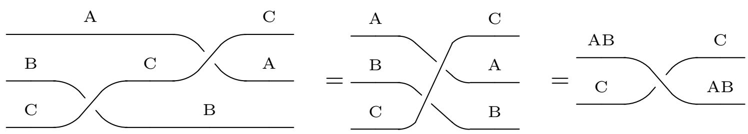
$$\mathsf{X} \times \mathsf{Y} \longleftrightarrow \mathsf{Y} \times \mathsf{X},$$

where the identification of outcomes (x, y) and (y, x) is intended.

Last, but not least, the braiding is also required to satisfy the *hexagon identities*:



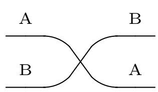
, (2.20a)



, (2.20b)

for all A, B, C $\in \mathsf{Sys}(\Theta)$. These identities encode the behaviour of braiding in the presence of composite systems. They state that when exchanging multiple systems, performing the exchange individually or with composite systems yields the same overall transformation.

An OPT in which the braiding satisfies $\mathscr{S}_{\mathrm{A,B}} = \mathscr{S}^{-1}_{\mathrm{A,B}}$ for all system pairs is called *symmetric*. In this case, the braiding reduces to a transposition—also referred to as the *swap* operation—and is represented diagrammatically as:



.

[4]In other words, requiring that the braiding satisfies the naturality property means that we are requiring it to be a natural transformation (Definition 68).

In the case of symmetric theories the two hexagon identities (2.20a) and (2.20b) are equivalent, as can be easily shown by direct calculation [169–171].

Observation 1

To date, all OPTs defined in the literature are symmetric [98, 133, 158, 159, 167, 172]. Nonetheless, since the framework permits non-symmetric structures, it would be of interest to construct an explicit example of a non-symmetric OPT. Such a construction could shed light on the role and physical significance of the symmetry assumption itself.
For the moment, there is no justification for whether non-symmetric theories should be allowed or not, beyond the desire for the framework to be formulated in its most general form possible. In this sense, making the symmetry assumption explicit allows one to recognize it as a genuine assumption, reducing the number of implicit biases introduced in the construction of a theory.

Remark 6 (On the structure of OPTs' diagrams)

Since the OPT framework is designed to model general physical experiments, one must find a way to compress the information associated with the full four-dimensional *spatio-temporal* configuration of an experiment into a two-dimensional diagram. One prominent example of this compression is the fact that unordered events occurring in a three-dimensional space must be arranged in a top-to-bottom order when represented via the parallel composition map, thereby introducing a degree of arbitrariness in the diagram's encoding. However, this arbitrariness is resolved thanks to the presence of the braiding operation. Since this operation is reversible, any ordering imposed for drawing purposes becomes irrelevant, as it can always be adjusted using braiding. As a result, even the representation of spatial relationships becomes flexible. When constructing a diagram, one is free to arrange its components in whatever layout provides the most clarity, without altering its formal meaning.
Regarding temporal relations, we recall that these are not explicitly represented in the diagrams. The only form of "temporal information" encoded is the input-output structure of the processes. For this to be well-defined, unconnected boxes must commute, as formalized by the sliding rule (2.17).
In summary, even though we are constrained to representing diagrams on a two-dimensional surface for the sake of communication, the representation remains highly flexible—in particular, it is *spatially isotopic*—allowing for

any reasonable rearrangement. Diagrams can be thought of as living in a three-dimensional environment without incurring in any contradiction. For a more formal treatment, see appendix A.3.

Relationship with category theory

The structure we have been discussing so far corresponds to that of a *braided strict monoidal category* [169–171]. In particular, the set of all events of an OPT, denoted $\mathsf{Event}(\Theta)$, forms a category of this kind, which we denote by Event_{Θ} [143]. The objects of this category are the systems of the theory: $\mathsf{Obj}(\mathsf{Event}_{\Theta}) = \mathsf{Sys}(\Theta)$, while the morphisms are given by the events: $\mathsf{Mor}(\mathsf{Event}_{\Theta}) = \mathsf{Event}(\Theta)$. The composition of morphisms is given by the associative binary operation $\square$, and the identity morphisms are given by the family $\{\mathscr{I}_{\mathrm{A}}\}_{\mathrm{A}\in\mathsf{Sys}(\Theta)}$. The *monoidal structure* arises naturally from the associative operation of parallel composition $\boxtimes$, and the *braided structure* is provided by the family of braiding operations $\{\mathscr{S}_{\mathrm{A},\mathrm{B}}\}_{\mathrm{A},\mathrm{B}\in\mathsf{Sys}(\Theta)}$.

One of the key advantages of establishing a well-defined connection with category theory is that the framework can directly inherit known results for categories. For instance, braided monoidal categories admit a sound and complete graphical calculus (Theorem 44)—a result that guarantees the well-definiteness of the diagrammatic representation in OPTs as a legitimate method of calculation.

For a more in-depth discussion of categorical notions, we refer the interested reader to appendix A, where we also explain what does it mean that Event_{Θ} is *strict*.

2.2.2 The probabilistic structure of a theory

The structure we have described so far is merely a descriptive tool. It allows one to model any experimental setup in the form of a Directed Acyclic Graph (DAG). A DAG is characterized by two sets: the finite set of *nodes* composing the graph, denoted $\mathsf{Nodes}(G)$, and the set of *directed edges*, $\mathsf{Edges}(G) \subseteq \mathsf{Nodes}(G) \times \mathsf{Nodes}(G)$, meaning that an edge is an ordered pair of nodes. Furthermore, the graph is *acyclic*, which means there is no way to start and end at the same node by traversing edges forward. In the case of OPTs diagrams, the nodes correspond to the test (or event) boxes, while the edges to the system wires.

To enable the OPT framework to make predictions about the outcomes of experiments, we must equip it with a *probabilistic structure*. Specifically, we require that any *closed circuit*—that is, an acyclic diagram whose overall input and output systems are both the trivial system, i.e., one that begins with a preparation-test and ends with an observation-test—is associated with a probability distribution over the outcome space.

In particular, to any event composing the overall test is associated a probability:

$$\mathsf{Event}(\mathrm{I} \to \mathrm{I}) \ni \mathrm{G}\left(\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z\right) := \mathrm{p}\left(x, y, \ldots, z \,|\, \mathrm{G}\left(\mathsf{T}_\mathsf{X}, \mathsf{G}_\mathsf{Y}, \ldots, \mathsf{F}_\mathsf{Z}\right)\right), \quad (2.21)$$

where $\mathrm{G}\left(\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z\right)$ is the notation we adopt to denote a closed circuit composed by the test $\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z$ within inline formulas.

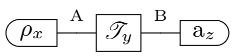
On the right-hand side of (2.21), we explicitly indicate the dependence of the probability distribution associated to a circuit on the explicit form of the graph. However, for clarity, we will almost always use the more compact notation

$$\mathrm{p}\left(x, y, \ldots, z \,|\, \mathsf{T}_\mathsf{X}, \mathsf{G}_\mathsf{Y}, \ldots, \mathsf{F}_\mathsf{Z}\right),$$

leaving the dependence on the structure of the circuit implicit.

In words, equation (2.21) provides the probability of occurrence of any sequence of events $\mathrm{G}\left(\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z\right)$, given that the experiment $\mathrm{G}\left(\mathsf{T}_\mathsf{X}, \mathsf{G}_\mathsf{Y}, \ldots, \mathsf{F}_\mathsf{Z}\right)$ is performed and the outcomes $x, y, \ldots, z$ are observed.

For example:



$$:= \mathrm{p}\left(x, y, z \,|\, \rho_\mathsf{X}, \mathsf{T}_\mathsf{Y}, \mathrm{a}_\mathsf{Z}\right), \quad (2.22)$$

where ρ_x is an event of the test ρ_X, and analogously for the others.

Within the framework we always assume that the zero element is included among the scalar events.

Summarising the requirements made on the probabilistic structure, for all $\mathrm{P}_\mathsf{X}^{\mathrm{I}\to\mathrm{I}} \in \mathsf{Test}(\mathrm{I} \to \mathrm{I})$ the following holds:

$$\forall \tilde{x} \in \mathsf{X}, \mathrm{p}_{\tilde{x}} \in [0,1] : \sum_{\tilde{x} \in \mathsf{X}} \mathrm{p}_{\tilde{x}} = 1, \quad 0 \in \mathsf{Event}(\mathrm{I} \to \mathrm{I}). \quad (2.23)$$

Remark 7

A direct consequence of (2.23) is that the deterministic scalar event is equal to 1. Since determinism is preserved under both sequential and parallel composition, it follows that composing deterministic preparations, transformations, and observations into a closed circuit yields the scalar 1. In other words, such a closed circuit corresponds to an event that occurs with certainty.

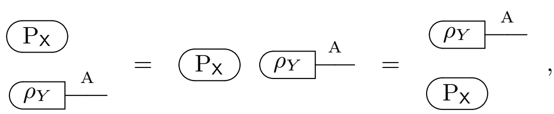
Composition of scalars

Following the definition of the probabilistic structure, the notion of scalar-test becomes central to the framework since it represents the bridge between the theoretical framework of OPTs and the real world. It is only the probability distributions of experiments that are overall described by scalar-tests that can be confronted with data from real-world experiments.

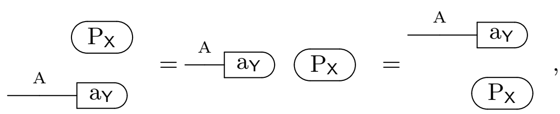
For this correspondence to be well-defined, independence of scalar-tests must be reflected at the probabilistic level. In particular, if two scalar-tests are independent, their joint probability distribution must factorise as the product of the probability distributions associated with the two scalar-tests taken individually.

This requirement does not follow from any other structural principle of the framework and must therefore be postulated.

The miracle of scalars To formally articulate this requirement, we must first understand how the composition of scalar tests behaves within the OPT framework. From the sliding rule (2.11) and the fact that the trivial system acts as the identity for parallel composition (2.15), it immediately follows that:

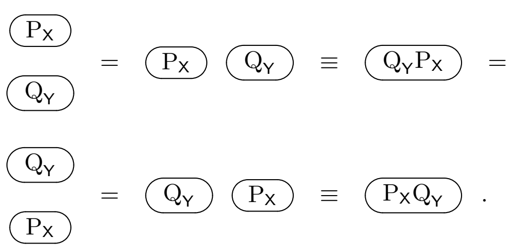


and



for any appropriate test.

In particular, when we restrict attention to scalar-tests alone, these identities yield the following equivalence:



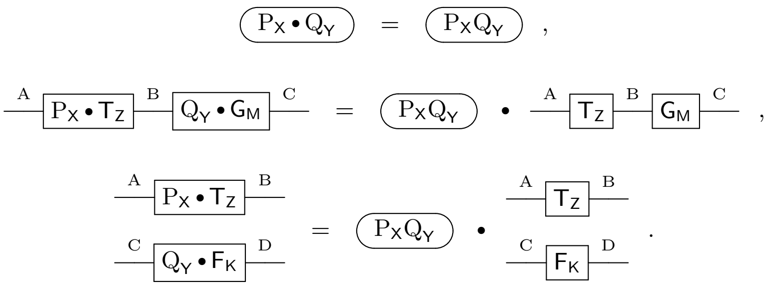
(2.24)

In words, for scalar-tests there is no distinction between sequential and parallel composition. Their combination yields the same scalar regardless of the order or manner in which they are composed.

These observations motivate the formal introduction of an operation called *scalar multiplication*:

$$\bullet : \mathsf{Test}(\mathrm{I} \to \mathrm{I}) \times \mathsf{Test}(\mathrm{A} \to \mathrm{B}) \to \mathsf{Test}(\mathrm{A} \to \mathrm{B}) \,, \tag{2.25}$$

which encodes the composition of general tests with scalar-tests. This operation satisfies the following properties:



Observation 2 (The miracle of scalars)

The title of this subsection—"the miracle of scalars"—refers to the name given by Abramsky and Coecke to the fact that, in any strict monoidal category, the set of endomorphisms of the monoidal unit $(\mathsf{Mor}(\mathrm{I},\mathrm{I})\,,\bullet)$ forms a *commutative monoid* [173][a].
This guarantees that scalar multiplication is a well-defined algebraic operation. Scalar-tests can be composed in any order, and that their composition is associative, commutative, and has an identity element.

[a] We recall that a *monoid* is a set equipped with a binary operation that is associative and has an identity element. A *commutative monoid* is a monoid where the binary operation is also commutative.

Probability of compound scalar-tests We now have all the elements to state that, in any OPT, for any pair of independent scalar-tests $\mathrm{P}_{\mathsf{X}}, \mathrm{Q}_{\mathsf{Y}} \in \mathsf{Test}(\mathrm{I} \to \mathrm{I})$ and any pair of outcomes $x \in \mathsf{X}$ and $y \in \mathsf{Y}$, their composition yields a scalar whose probability is given by the usual product of real numbers:

$$\mathrm{p}\left((x,y)\,|\,\mathrm{P}_{\mathsf{X}} \bullet \mathrm{Q}_{\mathsf{Y}}\right) \equiv \mathrm{p}_x \bullet \mathrm{q}_y \coloneqq \mathrm{p}_x \cdot \mathrm{q}_y \equiv \mathrm{p}\left(x\,|\,\mathrm{P}_{\mathsf{X}}\right) \cdot \mathrm{p}\left(y\,|\,\mathrm{Q}_{\mathsf{Y}}\right). \tag{2.26}$$

Here we used the identification $\mathrm{p}_x \equiv \mathrm{p}\left(x \,|\, \mathrm{P}_{\mathsf{X}}\right)$, and the operation of scalar multiplication to emphasise that the result holds independently of how the scalars are composed (see (2.24)).

In other words, this captures the fact that within the OPT framework different closed circuits represent distinct and uncorrelated experiments. As a consequence, their associated scalar events are independent, and their joint probability factorises as in (2.26).

Remark 8

Closed circuits—that is, scalars—are the only elements in the framework to which one can assign a probability independently of the experimental context in which they appear:

$$\mathrm{p}\left(x \,|\, \mathrm{P}_{\mathsf{X}} \bullet \mathrm{Q}_{\mathsf{Y}}\right) = \mathrm{p}\left(x \,|\, \mathrm{P}_{\mathsf{X}}\right) = \mathrm{p}\left(x \,|\, \mathrm{Q}_{\mathsf{Y}} \bullet \mathrm{P}_{\mathsf{X}}\right),$$

where $\mathrm{P}_{\mathsf{X}}, \mathrm{Q}_{\mathsf{Y}} \in \mathsf{Test}(\mathrm{I} \to \mathrm{I})$, and we are ignoring which outcome is obtained from Q_{Y}. As we will see in section 2.2.3, the framework allows one to ignore outcomes by means of the coarse-graining operation.
In contrast, for a generic (non-scalar) test, it is not possible to assign a probability to an event without reference to the complete circuit in which it appears. As shown in (2.22), one cannot assign a probability to an event of the test T_{Y} in isolation—one must also consider the preparation-test ρ_{X} and the observation-test a_{Y}[a].

[a]More precisely, we will see that in causal OPTs, it is also possible to assign meaningful probabilities to individual preparation events. For further details, see Remark 19.

2.2.3 Coarse-graining: filtering out the uninteresting

With the addition of the probabilistic structure, we have achieved the goal outlined at the beginning of this chapter: to construct a framework capable of describing general physical theories based solely on two primitives—the operations that can be performed in a laboratory, and the rules for making predictions about them. There is now one last requirement that we make on the structure of the framework related to the possibility of discarding information in an experiment.

Suppose we perform an experiment described by the test $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$. We already know that the information about the possible outcomes is encoded in the outcome space $\mathsf{X} \equiv \{x\}$. However, sometimes we are interested in ignoring part of this information. This is, for example, a standard procedure in particle physics, where experiments such as those conducted at CERN generate so much data that, without proper compression that retains only the relevant parts, storing it all would be infeasible[5].

This procedure of discarding part of the outcome information—by grouping certain outcomes together—is commonly referred to in the operationalist literature as *coarse-graining*. It belongs to a broader class of operations known as *classical post-processing*, as it involves post-experiment analysis of classical data, i.e., the outcomes.

Returning to our particle physics analogy, coarse-graining is the operation performed when histograms are made: different data points are aggregated into the same bin.

Let us now consider an explicit example. Take the outcome space of a die roll:

$$\mathsf{X} \equiv \{1, 2, 3, 4, 5, 6\}\,.$$

Suppose we are not interested in the precise number, but only in whether it is odd or even. We then define a *coarse-grained* outcome space:

$$\mathsf{X}' \equiv \{o, e\}\,,$$

and a function $\phi : X \to X'$ that maps each fine-grained outcome to a coarse-grained one:

$$\phi(x) = \begin{cases} o & \text{if } x \in \{1,3,5\}, \\ e & \text{if } x \in \{2,4,6\}. \end{cases}$$

This is the essence of coarse-graining: it allows us to reinterpret the outcomes of a test through a function that forgets some of the original information, replacing the fine-grained outcome space X with a compressed version X'. In doing so, we reduce the resolution of the description while preserving the meaningful information relevant to the task at hand.

More formally, given an outcome space X, let $\mathcal{K}(\mathsf{X}) = \left\{\mathsf{X}^{(k)}\right\}_{k\in\mathsf{K}}$ denote an arbitrary partition of X—that is, a collection of non-empty sets $\mathsf{X}^{(k)}$ such that

$$\bigcup_{k\in\mathsf{K}} \mathsf{X}^{(k)} = \mathsf{X} \quad \text{and} \quad \mathsf{X}^{(k_1)} \cap \mathsf{X}^{(k_2)} = \emptyset \quad \forall k_1 \neq k_2.$$

[5] As a curiosity, we note that even by performing an attentive selection on the information to maintain, CERN still collects and stores an incredible amount of data. The CERN Data Centre stores more than 30 petabytes of data per year from the LHC experiments, enough to fill about 1.2 million Blu-ray discs, i.e., 250 years of HD video. Over 100 petabytes of data are permanently archived on tape [174].

For any $\mathsf{X} \in \mathsf{Out}(\Theta)$, let $\mathsf{Part}(\mathsf{X})$ denote the set of all partitions of X. Given a test $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{B})$, where $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ are generic systems of the OPT, and a partition $\mathcal{K}(\mathsf{X}) = \left\{\mathsf{X}^{(k)}\right\}_{k\in\mathsf{K}}$, we define a *coarse-graining map*

$$\mathfrak{C}_{\mathcal{K}(\mathsf{X})} : \mathsf{Test}(\mathrm{A}\to\mathrm{B}) \to \mathsf{Test}(\mathrm{A}\to\mathrm{B}) \tag{2.27}$$

whose action is given by

$$\mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right) := [\![\mathscr{T}_{\mathsf{X}^{(k)}}]\!]_{k\in\mathsf{K}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{B}), \tag{2.28}$$

such that, if an element $\mathsf{X}^{(k)} \in \mathcal{K}(\mathsf{X})$ is a singleton set $\mathsf{X}^{(k)} = \left\{x^{(k)}\right\}$ for some $x^{(k)} \in \mathsf{X}^{(k)}$, then the corresponding coarse-grained event coincides with the original one associated with the outcome $x^{(k)} = \tilde{x} \in \mathsf{X}$, namely:

$$\mathscr{T}_{\left\{x^{(k)}\right\}} := \mathscr{T}_{\mathsf{x}^{(k)}} = \mathscr{T}_{\tilde{x}}, \quad \forall x^{(k)} \in \mathsf{X}, \quad \forall \mathcal{K}(\mathsf{X}) \in \mathsf{Part}(\mathsf{X}) : \left\{x^{(k)}\right\} \in \mathcal{K}(\mathsf{X}).$$

This implies the following:

- The coarse-graining map associated with the trivial partition leaves the test unchanged: no information is disregarded.

- For singleton tests, the only admissible coarse-graining is that associated with the trivial partition. Leaving aside the mathematical fact that a singleton set admits only the trivial partition, this result can also be understood intuitively: since singleton tests describe events that occur with certainty, they do not carry any information to begin with. Hence, there is nothing that could be disregarded via coarse-graining.

The right-hand side of (2.28) represents the test in which part of the outcome information has been discarded. Specifically, we no longer have access to the complete fine-grained outcome, but only to the label $k \in \mathsf{K}$ identifying the partition element to which the original outcome belonged. The outcomes $x \in \mathsf{X}$ and their associated events $\mathscr{T}_x$ are grouped together according to the partition $\mathcal{K}(\mathsf{X})$. Each fine-grained outcome $x \in \mathsf{X}$ is mapped to a new coarse-grained label $k \in \mathsf{K}$ whenever $x \in \mathsf{X}^{(k)}$. Accordingly, every event $\mathscr{T}_x$ is mapped to a *coarse-grained event* $\mathscr{T}_{\mathsf{X}^{(k)}} \in \mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right)$.

The test $\mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right)$ is called the *coarse-grained test of* $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}$ *with respect to the partition* $\mathcal{K}(\mathsf{X})$.

As for the operations of sequential and parallel composition, OPTs are required to be closed with respect to the coarse-graining operation.

What happens with multiple coarse-graining operations

There are two further aspects that must be clarified for the operation of coarse-graining to be completely well-defined. The first concerns the composition of multiple coarse-graining maps. The requirement we impose is that the final coarse-grained test depends only on the resulting partition, and not on the particular sequence of coarse-graining operations through which that partition is obtained.

Formally, let $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ be a generic test in an OPT, and let $\mathcal{K}(\mathsf{X}) = \left\{\mathsf{X}^{(k)}\right\}_{k\in\mathsf{K}}$ and $\mathcal{M}(\mathsf{X}) = \left\{\mathsf{X}^{(m)}\right\}_{m\in\mathsf{M}}$ be two partitions of the outcome space X. Consider now a partition $\mathcal{L}(\mathsf{K}) = \left\{\mathsf{K}^{(l)}\right\}_{l\in\mathsf{L}}$ of the set K such that $\mathcal{L}(\mathcal{K}(\mathsf{X})) = \mathcal{M}(\mathsf{X})$. We require that

$$\mathfrak{C}_{\mathcal{L}(\mathsf{K})}\left(\mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right)\right) = \mathfrak{C}_{\mathcal{M}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right).$$

We observe that this structure induces a partial ordering among coarse-grained tests, from more refined to less refined. The less refined a test is, the more information it retains from the original.

Furthermore, the following lemmas holds.

Lemma 4

The deterministic coarse-graining—that is, the case in which the partition is trivial and consists of the full outcome space as a single set—yields an unique singleton test, and the associated event is deterministic:

$$\mathfrak{C}_{\mathsf{X}}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right) = \left[\!\left[\mathscr{T}_{\{\mathsf{X}\}}\right]\!\right].$$

Lemma 5

Let $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ be a generic test in an OPT, and let $\mathcal{K}(\mathsf{X})$ and $\mathcal{L}(\mathsf{X})$ be two different partitions of the same outcome space X. Suppose there exist $\mathsf{X}^{(k)} \in \mathcal{K}(\mathsf{X})$ and $\mathsf{X}^{(l)} \in \mathcal{L}(\mathsf{X})$ such that

$$\mathsf{X}^{(k)} = \mathsf{X}^{(l)}.$$

Then, the corresponding coarse-grained events are equal:

$$\mathscr{T}_{\mathsf{X}^{(k)}} = \mathscr{T}_{\mathsf{X}^{(l)}},$$

where $\mathscr{T}_{\mathsf{X}^{(k)}}$ and $\mathscr{T}_{\mathsf{X}^{(l)}}$ are the events of the tests $\mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right)$ and $\mathfrak{C}_{\mathcal{L}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}\right)$, respectively [143].

In words, this lemma states that if two coarse-graining partitions share a common subset of the outcome space, then the coarse-grained event corresponding to that subset is independent of which partition was used to define it.

To better clarify the meaning of $\mathcal{L}(\mathcal{K}(\mathsf{X})) = \{\mathsf{X}^{(m)}\}_{m\in\mathsf{M}}$, let us consider an example. Take the outcome space of a die:

$$\mathsf{X} = \{1, 2, 3, 4, 5, 6\}.$$

We can define two different partitions:

$$\mathcal{K}(\mathsf{X}) = \{k_1, k_2, k_3\} \equiv \{\{1, 2\}, \{3, 4\}, \{5, 6\}\},$$
$$\mathcal{M}(\mathsf{X}) = \{m_1, m_2\} \equiv \{\{1, 2, 3, 4\}, \{5, 6\}\}.$$

Then, define the partition $\mathcal{L}(\mathsf{K})$ as:

$$\mathcal{L}(\mathsf{K}) = \{l_1, l_2\} \equiv \{\{k_1, k_2\}, \{k_3\}\}.$$

It is straightforward to verify that:

$$\mathcal{L}(\mathcal{K}(\mathsf{X})) = \{\{1, 2, 3, 4\}, \{5, 6\}\} = \mathcal{M}(\mathsf{X}).$$

Thus, we have composed two partitioning operations and reached the same final result.

Relationship with the compositional structure

The second requirement for the well-definedness of the coarse-graining operation concerns its compatibility with the sequential $\Box$ and parallel $\boxtimes$ composition maps.

In order to make this compatibility precise, we first have to distinguish between coarse-graining performed at the level of individual outcome spaces and coarse-graining performed at the level of composite outcome spaces.

Let $\mathcal{K}(\mathsf{X}) = \{\mathsf{X}^{(k)}\}_{k\in\mathsf{K}}$ and $\mathcal{L}(\mathsf{Y}) = \{\mathsf{Y}^{(l)}\}_{l\in\mathsf{L}}$ be two partitions of the outcome spaces X and Y, respectively. Starting from these partitions, two composite outcome spaces can be constructed.

- On the one hand, one can first coarse-grain the individual outcome spaces and then compose them, obtaining the family

  $$\{\mathsf{X}^{(k)} \times \mathsf{Y}^{(l)}\}_{(k,l)\in\mathsf{K}\times\mathsf{L}}.$$

  This corresponds to the outcome space of a composite test—either sequential or parallel—obtained by composing two individually coarse-grained tests.

- On the other hand, one can first form the composite outcome space and then induce a partition on it. For every $k \in \mathsf{K}$ and $l \in \mathsf{L}$, define
$$(\mathsf{X} \times \mathsf{Y})_{(k,l)} \coloneqq \left\{ (x, y) \in \mathsf{X} \times \mathsf{Y} \,\middle|\, x \in \mathsf{X}^{(k)},\ y \in \mathsf{Y}^{(l)} \right\}.$$
These sets define a partition of $\mathsf{X} \times \mathsf{Y}$, namely
$$\mathcal{KL}\left(\mathsf{X} \times \mathsf{Y}\right) \coloneqq \left\{ (\mathsf{X} \times \mathsf{Y})_{(k,l)} \right\}_{(k,l) \in \mathsf{K} \times \mathsf{L}}.$$

It is immediate to see that the two outcome spaces defined above admit a natural identification, given by
$$(\mathsf{X} \times \mathsf{Y})_{(k,l)} \longleftrightarrow \mathsf{X}^{(k)} \times \mathsf{Y}^{(l)}. \tag{2.29}$$
Indeed, these two sets contain precisely the same outcomes from the original outcome spaces X and Y for every pair $(k, l) \in \mathsf{K} \times \mathsf{L}$.

This is a first step that guarantees that the operation of sequential and parallel composition are compatible with the operation of coarse-graining. From the point of view of the outcome spaces, the order in which these operations are performed is not important. To conclude the argument, we therefore require that the coarse-graining map $\mathfrak{C}$ preserves both sequential and parallel composition under the induced composite partitions:
$$\begin{aligned}
\mathfrak{C}_{\mathcal{KL}(\mathsf{X} \times \mathsf{Y})}\left(\mathsf{G}_{\mathsf{Y}} \,\square\, \mathsf{T}_{\mathsf{X}}\right) &= \mathfrak{C}_{\mathcal{L}(\mathsf{Y})}\left(\mathsf{G}_{\mathsf{Y}}\right) \,\square\, \mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}\right), \\
\mathfrak{C}_{\mathcal{KL}(\mathsf{X} \times \mathsf{Y})}\left(\mathsf{T}_{\mathsf{X}} \boxtimes \mathsf{G}_{\mathsf{Y}}\right) &= \mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\mathsf{T}_{\mathsf{X}}\right) \boxtimes \mathfrak{C}_{\mathcal{L}(\mathsf{Y})}\left(\mathsf{G}_{\mathsf{Y}}\right).
\end{aligned}$$

At the level of events, this corresponds to:
$$\begin{aligned}
(\mathscr{G} \,\square\, \mathscr{T})_{(\mathsf{X} \times \mathsf{Y})_{(k,l)}} &= \mathscr{G}_{\mathsf{Y}^{(l)}} \,\square\, \mathscr{T}_{\mathsf{X}^{(k)}}, \\
(\mathscr{T} \boxtimes \mathscr{G})_{(\mathsf{X} \times \mathsf{Y})_{(k,l)}} &= \mathscr{T}_{\mathsf{X}^{(k)}} \boxtimes \mathscr{G}_{\mathsf{Y}^{(l)}}.
\end{aligned}$$

Operationally, this requirement expresses the idea that disregarding certain outcomes of individual experiments—whether composed sequentially or in parallel—is equivalent to ignoring the corresponding combinations of outcomes in the composite experiment.

The identification (2.29) guarantees the well-definiteness of the above expressions.

Note that, in general, given a composite outcome space $\mathsf{X} \times \mathsf{Y}$, there exist partitions $\mathcal{K}\left(\mathsf{X} \times \mathsf{Y}\right)$ that cannot be obtained as product partitions of the form $\mathcal{L}\left(\mathsf{X}\right) \times \mathcal{M}\left(\mathsf{Y}\right)$. For example, if
$$\begin{aligned}
\mathsf{X} &= \{1, 2\}, \\
\mathsf{Y} &= \{1, 2\},
\end{aligned}$$

then the partition

$$\mathcal{K}(\mathsf{X}\times\mathsf{Y}) = \{\{(1,1),(2,2)\},\ \{(1,2),(2,1)\}\}$$

cannot be written as a product of partitions of X and Y.

Finally, we observe that the coarse-graining map induces a binary operation $\curlyvee$ on events of the same test:

$$\mathscr{T}_x \curlyvee \mathscr{T}_{\tilde{x}} \coloneqq \mathscr{T}_{\{x\}\cup\{\tilde{x}\}}, \quad \forall \mathscr{T}_x, \mathscr{T}_{\tilde{x}} \in \mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}. \tag{2.30}$$

This operation is associative, commutative, and distributes over both sequential $\square$ and parallel $\boxtimes$ composition.

Relationship with the probabilistic structure

One particularly interesting aspect to highlight is how the coarse-graining operation interacts with the probabilistic structure of an OPT.

First of all, given the operational meaning of (2.21), it is immediate to interpret an expression of the form

$$\mathrm{G}\left(\mathscr{T}_x, \ldots, \bigcurlyvee_{j=1}^{n} \mathscr{G}_{y_j}, \ldots, \mathscr{F}_z\right)$$

as representing the total probability of mutually exclusive events to occur. That is, it corresponds to the probability that one of the outcomes $x, \ldots, y_i, \ldots, z$, with $i = 1, \ldots, n$, occurs, given that the experiment $\mathrm{G}(\mathsf{T}_{\mathsf{X}}, \ldots, \mathsf{G}_{\mathsf{Y}}, \ldots, \mathsf{F}_{\mathsf{Z}})$ is performed.

Secondly, since scalar events are associated with real numbers in the interval $[0,1]$ (i.e., probabilities), we can characterize the action of coarse-graining on them in purely numerical terms. In particular, when applied to scalar events, the coarse-graining operation $\curlyvee$ acts as the ordinary sum over real numbers.

Theorem 1 (Coarse-graining of scalar events is ordinary summation)

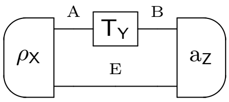
Let Θ be a generic OPT. The coarse-graining operation $\curlyvee$ on scalar events in Θ coincides with the usual sum of real numbers. For every scalar test $\mathrm{P}_{\mathsf{X}}^{\mathrm{I}\to\mathrm{I}} \in \mathsf{Test}(\mathrm{I}\to\mathrm{I})$ and every partition $\mathcal{K}(\mathsf{X}) \in \mathsf{Part}(\mathsf{X})$, the following identity holds [143]:

$$\bigcurlyvee_{x\in\mathsf{X}^{(k)}} \mathrm{p}_x = \mathrm{p}\left(k \,\middle|\, \mathfrak{C}_{\mathcal{K}(\mathsf{X})}(\mathrm{P}_{\mathsf{X}})\right) = \sum_{x\in\mathsf{X}^{(k)}} \mathrm{p}(x\,|\,\mathsf{P}_{\mathsf{X}}), \quad \forall \mathsf{X}^{(k)} \in \mathcal{K}(\mathsf{X}).$$

A more in-depth discussion of these ideas can be found in Ref. [143].

2.2.4 To quotient or not to quotient

Let us now consider a scenario similar to the one depicted in (2.22):



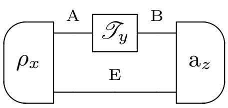
$$= [\![\mathrm{p}\,(x, y, z \,|\, \rho_\mathsf{X}, \mathsf{T}_\mathsf{Y}, \mathrm{a}_\mathsf{Z})]\!]_{(x,y,z)\in \mathsf{X}\times\mathsf{Y}\times\mathsf{Z}} .$$

One way to interpret this setup is to regard it as probing the behaviour of the test T_Y when it is composed between a preparation-test ρ_X and an observation-test a_Z. The inclusion of the ancillary system E guarantees that the analysis remains fully general, allowing for arbitrary correlations with an environment.

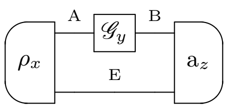
By probing the test T_Y against all possible pairs of preparation- and observation-tests, one can achieve a full tomography of the test T_Y. This is, in fact, the standard approach used to perform tomography of a quantum instrument.

What happens, then, if two tests T_Y and G_Y yield the same probabilities for all possible preparations and observations? We define them to be *probabilistically equivalent*.

Formally, for any systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and any two tests $\mathsf{T}_\mathsf{Y} \equiv [\![\mathscr{T}_y]\!]_{y\in\mathsf{Y}}$, $\mathsf{G}_\mathsf{y} \equiv [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{B})$ we define the equivalence relation $\mathsf{T}_\mathsf{Y} \sim \mathsf{G}_\mathsf{Y}$ if



$$=$$



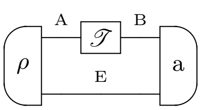
$$\forall y \in \mathsf{Y}, \tag{2.31}$$

for all $\mathrm{E} \in \mathsf{Sys}(\Theta)$, $\rho_x \in \rho_\mathsf{X} \in \mathsf{Event}(\mathrm{I}\to\mathrm{AE})$, and $\mathrm{a}_z \in \mathrm{a}_\mathsf{Z} \in \mathsf{Event}(\mathrm{BE}\to\mathrm{I})$. Meaning that two tests are operationally equivalent if they lead to the same probability distributions when inserted in the same circuit. The preparations and observations are assumed to vary over all admissible ones, where here by this we mean that they belong to a test of the theory.

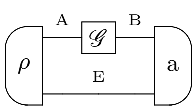
This reflects the principle that tests producing identical statistics in all experiments are *operationally indistinguishable*.

Clearly, if $\mathsf{T}_\mathsf{Y} = \mathsf{G}_\mathsf{y}$, then (2.32) holds. However, the converse is generally false. Hence, the need to define this as an equivalence relation.

The equivalence relation for tests (2.31) immediately induces an equivalence relation also for events. For any systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and any events $\mathscr{T}, \mathscr{G} \in \mathsf{Event}(\mathrm{A}\to\mathrm{B})$, we define the equivalence relation $\mathscr{T} \sim \mathscr{G}$ if



$$=$$



$$\tag{2.32}$$

for all $E \in \mathsf{Sys}(\Theta)$, $\rho \in \mathsf{Event}(I \to AE)$, and $a \in \mathsf{Event}(BE \to I)$.

Remark 9 (Tests with different outcome spaces)

In the equivalence relation (2.31), we have directly assumed that the two tests under study have the same outcome space. This follows from the assumption that tests with different outcome spaces cannot be equivalent. If one imagines a test as a black box with lights that flash according to the outcome, then differing numbers of possible outcomes make the tests distinguishable.

Remark 10

We highlight that in line with Assumption 2, test that differ only for the order of the events composing them are not mapped in the same equivalence class by (2.31).

If we *do not* quotient the sets of tests and events by the equivalence relations (2.32), we remain with what is known as an *unquotiented OPT* [175]. However, proper OPTs are required to be *quotiented* [143]. This follows from the operational *desiderata* that the framework captures only operationally accessible data, eliminating redundancies due to operationally indistinguishable structures.

The equivalence classes of events are called *transformations*, and we write:

$$\mathsf{Transf}(A \to B) := \mathsf{Event}(A \to B)/\sim.$$

In particular, the sets $\mathsf{St}(A) := \mathsf{Transf}(I \to A)$ and $\mathsf{Eff}(A) := \mathsf{Transf}(A \to I)$ define the *states* and *effects* of system A, respectively. The subsets of deterministic transformations, states, and effects are denoted $\mathsf{Transf}_1(A \to B)$, $\mathsf{St}_1(A)$, and $\mathsf{Eff}_1(A)$, respectively.

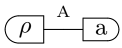
Likewise, tests from A to B become equivalence classes called *instruments*, and we define:

$$\mathsf{Instr}(A \to B) := \mathsf{Test}(A \to B)/\sim.$$

Finally, the collections of *preparation-instruments* and *observation-instruments* are denoted by $\mathsf{Prep}(A)$ and $\mathsf{Obs}(A)$, respectively.

The case of states and effects

In the special case of preparations and observations, the equivalence relation (2.32) simplifies to:



. (2.33)

This seemingly simple observation has a fundamental consequence: *states are separating for effects, and effects are separating for states.* Concretely, this means that for any pair of distinct states $\rho_1, \rho_2 \in \mathsf{St}(\mathrm{A})$ with $\rho_1 \neq \rho_2$, there exists an effect $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$ such that:

$$(\mathrm{a}|\,\rho_1)_{\mathrm{A}} \neq (\mathrm{a}|\,\rho_2)_{\mathrm{A}}\,,$$

and vice versa: for any two distinct effects, there exists a state that distinguishes them.

This property is often referred to in the literature as the fact that the set of states [effects] is *tomographically complete* for the set of effects [states] [176].

Furthermore, this also allows one to view states as a set of functionals that assign to each effect a probability in the interval $[0,1]$, i.e.,

$$\rho\colon \mathsf{Eff}(\mathrm{A}) \to [0,1], \quad \forall \rho \in \mathsf{St}(\mathrm{A})\,,$$

and effects as functionals acting on states:

$$\mathrm{a}\colon \mathsf{St}(\mathrm{A}) \to [0,1], \quad \forall \mathrm{a} \in \mathsf{Eff}(\mathrm{A})\,,$$

for all systems A of the theory. In this sense, the sets $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{A})$ are dual to each other under the pairing defined by the evaluation of probabilities.

A more in-depth discussion on the case of transformations

While states can be fully characterised by their action on effects—as a consequence of (2.33)—and, analogously, effects by their action on states, the complete characterisation of a transformation requires additional considerations. In particular, one must account for the possibility of correlations with an environment.

By taking into account both (2.32) and (2.33), a transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, with A, $\mathrm{B} \neq \mathrm{I}$, can only be fully characterised by examining its action on states $\rho \in \mathsf{St}(\mathrm{AE})$ for every possible ancillary system E. Equivalently, one may consider how it acts on effects $\mathrm{a} \in \mathsf{Eff}(\mathrm{BE})$.

The former perspective can be seen as a generalisation of the Schrödinger picture to the setting of OPTs, where transformations act on input states and produce output states. The latter, instead, corresponds to the Heisenberg picture, where transformations act on effects, propagating them backward from output to input.

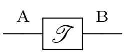
Following this description, and in analogy with the fact that states can be seen as functionals from effects to $[0, 1]$, transformations can be characterised as a family of maps of the form

$$\begin{aligned} \left\{ \widehat{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}} \right\}_{\mathrm{E} \in \mathsf{Sys}(\Theta)} &: \mathsf{St}(\mathrm{AE}) \to \mathsf{St}(\mathrm{BE}) \\ &\rho \mapsto \widehat{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}}\,(\rho) := (\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}) \,\Box\, \rho, \end{aligned} \tag{2.34}$$

or, equivalently, as

$$\begin{aligned} \left\{ \widecheck{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}} \right\}_{\mathrm{E} \in \mathsf{Sys}(\Theta)} &: \mathsf{Eff}(\mathrm{BE}) \to \mathsf{Eff}(\mathrm{AE}) \\ &\mathrm{a} \mapsto \widecheck{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}}\,(\mathrm{a}) := (\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}) \,\Box\, \mathrm{a}. \end{aligned} \tag{2.35}$$

As a consequence, in order to fully specify the action of a transformation in an OPT, one typically needs to determine the behaviour of a potentially infinite family of maps:



$$\longleftrightarrow \qquad \left\{ \widehat{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}} \right\}_{\mathrm{E} \in \mathsf{Sys}(\Theta)}.$$

However, there are special cases in which this exhaustive characterisation is not required. For example, we have already seen that in the case of states and effects, the ancillary system plays no role in (2.33).

More generally, this simplification also applies to all transformations of QT and CT. These theories fall into a special class of OPTs in which it is sufficient to consider the action of a transformation on the input system alone (in the Schrödinger picture), or on the output system (in the Heisenberg picture) [133]. This is due to the fact that such theories satisfy the property of *local discriminability*[6].

Moreover, there are theories in which it is sufficient to consider the action on a restricted set of ancillary systems. For instance, in the case of Bilocal Classical Theory (BCT), characterisation requires only one particular choice of ancillary system in addition to the trivial one[7].

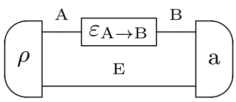
[6] For a more in-depth discussion of this property, please refer to section 3.4.1
[7] More details on BCT are provided in section 3.4.2

Remark 11

Strictly speaking, it would already have been possible to present all of the above results—namely, the characterisation of states, effects, and transformations as functionals—directly at the level of preparations, observations, and events [143].

The null-transformation

A consequence of the fact that $0 \in \mathsf{Transf}(\mathrm{I} \to \mathrm{I})$ always in an OPT (2.23), alongside the presence of parallel composition, the equivalence relation (2.32), and the factorisation assumption (2.26), imply the existence of the *null-transformation*. Formally, in any OPT Θ for every system $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the set of transformations $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ includes the null-transformation $\varepsilon_{\mathrm{A}\to\mathrm{B}}$, defined by the following relation:



$$= (\mathrm{a}|_{\mathrm{BE}} \, (\varepsilon_{\mathrm{A}\to\mathrm{B}} \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\rho)_{\mathrm{AE}} = 0, \tag{2.36}$$

for any system $\mathrm{E} \in \mathsf{Sys}(\Theta)$, $\rho \in \mathsf{St}(\mathrm{AE})$, and $\mathrm{a} \in \mathsf{Eff}(\mathrm{BE})$. In words, the null transformation is the transformation that always occur with null probability in any closed circuit. Clearly, transformations are invariant for coarse-graining with the null one

$$\mathscr{T} = \mathscr{T} \curlyvee \varepsilon_{\mathrm{A}\to\mathrm{B}}, \tag{2.37}$$

for any couple of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$.

A property of the null-transformation is that appending it to any instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ still yields a valid instrument, i.e.,

$$[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B}) \quad \Longleftrightarrow \quad \left[\!\left[\mathscr{T}_1, \ldots, \mathscr{T}_{|\mathsf{X}|}, \varepsilon_{\mathrm{A}\to\mathrm{B}}\right]\!\right] \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})\,. \tag{2.38}$$

Connections with contextuality

The distinction between quotiented and unquotiented OPTs is a relatively recent development [175]. It was introduced to bridge the framework of OPTs with the extensive literature on *contextuality* [41, 43, 46, 48, 152, 153, 177], which has typically been developed within the setting of Generalised Probabilistic Theories (GPTs).

The central difference is that unquotiented theories retain information about the *context* of an experiment—namely, aspects of the physical implementation that do not affect the statistical distribution of outcomes, and which are therefore erased by quotienting.

To illustrate this idea, consider an example from QT involving a measurement of a photon's polarisation. Suppose we implement two distinct detectors:

- M_1: a Polaroid film aligned along the $\hat{z}$-axis, followed by a photodetector;
- M_2: a birefringent crystal oriented to spatially separate vertical and horizontal polarisation components along the $\hat{z}$-axis, with a photodetector in the vertical output channel.

Although M_1 and M_2 implement distinct physical procedures—and thus correspond to different observation-tests—they are probabilistically equivalent: they yield identical outcome statistics for all preparations. In a quotiented theory, the equivalence relation (2.32) identifies such procedures, and both fall into the same equivalence class of instruments.

In general, any implementation detail that leaves outcome probabilities unchanged is regarded as part of the context, and is discarded upon quotienting. This is why contextuality can only be meaningfully analysed at the level of unquotiented theories.

One of the reasons for opening the framework of OPTs to the literature on contextuality is that it leads to a notion of classicality different from the one adopted in this thesis (section 3.8). A widely used criterion for deciding whether the predictions of an operational theory admit a classical explanation is whether its unquotiented version supports a *generalised-non-contextual ontological model* [48, 153].

An *ontological model* postulates a set of *ontic states* that fully specify the system, together with *epistemic states* representing an agent's (possibly partial) knowledge of the ontic state. This framework generalises hidden-variable models and has played a central role in recent developments in quantum foundations. For example, Spekkens refined the Bell–Kochen–Specker theorem [28, 41], showing that quantum theory is contextual even for two-dimensional systems [48]. He also introduced a toy theory that is generalised-non-contextual—admitting such an ontological model—yet reproduces several quantum-like features via an epistemic restriction limiting accessible information [128].

We highlight that it has been proven that an unquotiented theory admits a generalised-non-contextual ontological model if and only if its quotiented version admits an ontological model. In the case of the quotiented theory, it is not meaningful to ask whether the ontological model is quotiented or not, since—at that level—all distinctions arising from context have already been erased. Everything is quotiented. In this settings, an ontological model is formally defined as a diagram-preserving map into CT both for the quotiented and unquotiented theories. Then, such a model satisfies *generalised-non-contextuality* whenever any pair of events that are probabilistically equivalent in the unquotiented theory are mapped to the same element of CT [152, 153].

Observation 3 (Contextuality of tests)

Recent work has shown that tests themselves do not provide context [175]. In strongly causal[a] OPTs, events are always mapped to the same element in ontological models, independently of the test to which they belong.

[a]See section 3.6 for the definition of strongly causal OPTs.

Remark 12 (Our notion of classicality)

We stress that the notion of classicality adopted in this thesis is distinct—though related—to that of admitting a generalised-non-contextual ontological model. Our precise definition is given in section 3.8 and refers to the geometric structure of the state spaces of a theory. In particular, the vertices of a state space have to form a simplex and need to be perfectly discriminable.
More information about the relationship between our notion of classicality and that based on contextuality can be found in section 3.8.2.

2.3 Additional Structure

With the introduction of the quotienting operation, the framework of OPTs is complete. We have constructed a well-defined and coherent structure that enables the description of generic physical theories from the point of view of their information-processing capabilities, starting from just three simple primitives.

However, if one considers only the structure defined thus far, the framework might appear somewhat abstract and, at times, difficult to work with. For instance, defining a consistent rule that allows one to explicitly calculate the probability distributions associated with circuits of tests could prove challenging without invoking further structure. Consider, for example, attempting to do this for QT without reference to the underlying Hilbert space formalism[8]. Moreover, there are certain questions that the current state of the framework is simply not equipped to answer.

[8]For completeness, we note that such approaches have been successfully carried out in specific contexts—see, for example, Refs. [178, 179].

One might ask, for example, how different two transformations are. To address such a question, one would need a notion of distance between transformations, which in turn requires the framework to possess a topological structure.

Fortunately, the framework is so robust that the structure developed so far naturally gives rise to additional layers of structure. In particular, it induces both a *linear* and a *topological* structure, which further enrich the operational framework.

2.3.1 Linear Structure

To begin with, we will show how the fact that states and effects are separating for each other induces an $\mathbb{R}$-linear structure within the framework. This construction can then be extended to the case of transformations and instruments. For a complete treatment, we refer the interested reader to Ref. [143]. We also note that the argument underpinning the derivation for states and effects is closely related to *Ludwig's embedding theorem* [149, 180].

The derivation of the linear structure fundamentally relies on (2.33) and its consequences:

- States can be seen as linear functionals on effects with values in $[0, 1]$, and vice versa.
- States are separating for effects, and effects are separating for states:

$$\begin{aligned} &\forall \rho_1, \rho_2 \in \mathsf{St}(\mathrm{A}), \quad \rho_1 = \rho_2 \iff (\mathrm{a}|\, \rho_1)_{\mathrm{A}} = (\mathrm{a}|\, \rho_2)_{\mathrm{A}} \quad \forall \mathrm{a} \in \mathsf{Eff}(\mathrm{A}), \\ &\forall \mathrm{a}_1, \mathrm{a}_2 \in \mathsf{Eff}(\mathrm{A}), \quad \mathrm{a}_1 = \mathrm{a}_2 \iff (\mathrm{a}_1|\, \rho)_{\mathrm{A}} = (\mathrm{a}_2|\, \rho)_{\mathrm{A}} \quad \forall \rho \in \mathsf{St}(\mathrm{A}), \end{aligned} \tag{2.39}$$

  for any system A of the OPT.

These relationships can be extended to generic real linear combinations of states and effects.

Let $\{\rho_1, \rho_2\} \subset \mathsf{St}(\mathrm{A})$ and $\{\mathrm{a}_1, \mathrm{a}_2\} \subset \mathsf{Eff}(\mathrm{A})$ be two sets of states and effects, respectively. We then define their real linear combinations as

$$\begin{aligned} \tilde{\rho} &:= \mathrm{r}_1 \rho_1, \\ \tilde{\rho}' &:= \mathrm{r}_1 \rho_1 + \mathrm{r}_2 \rho_2, \end{aligned} \tag{2.40}$$

and

$$\begin{aligned} \tilde{\mathrm{a}} &:= \mathrm{r}_1 \mathrm{a}_1, \\ \tilde{\mathrm{a}}' &:= \mathrm{r}_1 \mathrm{a}_1 + \mathrm{r}_2 \mathrm{a}_2, \end{aligned} \tag{2.41}$$

with $\mathrm{r}_1, \mathrm{r}_2 \in \mathbb{R}$.

An attentive reader might wonder what the meaning of these newly defined objects is. After all, we have never defined what it means to "multiply" a transformation by a real number or to sum two of them. The closest notions available so far are scalar multiplication (2.25) and coarse-graining (2.30). These are certainly present, but not sufficient. However, we will later see how the linear operations introduced here are closely related to the operational ones already in the framework.

For the moment, the linear combinations above are defined via the following relations:

$$\begin{aligned} (\mathrm{b}|\, \tilde{\rho})_{\mathrm{A}} &\coloneqq \mathrm{r}_1\, (\mathrm{b}|\, \rho_1)_{\mathrm{A}}\,, \\ \left(\mathrm{b}\middle|\, \tilde{\rho}'\right)_{\mathrm{A}} &\coloneqq \mathrm{r}_1\, (\mathrm{b}|\, \rho_1)_{\mathrm{A}} + \mathrm{r}_2\, (\mathrm{b}|\, \rho_2)_{\mathrm{A}}\,, \end{aligned} \tag{2.42}$$

for all effects $\mathrm{b} \in \mathsf{Eff}(\mathrm{A})$, and

$$\begin{aligned} (\tilde{\mathrm{a}}|\, \sigma)_{\mathrm{A}} &\coloneqq \mathrm{r}_1\, (\mathrm{a}_1|\, \sigma)_{\mathrm{A}}\,, \\ \left(\tilde{\mathrm{a}}'\middle|\, \sigma\right)_{\mathrm{A}} &\coloneqq \mathrm{r}_1\, (\mathrm{a}_1|\, \sigma)_{\mathrm{A}} + \mathrm{r}_2\, (\mathrm{a}_2|\, \sigma)_{\mathrm{A}}\,, \end{aligned} \tag{2.43}$$

for all states $\sigma \in \mathsf{St}(\mathrm{A})$. This defines the new objects introduced in (2.40) and (2.41) as functionals from $\mathsf{Eff}(\mathrm{A})$ and $\mathsf{St}(\mathrm{A})$, respectively, to $\mathbb{R}$. These functionals result to be well-defined, since the pairing $(\cdot|\, \cdot)_{\mathrm{A}}$ on the right-hand side of the definitions always involves well-defined states and effects within the theory, what remains then is just a linear combination of real numbers.

From now on, we refer to these new functionals as defined in (2.40) and (2.41) as *generalised states*, denoted by $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$, and *generalised effects*, denoted by $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$, respectively. By construction, these sets coincide with the real linear spans of the original sets:

$$\mathsf{St}_{\mathbb{R}}(\mathrm{A}) \equiv \mathsf{Span}_{\mathbb{R}}(\mathsf{St}(\mathrm{A}))\,, \qquad \mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) \equiv \mathsf{Span}_{\mathbb{R}}(\mathsf{Eff}(\mathrm{A}))\,,$$

as understood via (2.42) and (2.43).

The existence of the null-transformation (2.36) for every pair of systems in the theory also implies the existence of a *null-state* and a *null-effect*, defined as the zero functionals in the respective vector spaces. For any system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, the following identities hold:

$$\begin{aligned} 0\,\tilde{\rho} &= \varepsilon, \\ \tilde{\rho} + \varepsilon &= \tilde{\rho}, \end{aligned} \qquad \forall \tilde{\rho} \in \mathsf{St}_{\mathbb{R}}(\mathrm{A})\,,$$

and

$$\begin{aligned} 0\,\tilde{\mathrm{a}} &= \varepsilon, \\ \tilde{\mathrm{a}} + \varepsilon &= \tilde{\mathrm{a}}, \end{aligned} \qquad \forall \tilde{\mathrm{a}} \in \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})\,.$$

To summarise, for every system A in the theory, one can associate a real vector space of states $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and a real vector space of effects $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$. The zero element of each of these spaces corresponds to the null-state and null-effect, respectively.

The set of effects $\mathsf{Eff}(\mathrm{A})$ remains separating for the generalised states $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$, as can be verified by direct calculation. The same holds for the set of states with respect to the generalised effects $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$. In addition, both $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ consist of $\mathbb{R}$-valued functionals on each other: the action of any functional $\tilde{\rho} \in \mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\tilde{a} \in \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$, originally defined on $\mathsf{Eff}(\mathrm{A})$ and $\mathsf{St}(\mathrm{A})$ respectively, uniquely extend by linearity to the entire spaces $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$. Together, these facts imply that each space is contained in the algebraic dual of the other:

$$\mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) \subseteq \mathsf{St}_{\mathbb{R}}(\mathrm{A})^{\vee}, \qquad \mathsf{St}_{\mathbb{R}}(\mathrm{A}) \subseteq \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})^{\vee}.$$

In the particular case in which these vector spaces are finite-dimensional, the inclusions above become equalities. That is, if $\mathsf{dim}(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) < +\infty$ (or $\mathsf{dim}(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})) < +\infty$), then

$$\mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) = \mathsf{St}_{\mathbb{R}}(\mathrm{A})^{\vee}, \qquad \mathsf{St}_{\mathbb{R}}(\mathrm{A}) = \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})^{\vee},$$

which clearly implies that $\mathsf{dim}(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = \mathsf{dim}(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A}))$.

In general, the vector space $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ associated with a system A allows one to define a notion of dimension for the system. We define

$$D_{\mathrm{A}} := \mathsf{dim}(\mathsf{St}_{\mathbb{R}}(\mathrm{A}))$$

as the *dimension* (or *size*) of the system A[9].

The quantity D_{A} represents the number of real parameters (i.e., probabilities) that must be specified to fully characterise a state of system A when expressed as a vector in $\mathbb{R}^{D_{\mathrm{A}}}$. For instance, in quantum theory, a system A described by a Hilbert space $\mathcal{H}_{\mathrm{A}}$ of dimension d_{A} has size $D_{\mathrm{A}} = d_{\mathrm{A}}^2$.

We now present the third assumption made in this thesis.

Assumption 3: Finite dimensional vector spaces

The dimension D_{A} of every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ in any OPT Θ is assumed to be finite, that is, $D_{\mathrm{A}} < +\infty$.

[9] The construction described so far extends to the infinite-dimensional case. However, particular care must be taken regarding the duality relations in that setting.

The case of transformations

The construction we have just introduced for defining the linear spaces of states and effects can be naturally extended to the case of transformations.

To do so, we begin by observing that transformations can be treated as functionals over states and effects. In particular, we are interested in understanding how transformations act on *generalised* states and effects. What can be shown is the following:

Theorem 2 (Transformations as families of linear maps)

Let Θ be a generic OPT, and let $\mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta)$ and $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$. Then there exists a unique linear map $\widehat{\mathscr{T}} : \mathsf{St}_{\mathbb{R}}(\mathrm{A}) \to \mathsf{St}_{\mathbb{R}}(\mathrm{B})$ and a unique linear map $\widecheck{\mathscr{T}} : \mathsf{Eff}_{\mathbb{R}}(\mathrm{B}) \to \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ such that, for every $\rho \in \mathsf{St}(\mathrm{A})$ and $\mathrm{a} \in \mathsf{Eff}(\mathrm{B})$, the following identities hold [143]:

$$\left|\widehat{\mathscr{T}}(\rho)\right)_{\mathrm{A}} = \mathscr{T}\,|\rho)_{\mathrm{A}}\,,$$
$$\left(\widecheck{\mathscr{T}}(\mathrm{a})\right|_{\mathrm{B}} = (\mathrm{a}|_{\mathrm{B}}\,\mathscr{T},$$
$$\widehat{\mathscr{T}} = \widecheck{\mathscr{T}}.$$

The proof of this theorem relies on the fact that the action of a transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is defined on $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{B})$, which form bases for $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and $\mathsf{Eff}_{\mathbb{R}}(\mathrm{B})$, respectively. Therefore, the transformation $\mathscr{T}$ admits a unique linear extension as a map from $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ to $\mathsf{St}_{\mathbb{R}}(\mathrm{B})$, and dually from $\mathsf{Eff}_{\mathbb{R}}(\mathrm{B})$ to $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$.

This result refines the discussion in (2.34) and (2.35). To any transformation, one can associate a unique family of linear maps:

$$\left\{\widehat{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}}\right\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)} \equiv \left\{\widecheck{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}}\right\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)},$$

where

$$\widehat{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}} : \mathsf{St}_{\mathbb{R}}(\mathrm{AE}) \to \mathsf{St}_{\mathbb{R}}(\mathrm{BE})\,,$$
$$\widecheck{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}} : \mathsf{Eff}_{\mathbb{R}}(\mathrm{BE}) \to \mathsf{Eff}_{\mathbb{R}}(\mathrm{AE})\,,$$

for any system $\mathrm{E} \in \mathsf{Sys}(\Theta)$.

Following the same construction as for states (2.40) and effects (2.41), we define the real vector space $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, whose elements are called *generalised transformations*. These are defined by linear combinations of the corresponding families of linear maps associated with each transformation. The operations of sequential and parallel composition can then be extended to generalised transformations—as well as to generalised states and effects—by linearity, thereby preserving the structure of the theory.

Actually, we can say a bit more about the rules governing sequential and parallel composition.

First, in the case of sequential composition, the map defining the composition of the associated families of linear transformations coincides with the standard composition of linear maps:

$$\overset{\mathrm{A}}{\longrightarrow}\boxed{\mathscr{G} \,\square\, \mathscr{T}}\overset{\mathrm{B}}{\longrightarrow} \quad \longleftrightarrow \quad \left\{ \widehat{\mathscr{G} \boxtimes \mathscr{I}_{\mathrm{E}}} \circ \widehat{\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}} \right\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)} .$$

On the other hand, the map corresponding to parallel composition does not, in general, coincide with the standard algebraic tensor product $\otimes$. In fact, for a generic OPT, we have:

$$\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B}) \boxtimes \mathsf{Transf}_{\mathbb{R}}(\mathrm{C} \to \mathrm{D}) \subseteq \mathsf{Transf}_{\mathbb{R}}(\mathrm{AC} \to \mathrm{BD})\,, \tag{2.44}$$

with equality holding if and only if the parallel composition $\boxtimes$ coincides with the tensor product $\otimes$.

Remark 13 (The composition rule for quantum systems is not the only possible one)

While the fact that the equality in (2.44) does not hold for generic OPTs might at first appear to be a mere mathematical curiosity, this possibility has led to remarkably insightful results. In particular, it has recently been shown that one can construct a class of theories, called Latent Quantum Theories (LQTs), which are indistinguishable from QT in every operational aspect, except for the compositional rule [172]. This demonstrates that the use of the tensor product to compose systems in QT is not a consequence of the other axioms of the theory.

Furthermore, these alternative theories are proven to be empirically indistinguishable from QT via Bell-type experiments.

Formally, following the axiomatic reconstruction of QT proposed in Ref. [133], LQTs differ from QT only in that they do not satisfy the axiom of *local discriminability*. They still satisfy the other five axioms:

I) Atomicity of composition,

II) Perfect discriminability,

III) Ideal compression,

IV) Causality,

V) Purification.

Another illuminating example is provided by Bilocal Classical Theory (BCT) [167]—a classical theory that exhibits entanglement. In this case, the presence of entanglement arises precisely from the fact that the compositional rule deviates from the usual one.

For a more in-depth discussion and the full derivation of the results above presented, we refer the interested reader to Ref. [143].

What about coarse-graining?

Another valuable consequence of introducing a linear structure is the possibility of formalising the rule for coarse-graining (2.30) in a more explicit and intuitive way.

Theorem 3 (Coarse-graining is just the sum)

Let Θ be an OPT. Then, the coarse-graining operation $\curlyvee$ on the transformations of Θ is given by their sum. Equivalently, for all $\mathsf{T}_{\mathsf{X}} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ and for every partition $\mathcal{K}(\mathsf{X}) \in \mathsf{Part}(\mathsf{X})$, the following holds [143]:

$$\curlyvee_{x\in\mathsf{X}^{(k)}} \mathscr{T}_x = \sum_{x\in\mathsf{X}^{(k)}} \mathscr{T}_x, \qquad \forall \mathsf{X}^{(k)} \in \mathcal{K}(\mathsf{X}). \tag{2.45}$$

Thanks to the linear structure, we have thus generalised the result of Theorem 1, which stated that the coarse-graining of scalars corresponds to the standard sum of real numbers.

This leads to a conceptual reinterpretation of the coarse-graining operation within the framework:

Sum represents the operation of ignoring occurrences, namely, a compound of mutually exclusive events.

Scalar multiplication

Analogously, the operation of scalar multiplication (2.25) becomes the usual scalar multiplication in real vector spaces:

$$(\mathrm{p}) \bullet \;\xrightarrow{\mathrm{A}}\boxed{\mathscr{T}}\xrightarrow{\mathrm{B}} \;=\; \xrightarrow{\mathrm{A}}\boxed{\mathrm{p}\mathscr{T}}\xrightarrow{\mathrm{B}} \quad .$$

This, too, admits a clear operational interpretation:

Scalar multiplication represents the parallel generation of a probability, namely, a randomisation process.

Observation 4

In a sense, the introduction of the linear structure makes the framework of OPTs practically “usable.” It guarantees that, regardless of how abstract or complex the constituents of a theory may be, it is always possible—at least in principle—to embed them into a real vector space; a mathematical object that physicists know and love, and know how to work with effectively.
Nevertheless, we remark that while such an embedding always exists in principle, constructing it explicitly can be highly non-trivial. Recent investigations, in collaboration with Dr. Elie Wolfe[a], explored the possibility of formulating an OPT based on the network of correlations introduced in Ref. [178]. However, the project turned out to be significantly more challenging than initially expected, primarily due to computational complexity. Even under the assumption of strong symmetries in the sources and measurement devices, the number of parameters required to characterise the system proved to be intractable on a standard laptop.
Further progress may be possible with access to greater computational resources.

[a] Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, N2L 2Y5; Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

The case of instruments

We can now present one of the novel result that we have introduced in Ref. [158]: how to associate a linear vector space to instruments.

Apart from the subtlety of dealing with families of linear maps when considering transformations—as discussed above—associating a vector space to the set of transformations is relatively straightforward. The case of instruments, however, requires a bit more manipulation, due to the fact that they are *ordered families* of transformations (Assumption 2). This implies that, in general,

$$\mathsf{Instr}(\mathrm{A} \to \mathrm{B}) \in \mathscr{P}_{\mathrm{ord}}\left\{\mathsf{Transf}(\mathrm{A} \to \mathrm{B})\right\},$$

for all $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, where $\mathscr{P}_{\mathrm{ord}}\left\{A\right\}$ denotes the set of all *ordered* subsets of S.

Given this, the most natural approach is to associate to $\mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ a *family* of linear spaces:

$$\mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B}) := \left\{\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})\right\}_{N \in \mathbb{N}}, \tag{2.46}$$

with

$$\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B}) := \bigoplus_{n=1}^{N} \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B}), \tag{2.47}$$

where N is the cardinality of the outcome space. In other words, once the number of outcomes of an instrument is fixed, it is natural to construct the vector space defined in (2.47). However, this construction captures only the instruments with a specific number of outcomes. Hence, $\mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ is defined as the union of all such vector spaces over all finite N.

To align with our previous nomenclature, we define a *generalised instrument* as an element $\{\mathscr{T}_x\}_{x \in \mathsf{X}} \in \mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, that is, a vector in $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ for some $N \in \mathbb{N}$. Equivalently, a generalised instrument can be regarded as an indexed family of generalised transformations[10].

Throughout this thesis, we assume that outcome spaces of tests and instruments have finite cardinality (see Assumption 1). Under this assumption, the operation of direct sum is always well-defined. In this setting, each space $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ is a vector space for any $N \in \mathbb{N}$, with the usual operations of scalar multiplication:

$$\mathrm{p}\left\{\mathscr{T}_1, \ldots, \mathscr{T}_N\right\} = \left\{\mathrm{p}\mathscr{T}_1, \ldots, \mathrm{p}\mathscr{T}_N\right\},$$

and element-wise addition:

$$\left\{\mathscr{T}_1, \ldots, \mathscr{T}_N\right\} + \left\{\mathscr{G}_1, \ldots, \mathscr{G}_N\right\} = \left\{\mathscr{T}_1 + \mathscr{G}_1, \ldots, \mathscr{T}_N + \mathscr{G}_N\right\},$$

where scalar multiplication $\mathrm{p}\mathscr{T}_1$ with $\mathrm{p} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{I})$, and addition $\mathscr{T}_1 + \mathscr{G}_1$ are those defined in $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$.

[10] To emphasise the distinction, we adopt the notation $\{\mathscr{T}_x\}_{x \in \mathsf{X}}$ for generalised instruments, reserving $[\![\mathscr{T}_x]\!]_{x \in \mathsf{X}}$ for proper instruments of an OPT (Remark 3).

Remark 14

The definition of generalised instruments can be extended to outcome spaces with arbitrary (possibly infinite) cardinality. In such a case, (2.47) would be generalised by replacing the direct sum with a direct product. This is necessary to accommodate generalised instruments defined by (possibly infinite) sequences of transformations that are not all null. The definition remains compatible with the finite case, since direct sum and direct product coincide for finite families of vector spaces.

The generalised spaces of instruments satisfy an interesting structural property: for any $N < M$, the space $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ can be seen as a subspace of $\mathsf{Instr}_{\mathbb{R}}^{(M)}(\mathrm{A} \to \mathrm{B})$. This follows from the fact that any instrument with N outcomes can be regarded as an instrument with M outcomes, in which the additional $M-N$ outcomes occur with zero probability. Formally,

$$\{\mathscr{T}_n\}_{n\in\mathsf{N}} \longrightarrow \{\mathscr{T}_{x_1}, \dots, \mathscr{T}_{x_N}, \underbrace{\varepsilon_{\mathrm{A}\to\mathrm{B}}, \dots, \varepsilon_{\mathrm{A}\to\mathrm{B}}}_{M-N}\}, \tag{2.48}$$

where $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ and

$$\{\mathscr{T}_{x_1}, \dots, \mathscr{T}_{x_N}, \underbrace{\varepsilon_{\mathrm{A}\to\mathrm{B}}, \dots, \varepsilon_{\mathrm{A}\to\mathrm{B}}}_{M-N}\} \in \mathsf{Instr}_{\mathbb{R}}^{(M)}(\mathrm{A} \to \mathrm{B}).$$

.

Coarse-graining The coarse-graining operation (2.27) on generalised instruments is naturally extended by linearity from the coarse-graining defined on the allowed instruments of a theory. In formulae, given a generalised instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ and a partition $\mathcal{K}(\mathsf{X}) \in \mathsf{Part}(\mathsf{X})$, we still have

$$\mathfrak{C}_{\mathcal{K}(\mathsf{X})}\left(\{\mathscr{T}_x\}_{x\in\mathsf{X}}\right) = \{\mathscr{T}_{\mathsf{X}^{(k)}}\}_{k\in\mathsf{K}},$$

where the coarse-grained transformations $\mathscr{T}_{\mathsf{X}^{(k)}}$ are generalised transformations given by

$$\mathscr{T}_{\mathsf{X}^{(k)}} = \sum_{x\in\mathsf{X}^{(k)}} \mathscr{T}_x, \quad \forall \mathsf{X}^{(k)} \in \mathcal{K}(\mathsf{X}).$$

This makes the coarse-graining operation a linear map from $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ into $\mathsf{Instr}_{\mathbb{R}}^{(M)}(\mathrm{A} \to \mathrm{B})$, with $M < N$.

If one thinks of an instrument as a collection of N matrices, each of dimension $D_{\mathrm{AE}} \cdot D_{\mathrm{BE}}$, then the coarse-graining map is a linear transformation from a space of dimension $D_{\mathrm{AE}} \cdot D_{\mathrm{BE}} \cdot N$ to one of dimension $D_{\mathrm{AE}} \cdot D_{\mathrm{BE}} \cdot M$. In other words, coarse-graining prescribes which of the N matrices must be summed together to yield an instrument with M outcomes.

2.3.2 Topological structure

As briefly discussed earlier, within the framework we want to be able to make statements about the *similarity* between different instruments and transformations. This would allow us, for instance, to formally address the question of how likely an agent is to successfully distinguish between two transformations.

To achieve this, we introduce a function that captures the *operational distance* between instruments and transformations, such that the more similar two objects are, the smaller their operational distance and vice versa.

This function, in turn, induces a *topological structure* on the spaces of generalised transformations $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ and generalised instruments $\mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ for any pair of systems in the OPT. In particular, we will exploit the linear structure just defined to construct this function.

Operational norm for states

Let us start with the case of states. We will then extend the notion of distance between states to define the distance between transformations.

Suppose an agent (Alice) can prepare two deterministic states ρ_0 and $\rho_1 \in \mathsf{St}_1(\mathrm{A})$ and send them to another agent (Bob), whose goal is to determine which preparation Alice has chosen. The most general strategy that Bob can adopt is to perform a dichotomic measurement, described by an appropriate observation-instrument $[\![\mathrm{a}_0, \mathrm{a}_1]\!] \in \mathsf{Obs}(\mathrm{A})$, with outcomes 0 and 1 corresponding to guessing ρ_0 and ρ_1, respectively. Performing a measurement with more than two outcomes would be equivalent—by coarse-graining—to the dichotomic case, while introducing a non-trivial output system would retain information we are not interested in. The best probability of successful discrimination is given by:

$$\mathrm{p}_{\mathrm{succ}} = \mathrm{p}_0 \left(\mathrm{a}_0 \middle| \, \rho_0\right)_{\mathrm{A}} + \mathrm{p}_1 \left(\mathrm{a}_1 \middle| \, \rho_1\right)_{\mathrm{A}} ,$$

where we also assume that Bob, knowing Alice, assigns prior probabilities p_0 and $\mathrm{p}_1 = 1 - \mathrm{p}_0$ to the preparations ρ_0 and ρ_1, respectively.

Defining $\mathrm{e}_k := \mathrm{a}_0 + \mathrm{a}_1 \in \mathsf{Eff}_1(\mathrm{A})$ as the deterministic effect corresponding to the coarse-graining of Bob's measurement, we can rewrite the probability of success as:

$$\begin{aligned} p_{\text{succ}} &= (e_k|\ p_0\rho_0)_A + (a_1|\ p_1\rho_1 - p_0\rho_0)_A = p_0 + (a_1|\ p_1\rho_1 - p_0\rho_0)_A \\ &= (e_k|\ p_1\rho_1)_A - (a_0|\ p_1\rho_1 - p_0\rho_0)_A = p_1 - (a_0|\ p_1\rho_1 - p_0\rho_0)_A\,, \end{aligned}$$

where we used the fact that both ρ_0 and ρ_1 are deterministic, i.e., $(e_k|\ \rho_x)_A = 1$ for $x = 0, 1$. Combining these expressions, we find:

$$p_{\text{succ}} = \frac{1}{2}\left(1 + (a_1 - a_0|\ p_1\rho_1 - p_0\rho_0)_A\right).$$

Since Bob seeks the optimal strategy, he will optimise over all possible observation-instruments $[\![a_0, a_1]\!] \in \mathsf{Obs}(A)$. The optimal probability of success is thus given by:

$$p^{(\text{opt})}_{\text{succ}} = \sup_{[\![a_0,a_1]\!]\in\mathsf{Obs}(A)} \frac{1}{2}\left(1 + (a_1 - a_0|\ p_1\rho_1 - p_0\rho_0)_A\right). \tag{2.49}$$

It is straightforward to verify that the above quantity always lies between 0 and 1.

From (2.49), we can define a linear functional $\|\cdot\|_{op}$ on $\mathsf{St}_{\mathbb{R}}(A)$ such that

$$p^{(\text{opt})}_{\text{succ}} = \frac{1}{2}\left(1 + \|p_1\rho_1 - p_0\rho_0\|_{op}\right).$$

More generally, for any $\tau \in \mathsf{St}_{\mathbb{R}}(A)$, we define the *operational norm* for states as

$$\|\tau\|_{op} := \sup_{[\![a_0,a_1]\!]\in\mathsf{Obs}(A)} (a_1 - a_0|\ \tau)_A\,. \tag{2.50}$$

We note that, due to the linearity of the functional, the operational norm of a generalised state is always finite. To see this, it suffices to express the generalised state as a linear combination of states in $\mathsf{St}(A)$ and use the fact that all such states yield bounded values when paired with effects in $\mathsf{Eff}(A)$.

By direct calculation, one can then verify that the functional defined in (2.50) satisfies the properties of a norm, which we recall here for completeness.

Definition 2 (Norm (over a real vector space))

A *norm* on a real vector space V is a function

$$\|\cdot\| : V \to \mathbb{R},$$

that satisfies the following properties[a]:

- *Triangle inequality*:

$$\|v + w\| \leq \|v\| + \|w\| \, ,$$

for all $v, w \in V$.

- *Absolute homogeneity*:

$$\|rv\| = |r| \, \|v\| \, ,$$

for all $r \in \mathbb{R}$ and $v \in V$.

- *Point-separation*: if $v \in V$ and $\|v\| = 0$, then $v = 0$.

[a] We remind the reader that together the triangle inequality and absolute homogeneity imply the property of *non-negativity*.

From the operational norm, we can define the *operational distance* between two states as

$$d_{op}(\rho, \sigma) := \|\rho - \sigma\|_{op} \, ,$$

for any $\rho, \sigma \in \mathsf{St}_{\mathbb{R}}(\mathrm{A})$ and any system A of the theory.

This distance satisfies the *desideratum* expressed at the beginning of this section: there is a clear relationship between the distinguishability of two states and their distance. The further apart they are, the easier it is to discriminate between them, and vice versa. Indeed, the strategy implemented by Bob is the optimal one for distinguishing between deterministic states.

Observation 5

It can be shown that the operational norm defined in (2.49) coincides with the *trace norm* $\|\rho\|_1 = \mathsf{Tr}\left[\sqrt{\rho\rho^\dagger}\right]$ in QT [133].

Operational norm for transformations

Starting from the definition of the operational norm for states (2.50), it is possible to define a norm on the spaces of generalised transformations $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ as well.

Let A and B be two generic systems of an OPT, and let $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$.

We define:

$$\|\mathscr{T}\|_{op} := \sup_{\mathrm{E}\in\mathsf{Sys}(\Theta)} \sup_{\rho\in\mathsf{St}_1(\mathrm{AE})} \left\|\mathscr{T}\,|\rho)_{\mathrm{AE}}\right\|_{op}, \tag{2.51}$$

where the norm on the right-hand side is the operational norm for states, as defined in (2.50). In words, the operational norm of a transformation is defined as the supremum of the operational norms of the output states $\mathscr{T}\,|\rho)_{\mathrm{AE}}$, where the input is any deterministic joint state $\rho \in \mathsf{St}_1(\mathrm{AE})$ with E an arbitrary system of the theory.

By direct calculation, one can verify that the quantity defined in (2.51) satisfies the axioms of a norm, as per Definition 2.

This definition naturally leads to a corresponding notion of distance between transformations:

$$d_{op}(\mathscr{T},\mathscr{G}) := \|\mathscr{T}-\mathscr{G}\|_{op}.$$

Exactly what we were looking for.

Is this the best we can do? In the case of states, the operational norm (2.50) and the associated distance are directly related to the optimal discrimination strategy that one can implement. For transformations, this is no longer the case. In Ref. [143], the authors propose an alternative definition of the operational norm for transformations, derived from the analysis of optimal discrimination strategies. The idea is based on the use of "tester" circuits, like to those used in defining equivalence of tests and events (2.31).

To optimally discriminate two transformations, one must optimise over all possible preparation- and observation-instruments that form a closed circuit. The norm defined in (2.51) corresponds to a particular case of this general strategy, namely when the preparation-instrument is a singleton and the observation-instrument is dichotomic, i.e., has only two outcomes. Therefore, it provides a lower bound on the discrimination probability.

Nevertheless, this distinction is ultimately unimportant for our purposes. This is due to two reasons:

I) The results derived in this thesis regarding the operational norm rely on the fact that it satisfies the property of *monotonicity* (Lemma 6), a condition fulfilled in both its optimal and our less optimal formulations.

II) We work under the assumption that all vector spaces of generalised states are finite-dimensional (Assumption 3), which implies that the generalised spaces of transformations are also finite-dimensional. In such spaces, all norms are *equivalent*—that is, equal up to multiplicative constants. Hence, all results we will derive hold true independently of the specific choice of norm.

We conclude by recalling the formal definition of equivalent norms:

Definition 3 (Equivalent norms)

Two norms p and q on a vector space V are said to be *equivalent* if there exist two positive real constants c and C such that for every vector $v \in V$:

$$cq(v) \leq p(v) \leq Cq(v).$$

The relation is symmetric: if p and q are equivalent, then

$$\frac{1}{C}p(v) \leq q(v) \leq \frac{1}{c}p(v).$$

Since it is reflexive, symmetric, and transitive, norm equivalence defines an equivalence relation on the space of norms over V.

Sup norm for transformations

Ignoring for a moment the equivalence of norms, when considering generic OPTs, the operational norm is generally not sufficient for all applications of interest within the framework. For instance, it is generally not possible to determine the relationship between $\|\mathscr{T}\mathscr{G}\|_{op}$ and $\|\mathscr{G}\|_{op}\|\mathscr{T}\|_{op}$. Furthermore, the operational norm is not suitable for the construction of Cellular Automata (CAs) within the framework [134]. For these reasons, a stronger norm is introduced.

Definition 4 (Sup norm)

Given a generic OPT, the *sup norm* $\|\mathscr{T}\|_{sup}$ of $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ is defined as [134]

$$\begin{aligned} &\|\mathscr{T}\|_{sup} := \inf J(\mathscr{T}), \\ &J(\mathscr{G}) := \{\lambda | \exists \mathscr{A} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B}), \lambda\mathscr{A} \succcurlyeq \mathscr{G} \succcurlyeq -\lambda\mathscr{A}\}. \end{aligned} \tag{2.52}$$

Here, $\mathscr{T} \succcurlyeq \mathscr{G}$ denotes the partial order induced by a cone structure on the space $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, defined as follows:

$$K(\mathrm{A} \to \mathrm{B}) := \{\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B}) \,|\, \mathscr{T} \succcurlyeq 0\},$$

where we write $\mathscr{T} \succcurlyeq 0$ when the following condition holds: for every system E in the theory and every state $|\rho)_{\mathrm{AE}} \in \mathsf{St}_{+}(\mathrm{AE})$, the output state

$$(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}}$$

belongs to $\mathsf{St}_{+}(\mathrm{BE})$ [134]. The set $\mathsf{St}_{+}(\mathrm{A})$, as will be discussed in detail in section 3.1, is defined as

$$\mathsf{St}_{+}(\mathrm{A}) \coloneqq \mathsf{Span}_{+}(\mathsf{St}(\mathrm{A})) \equiv \mathsf{ConicHull}(\mathsf{St}(\mathrm{A}))$$

that is, the conic hull of the set of states, where we recall that the conic hull of a set S is defined as

$$\mathsf{ConicHull}(S) \coloneqq \left\{ \sum_{i=1}^{k} \alpha_i s_i \,\middle|\, s_i \in S, \alpha_i \in \mathbb{R}_{\geq 0}, k \in \mathbb{N} \right\}. \tag{2.53}$$

A note on Cellular Automata As discussed at the beginning of this thesis, the goal of the OPT framework is to describe physical theories as theories of information processing, starting from minimal assumptions. As we will see, this approach is quite powerful and enables the derivation of many results that characterise the relationships between distinct physical properties—for instance, the interplay between complementarity and uncertainty, which will be discussed in detail later in this thesis.

One consequence of this approach—in accordance with the operational ideology—is the absence of a direct connection with traditional physical concepts such as mass, energy, position, and space-time. Within the framework of OPTs, systems can be interpreted as memory cells, and the transformations as updates to their state.

A recent proposal to bridge this gap and reintroduce more "traditional" physical structure is based on the idea that physical laws can be understood as algorithms that evolve systems by changing their states. Extending the philosophy behind operational frameworks, physical laws become algorithms acting on strings of memory cells—in essence, they are treated analogously to computer programs.

This is the approach underlying Cellular Automata (CAs).

The theory of CA is a well-established branch of computer science, which has been recently extended to the quantum realm by defining the notion of [31, 181–186]. This line of research has led to remarkable results, including the successful reconstruction of fundamental equations such as those of Weyl, Dirac, and Maxwell [187–190].

In Ref. [134], this formalism is further extended to encompass generic CAs defined over arbitrary OPTs, paving the way for the study of more "traditional" physical concepts for generic theories of information processing.

Operational distance between effects

Last but not least, from the definition of the operational norm for transformations (2.51), one can also induce a corresponding norm for effects. For all effects $\mathrm{a} \in \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$, we define:

$$\|\mathrm{a}\|_{op} := \sup_{\rho \in \mathsf{St}_1(\mathrm{A})} \left| (\mathrm{a}|\, \rho)_{\mathrm{A}} \right| .$$

Analogously, the *operational distance* between two generalised effects is defined as:

$$d_{op}(\mathrm{a}, \mathrm{b}) := \|\mathrm{a} - \mathrm{b}\|_{op} .$$

The case of instruments

Following Ref. [119], we can now introduce a notion of norm also for instruments.

Definition 5 (Norm for instruments)

Let Θ be an OPT in which the spaces of transformations are equipped with a norm $\|\cdot\|_{gen}$, and let $\{\mathscr{T}_n\}_{n\in N} \in \mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ be a generalised instrument. We define its norm as

$$\left\| \{\mathscr{T}_n\}_{n\in N} \right\|_{gen} := \sum_{n\in N} \|\mathscr{T}_n\|_{gen} , \tag{2.54}$$

for any $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and any outcome-space cardinality $N \in \mathbb{N}$ [158].

As in the case of transformations, different norms can be defined over the space of generalised instruments. Here, we have chosen to define the norm of an instrument in terms of the norms of its constituent transformations. This approach has the advantage that the convergence of sequences of generalised instruments is closely tied to the convergence of the transformations composing them: one implies the other and vice versa (see Theorem 7).

More precisely, the above definition describes a *family* of norms, parametrised by the choice of norm on transformations. For instance, if the operational norm $\|\cdot\|_{op}$ is chosen for the transformations, we obtain the *operational norm for instruments*, given by

$$\left\| \{\mathscr{T}_n\}_{n\in N} \right\|_{op} = \sum_{n\in N} \|\mathscr{T}_n\|_{op} ,$$

for any $\{\mathscr{T}_n\}_{n\in N} \in \mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$.

The norm defined in (2.54) is generally not the optimal one for discriminating between two instruments. Defining such a norm would require solving a minimax problem, whose analysis is often complex and only tractable under special assumptions. For this reason, the development of such an optimal norm is left for future work.

In the setting of this thesis, however, the distinction between different norms on instruments is ultimately *meaningless*. Since we work with finite-dimensional vector spaces (Assumption 3), all norms are equivalent (Definition 3).

The *operational distance* between two generalised instruments is defined as:

$$d_{op}\left(\{\mathscr{T}_x\}_{x\in\mathsf{X}}, \{\mathscr{G}_x\}_{x\in\mathsf{X}}\right) := \sum_{x\in\mathsf{X}} \|\mathscr{T}_x - \mathscr{G}_x\|_{op},$$

where we remark that the distance between instruments can be defined only for instruments with the same outcome-space cardinality. However, as shown in (2.48), any two instruments can be compared by suitably adding a sufficient number of null transformations.

We observe that this definition can be generalised by considering any suitable norm in place of the operational one.

Operational completeness

The definition of the topological structure for OPTs allows us to formalise an operational *desiderata*:

Assumption 4: Operational completeness

For any OPTs Θ, the sets of instruments $\mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ and transformations $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ are supposed to be Cauchy-complete with respect to the operational norm, for any pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$.
More specifically, in the case of instruments, we require that the sets $\mathsf{Instr}^{(N)}(\mathrm{A} \to \mathrm{B})$ are Cauchy-complete for any $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and for any outcome-space cardinality $N \in \mathbb{N}$. Here, $\mathsf{Instr}^{(N)}(\mathrm{A} \to \mathrm{B})$ denotes the subset of $\mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ consisting of instruments whose outcome space has cardinality N.

In words, we are requiring that if there exists a generalise transformation [instrument] that can be arbitrarily approximated by a proper transformation [instrument] of the theory, then the generalised transformation [instrument] is an actual transformation [instrument] of the theory.

The relationship with closure In general, the requirement that a subspace of a topological space be Cauchy-complete is stronger than the requirement that it be closed. While Cauchy-completeness implies closure, the converse is not true in general. A classical counterexample is the set of rational numbers $\mathbb{Q}$, which is closed in itself but not Cauchy-complete[11].

The two notions become equivalent when the subspace in question is embedded in a complete topological space. This is precisely the case in this thesis. The vector spaces of instruments and transformations we consider are all isomorphic to $\mathbb{R}^d$ for some $d \in \mathbb{N}$, which is Cauchy-complete. Therefore, the spaces $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ are Cauchy-complete for all $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and $N \in \mathbb{N}$.

As a consequence, in the present setting it makes no difference whether one proves that a space is closed or Cauchy-complete. Nonetheless, for generality, we will formulate our arguments in terms of Cauchy sequences whenever possible.

2.4 Summary of the assumptions

With the introduction of the topological structure, we have completed the construction of the framework, along with all its additional structures.

To conclude the chapter, we now provide a summary of all the assumptions made throughout this about the framework.

Assumption 1 : Finite outcome spaces

For any test $\mathsf{T}_{\mathsf{X}} \in \mathsf{Test}(\Theta)$ in an OPT Θ, the cardinality of the outcome space X is assumed to be finite, i.e., $|\mathsf{X}| < \infty$.

Assumption 2 : The order matters

If two tests consist of the same events in a different order—e.g., $\mathsf{T}_{\mathsf{X}} \equiv [\![\mathscr{T}_1, \mathscr{T}_2]\!]$ and $\mathsf{T}'_{\mathsf{X}} \equiv [\![\mathscr{T}_2, \mathscr{T}_1]\!]$—they are *not* equal. The only exception is when all the events of the tests coincide.

Assumption 3 : Finite dimensional vector spaces

The dimension D_{A} of every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ in an OPT Θ is assumed to be finite, that is, $D_{\mathrm{A}} < +\infty$.

[11] That is, as a subspace of itself $\mathbb{Q}$ is trivially closed—but not complete.

Assumption 4 : Operational completeness

For any OPTs Θ, the sets of instruments $\mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ and transformations $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ are supposed to be Cauchy-complete with respect to the operational norm, for any pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$.
More specifically, in the case of instruments, we require that the sets $\mathsf{Instr}^{(N)}(\mathrm{A} \to \mathrm{B})$ are Cauchy-complete for any $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and for any outcome-space cardinality $N \in \mathbb{N}$. Here, $\mathsf{Instr}^{(N)}(\mathrm{A} \to \mathrm{B})$ denotes the subset of $\mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ consisting of instruments whose outcome space has cardinality N.

3 Properties of OPTs

In the previous chapter, we mentioned various properties that OPTs may satisfy, among which *causality* appeared most prominently. In the present chapter, we develop a systematic treatment of several such properties.

We begin by analysing the geometrical structure of OPTs, introducing the notion of *atomicity*, which will play a pivotal role in the developments to follow. We then establish several properties of the operational and sup norms, and examine how the convergence of sequences of generalised transformations and instruments is intertwined.

Next, we introduce the different degrees of discriminability that an OPT may fulfil, and subsequently the concepts of *causal* and *strongly causal* OPTs, also showing how every causal OPT can be extended to a strongly causal one.

The chapter concludes with a discussion of the notion of *classicality* adopted here and of the properties that follow from it.

3.1 Atomicity, extremality, and more

The coarse-graining operation (2.30) provides a natural way to classify transformations based on how they can be decomposed into combinations of other transformations within the theory. This leads to the formalisation of familiar notions such as *purity*, *atomicity*, and *mixedness*.

Atomic transformations Let's start by defining

Definition 6 (Refinement set)

Let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$. A *refinement* of $\mathscr{T}$ is a collection $\{\mathscr{T}_{\tilde{x}}\}_{\tilde{x}\in\tilde{\mathsf{X}}} \subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ of transformations such that $\sum_{\tilde{x}\in\tilde{\mathsf{X}}} \mathscr{T}_{\tilde{x}} = \mathscr{T}$. In words, a refinement of $\mathscr{T}$ is a collection of transformations whose coarse-graining yields $\mathscr{T}$. The *refinement set* of $\mathscr{T}$, denoted with $\mathsf{Ref}(\mathscr{T})$ is the union of all the refinements of $\mathscr{T}$.

and subsequently

Definition 7 (Atomic transformation)

A transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is *atomic* if its refinement set $\mathsf{Ref}(\mathscr{T})$ is trivial, namely, every element in the set is proportional to $\mathscr{T}$. That is, for all $\mathscr{G} \in \mathsf{Ref}(\mathscr{T})$ it holds that $\mathscr{G} \propto \mathscr{T}$. In other words, $\mathscr{T}$ is atomic if, given $\mathscr{T}_1, \mathscr{T}_2 \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, one has the following implication:

$$\mathscr{T} = \mathscr{T}_1 + \mathscr{T}_2 \implies \mathscr{T}_1, \mathscr{T}_2 \propto \mathscr{T}. \tag{3.1}$$

The notion of atomic transformations captures the idea of "indecomposable" events from a conic point of view. Formally, these are the transformations that generate the extremal rays of the cone associated with the transformation set:

$$\mathsf{Transf}_+(\mathrm{A} \to \mathrm{B}) \equiv \mathsf{ConicHull}(\mathsf{Transf}(\mathrm{A} \to \mathrm{B}))\,, \tag{3.2}$$

where we recall that for us the conic hull of a set S is defined as

$$\mathsf{ConicHull}(S) := \left\{ \sum_{i=1}^{k} \alpha_i s_i \,\middle|\, s_i \in S,\ \alpha_i \in \mathbb{R}_{\geq 0},\ k \in \mathbb{N} \right\}. \tag{2.53}$$

We highlight that we always include the zero vector in the conic hull.

In this thesis, the definition of atomic transformations is based solely on the linear structure of an OPT. In (3.1), there are no additional constraints on $\mathscr{T}_1$ and $\mathscr{T}_2$ beyond requiring that they belong to the same transformation space as $\mathscr{T}$. However, one could alternatively define a stricter notion of refinement—and consequently of atomicity—in which the transformations are required to belong to the same instrument [143]. Such a notion could be formulated entirely in terms of the coarse-graining operation (2.30), without reference to the underlying linear structure. Our definition, instead, is meaningful only when the transformations in (3.1) are interpreted as elements of $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, where the sum operation is always well-defined, regardless of whether the transformations are part of the same instrument.

Extremal transformations A similar line of reasoning can be applied to convex combinations.

Definition 8 (Convex-refinement set)

Let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$. A *convex-refinement* of $\mathscr{T}$ is a collection $\{\mathscr{T}_{\tilde{x}}\}_{\tilde{x} \in \tilde{\mathsf{X}}} \subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ of transformations such that

$$\mathscr{T} = \sum_{\tilde{x} \in \tilde{\mathsf{X}}} \mathrm{p}_{\tilde{x}} \mathscr{T}_{\tilde{x}},$$

for some probability distribution $[\![\mathrm{p}_{\tilde{x}}]\!]_{\tilde{x} \in \tilde{\mathsf{X}}}$. In words, a convex-refinement of $\mathscr{T}$ is a collection of transformations whose convex combination yields $\mathscr{T}$. The *convex-refinement set* of $\mathscr{T}$, denoted by $\mathsf{ConvRef}(\mathscr{T})$, is the union of all convex-refinements of $\mathscr{T}$.

Definition 9 (Extremal transformations)

A transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is called *extremal* if, for all $\mathscr{T}_1, \mathscr{T}_2 \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and $\mathrm{p} \in (0, 1)$, the condition

$$\mathscr{T} = \mathrm{p}\mathscr{T}_1 + (1 - \mathrm{p})\, \mathscr{T}_2$$

implies that $\mathscr{T}_1 = \mathscr{T}_2 = \mathscr{T}$. In other words, the convex-refinement of $\mathscr{T}$ is trivial: $\mathsf{ConvRef}(\mathscr{T}) \equiv \{\mathscr{T}\}$.

Extremal transformations capture the idea of *extreme points* in convex sets. This concept is especially relevant in a particular class of theories in which each set of transformations is itself convex.

Definition 10 (Convex OPTs)

An OPT Θ is said to be *convex* [133, 167] if for every pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the set $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ coincides with its convex hull, where $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is seen as a set in $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$.

In general, the notions of atomicity and extremality are *independent*: a trans-

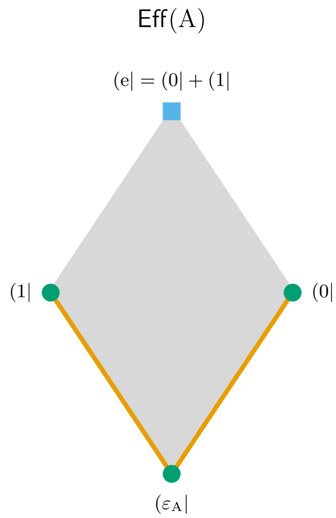


Figure 3.1: Example of a set of effects for a system A of size $D_{\mathrm{A}} = 2$. For example, this would represent the set of all measurements that can be performed on a bit. The elements depicted in orange are the atomic ones, while those in green—the null effect $(\varepsilon_{\mathrm{A}}|$, $(1|$ and $(0|$—are both atomic and extremal. The element in blue—$(\mathrm{e}| = ((0| + (1|)$—is the pure effect of $\mathsf{Eff}(\mathrm{A})$, which is deterministic and extremal but not atomic. The elements in light grey are the mixed effects of A.

formation can be extremal but not atomic, or atomic but not extremal. A notable example is the deterministic effect of a system in CT or QT [133], which is an extremal point in the set of effects but can nevertheless be obtained as the coarse-graining of many different observation-instruments. This can be easily seen from the geometric representation of effect spaces; for instance, one may consider the effect space of a bit shown in figure 3.1.

Pure and mixed transformations Departing slightly from the terminology commonly used in QT, and following Ref. [167], we adopt the following definitions:

Definition 11 (Pure and mixed transformations)

A transformation is called *pure* if it is both extremal and deterministic. Conversely, a transformation is called *mixed* if it is neither atomic nor extremal.

In the following, the set of pure states of a system A will be denoted by $\mathsf{PurSt}(\mathrm{A})$.

All of these notions admit a clear geometric interpretation, as illustrated in figure 3.1 and figure 3.2, where the effect and state spaces of a system with $D_{\mathrm{A}} = 2$, such as a classical bit, are depicted, respectively. These figures highlights the different decompositional properties of the elements of $\mathsf{Eff}(\mathrm{A})$ and $\mathsf{St}(\mathrm{A})$.

3.1.1 Amygdaloidal OPTs

From an informational point of view, atomic transformations represent operations in which no information has been discarded. They correspond to events that cannot be further refined and express the maximal knowledge attainable about the physical process occurring in the experiment. In this sense, atomic transformations embody the maximum level of control achievable in a laboratory setting.

Being accustomed to the structure of QT and CT, it seems natural to assume that atomic states and effects always exist. Consider, for example, the pure states of a quantum system. Without them, one would be forced to accept the existence of theories in which it is occasionally impossible to prepare a system in a specific state.

Surprisingly, in the rich and exotic zoo of OPTs, such theories do exist. These are called *amygdaloidal OPTs*. Formally:

Definition 12 (Amygdaloidal OPT)

An *amygdaloidal OPT* is an OPT in which $\mathsf{St}_{+}(\mathrm{A}) \subset \overline{\mathsf{St}_{+}(\mathrm{A})}$ and $\mathsf{Eff}_{+}(\mathrm{A}) \subset \overline{\mathsf{Eff}_{+}(\mathrm{A})}$ for every system A of the theory. In words, an amygdaloidal OPT is an OPT in which the cone spanned by the set of states and effects is not closed.

From this definition, it is clear that such theories are highly peculiar. They contain states and effects that cannot be accessed—not even in principle. These limitations are *structural*, rather than *operational*: they arise not from imperfections or noise, but from the very structure of the theory itself.

To make this idea more concrete, let us consider a familiar example: the state space of a classical bit. A representation of it is shown in figure 3.2.

In figure 3.3, we depict a noisy version of the state space of a classical bit.

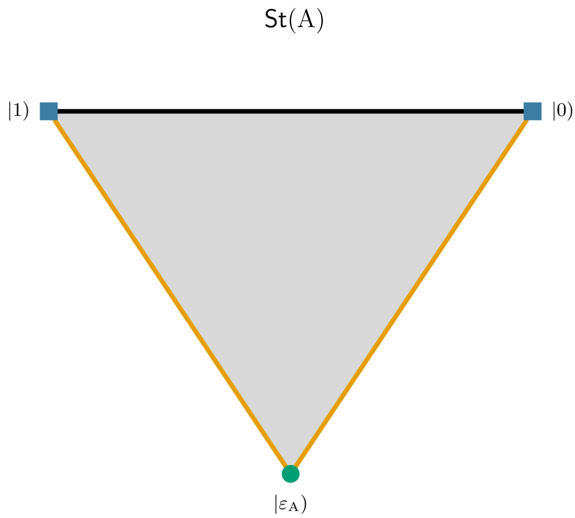


Figure 3.2: Example of a set of states for a system A of size $D_{\mathrm{A}} = 2$. $\mathsf{St}(\mathrm{A})$ is convex and represents the state space of a classical bit. The elements depicted in orange are the atomic ones, while the green point—the null state $|\varepsilon_{\mathrm{A}})$—is both atomic and extremal. The elements in blue—$|0)$ and $|1)$—are the pure states of $\mathsf{St}(\mathrm{A})$. A darker shade of blue is used to indicate that they are not only extremal, but also atomic. The elements in light grey are the mixed states of A, while those in black are the deterministic ones ($\mathsf{St}_1(\mathrm{A})$).

In this case, it is not possible to prepare the system in any state other than the maximally mixed one. Substantially, we are modelling a coin that can only be thrown, but never set to a specific face. However, this is *not* an amygdaloidal OPT, since—in principle—removing all noise from the preparation process would make it possible to realise the pure states $|0)$ and $|1) \in \mathsf{PurSt}(\mathrm{A})$.

In contrast, the state space of an amygdaloidal OPT is represented in figure 3.4. This state space satisfies all the structural requirements of the framework. Since it is closed—and hence complete, given that $D_{\mathrm{A}} < \infty$—, it satisfies Assumption 4. However, it contains no atomic state other than the null one—not even in principle.

In this case, the cone $\mathsf{St}_+(\mathrm{A})$ is an open set whose closure is the full cone of the classical bit:

$$\overline{\mathsf{St}_+(\mathrm{A})} = \mathsf{ConicHull}(|0)\,, |1)\,, \varepsilon_{\mathrm{A}})$$

At first glance, this might appear to be a mere curiosity, a rather odd but incon-

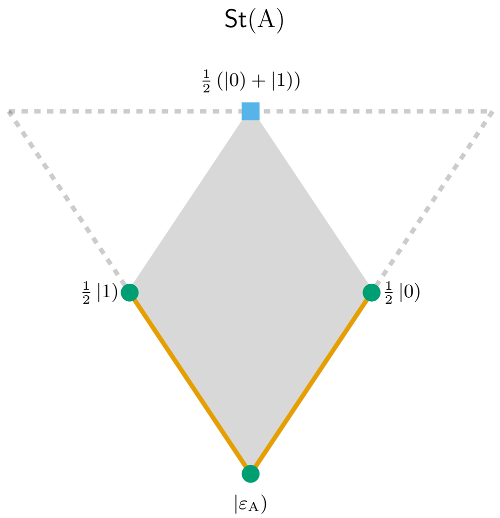


Figure 3.3: Example of a set of states for a system A of size $D_A = 2$. $\mathsf{St}(A)$ is convex and contained within the state space of a classical bit. It represents a noisy version of a classical bit. The elements depicted in orange are the atomic ones, while those in green—the null state $|\varepsilon_A)$, $\frac{1}{2}|1)$, and $\frac{1}{2}|0)$—are both atomic and extremal. The element in blue—$\frac{1}{2}(|0) + |1))$—is the pure state of $\mathsf{St}(A)$, which is deterministic and extremal, but not atomic. The elements in light grey are the mixed states of A.

sequential corner of the theoretical landscape. However, this is not the case. Many results within the framework rely on the existence of atomic states or transformations. The fact that these objects cannot be decomposed makes them a powerful tool in proofs. Therefore, it is important to find ways either to circumvent this issue or to identify sufficient conditions guaranteeing the existence of atomic states.

One way to carry out proofs even in the absence of atomic states in a theory is to exploit *generalised states*. In particular, one can consider the generalised states lying on the extremal rays of the closed cone spanned by $\mathsf{St}(A)$, which are atomic states of $\overline{\mathsf{St}_+(A)}$ by definition.

This strategy is adopted in the proof of Theorem 26, where we provide a generalisation of the original argument presented in Ref. [191], which explicitly relies on the existence of atomic states and effects.

An alternative approach consists in working under assumptions that guarantee

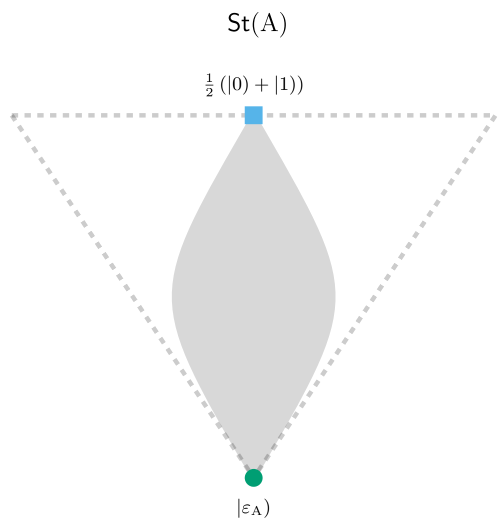


Figure 3.4: Example of an amygdaloidal set of states for a system A of size $D_A = 2$. $\mathsf{St}(A)$ is convex and contained within the state space of a classical bit. It represents an amygdaloidal version of a classical bit. The element depicted in green—the null state $|\varepsilon_A)$—is both atomic and extremal. The element in blue—$\frac{1}{2}(|0) + |1))$—is the pure state of $\mathsf{St}(A)$, which is deterministic and extremal, but not atomic. The elements in light grey are the mixed states of A.

the existence of (non-generalised) atomic states. For instance, this is the case when restricting attention to strongly causal OPTs (Definition 21) .

Strongly causal OPTs are characterised by the property that any state is proportional to a deterministic one (Theorem 20). In particular, in such theories the set of normalised states $\mathsf{St}_1(A)$ is bounded. As a consequence, one has

$$\mathsf{ConicHull}(\mathsf{St}(A)) = \mathsf{ConicHull}(\mathsf{St}_1(A))\,.$$

This, in turn, guarantees that the theory is not amygdaloidal, i.e., that its state cone is closed, as follows from the next theorem.

Theorem 4

The cone generated by a bounded, closed set B with non-zero distance from the origin is closed.

Remark 15

We remark that Theorem 4 holds even in the infinite-dimensional case.

The proof of Theorem 4 is provided in appendix B.

Atomic states in proofs Above we mentioned that atomic states play a central role in many proofs within the OPT framework, due to their *indecomposability*. We now explain in detail why this property makes them particularly useful.

A recurring task in OPT proofs is to establish the equality between two instruments. By definition, this requires showing that the two objects act identically on *all* states of their input system—or, equivalently, all effects of their output system. However, in a generic OPT there is no explicit characterisation of either the set of states or the set of transformations. As a result, such a verification is, in general, highly non-trivial.

A standard strategy when working within abstract frameworks to simplify this task is to identify a spanning set of states, and to verify the desired equality only on that set. Linearity then guarantees equality on all states. In this context, atomic states provide a particularly convenient choice, as guaranteed by the following result.

Theorem 5

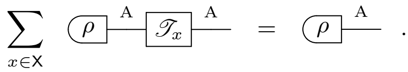
In an OPT, the set $\overline{\mathsf{St}_+(\mathrm{A})}$ is the convex hull of its extremal rays together with the null state.

In words, this means that any generalised state $\rho \in \overline{\mathsf{St}_+(\mathrm{A})}$ can be written as a convex combination of generalised states belonging to the extremal rays of the cone. The generalised states associated with the extremal rays of the closed cone are, by definition, atomic in $\overline{\mathsf{St}_+(\mathrm{A})}$. Therefore, when working within $\overline{\mathsf{St}_+(\mathrm{A})}$, proofs can be carried out by exploiting atomic generalised states, without requiring the existance of atomic states.

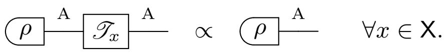
Whenever the theory admits atomic (non-generalised) states, these necessarily lie on the extremal rays of the state cone.

Although Theorem 5 might appear intuitive, it is in fact a consequence of a non-trivial mathematical result, namely Klee's theorem (Theorem 45). The proof of the theorem is provided in appendix B.

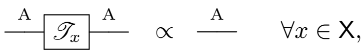
The usefulness of atomic states becomes especially apparent when proving that an instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{A})$ coincides with the identity test $\mathscr{I}_{\mathrm{A}}$. Consider an atomic (possibly generalised) state ρ of system A. For the instrument to be equal to the identity, it must satisfy



Since the right-hand side is atomic, the indecomposability of ρ implies that each term in the sum must be proportional to it, namely



If this proportionality condition holds for all atomic states, which form a spanning set, then it follows that



which is an extremely strong and easily exploitable constraint on the instrument.

This reasoning constitutes the main technical tool underlying the proofs of Theorem 38 and Theorem 41.

3.2 Properties of the norms

In our discussion on the topological structure of OPTs (section 2.3.2), we hinted at some important properties satisfied by the norms introduced therein. In particular, the operational norm (2.51) and the sup norm (2.52) exhibit a number of interesting features. Moreover, the definition of the norm for instruments (2.54)—as a sum over the norms of the constituent transformations—entails a symmetry in the convergence behaviour of sequences of generalised instruments and of the transformations composing them.

In this section, we explicitly present these results.

3.2.1 Properties of the operational norm

Let us begin with the case of the operational norm (2.51).

Lemma 6 (Monotonicity of the operational norm)

In every OPT, the operational norm of a generalised transformation $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{B} \to \mathrm{C})$ satisfies

$$\|\mathscr{E}\mathscr{T}\mathscr{C}\|_{op} \leq \|\mathscr{T}\|_{op},$$

where $\mathscr{E} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$ and $\mathscr{C} \in \mathsf{Transf}_1(\mathrm{C} \to \mathrm{D})$ are deterministic transformations.
Equality holds if both $\mathscr{E}$ and $\mathscr{C}$ are reversible [133].

Proof. To prove this result, let us recall the definition of the operational norm for transformations (2.51):

$$\|\mathscr{T}\|_{op} = \sup_{\mathrm{E} \in \mathsf{Sys}(\Theta)} \sup_{\rho \in \mathsf{St}_1(\mathrm{BE})} \sup_{[\![\mathrm{a}_0, \mathrm{a}_1]\!] \in \mathsf{Obs}(\mathrm{CE})} (\mathrm{a}_0 - \mathrm{a}_1|_{\mathrm{CE}} \left(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\right) |\rho)_{\mathrm{BE}} .$$

The action of composing $\mathscr{T}$ with $\mathscr{E}$ and $\mathscr{C}$ on either side corresponds to restricting the set of states and effects over which the supremum is taken. Indeed, any effect of the form

$$(\mathrm{a}_x|_{\mathrm{CE}} \left(\mathscr{E} \boxtimes \mathscr{I}_{\mathrm{E}}\right)$$

is a special case of an effect on CE, and likewise, any state of the form

$$\left(\mathscr{C} \boxtimes \mathscr{I}_{\mathrm{E}}\right) |\rho)_{\mathrm{BE}}$$

is a special case of a state on BE.

As a consequence, we are optimizing over a subset of the original domain, and thus:

$$\|\mathscr{E}\mathscr{T}\mathscr{C}\|_{op} \leq \|\mathscr{T}\|_{op} .$$

It is crucial that $\mathscr{E}$ and $\mathscr{C}$ are deterministic: this ensures that, when composed with instruments (either preparation- or observation-instruments), the resulting transformations remain within the allowed sets over which the norm is defined.

To prove the reverse inequality under the assumption that both $\mathscr{E}$ and $\mathscr{C}$ are reversible, observe that:

$$\|\mathscr{T}\|_{op} = \left\|\mathscr{E}^{-1} \left(\mathscr{E}\mathscr{T}\mathscr{C}\right) \mathscr{C}^{-1}\right\|_{op} \leq \|\mathscr{E}\mathscr{T}\mathscr{C}\|_{op} ,$$

where the inequality follows from the monotonicity just established. □

Using monotonicity, one can then show that the operational norm is invariant if arbitrary ancillary systems are considered.

Lemma 7 (Invariance of the norm in presence of ancillary systems)

Given an OPT Θ, for any systems $\mathrm{A}, \mathrm{B}, \mathrm{E} \in \mathsf{Sys}(\Theta)$ and any generalised transformation $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, it holds that [98]

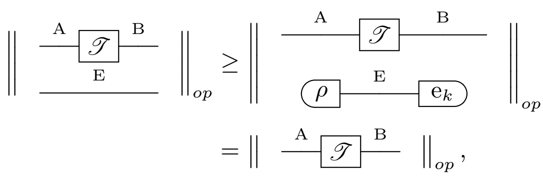
$$\|\mathscr{T}\|_{op} = \|\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\|_{op} .$$

Proof. From the definition of the operational norm (2.51), one immediately sees that

$$\|\mathscr{T}\|_{op} \geq \|\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\|_{op} ,$$

since on the right-hand side the supremum is taken over a restricted set of ancillary systems—namely, those of the form EE', with E fixed and $\mathrm{E}' \in \mathsf{Sys}(\Theta)$.

To prove the converse inequality, we apply the monotonicity of the operational norm. Consider the following:



where we have taken the two deterministic transformations to be:

- $\mathscr{I}_{\mathrm{A}} \boxtimes \rho$, with $\rho \in \mathsf{St}_1(\mathrm{E})$;
- $\mathscr{I}_{\mathrm{B}} \boxtimes \mathrm{e}_k$, with $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{E})$.

The final equality follows from the fact that applying a deterministic effect to a deterministic state yields probability one. Therefore,

$$\|\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\|_{op} \geq \|\mathscr{T}\|_{op} ,$$

and the result follows. □

Remark 16

The invariance of the operational norm under the inclusion of ancillary systems implies that the presence of such systems does not enhance the distinguishability of transformations. This is not as intuitive as one might initially expect. Indeed, in the framework of OPTs, there exist highly

"non-local" theories—such as the previously mentioned BCT and LQTs—in which the action of a transformation can influence eventual ancillary systems.
What the present result tells us, precisely, is that since the definition of the operational norm (2.51) already takes in consideration ancillary systems, the ability to distinguish transformations remains unaffected by whether or not such systems are physically incorporated into the discrimination task.

Furthermore, the operational norm is always bounded by 1 on the physically meaningful elements of an OPT—namely, states, effects, and transformations.

Proposition 1

Let Θ be a generic OPT, and let $\rho \in \mathsf{St}(\mathrm{A})$ be a state of system A. Then

$$\|\rho\|_{op} \leq 1.$$

Being deterministic is a sufficient, but not necessary condition for the norm to be equal to 1.

Proof. The result follows from the general bound [133]

$$\|\rho\|_{op} \leq \sup_{\mathrm{a}_0 \in \mathsf{Eff}(\mathrm{A})} (\mathrm{a}_0|\, \rho)_{\mathrm{A}} - \inf_{\mathrm{a}_1 \in \mathsf{Eff}(\mathrm{A})} (\mathrm{a}_1|\, \rho)_{\mathrm{A}}\,.$$

Since ρ is a valid state, one has

$$0 \leq (\mathrm{a}|\, \rho)_{\mathrm{A}} \leq 1$$

for all $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$. Hence, the supremum is bounded above by 1 and the infimum is bounded below by 0, which yields

$$\|\rho\|_{op} \leq 1.$$

Now suppose that $\rho \in \mathsf{St}_1(\mathrm{A})$. By monotonicity of the operational norm we obtain

$$1 = (\mathrm{e}_k|\, \rho)_{\mathrm{A}} \leq \|\rho\|_{op} \leq 1,$$

for every deterministic effect $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})$, where the upper bound has already been established above. This chain of inequalities shows that $\|\rho\|_{op} = 1$, as claimed. □

Lemma 8

Let Θ be a generic OPT, and let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ be a proper transformation. Then

$$\|\mathscr{T}\|_{op} \leq 1. \tag{3.3}$$

Being deterministic is a sufficient, but not necessary condition for the norm to be equal to 1.

Proof. The result follows directly from Proposition 1, together with the definition of the operational norm for transformations (2.51) since the operational norm for transformations is defined starting from that for states.

Let then $\mathscr{T} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$. By monotonicity of the operational norm we obtain

$$1 = (\mathrm{e}_k|_{\mathrm{B}}\, \mathscr{T}\, |\rho)_{\mathrm{A}} \leq \|\mathscr{T}\|_{op} \leq 1,$$

for every deterministic effect $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})$ and deterministic state $\rho \in \mathsf{St}_1(\mathrm{A})$, where the upper bound has already been established above. This chain of inequalities shows that $\|\mathscr{T}\|_{op} = 1$, as claimed. □

3.2.2 Discussion around necessary and sufficient conditions for determinism

The result reported in Lemma 8 opens the way for a discussion of the necessary and sufficient conditions for determinism of a transformation. Indeed, Lemma 8 provides only a *necessary* condition. In general, a transformation may have operational norm equal to one without being deterministic.

For instance, consider the decomposition of the identity transformation for a bidimensional system in CT:

$$\left[\!\left[\ \stackrel{\mathrm{A}}{—}\boxed{0}) \ (\boxed{0}\stackrel{\mathrm{A}}{—} \ , \ \stackrel{\mathrm{A}}{—}\boxed{1}) \ (\boxed{1}\stackrel{\mathrm{A}}{—} \ \right]\!\right].$$

Due to the existence of the zero state $0 \in \mathsf{St}_1(\mathrm{A})$ and the deterministic effect $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{A})$, it holds that

$$\left\| \ \stackrel{\mathrm{A}}{—}\boxed{0}) \ (\boxed{0}\stackrel{\mathrm{A}}{—} \ \right\|_{op} = 1,$$

even though this transformation is clearly not deterministic.

Therefore, since the norm is not a sufficient condition, a different characterisation is needed to assess whether a transformation is deterministic.

Theorem 6

Let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ be a transformation between two generic systems of an OPT. Then, $\mathscr{T}$ is deterministic if and only if

$$\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}} : \mathsf{St}_1(\mathrm{AE}) \to \mathsf{St}_1(\mathrm{BE})$$

for every $\mathrm{E} \in \mathsf{Sys}(\Theta)$. Equivalently, $\mathscr{T}$ is deterministic if and only if

$$\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}} : \mathsf{Eff}_1(\mathrm{BE}) \to \mathsf{Eff}_1(\mathrm{AE})$$

for every $\mathrm{E} \in \mathsf{Sys}(\Theta)$ [167].

Proof. $\Longrightarrow$) If $\mathscr{T}$ is deterministic, then so is $\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}$, and the mapping preserves normalised states (or effects) by definition.

$\Longleftarrow$) Suppose that $\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}$ maps $\mathsf{St}_1(\mathrm{AE})$ into $\mathsf{St}_1(\mathrm{BE})$ for every $\mathrm{E} \in \mathsf{Sys}(\Theta)$, but that $\mathscr{T}$ is not deterministic. Then there exists an instrument

$$[\![\mathscr{G}_1, \ldots, \mathscr{G}_{|\mathsf{Z}|}, \mathscr{T}]\!] \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$$

with at least one $\mathscr{G}_z \neq \varepsilon_{\mathrm{A}\to\mathrm{B}}$ to which $\mathscr{T}$ belongs. By construction $\mathscr{T} + \sum_{z\in\mathsf{Z}} \mathscr{G}_z$ is deterministic.

Let then $[\![\rho_x]\!]_{x\in\mathsf{X}} \in \mathsf{Prep}(\mathrm{AE})$ and $[\![\mathrm{a}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{BE})$ be two arbitrary preparation- and observation-instruments. The full coarse-graining of instruments is always deterministic, hence, by hypothesis, it holds that

$$(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}} \in \mathsf{St}_1(\mathrm{BE})\,,$$

where

$$\rho = \sum_{x\in\mathsf{X}} \rho_x \in \mathsf{St}_1(\mathrm{AE})\,.$$

Accordingly,

$$(\mathrm{e}_k|_{\mathrm{BE}}\, (\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}} = 1,$$

where

$$(\mathrm{e}_k|_{\mathrm{BE}} = \sum_{y\in\mathsf{Y}} \mathrm{a}_y \in \mathsf{Eff}_1(\mathrm{BE})\,.$$

However, it also holds that

$$(\mathrm{e}_k|_{\mathrm{BE}}\, (\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}} + \sum_{z\in\mathsf{Z}} (\mathrm{e}_k|_{\mathrm{BE}}\, (\mathscr{G}_z \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}} = 1,$$

which, in turn, implies that

$$(\mathrm{e}_k|_{\mathrm{BE}} (\mathscr{G}_z \boxtimes \mathscr{I}_{\mathrm{E}}) |\rho)_{\mathrm{AE}} = 0 \tag{3.4}$$

for all $x \in \mathsf{X}$, $z \in \mathsf{Z}$, and $y \in \mathsf{Y}$, since $(\mathrm{e}_k|_{\mathrm{BE}} (\mathscr{G}_z \boxtimes \mathscr{I}_{\mathrm{E}}) |\rho)_{\mathrm{AE}} \in [0,1]$ for all $x \in \mathsf{X}$, $z \in \mathsf{Z}$, and $y \in \mathsf{Y}$.

The equality (3.4) holds for any preparation- and observation-instrument of the theory. Therefore, $\mathscr{G}_z = \varepsilon_{\mathrm{A}\to\mathrm{B}}$ for all $z \in \mathsf{Z}$.

Given that transformation are invariant under coarse-graining with the null-transformation (2.37), one concludes that

$$\mathscr{T} = \mathscr{T} + \sum_{z\in\mathsf{Z}} \varepsilon_{\mathrm{A}\to\mathrm{B}} \in \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})\,.$$

Hence, $\mathscr{T}$ must itself be deterministic. This concludes the proof. □

Corollary 1

A state $\rho \in \mathsf{St}(\mathrm{A})$ is deterministic if and only if it yields probability 1 when paired with any deterministic effect of the theory. That is,

$$(\mathrm{e}_k|\,\rho)_{\mathrm{A}} = 1$$

for all $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})$.

The previous corollary allows us to prove the following lemma.

Lemma 9

In a causal OPT, a state $\rho \in \mathsf{St}(\mathrm{A})$ is deterministic if and only if $\|\rho\|_{op} = 1$.

Proof. Whenever we are working with a proper state of an OPT, the operational norm (2.50) can be computed as

$$\|\rho\|_{op} = \sup_{[\![\mathrm{a}_0,\mathrm{a}_1]\!]\in\mathsf{Obs}(\mathrm{A})} (\mathrm{a}_0 - \mathrm{a}_1|\,\rho)_{\mathrm{A}} = \sup_{\mathrm{e}\in\mathsf{Eff}_1(\mathrm{A})} (\mathrm{e}|\,\rho)_{\mathrm{A}}\,.$$

This is due to the fact that the supremum is always attained when $(\mathrm{a}_1|\,\rho)_{\mathrm{A}} = 0$, i.e., when $\mathrm{a}_1 = \varepsilon_{\mathrm{A}}$. This implies that the effect a_0 in the observation-instrument $[\![\mathrm{a}_0,\mathrm{a}_1]\!]$ must be deterministic (see the argument at the end of the proof of Theorem 6).

Since a causal OPT is one in which the deterministic effect e is unique for every system of the theory (Theorem 14), it follows that

$$\|\rho\|_{op} = (\mathrm{e}|\,\rho)_{\mathrm{A}}\,.$$

On the one hand, if $\rho \in \mathsf{St}_1(\mathrm{A})$, then by definition $(e|\,\rho)_\mathrm{A} = 1$, and thus $\|\rho\|_{op} = 1$.

On the other hand, if $\|\rho\|_{op} = 1$, then the pairing with the unique deterministic effect must be equal to 1, and thus $\rho \in \mathsf{St}_1(\mathrm{A})$. By Corollary 1, this implies that the state is deterministic. □

3.2.3 Properties of the sup norm

We now go back to the study of the properties of the norms. We here report the one of relevance for this thesis concerning the sup norm (Definition 4). For a more in depth discussion of this norm and its properties, we refer the interest reader to Ref. [134].

Lemma 10

Let $\mathscr{T} \in \mathsf{Transf}_\mathbb{R}(\mathrm{A} \to \mathrm{B})$ and $\mathscr{G} \in \mathsf{Transf}_\mathbb{R}(\mathrm{B} \to \mathrm{C})$ be two arbitrary transformations of an OPT. Then [134],

$$\|\mathscr{G}\mathscr{T}\|_{sup} \leq \|\mathscr{G}\|_{sup}\,\|\mathscr{T}\|_{sup}\,.$$

Proof. Let $x \in J(\mathscr{T})$ and $y \in J(\mathscr{G})$. By definition, there exist $\mathscr{A}$ and $\mathscr{B}$ such that $x\mathscr{A} \pm \mathscr{T} \succeq 0$ and $y\mathscr{B} \pm \mathscr{G} \succeq 0$.

Now, we have

$$\frac{1}{2}\left[(x\mathscr{A} + \mathscr{T})\,(y\mathscr{B} - \mathscr{G}) + (x\mathscr{A} - \mathscr{T})\,(y\mathscr{B} + \mathscr{G})\right] = xy\mathscr{B}\mathscr{A} - \mathscr{G}\mathscr{T} \succeq 0$$
$$\frac{1}{2}\left[(x\mathscr{A} + \mathscr{T})\,(y\mathscr{B} + \mathscr{G}) + (x\mathscr{A} - \mathscr{T})\,(y\mathscr{B} - \mathscr{G})\right] = xy\mathscr{B}\mathscr{A} + \mathscr{G}\mathscr{T} \succeq 0,$$

which implies that $xy \in J(\mathscr{G}\mathscr{T})$. Therefore,

$$\|\mathscr{G}\mathscr{T}\|_{sup} \leq \|\mathscr{G}\|_{sup}\,\|\mathscr{T}\|_{sup}\,.$$

□

3.2.4 Properties of Cauchy sequences

We now analyse the properties of Cauchy sequences of instruments and transformations and how they are related within an OPT in terms of a generic norm.

Cauchy sequences for transformations are quite straightforward. An ensemble of generalised transformations

$$\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})\,,$$

is a Cauchy sequence if $\forall \varepsilon > 0$ there exists $\tilde{n} \in \mathbb{N}$ such that for all $n, m \geq \tilde{n}$

$$\|\mathscr{T}^n - \mathscr{T}^m\|_{gen} \leq \varepsilon.$$

On the contrary, to define a Cauchy sequence for instruments, one must pay attention to a detail. When considering an ensemble of generalised instruments

$$\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})\,,$$

in the most generic case, the cardinality of the outcome spaces can vary with n. This opens the possibility that, as $n \to \infty$, the cardinality of the outcome space diverges. Formally, it is possible that $\nexists N \in \mathbb{N}$ such that $|\mathsf{X}^n| \leq N$ for all $n \in \mathbb{N}$. This possibility cannot be treated within the hypotheses currently assumed in this thesis. For two instruments to be compared, their outcome spaces must have the same cardinality—see (2.54). This also implies that the cardinality of the outcome spaces of the instruments in the sequence stabilises to a fixed value. Furthermore, this also guarantees that the cardinality of the outcome space of the limit instrument is finite—in accordance with Assumption 1. We summarise this in the following assumption.

Assumption 5: Stabilisation of the outcome space cardinality in Cauchy sequences of instruments

For any Cauchy sequence of generalised instruments

$$\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$$

in an OPT Θ, where A, B are two generic systems of the theory, the cardinality of the outcome spaces $\{\mathsf{X}^n\}_{n\in\mathbb{N}}$ stabilises as $n \to \infty$. Formally, there exists $N \in \mathbb{N}$ such that $|\mathsf{X}^n| \leq N$ for all $n \in \mathbb{N}$.

Under Assumption 5, it is then possible to completely treat sequences of instruments. Every instrument of a sequence can be embedded into the space of generalised instruments $\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$, where

$$N = \max_{n\in\mathbb{N}} |\mathsf{X}^n|.$$

In this setting, the distance between two instruments (2.54) is always well defined. Furthermore, we can also simplify the notation, dropping the dependence of the outcome spaces on the index n:

$$\left\{\{\mathscr{T}_x\}^n_{x\in\tilde{\mathsf{X}}}\right\}_{n\in\mathbb{N}} \in \mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})\,,$$

where $|\tilde{\mathsf{X}}| = N$.

We can now state the anticipated result that strictly connects the convergence of Cauchy sequences of instruments with that of the transformations composing them [158].

Theorem 7

Given a generic OPT, a sequence $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ is Cauchy with respect to the norm $\|\cdot\|_{gen}$ if and only if each sequence of transformations $\{\mathscr{T}^n_x\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ is Cauchy with respect to the same norm for all $x\in\mathsf{X}$. Furthermore [158],

$$\lim_{n\to\infty}\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\} = \left\{\lim_{n\to\infty}\mathscr{T}^n_x\right\}_{x\in\mathsf{X}}.$$

Proof. $\Longrightarrow$) Let $\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}$ be a Cauchy sequence. Then for every $\varepsilon>0$, there exists $\tilde{n}$ such that for all $n,m\geq\tilde{n}$,

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}} - \{\mathscr{T}_x\}^m_{x\in\mathsf{X}}\right\|_{gen} = \sum_{x\in\mathsf{X}}\|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{gen} \leq \varepsilon.$$

This implies that for every $x\in\mathsf{X}$, the sequence $\{\mathscr{T}^n_x\}_{n\in\mathbb{N}}$ is Cauchy:

$$\|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{gen} \leq \varepsilon.$$

$\Longleftarrow$) Conversely, suppose that each sequence $\{\mathscr{T}^n_x\}_{n\in\mathbb{N}}$ is Cauchy, for all $x\in\mathsf{X}$. Then for every $\varepsilon>0$, there exists $\tilde{n}$ such that for all $n,m\geq\tilde{n}$,

$$\|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{gen} \leq \varepsilon \quad \forall x\in\mathsf{X}.$$

We emphasise that $\tilde{n}$ can be taken independently of x. Indeed, up to taking the maximum of the bounds for each individual sequence, the condition above can be enforced uniformly for all $x\in\mathsf{X}$. Then, computing the norm of the difference between instruments, it follows that

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}} - \{\mathscr{T}_x\}^m_{x\in\mathsf{X}}\right\|_{gen} = \sum_{x\in\mathsf{X}}\|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{gen} \leq |\mathsf{X}|\varepsilon.$$

Since $|\mathsf{X}| = N < \infty$, the sequence $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}$ is Cauchy.

To conclude, let $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ denote the limit of the Cauchy sequence $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}$—which exists due to the completeness of instruments. Then for every $\varepsilon > 0$, there exists $\tilde{n}$ such that for all $n \geq \tilde{n}$,

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}} - \{\mathscr{T}_x\}_{x\in\mathsf{X}}\right\|_{gen} = \sum_{x\in\mathsf{X}} \|\mathscr{T}^n_x - \mathscr{T}_x\|_{gen} \leq \varepsilon.$$

This implies that each sequence $\{\mathscr{T}^n_x\}_{n\in\mathbb{N}}$ converges to $\mathscr{T}_x$, completing the proof. $\square$

Corollary 2

Given a generic OPT and a Cauchy sequence $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ with respect to a norm $\|\cdot\|_{gen}$, each sequence of coarse-grainings

$$\left\{\sum_{x\in\mathsf{X}'} \mathscr{T}^n_x\right\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$$

is also Cauchy with respect to that norm for all $\mathsf{X}' \subseteq \mathsf{X}$. Furthermore [158],

$$\lim_{n\to\infty}\left(\sum_{x\in\mathsf{X}'} \mathscr{T}^n_x\right) = \sum_{x\in\mathsf{X}'}\left(\lim_{n\to\infty} \mathscr{T}^n_x\right).$$

Proof. By Theorem 7, each sequence $\{\mathscr{T}^n_x\}_{n\in\mathbb{N}}$ is Cauchy. Hence, for every $\varepsilon > 0$, there exists $\tilde{n}$ such that for all $n, m \geq \tilde{n}$ and for all $x \in \mathsf{X}$,

$$\|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{op} \leq \varepsilon.$$

We emphasise that $\tilde{n}$ can be taken independently of x. Indeed, up to taking the maximum of the bounds for each individual sequence, the condition above can be enforced uniformly for all $x \in \mathsf{X}$. Then,

$$\left\|\sum_{x\in X'} (\mathscr{T}^n_x - \mathscr{T}^m_x)\right\|_{op} \leq \sum_{x\in\mathsf{X}'} \|\mathscr{T}^n_x - \mathscr{T}^m_x\|_{op} \leq |\mathsf{X}'|\varepsilon.$$

As $|\mathsf{X}'| \leq |\mathsf{X}| = N < \infty$, the sequence is Cauchy. The second claim follows from the fact that summation and limits can be exchanged for finite sums. $\square$

3.2.5 Cauchy sequences and determinism

We have seen that, in general, it is not possible to relate the operational norm of a transformation to the fact that it is deterministic (except in the special case of states in causal theories). However, if we consider Cauchy sequences of deterministic transformations, determinism is preserved in the limit. That is, the limit of a Cauchy sequence of deterministic transformations is itself deterministic.

Lemma 11

In a generic OPT, the limit (with respect to the operational norm) of a Cauchy sequence of deterministic transformations is itself deterministic. In other words, for any $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the set $\mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$ is closed with respect to the operational norm [158].

Proof. We begin by showing that the set of deterministic states is closed with respect to the operational norm.

Let $\rho = \lim_{n\to\infty} \rho^n$ be the limit of a Cauchy sequence of deterministic states $\{\rho^n\}_{n\in\mathbb{N}} \subset \mathsf{St}_1(\mathrm{A})$. We aim to show that also ρ is deterministic, i.e., that it gives unit probability when evaluated on any deterministic effect $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})$ (Corollary 1).

Using the monotonicity of the operational norm (Lemma 6), we have:

$$\left\| (\mathrm{e}_k |\, \rho)_{\mathrm{A}} - (\mathrm{e}_k |\, \rho^n)_{\mathrm{A}} \right\|_{op} \le \|\rho - \rho^n\|_{op}\,, \quad \forall \mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})\,,$$

which also implies that

$$\lim_{n\to\infty} (\mathrm{e}_k |\, \rho^n)_{\mathrm{A}} = (\mathrm{e}_k |\, \rho)_{\mathrm{A}}\,, \quad \forall \mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})\,.$$

Since each ρ^n is deterministic, the left-hand side is always equal to 1. Thus, we conclude that $(\mathrm{e}_k |\, \rho)_{\mathrm{A}} = 1$ for all $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{A})$, meaning that the limit state is deterministic.

Now consider a sequence of deterministic transformations

$$\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})\,.$$

By Theorem 6, for any system E and any deterministic state $\rho \in \mathsf{St}_1(\mathrm{AE})$, we have:

$$\left\{ (\mathscr{T}^n \boxtimes \mathscr{I}_{\mathrm{E}})\, |\rho)_{\mathrm{AE}} \right\}_{n\in\mathbb{N}} \subset \mathsf{St}_1(\mathrm{BE})\,. \tag{3.5}$$

This is still a Cauchy sequence by Lemma 6 and Lemma 7: adding ancillary systems preserves the norm, while composition with a deterministic state is norm-decreasing.

Since we have just shown that deterministic states form a closed set, the sequence (3.5) converges to a deterministic state:

$$(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}) \,|\rho)_{\mathrm{AE}} = \lim_{n\to\infty} (\mathscr{T}^n \boxtimes \mathscr{I}_{\mathrm{E}}) \,|\rho)_{\mathrm{AE}} \in \mathsf{St}_1(\mathrm{BE})\,. \tag{3.6}$$

Given that this holds for arbitrary systems E and deterministic states $\rho \in \mathsf{St}_1(\mathrm{AE})$, we conclude that the limit transformation $\mathscr{T}$ is deterministic.

The fact that the limit transformation is precisely the one presented in (3.6) derives from the fact that

$$\left\|(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}) \,|\rho)_{\mathrm{AE}} - (\mathscr{T}^n \boxtimes \mathscr{I}_{\mathrm{E}}) \,|\rho)_{\mathrm{AE}}\right\|_{op} \leq \left\|\mathscr{T} - \mathscr{T}^n\right\|_{op},$$

where the inequality still follows from Lemma 6 and Lemma 7. □

Lemma 12

In a generic OPT the limit of a Cauchy sequence, in the operational norm, of compound-local-transformations $\{\mathscr{T}^n \boxtimes \mathscr{G}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B}) \boxtimes \mathsf{Transf}_1(\mathrm{C}\to\mathrm{D})$ is still a deterministic compound-local-transformations. In other words, the set of deterministic compound-local-transformations is closed with respect to the operational norm [158].

Proof. By Lemma 11, the sequence converges to a deterministic transformation. Therefore, it is sufficient to show that the limit is a compound-local-transformations. The first step is to show that also the sequences $\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})$ and $\{\mathscr{G}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_1(\mathrm{C}\to\mathrm{D})$ are Cauchy. This can be done by direct calculation:

$$\begin{aligned}\left\|\mathscr{T}^n \boxtimes \mathscr{G}^n - \mathscr{T}^m \boxtimes \mathscr{G}^m\right\|_{op} &\geq \left\|\mathscr{T}^n \boxtimes (\mathrm{e}_k|_{\mathrm{D}}\, \mathscr{G}^n\, |\rho)_{\mathrm{C}} - \mathscr{T}^m \boxtimes (\mathrm{e}_k|_{\mathrm{D}}\, \mathscr{G}^m\, |\rho)_{\mathrm{C}}\right\|_{op} \\ &= \left\|\mathscr{T}^n - \mathscr{T}^m\right\|_{op},\end{aligned}$$

where $\rho \in \mathsf{St}_1(\mathrm{C})$ and $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{D})$ and the inequality follows from Lemma 6. A similar calculation can also be carried out for $\mathscr{G}^n$.

Defining $\mathscr{T} := \lim_{n\to\infty} \mathscr{T}^n$ and $\mathscr{G} = \lim_{n\to\infty} \mathscr{G}^n$, and exploiting Lemma 6, one can prove that

$$\begin{aligned}\|\mathscr{T} \boxtimes \mathscr{G} - \mathscr{T}^n \boxtimes \mathscr{G}^n\|_{op} &= \|\mathscr{T} \boxtimes \mathscr{G} - \mathscr{T} \boxtimes \mathscr{G}^n + \mathscr{T} \boxtimes \mathscr{G}^n - \mathscr{T}^n \boxtimes \mathscr{G}^n\|_{op} \\ &\leq \|\mathscr{T} \boxtimes \mathscr{G} - \mathscr{T} \boxtimes \mathscr{G}^n\|_{op} + \|\mathscr{T} \boxtimes \mathscr{G}^n - \mathscr{T}^n \boxtimes \mathscr{G}^n\|_{op} \\ &= \|\mathscr{T} \boxtimes (\mathscr{G} - \mathscr{G}^n)\|_{op} + \|(\mathscr{T} - \mathscr{T}^n) \boxtimes \mathscr{G}^n\|_{op} \\ &\leq \|\mathscr{T} - \mathscr{T}^n\|_{op} + \|\mathscr{G} - \mathscr{G}^n\|_{op},\end{aligned}$$

which implies

$$\mathscr{T} \boxtimes \mathscr{G} = \lim_{n\to\infty} \mathscr{T}^n \boxtimes \mathscr{G}^n.$$

□

Lemma 13

In a generic OPT, let $\{\mathrm{p}^n \mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ be a Cauchy sequence, where for every $n \in \mathbb{N}$ one has $\mathrm{p}^n \in [0,1]$ and $\mathscr{T}^n \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$, with A and B arbitrary systems of the theory. Then both $\{\mathrm{p}^n\}_{n\in\mathbb{N}}$ and $\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ are Cauchy, and

$$\mathrm{p}\,\mathscr{T} = \lim_{n\to\infty} \mathrm{p}^n \mathscr{T}^n,$$

where

$$\mathrm{p} = \lim_{n\to\infty} \mathrm{p}^n,$$
$$\mathscr{T} = \lim_{n\to\infty} \mathscr{T}^n.$$

Proof. Since the transformations are deterministic—hence $\|\mathscr{T}^n\|_{op} = 1$ by Lemma 8—we have $\|\mathrm{p}^n \mathscr{T}^n\|_{op} = \mathrm{p}^n$ for every $n \in \mathbb{N}$.

This, in conjunction with the reverse triangle inequality—analogously one could use the monotonicity of the operational norm—, for all $m, n \in \mathbb{N}$,

$$|\mathrm{p}^n - \mathrm{p}^m| = \left|\|\mathrm{p}^n \mathscr{T}^n\|_{op} - \|\mathrm{p}^m \mathscr{T}^m\|_{op}\right| \leq \|\mathrm{p}^n \mathscr{T}^n - \mathrm{p}^m \mathscr{T}^m\|_{op}.$$

Since $\{\mathrm{p}^n \mathscr{T}^n\}_{n\in\mathbb{N}}$ is Cauchy by hypothesis, it follows that $\{\mathrm{p}^n\}_{n\in\mathbb{N}}$ is Cauchy, hence $\mathrm{p} := \lim_{n\to\infty} \mathrm{p}^n$ exists in $[0,1]$.

Moreover,

$$\lim_{n\to\infty} \mathrm{p}^n \mathscr{T}^n = \varepsilon_{\mathrm{A}\to\mathrm{B}} \iff \lim_{n\to\infty} \|\mathrm{p}^n \mathscr{T}^n\|_{op} = 0 \iff \lim_{n\to\infty} \mathrm{p}^n = 0.$$

If the sequence does not converge to the null transformation, then $\mathrm{p} > 0$. This implies that there exist $\varepsilon > 0$ and $\tilde{n} \in \mathbb{N}$ such that $\mathrm{p}^n \geq \varepsilon$ for all $n \geq \tilde{n}$. This allows us to conclude that, for $n, m \geq \tilde{n}$,

$$\begin{aligned}
\|\mathscr{T}^n - \mathscr{T}^m\|_{op} &= \left\| \frac{1}{\mathrm{p}^n} \left(\mathrm{p}^n \mathscr{T}^n - \mathrm{p}^m \mathscr{T}^m\right) + \left(\frac{1}{\mathrm{p}^n} - \frac{1}{\mathrm{p}^m}\right) \mathrm{p}^m \mathscr{T}^m \right\|_{op} \\
&\leq \frac{1}{\mathrm{p}^n} \|\mathrm{p}^n \mathscr{T}^n - \mathrm{p}^m \mathscr{T}^m\|_{op} + \left| \frac{1}{\mathrm{p}^n} - \frac{1}{\mathrm{p}^m} \right| \|\mathrm{p}^m \mathscr{T}^m\|_{op} \\
&= \frac{1}{\mathrm{p}^n} \|\mathrm{p}^n \mathscr{T}^n - \mathrm{p}^m \mathscr{T}^m\|_{op} + \frac{1}{\mathrm{p}^m} |\mathrm{p}^m - \mathrm{p}^n| \\
&\leq \frac{1}{\varepsilon} \|\mathrm{p}^n \mathscr{T}^n - \mathrm{p}^m \mathscr{T}^m\|_{op} + \frac{1}{\varepsilon} |\mathrm{p}^n - \mathrm{p}^m| \,.
\end{aligned}$$

Both terms on the right-hand side vanish as $n, m \to \infty$ because $\{\mathrm{p}^n \mathscr{T}^n\}_{n\in\mathbb{N}}$ and $\{\mathrm{p}^n\}_{n\in\mathbb{N}}$ are Cauchy. Hence $\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ is Cauchy, and we can set $\mathscr{T} \coloneqq \lim_{n\to\infty} \mathscr{T}^n$.

Finally, for $n \geq \tilde{n}$,

$$\begin{aligned}
\|\mathrm{p}^n \mathscr{T}^n - \mathrm{p}\,\mathscr{T}\|_{op} &\leq |\mathrm{p}^n - \mathrm{p}| \; \|\mathscr{T}^n\|_{op} + |\mathrm{p}| \; \|\mathscr{T}^n - \mathscr{T}\|_{op} \\
&= |\mathrm{p}^n - \mathrm{p}| + |\mathrm{p}| \; \|\mathscr{T}^n - \mathscr{T}\|_{op} \,,
\end{aligned}$$

which tends to 0 as $n \to \infty$. Therefore,

$$\lim_{n\to\infty} \mathrm{p}^n \mathscr{T}^n = \mathrm{p}\,\mathscr{T}.$$

This concludes the proof. □

3.3 Coherently completing a theory

By definition, an OPT is complete under both sequential and parallel composition, and all sets of instruments and transformations are Cauchy-complete with respect to the topology induced by the operational norm. However, these two conditions are not automatically satisfied when a theory is *constructed* element by element. In such cases, it is necessary to verify the consistency between the compositional and topological structures of the framework. We show here that, whenever the operational and sup norms are equivalent, an OPT can always be coherently Cauchy-completed. The resulting theory remains closed under sequential and parallel composition, as well as coarse-graining [158].

Let us denote by Θ the OPT obtained from a tentative theory $\tilde{\Theta}$ through the operation of Cauchy completion. Thanks to Corollary 2, we know that the spaces of transformations of Θ are still closed under the operation of coarse-graining. We now turn to the question of closure under sequential and parallel composition.

Let us begin with the former. We present the argument for transformations, noting that it extends immediately to instruments by invoking Theorem 7.

Let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and $\mathscr{G} \in \mathsf{Transf}(\mathrm{B} \to \mathrm{C})$ be two transformations in Θ, obtained as limits of Cauchy sequences of transformations in $\tilde{\Theta}$:

$$\mathscr{T} = \lim_{n\to\infty} \mathscr{T}^n,$$
$$\mathscr{G} = \lim_{m\to\infty} \mathscr{G}^m.$$

We aim to show that the composition $\mathscr{G} \,\square\, \mathscr{T}$ is a transformation in Θ, i.e., that it can be obtained as the limit of a Cauchy sequence of transformations in $\tilde{\Theta}$.

To prove the result, it suffices to show that $\mathscr{G} \,\square\, \mathscr{T}$ is the limit of a Cauchy sequence of transformations in $\tilde{\Theta}$. A natural candidate is

$$\mathscr{G} \,\square\, \mathscr{T} = \lim_{m\to\infty} \lim_{n\to\infty} \mathscr{G}^m \,\square\, \mathscr{T}^n.$$

To show that this equality holds, we exploit the fact that the operational and sup norms are equivalent. Whenever these two norms are equivalent—as they are throughout this thesis due to the finite dimensionality of all vector spaces (Assumption 3)—the following inequality holds:

$$\begin{aligned} \|\mathscr{G}\mathscr{T}\|_{op} &\le C\,\|\mathscr{G}\mathscr{T}\|_{sup} \\ &\le C\,\|\mathscr{G}\|_{sup}\,\|\mathscr{T}\|_{sup} \\ &\le \frac{C}{c^2}\,\|\mathscr{G}\|_{op}\,\|\mathscr{T}\|_{op}\,, \end{aligned} \tag{3.7}$$

where $c, C > 0$ are suitable constants (Definition 3), and the second inequality follows from Lemma 10. Using this bound, we estimate:

$$\begin{aligned} &\|\mathscr{G}^m\mathscr{T}^n - \mathscr{G}\mathscr{T}\|_{op} \\ &= \|(\mathscr{G}^m\mathscr{T}^n - \mathscr{G}^m\mathscr{T}) + (\mathscr{G}^m\mathscr{T} - \mathscr{G}\mathscr{T})\|_{op} \\ &\le \|\mathscr{G}^m(\mathscr{T}^n - \mathscr{T})\|_{op} + \|(\mathscr{G}^m - \mathscr{G})\,\mathscr{T}\|_{op} \\ &\le \frac{C}{c^2}\,\|\mathscr{G}^m\|_{op}\,\|\mathscr{T}^n - \mathscr{T}\|_{op} + \frac{C}{c^2}\,\|\mathscr{G}^m - \mathscr{G}\|_{op}\,\|\mathscr{T}\|_{op} \\ &\le \frac{C}{c^2}\left(\|\mathscr{T}^n - \mathscr{T}\|_{op} + \|\mathscr{G}^m - \mathscr{G}\|_{op}\right), \end{aligned}$$

where, in the last step, we used the fact that the operational norm of any transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is bounded by one (Lemma 8). In the case of instruments, the same reasoning holds, up to a finite multiplicative constant given by the cardinality of the outcome space.

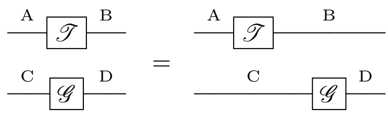
Since both $\{\mathscr{G}^m\}_{m\in\mathbb{N}}$ and $\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ are Cauchy sequences, the preceding chain of inequalities shows that the double-indexed sequence

$$\{\mathscr{G}^m \,\Box\, \mathscr{T}^n\}_{n,m\in\mathbb{N}}$$

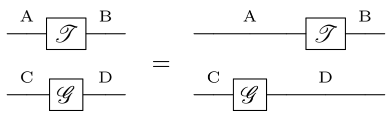
converges to $\mathscr{G} \,\Box\, \mathscr{T}$, and is therefore itself Cauchy.

This concludes the proof of closure under sequential composition.

For closure under parallel composition, the result follows from the fact that the parallel composition of two transformations can always be expressed as a sequential composition involving identities. Specifically, we recall that (2.17) gives:



or, equivalently,



.

Recalling now the invariance of the operational norm under composition with identities (Lemma 7), we can reduce the problem of closure under parallel composition to that of sequential composition, which has already been established.

In summary, we have shown that completing a tentative OPT $\tilde{\Theta}$—sometimes referred to as a *pre-OPT*—by including limits of Cauchy sequences in the spaces of instruments and transformations yields a well-defined OPT Θ, provided that the operational and sup norms are equivalent. This leads us to the following result.

Theorem 8

Let $\tilde{\Theta}$ be a pre-OPT whose spaces of instruments and transformations are not Cauchy complete. Then, the theory Θ obtained by Cauchy completing these spaces is a well-defined OPT, provided that the operational and sup norms are equivalent [158].

The argument we have just presented also establishes another important property that holds in OPTs whenever the operational and sup norms are equivalent.

Theorem 9

Let Θ be an OPT in which the operational and sup norms are equivalent. Consider two Cauchy sequences of instruments, $\left\{[\![\mathscr{T}_x]\!]^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and $\left\{[\![\mathscr{G}_y]\!]^n_{y\in\mathsf{Y}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$, with respective limits $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} = \lim_{n\to\infty}[\![\mathscr{T}_x]\!]^n_{x\in\mathsf{X}}$ and $[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} = \lim_{n\to\infty}[\![\mathscr{G}_y]\!]^n_{y\in\mathsf{Y}}$. Then, the sequential composition of the limits is the limit of the sequential compositions [158]:

$$[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \,\square\, [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} = \lim_{m\to\infty}\lim_{n\to\infty}[\![\mathscr{G}_y]\!]^m_{y\in\mathsf{Y}} \,\square\, [\![\mathscr{T}_x]\!]^n_{x\in\mathsf{X}}\,.$$

Theorem 10

Let Θ be an OPT in which the operational and sup norms are equivalent. Consider two Cauchy sequences of transformations, $\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}(\mathrm{A}\to\mathrm{B})$ and $\{\mathscr{G}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}(\mathrm{B}\to\mathrm{C})$, with respective limits $\mathscr{T} = \lim_{n\to\infty}\mathscr{T}^n$ and $\mathscr{G} = \lim_{n\to\infty}\mathscr{G}^n$. Then, the sequential composition of the limits is the limit of the sequential compositions [158]:

$$\mathscr{G} \,\square\, \mathscr{T} = \lim_{m\to\infty}\lim_{n\to\infty}\mathscr{G}^m \,\square\, \mathscr{T}^n.$$

These results can be immediately extended to generalised instruments and transformations, as well as to the case of parallel composition.

3.4 Discriminability

Sometimes, when constructing a physical model, we make assumptions that seem so natural that they do not even feel like assumptions. *Local discriminability* is one such property. It states that composite systems can always be completely characterised by performing only local measurements. For instance, in QT, if two agents, Alice and Bob, share an entangled state, it does not matter where they are located. One could be on Earth and the other on the Moon. By performing local measurements and subsequently sharing their outcomes, they can reconstruct the full entangled state via tomography.

However, in the vast landscape of all OPTs, this property does not always hold. There exist theories in which local measurements are not sufficient to fully characterise the states of composite systems. In the most general case, given an m-partite system, it may be necessary to perform measurements on subsets of up to $n < m$ parties in order to fully reconstruct the global state. Such theories are said to satisfy *n-local discriminability*.

Definition 13 (N-local discriminability)

An OPT Θ satisfies the property of *n-local discriminability* if, for any m-partite system $A = A_1, \ldots, A_m$, the set of at most n-partite effects (with $n < m$) on A is separating for $\mathsf{St}(A)$. An effect is said to be n-partite if it belongs to the set of effects of a system composed of n subsystems[a] [159].

[a]For example, in QT, a 2-partite effect would be a Bell measurement over a pair of qubits.

Of particular interest are the cases $n = 1$ and $n = 2$. The former corresponds to *local discriminability*, the notion discussed above, whereby local measurements suffice to completely characterise composite systems. The case $n = 2$, instead, requires access to measurements on pairs of subsystems, and is known as *bilocal discriminability*. Both notions have been extensively studied in the literature, and are analysed in detail in the following.

Remark 17 (N-local tomography)

In the literature, alongside the property of n-local discriminability, one also finds the notion of a theory being *n-local-tomographic* [153, 176]. The two notions are equivalent [133], yet, the nomenclature highlights how these two notions capture slightly different properties of an operational theory.
A *discrimination* task refers to a one-shot scenario: after performing a measurement, an agent assigns a probability that, given the observed outcome, a particular state was prepared. This is precisely the setting used to introduce the operational norm for states. *Tomography*, on the other hand, refers to the statistical reconstruction of a state by performing multiple measurements on many equally prepared systems.
In the particular cases of theories satisfying local and bilocal discriminability, the nomenclature of *locally tomographic* and *bilocal-tomographic* is also used, respectively.

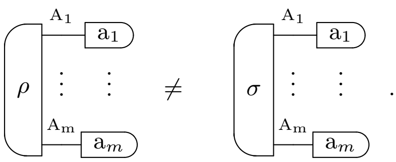
Remark 18

While not yet formally published, there are known examples of OPTs that do not satisfy n-local discriminability for any value of n [192].

3.4.1 Local discriminability

Definition 14 (Local discriminability)

An OPT that satisfies n-local discriminability for $n = 1$ is said to satisfy the property of *local discriminability*. This means that the states of composite systems can be completely characterised by performing only local measurements [133].
Formally, let $\mathrm{A} = \mathrm{A}_1, \ldots, \mathrm{A}_m$ be a composite system. Two states $\rho, \sigma \in \mathsf{St}(\mathrm{A})$ are different, i.e., $\rho \neq \sigma$, if and only if there exists an effect of the form $\mathrm{a}_1 \boxtimes \cdots \boxtimes \mathrm{a}_m \in \mathsf{Eff}(\mathrm{A})$, with $\mathrm{a}_i \in \mathsf{Eff}(\mathrm{A_i})$ for all $i = 1, \ldots, m$, such that

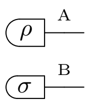


A useful characterisation of local discriminability can be given in terms of system dimensions.

In any OPT, it always holds that

$$D_{\mathrm{AB}} \geq D_{\mathrm{A}} D_{\mathrm{B}},$$

since product states of the form



are always included in $\mathsf{St}(\mathrm{AB})$. However, in theories satisfying local discriminability, the above inequality becomes an equality.

Theorem 11 (Composition rule for system sizes in locally tomographic theories)

An OPT Θ satisfies the property of local discriminability if and only if

$$D_{\mathrm{AB}} = D_{\mathrm{A}} D_{\mathrm{B}} \tag{3.8}$$

for all systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ [167].

Proof. $\Longrightarrow$) Suppose the theory satisfies local discriminability. Then, the parallel composition of local effects is separating for the set of states of the bipartite system. Formally, the set

$$T = \{\mathrm{a} \boxtimes \mathrm{b} \,|\, \mathrm{a} \in \mathsf{Eff}(\mathrm{A})\,,\, \mathrm{b} \in \mathsf{Eff}(\mathrm{B})\}$$

is separating for $\mathsf{St}(\mathrm{AB})$ and therefore spanning for $\mathsf{Eff}_{\mathbb{R}}(\mathrm{AB})$. Since $\mathsf{dim}(\mathsf{Span}_{\mathbb{R}}(T)) = D_{\mathrm{A}} D_{\mathrm{B}}$ and the state and effect spaces are dual, it follows that $D_{\mathrm{AB}} = D_{\mathrm{A}} D_{\mathrm{B}}$.

$\Longleftarrow$) Conversely, suppose Eq. (3.8) holds. Then the set of product effects spans the vector space $\mathsf{Eff}_{\mathbb{R}}(\mathrm{AB})$ and is therefore separating for $\mathsf{St}(\mathrm{AB})$. This implies that local discriminability holds. □

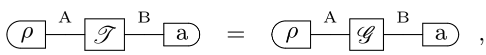
Examples

The typical examples of locally tomographic theories are CT and QT.

Properties of theories satisfying local discriminability

Local discriminability implies a couple of nice properties.

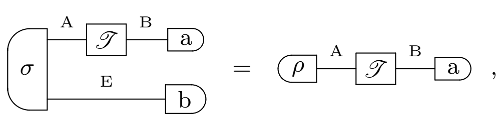
First of all, the condition for transformations to be operationally equivalent—as defined in (2.32)—is greatly simplified. In a locally tomographic theory, two transformations $\mathscr{T}, \mathscr{G} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ are equivalent, $\mathscr{T} \sim \mathscr{G}$, if and only if



(3.9)

for all states $\rho \in \mathsf{St}(\mathrm{A})$ and effects $\mathrm{a} \in \mathsf{Eff}(\mathrm{B})$. This condition naturally extends to instruments by modifying in the same way the equivalence condition (2.31).

The reason why (3.9) holds is that, if local discriminability holds, then states of the form $|(\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}})\,\sigma)_{\mathrm{BE}}$ are fully characterised by product effects in $\mathsf{Eff}(\mathrm{B}) \boxtimes \mathsf{Eff}(\mathrm{E})$. By substitution,



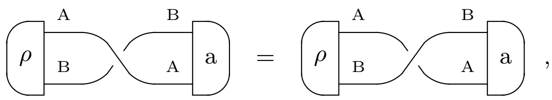
for some appropriate $\rho \in \mathsf{St}(\mathrm{A})$, and the new equivalence condition (3.9) follows directly.

The second one is a nice result that we stumbled upon almost by chance.

Proposition 2 (Local discriminability implies symmetry)

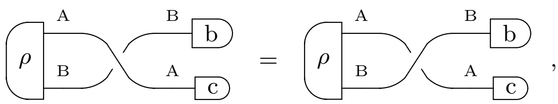
Every OPT satisfying the property of local discriminability is symmetric.

Proof. To prove our claim, we need to show that



for every state $\rho \in \mathsf{St}(\mathrm{AB})$ and effect $\mathrm{a} \in \mathsf{Eff}(\mathrm{BA})$.

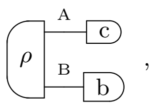
By local discriminability, it is sufficient to check that



(3.10)

for all effects $\mathrm{b} \in \mathsf{Eff}(\mathrm{B})$ and $\mathrm{c} \in \mathsf{Eff}(\mathrm{A})$.

By naturality of the braiding (2.19), both sides of (3.10) reduce to



which proves the desired equality. □

3.4.2 Bilocal discriminability

Definition 15 (Bilocal discriminability)

An OPT that satisfies n-local discriminability for $n = 2$ is said to satisfy the property of *bilocal discriminability* [167].

In the case of bilocal discriminability, the dimensionality condition takes on a more intricate form. Let us begin with the most general result, valid in any OPT.

Theorem 12 (Composition rules for system sizes)

Let Θ be a generic OPT, and let $\mathrm{ABC} \in \mathsf{Sys}(\Theta)$ be an arbitrary tripartite system. Then the following identity holds:

$$D_{\mathrm{ABC}} = D_{\mathrm{A}} D_{\mathrm{B}} D_{\mathrm{C}} + \Delta^{\{2\}}_{\mathrm{AB}} D_{\mathrm{C}} + \Delta^{\{2\}}_{\mathrm{BC}} D_{\mathrm{A}} + \Delta^{\{2\}}_{\mathrm{AC}} D_{\mathrm{B}} + \Delta^{\{3\}}_{\mathrm{ABC}},$$

where

$$\Delta^{\{2\}}_{\mathrm{AB}} := D_{\mathrm{AB}} - D_{\mathrm{A}} D_{\mathrm{B}},$$

and $\Delta^{\{3\}}_{\mathrm{ABC}}$ accounts for any contribution not captured by the preceding terms [167].

In the specific case of bilocal-tomographic theories, this expression simplifies significantly.

Theorem 13 (Composition rule for system sizes in bilocal-tomographic theories)
An OPT Θ satisfies the property of bilocal discriminability if and only if, for every tripartite system $\mathrm{ABC} \in \mathsf{Sys}(\Theta)$, the following identity holds:

$$D_{\mathrm{ABC}} = D_{\mathrm{A}} D_{\mathrm{B}} D_{\mathrm{C}} + \Delta^{\{2\}}_{\mathrm{AB}} D_{\mathrm{C}} + \Delta^{\{2\}}_{\mathrm{BC}} D_{\mathrm{A}} + \Delta^{\{2\}}_{\mathrm{AC}} D_{\mathrm{B}},$$

or, equivalently, $\Delta^{\{3\}}_{\mathrm{ABC}} = 0$ [167].

We refer the interested reader to Ref. [167] for the proofs of these results.

Examples

There are several examples in the literature of theories that satisfy bilocal discriminability.

Real Quantum Theory Historically, the first theory introduced as such—in a seminal paper that also recognised for the first time the possibility of violating local discriminability—is Real Quantum Theory (RQT) [147]. As the name suggests, this theory corresponds to standard QT restricted to real, rather than complex, Hilbert spaces.

The formal way to verify that this theory satisfies bilocal discriminability is through Theorem 13, by computing the dimension of composite systems. This can be done by recalling that the dimension of a system associated with a d_{A}-

dimensional Hilbert space is

$$D_{\mathrm{A}} = \frac{d_{\mathrm{A}}\left(d_{\mathrm{A}}+1\right)}{2},$$

namely, the dimension of the space of real symmetric matrices.

We present here an intuitive argument for why this is the case. Consider the smallest non-trivial system of RQT. This system is usually called a *rebit* and has dimension 3.

Because we are working over real Hilbert spaces, the basis used to describe the states of a rebit includes only the real Pauli matrices, rather than the full complex set. Explicitly:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

However, when considering composite systems, restricting to states generated only from these real matrices is not sufficient to span the full state space. The matrix

$$\sigma_y \otimes \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

is also real, even though σ_y itself is not. This shows that $\mathsf{St}_{\mathbb{R}}(\mathrm{A}) \otimes \mathsf{St}_{\mathbb{R}}(\mathrm{B}) \subset \mathsf{St}_{\mathbb{R}}(\mathrm{AB})$: the product of local state spaces does not exhaust the global one.

As a consequence, local measurements are no longer sufficient to fully characterise the global state. Instead, tomography requires performing a genuinely joint measurement on the entire composite system. In the pictorial scenario where Alice is on Earth and Bob on the Moon, this would correspond to a super-device encompassing both celestial bodies.

Observation 6

Interestingly, RQT was at the centre of a major discovery in 2021. It was shown that there exists a physical scenario in which RQT and standard QT yield different predictions and can, therefore, be experimentally distinguished [54]. In other words, not all quantum predictions can be reproduced within RQT.
Subsequent experiments were performed, confirming that standard QT provides a more accurate description of our physical world than RQT [193–195]. Although these tests have not yet achieved fully loophole-

free conditions, new proposals aiming to close the remaining gaps are currently under development [196].

Observation 7

An interesting perspective on the study of RQT has been recently proposed in Ref. [197]. There, RQT is recovered as the sub-theory of QT that arises by symmetrizing with respect to the collective action of a time-reversal symmetry.

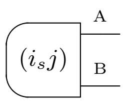
Bilocal Classical Theory There is then Bilocal Classical Theory (BCT)[1] [167]. This is a variant of CT explicitly constructed to satisfy bilocal discriminability.

The composition rule for system sizes in BCT is given by

$$D_{\mathrm{AB}} = D_{\mathrm{BA}} = \begin{cases} 2D_{\mathrm{A}}D_{\mathrm{B}} & \text{if } \mathrm{A} \neq \mathrm{I} \neq \mathrm{B}, \\ D_{\mathrm{A}} & \text{if } \mathrm{B} = \mathrm{I}. \end{cases}$$

This rule can be shown, by direct calculation, to satisfy the condition of Theorem 13.

As in the case of RQT, there is an intuitive way to understand why the theory is bilocally tomographic. The pure states of composite systems in BCT take the form



,

with $s = \pm$. The label s encodes a global degree of freedom that cannot be accessed by performing only local measurements on systems A and B. The only way to measure it is by performing a global measurement on the composite system AB.

Fermionic Quantum Theory To conclude, we mention the existence of Fermionic Quantum Theory (FQT) [198, 199].

This is an OPT in which systems are composed of *local fermionic modes*. The theory is constructed to satisfy the antisymmetry requirements characteristic of fermionic systems. For instance, FQT does not allow superposition of states with different fermionic parity.

[1] This theory was already briefly introduced in Remark 13.

For a detailed discussion of the reason why this theory satisfies bilocal discriminability, we refer the interested reader to Ref. [199].

3.5 Causal OPTs

Few topics in physics and philosophy are as divisive as the notion of causality. What does it truly mean for one event to cause another? Can we ever rigorously prove a causal relation between two events? These are questions that may remain unresolved indefinitely.

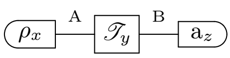
Nonetheless, certain aspects of causation can be rigorously defined and subjected to formal analysis. For example, it is possible to determine whether two occurrences are space-like separated—that is, whether any influence between them would require superluminal signalling.

When modelling such scenarios from an operational perspective, we say that two parties satisfy the *no-signalling* condition if neither can influence the other by sending signals. In frameworks where no underlying spacetime structure is assumed, this condition is formalised in terms of independencies in marginal probabilities. Specifically, there is no signalling from B to A if

$$p\,(A\,|\,B) = p\,(A)\,.$$

In the language of OPTs, the no-signalling condition is captured by the notion of *causality*.

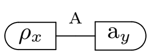
Even though circuits in the framework embed a clear relational structure, this structure, however, does not in general encode causal constraints. Indeed, the probabilities assigned to circuits depend on all the tests performed within them:



$$:= \mathrm{p}\,(x, y, z\,|\,\rho_{\mathsf{X}}, \mathsf{T}_{\mathsf{Y}}, \mathrm{a}_{\mathsf{Z}})\,, \tag{2.22}$$

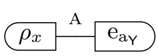
Hence, in a generic OPT, one cannot in general identify the flow of information with the input-output direction.

This can be seen already in the simplest scenario: a state preparation followed by a measurement:



$$= \mathrm{p}\,(x, y\,|\,\rho_{\mathsf{X}}, \mathrm{a}_{\mathsf{Y}})\,. \tag{3.11}$$

By marginalising over the measurement outcome in (3.11), one obtains



$$= \mathrm{p}\,(x\,|\,\rho_{\mathsf{X}}, \mathrm{a}_{\mathsf{Y}})\,, \tag{3.12}$$

where $\mathrm{e}_{\mathrm{a}_{\mathsf{Y}}} \in \mathsf{Eff}_1(\mathrm{A})$ is the deterministic effect obtained by coarse-graining the observation-instrument. Crucially, the deterministic effect depends on the chosen observation-instrument. Hence, different instruments can lead to different probabilities even when applied to the same state.

An example of an OPT where this dependence is present is Dual Classical Theory (DCT). This theory is simply the standard CT, but with the spaces of states and effects exchanged. The state space of a dual-bit is then the one represented in figure 3.1, while its effects correspond to those in figure 3.2. The only deterministic state of the theory is $|0) + |1)$, while there are two deterministic effects: $(0|$ and $(1|$. These two deterministic effects yield different probability distribution when applied to the preparation-instrument

$$\llbracket |0) , |1) \rrbracket .$$

An OPT is defined as causal if this cannot happen.

Definition 16 (Causality condition)

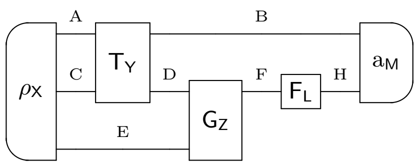
The *causality condition* states that the probability of the outcome of a preparation is independent of the choice of observation. In formulae, for any preparation-instrument $\llbracket \rho_x \rrbracket_{x\in\mathsf{X}} \in \mathsf{Prep}(\mathrm{A})$ and any observation-instrument $\llbracket \mathrm{a}_y \rrbracket_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$, for any system A, it holds that [133]

$$\sum_{y\in\mathsf{Y}} \mathrm{p}\left(x, y \mid \rho_{\mathsf{X}}, \mathrm{a}_{\mathsf{Y}}\right) = \mathrm{p}\left(x \mid \rho_{\mathsf{X}}, \mathrm{a}_{\mathsf{Y}}\right) = \mathrm{p}\left(x \mid \rho_{\mathsf{X}}\right). \tag{3.13}$$

The condition we just presented might at first seem too limited, as it applies only to probability distributions in circuits composed exclusively of preparation- and observation-instruments. However, this is not actually the case. The condition regulates the joint probability distribution of any closed circuit, since any such circuit can always be regarded as the composition of a preparation- and an observation-instrument, obtained by *slicing* the circuit.

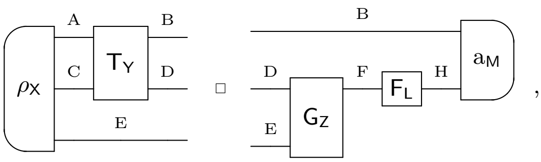
Here, slicing a circuit means partitioning it along a *slice*—that is, a maximal set of independent systems, meaning that no instrument connects any two of them.

For example, consider the following circuit:

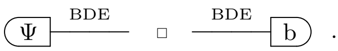


,

one way it can be split is the following



which we can treat as the composition of a preparation- and an observation-instrument:



If we now calculate the marginal probability of obtaining a particular outcome $y \in \mathsf{Y}$, we find

$$\mathrm{p}\left(y \mid \Psi, \mathrm{b}\right) = \mathrm{p}\left(y \mid \Psi\right) = \mathrm{p}\left(y \mid \rho_{\mathsf{X}}, \mathsf{T}_{\mathsf{Y}}\right), \tag{3.14}$$

where the first equality follows from the causality condition.

Therefore, we have shown that causality implies that the probability distribution of outcomes for an instrument within a circuit is not affected by instruments that are not connected to its input systems.

We can now give the precise definition of a causal OPT.

Definition 17 (Causal OPT)

An OPT is *causal* if all closed circuits in the theory satisfy the causality condition for marginal probability distributions [133].

Remark 19

An immediate consequence of (3.13) is that, in causal OPTs, one can consistently assign probabilities to states. It remains true, however, that no such assignment is possible for transformations: even under the causality condition, their probabilities would still depend on the particular states—and preparation-instruments—on which they act.

3.5.1 No-signalling from the future

When calculating (3.14), the key point we highlighted is that the probability distribution of outcomes for an instrument within a circuit is not affected by instruments that are not connected to its input systems. This effectively imposes a direction of information flow within framework circuits—which, in a sense, also establishes an arrow of time.

We say that an instrument T *precedes* another instrument G if some output of T connects to some input of G. This notion extends by transitivity: we say that T precedes G if there exists a chain of instruments, each preceding the next, such that T and G are the head and tail of the chain, respectively. Equivalently, we say that G *follows* T if T precedes G.

This nomenclature can be immediately extended to the case of transformations.

Treating circuits as DAGs, we say that an instrument precedes another if it is among its ancestors, assuming that the edges of the DAG follow the input-output direction. Analogously, an instrument follows another if it is one of its descendants.

Given this interpretation, we can restate the causality condition in the following way [133]:

Definition 18 (No-signalling from the future)

An OPT is *causal* if, for any instrument T that does not follow another instrument G, the marginal probability distribution of T is independent of the choice of G [133].

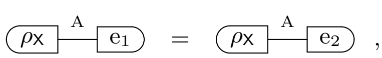
3.5.2 Unique deterministic effect

The main way to characterise whether an OPT is causal is by analysing its space of effects [133].

Theorem 14 (Unique waste bin)

An OPT Θ is causal if and only if, for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, there exists a unique deterministic effect, i.e., $\mathsf{Eff}_1(\mathrm{A}) = \{\mathrm{e}\}$ [133].

Proof. $\Longrightarrow$) Suppose that there exist two distinct deterministic effects $\mathrm{e}_1, \mathrm{e}_2 \in \mathsf{Eff}_1(\mathrm{A})$. Then, by causality,



for all preparation-instruments of system A. This implies that the two effects yield the same probability on all states. Since states are separating for effects, we conclude that $e_1 = e_2$, contradicting the assumption. Hence, the deterministic effect must be unique.

$\Longleftarrow$) Conversely, suppose that there exists a unique deterministic effect $e \in \mathsf{Eff}_1(A)$. Then, for any preparation-instrument $\rho_X \in \mathsf{Prep}(A)$ and any observation-instrument $a_Y \in \mathsf{Obs}(A)$, it holds that

$$\sum_{y\in \mathsf{Y}} \left(\rho_{\mathsf{X}} \,\overset{A}{\text{—}}\, a_{\mathsf{Y}} \right) = p\,(x \,|\, \rho_{\mathsf{X}}, a_{\mathsf{Y}}) = \left(\rho_{\mathsf{X}} \,\overset{A}{\text{—}}\, e \right) = p\,(x \,|\, \rho_{\mathsf{X}})\,,$$

where the last equality follows from the fact that e is the only deterministic effect in $\mathsf{Eff}_1(A)$. Therefore, the probability associated with a preparation-instrument is independent of the choice of observation-instrument, which is precisely the causality condition. □

Remark 20

When proving results related to causality, a common strategy is to argue by contradiction, assuming the existence of more than one deterministic effect. This condition, however, is only meaningful for systems of dimension at least two. Indeed, a one-dimensional system can, by construction, admit only a single deterministic effect—and likewise a unique deterministic state. To visualise this, consider the trivial system: it has dimension one, and its state and effect spaces are subsets of the $[0, 1]$ segment.
Consequently, when proving results about causal OPTs, it is often implicitly assumed that the theory contains systems of dimension greater than one.
An OPT composed solely of one-dimensional systems[a] is trivially causal.

[a] OPTs of this kind are sometimes referred to as *trivial OPTs* (*Definition 49*).

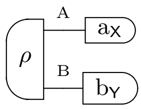
3.5.3 No-signalling without interaction

The causality condition also implies the impossibility of sending information between non-interacting parties [133].

Theorem 15 (No-signalling without interaction)

In a causal OPT, it is impossible to transmit information by performing only local operations [133].

Proof. Consider a general scenario involving two experimenters, Alice and Bob, who perform local experiments on a shared multipartite state:



.

The joint probability of outcomes x and y is given by

$$\mathrm{p}\left(x, y \,|\, \rho, \mathrm{a}_{\mathsf{X}}, \mathrm{b}_{\mathsf{Y}}\right),$$

with marginals

$$\begin{aligned} &\mathrm{p}_{\mathrm{A}}\left(x \,|\, \rho, \mathrm{a}_{\mathsf{X}}\right), \\ &\mathrm{p}_{\mathrm{B}}\left(y \,|\, \rho, \mathrm{b}_{\mathsf{Y}}\right). \end{aligned}$$

The causality condition guarantees that each marginal probability is independent of the other party's choice of observation-instrument. In particular, the marginal distribution observed by Alice does not depend on Bob's choice of measurement, and vice versa. Hence, no information can be transmitted between the two parties by acting locally on their respective systems.

This establishes the impossibility of signalling without interaction in a causal OPT. □

Observation 8

The structure we have delineated in causal OPTs coincides with the so-called *Einstein locality*. The causal cone structure (the Minkowskian causal cone) characteristic of Einstein's construction is induced in causal OPTs by the precedes/follows relation, which defines the forward and backward cones. Two events belonging to the same causal chain possess a definite time ordering. Conversely, in accordance with the Minkowskian view, if two systems do not interact, then the evolution of one system cannot be affected by any operation performed on the other [12, 133, 200].

3.5.4 Proportionality

Another way to characterise whether an OPT is causal is by analysing its states.

Theorem 16

An OPT in which every state of every system is proportional to a deterministic one is causal [133].

Proof. Suppose that an OPT satisfies the hypothesis of the theorem and admits two distinct deterministic effects, e_1 and e_2. For any state ρ in the theory, there exists a positive constant $k \in [0, 1]$ such that $\rho = k\tilde{\rho}$, where $\tilde{\rho}$ is deterministic. Then,

$$(e_1|\,\rho) = k = (e_2|\,\rho)\,.$$

Since this equality holds for all states, and states are separating for effects, it follows that $e_1 = e_2$. By Theorem 14, the theory is causal. □

3.5.5 Non-causal theories

In our introduction to the property of causality, we already mentioned that a simple way to construct a non-causal theory is to take CT and exchange the spaces of states and effects, obtaining the so-called DCT. However, there is much more to non-causal theories than this example.

By explicitly treating causality as an assumption in the construction of a physical theory, we are naturally led to ask what happens when this assumption is lifted.

Since the conception of the first process lacking a definite causal—or temporal—order, namely the *quantum switch* [201], the study of processes with indefinite causal structure has become an important line of investigation in quantum foundations.

The quantum switch is a device in which the order in which two quantum channels are performed is correlated with the state of a control qubit. As depicted in figure 3.5, the quantum switch is a *higher-order quantum operation* that takes as input two quantum channels (say, $\mathscr{T}$ and $\mathscr{G}$) and produces a new process where the order in which the channels are applied is placed in quantum superposition. Concretely:

- If the control qubit is in the state $|0\rangle$, the operation performed is $\mathscr{T}\mathscr{G}$.
- If the control qubit is in the state $|1\rangle$, the operation performed is $\mathscr{G}\mathscr{T}$.
- If the control is in a superposition $\alpha\,|0\rangle + \beta\,|1\rangle$, the order of operations is in a coherent superposition as well.

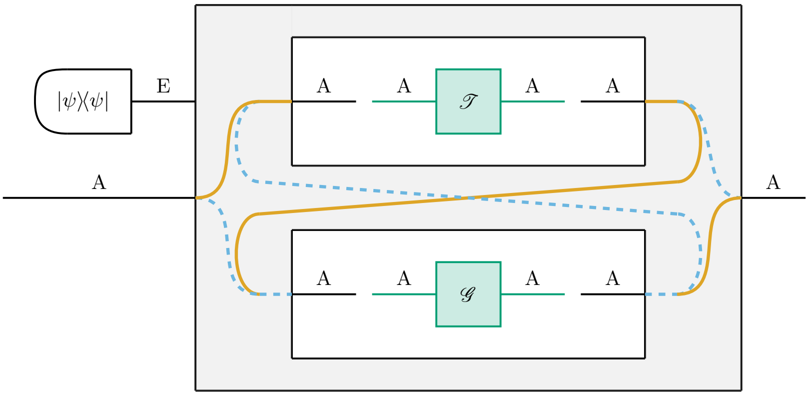


Figure 3.5: Representation of a quantum switch. The quantum switch is a higher-order quantum map that takes as input two quantum channels, $\mathscr{T}$ and $\mathscr{G}$, and produces a new process given by their coherent composition. The order in which the channels are applied is entangled with the state of a control qubit, here denoted by $|\psi\rangle\langle\psi|$. The two possible compositions, $\mathscr{T}\mathscr{G}$ and $\mathscr{G}\mathscr{T}$, are depicted by the solid orange line and the dashed blue line, respectively.

The idea of considering higher-order maps has been further explored, leading to the formulation of the frameworks of *higher-order theories* [133, 202–204] and *process matrices* [205–207]. These developments also enabled the derivation of Bell-type inequalities for processes with indefinite causal order [207–213].

In particular, within the framework of OPTs, the construction of higher-order theories is of special interest, as it provides a rich source of counterexamples.

Determinism without causality

For example, consider the first-order higher-order theory constructed from CT—that is, a theory where the states are the transformations of CT, and the transformations are all the maps that take a proper transformation of CT to another proper transformation. Exploiting this theory, one can show that *causality* and *determinism* are distinct properties of a physical theory. This stands in contrast to a common misconception that conflates the two under the notion of *causal determinism*. Famous is the quote by Planck:

> An event is causally determined if it can be predicted with certainty. [214]

Within the framework of OPTs, it is possible to provide a precise definition of what it means for a theory to be deterministic.

Definition 19 (Deterministic OPT)

An OPT is *deterministic* if the probabilities associated with all events in any closed circuit that can be constructed in the theory are either 0 or 1 [215].

This definition might seem somewhat reductive, yet it accurately captures the nature of deterministic phenomena. For instance, while CT, as it is usually defined operationally, is not deterministic, there is no obstruction to constructing a deterministic variant of it within the framework. The reason the "standard" formulation is not deterministic lies in the desire to easily describe probabilistic phenomena. However, the fact that a deterministic version can be consistently constructed is consistent with the fact that the probabilistic nature of CT is epistemic rather than ontological. In contrast, in theories where probabilistic behaviour is intrinsic—such as QT—no deterministic version can be constructed. The ability to measure the $|+\rangle$ and $|-\rangle$ states of a qubit in the computational basis demonstrates the existence of circuits whose outcome probabilities are genuinely different from 0 and 1.

So, how can we separate causality and determinism? With two counterexamples:

- QT is the simplest example of a theory that is causal but not deterministic.
- A deterministic version of the first-order higher-order CT is an example of a deterministic but non-causal theory [215]. The same applies to a deterministic variant of DCT.

To conclude this part, we highlight a property of OPTs that follows from Cauchy completeness.

Theorem 17 (Determinism vs full randomness)

An OPT is either deterministic, or the set of probabilities achievable in closed circuits coincides with the entire interval $[0, 1]$. In other words, if an OPT is not deterministic, then it is impossible to restrict the set of attainable probabilities to a proper subset of $[0, 1]$.

The proof follows from the fact that, if there exists a probability $p \in (0, 1)$, then by scalar multiplication and coarse-graining one can generate a set of probabilities that is dense in $[0,1]$. By Cauchy completeness, the closure of this set is the whole interval $[0,1]$.

3.6 Classical conditioning and strong causality

One aspect that we have not yet discussed is whether it is possible to perform certain operations based on the outcomes of an experiment. Indeed, it is: one can condition which test to perform depending on the result of a previous test. For example, one may choose to perform a specific measurement based on the outcome of a random generator, as is done in Bell-type experiments.

Formally, this is captured by the notion of a *conditional instrument*.

Definition 20 (Conditional instruments)

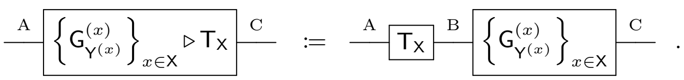
Let Θ be an OPT, $\mathsf{T}_\mathsf{X} = [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ a test of the theory, and $\left\{\mathsf{G}^{(x)}_{\mathsf{Y}^{(x)}} = \left[\!\!\left[\mathscr{G}^{(x)}_{y^{(x)}}\right]\!\!\right]_{y^{(x)}\in\mathsf{Y}^{(x)}}\right\}_{x\in\mathsf{X}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$ a labelled collection of instruments. The *conditioning operation* is a binary operation:

$$\triangleright : \mathsf{Instr}(\mathrm{B}\to\mathrm{C})^{|\mathsf{X}|} \times \mathsf{Instr}(\mathrm{A}\to\mathrm{B}) \to \mathsf{Instr}(\mathrm{A}\to\mathrm{C})\,,$$

where A, B, and $\mathrm{C} \in \mathsf{Sys}(\Theta)$ and X is the outcome space of the instrument in $\mathsf{Instr}(\mathrm{A}\to\mathrm{B})$. The operation is defined as:

$$\left\{\mathsf{G}^{(x)}_{\mathsf{Y}^{(x)}}\right\}_{x\in\mathsf{X}} \triangleright \mathsf{T}_\mathsf{X} := \left[\!\!\left[\mathscr{G}^{(x)}_{y^{(x)}} \,\square\, \mathscr{T}_x\right]\!\!\right]_{\left(x,y^{(x)}\right)\in\mathsf{X}\times\mathsf{Y}^{(x)}}\,. \tag{3.15}$$

The object on the right-hand side of the definition is called *conditional instrument*. With a slight abuse of notation, we diagrammatically represent conditional instruments as



Remark 21

Although the outcome space of the conditional instrument may depend on x, this dependence can always be eliminated—thanks to (2.48)—by considering a common outcome space Y with the largest cardinality. Therefore, we can rewrite (3.15) as

$$\left\{\mathsf{G}_{\mathsf{Y}}^{(x)}\right\}_{x\in\mathsf{X}} \triangleright \mathsf{T}_{\mathsf{X}} := \left[\!\left[\mathscr{G}_{y}^{(x)} \,\Box\, \mathscr{T}_{x}\right]\!\right]_{(x,y)\in\mathsf{X}\times\mathsf{Y}}.$$

Observation 9

The operation of conditioning is sometimes referred to as *classical conditioning*, since the conditioning is performed on classical outcomes. Furthermore, families of instruments that depend on a classical parameter—such as $\left\{\mathsf{G}_{\mathsf{Y}^{(x)}}^{(x)} = \left[\!\left[\mathscr{G}_{y^{(x)}}^{(x)}\right]\!\right]_{y^{(x)}\in\mathsf{Y}^{(x)}}\right\}_{x\in\mathsf{X}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$—are also referred to in the literature as *multimeters* [95, 216, 217].

In general, the object defined in (3.15) may not belong to $\mathsf{Instr}(\Theta)$, meaning it may represent an operation that is not implementable within the theory. An example of a theory where not every conditional instrument corresponds to an actual instrument is Minimal Classical Theory (MCT), introduced in Ref. [98] and here discussed in section 5.5.

In a theory where all such conditioning operations are physically implementable is said to satisfy the property of *strong causality*.

Definition 21 (Strongly causal OPTs)

An OPT Θ is *strongly causal* if every conditional instrument (3.15) belongs to $\mathsf{Instr}(\Theta)$, that is if every conditional instrument is an instrument of the theory [133].

In the literature on quantum theory the property of strong causality is also referred to as *classical control on outcomes* and *post-processing* [218].

As the name suggests, a strongly causal OPT is, in particular, a causal OPT.

Theorem 18

Strong causality implies causality [133].

Proof. Let A be a generic system of the theory, and suppose that its space of effects admits two distinct deterministic effects $\mathrm{e}_1, \mathrm{e}_2 \in \mathsf{Eff}_1(\mathrm{A})$. Since $\mathrm{e}_1 \neq \mathrm{e}_2$, there must exist a state ρ such that

$$(\mathrm{e}_1|\,\rho)_{\mathrm{A}} \neq (\mathrm{e}_2|\,\rho)_{\mathrm{A}}\,.$$

Without loss of generality, we furthermore assume that

$$(\mathrm{e}_1|\,\rho)_{\mathrm{A}} - (\mathrm{e}_2|\,\rho)_{\mathrm{A}} > 0.$$

Consider then $[\![\rho, \sigma]\!] \in \mathsf{Prep}(\mathrm{A})$, a preparation-instrument with ρ as one of its outcomes, and perform the following conditioning procedure: if the outcome corresponds to ρ, measure e_1; otherwise—i.e., if σ is prepared—measure e_2.

The total probability of all possible occurrences in this experiment is

$$(\mathrm{e}_1|\,\rho)_{\mathrm{A}} + (\mathrm{e}_2|\,\sigma)_{\mathrm{A}}\,.$$

Since $\rho + \sigma \in \mathsf{St}_1(\mathrm{A})$, it follows that

$$(\mathrm{e}_2|\,\sigma)_{\mathrm{A}} = 1 - (\mathrm{e}_2|\,\rho)_{\mathrm{A}}\,,$$

so the sum becomes

$$1 + \left[(\mathrm{e}_1|\,\rho)_{\mathrm{A}} - (\mathrm{e}_2|\,\rho)_{\mathrm{A}}\right] > 1,$$

which yields a probability greater than 1. This contradiction shows that a strongly causal theory cannot have two distinct deterministic effects, and is therefore causal. □

A counterexample to the converse of Theorem 18 is provided by MCT. As already discussed, this theory is not strongly causal. Nevertheless, it is causal, since its spaces of effects coincide with those of CT.

An intuitive way to understand why strong causality also fixes a direction for the flow of information is to consider how information is propagated under conditioning. By conditioning, one specifies which instrument is conditioned on which, thereby determining the order in which experiments are performed and fixing the direction of information flow.

3.6.1 Properties of strongly causal OPTs

Strongly causal OPTs satisfy a series of interesting properties.

Theorem 19

Let Θ be a strongly causal OPT. For any systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the set of transformations $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is convex as a subset of $\mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$. In other words, strongly causal OPTs are convex.

Proof. It suffices to show that for any $\mathscr{T}, \mathscr{G} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and any $\mathrm{p} \in [0,1]$, the convex combination

$$\mathrm{p}\mathscr{T} + (1-\mathrm{p})\mathscr{G}$$

is also in $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$.

Let $\left[\!\left[\mathscr{T}, \overline{\mathscr{T}}\right]\!\right]$ and $\left[\!\left[\mathscr{G}, \overline{\mathscr{G}}\right]\!\right] \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ be two instruments containing the transformations of interest.

By conditioning on the scalar-instrument $[\![\mathrm{p}, 1-\mathrm{p}]\!]$—i.e., a probability distribution—, one still obtains an instrument of the theory, as guaranteed by strong causality:

$$\left[\!\left[\mathrm{p}\mathscr{T}, \mathrm{p}\overline{\mathscr{T}}, (1-\mathrm{p})\mathscr{G}, (1-\mathrm{p})\overline{\mathscr{G}}\right]\!\right].$$

Closure under coarse-graining then guarantees that

$$\mathrm{p}\mathscr{T} + (1-\mathrm{p})\mathscr{G} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B}),$$

which is the desired result. □

Theorem 20

In every strongly causal OPT, each state is proportional to a deterministic one [142, Section 4.1.4].

Proof. Let ρ be a generic state of a system in the theory. Define the normalised state

$$\tilde{\rho} := \frac{\rho}{(\mathrm{e}|\, \rho)}.$$

An approximate preparation procedure for $\tilde{\rho}$ is as follows:

I) Pick an arbitrary binary preparation-instrument $[\![\rho_1, \rho_2]\!]$ such that $\rho_1 = \rho$.

II) Perform it N times, obtaining a sequence of outcomes $\{x_1, \ldots, x_N\}$.

III) Perform a conditional test that discards $N-1$ systems, keeping only a system i such that $x_i = 1$, if such a system exists; otherwise, keep the first system.

IV) Coarse-grain over all outcomes. The resulting deterministic state is

$$\rho_N := (1 - \mathrm{p}_N)\, \tilde{\rho} + \mathrm{p}_N \tilde{\rho_0}, \quad \text{where} \quad \mathrm{p}_N = (\mathrm{e}|\, \rho_0)^N .$$

As $N \to \infty$, we have $\mathrm{p}_N \to 0$ and thus $\rho_N \to \tilde{\rho}$. Since the set of states is assumed to be closed, $\tilde{\rho}$ is a valid state of the theory. □

By Theorem 16, Theorem 20 also provides an alternative proof that every strongly causal OPT is causal.

Remark 22

At first sight, the procedure in Theorem 20 might seem to violate the requirements for studying the convergence of sequences of instruments presented in section 3.2.4, since as $N \to \infty$ one is composing in parallel an unbounded number of preparation-tests, apparently leading to an outcome space of infinite cardinality. This is not the case: the coarse-graining step ensures that the transformation under study is deterministic and thus has a singleton outcome space. In other words, we are in fact studying the convergence of a sequence of deterministic transformations (or equivalently, of instruments with singleton outcome space).

A consequence of Theorem 20 is that strongly causal OPTs admit all atomic states. In other words, they cannot be amygdaloidal OPTs.

Corollary 3

A strongly causal OPT admits all its atomic states. Equivalently, $\mathsf{St}_+(\mathrm{A}) = \overline{\mathsf{St}_+(\mathrm{A})}$ for every system A of the theory.

Proof. The set of deterministic states $\mathsf{St}_1(\mathrm{A})$ is a bounded, closed subset of $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ with non-zero distance from the origin. By Theorem 4, the cone it spans is closed. Since every state is proportional to a deterministic one (Theorem 20), we have

$$\mathsf{St}_+(\mathrm{A}) = \mathsf{ConicHull}(\mathsf{St}_1(\mathrm{A})) = \overline{\mathsf{ConicHull}(\mathsf{St}_1(\mathrm{A}))} = \overline{\mathsf{St}_+(\mathrm{A})}.$$

This is the desired result. □

This feature will play a crucial role in the forthcoming discussion. In particular, it will be essential to prove Theorem 41.

3.6.2 Completeness with respect to strong causality

In the following theorem we show how to complete a causal theory under the assumption of strong causality (recall that causality is a necessary, but not sufficient, condition for strong causality).

Theorem 21 (Completion under strong causality)

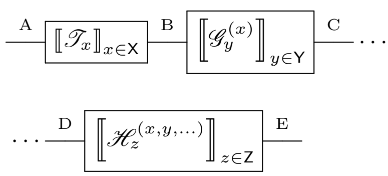
Let Θ be a causal OPT. If one constructs a new theory $\tilde{\Theta}$ by adding all instruments obtainable by conditioning a finite number of times on the outcomes of preceding instruments, and then taking the Cauchy completion of the spaces of transformations and instruments with respect to the operational norm, then $\tilde{\Theta}$ satisfies strong causality [158].

The proof is constructive, consisting of the following extension procedure, followed by a consistency check.

Procedure 1: How to make an OPT strongly causal

The procedure to make an OPT strongly causal proceeds in three steps:

I) *Addition of all conditional instruments.* Addition of all instruments obtained by conditioning a finite number of times on the outcomes of preceding instruments. Explicitly, one adds instruments of the form



(3.16)

where $(x, y, \dots)$ denotes the Cartesian product of the outcome spaces of the preceding instruments.

II) *Closure under operations.* Close the new spaces of instruments and transformations under sequential composition, parallel composition, and coarse-graining.

III) *Cauchy completion.* Take the Cauchy completion of the spaces of instruments and transformations with respect to the operational norm.

Remark 23

Consistently with the requirements stated at the beginning of section 3.2.4, the procedure described above does not introduce into $\tilde{\Theta}$ any instruments obtained by conditioning an *infinite* number of times. Otherwise, one could generate sequences of instruments whose outcome spaces have unbounded cardinality.

Remark 24

In the first step of Procedure 1 we add all instruments obtained through a *finite* number of conditioning steps, rather than restricting ourselves to the instruments of the form (3.15)—i.e., those with a single conditioning step—, as required by the definition of strongly causal OPTs (Definition 21). We do this to avoid recursion in the construction: if we added only single-step conditional instruments, then instruments obtained by conditioning on these conditional instruments would have to be included in subsequent iterations, and so on. Instead, the requirement (3.16) already encompasses all instruments that could be obtained through such an iterative procedure.
To illustrate this, consider conditioning of conditional instruments—the argument extends straightforwardly to any number of iterations. Let

$$\left[\!\left[\mathscr{G}_y^{(x)} \Box \mathscr{T}_x\right]\!\right]_{(x,y)\in\mathsf{X}\times\mathsf{Y}},$$

be an instrument in $\mathsf{Instr}(\mathrm{A}\to\mathrm{E})$, where

$$\begin{aligned}\left[\!\left[\mathscr{T}_x\right]\!\right]_{x\in\mathsf{X}} &= \left[\!\left[\mathscr{T}_{x',x''}\right]\!\right]_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''}\\ &= \left[\!\left[\mathscr{T}'_{x''}{}^{(x')} \Box \mathscr{T}'_{x'}\right]\!\right]_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''},\end{aligned}$$

is an instrument in $\mathsf{Instr}(\mathrm{A}\to\mathrm{C})$, and

$$\begin{aligned}&\left\{\left[\!\left[\mathscr{G}_y^{(x)}\right]\!\right]_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}}\\ &= \left\{\left[\!\left[\mathscr{G}_{y',y''}^{(x',x'')}\right]\!\right]_{(y',y'')\in\mathsf{Y}'\times\mathsf{Y}''}\right\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''}\\ &= \left\{\left[\!\left[\mathscr{G}''_{y''}{}^{(y',x',x'')} \Box \mathscr{G}'_{y'}{}^{(x',x'')}\right]\!\right]_{(y',y'')\in\mathsf{Y}'\times\mathsf{Y}''}\right\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''},\end{aligned}$$

is a labelled collection of instrument in $\mathsf{Instr}(\mathrm{C}\to\mathrm{E})$.

The overall conditional instrument is the given by

$$\left[\!\left[\mathscr{G}''^{(y',x',x'')}_{y''} \,\square\, \mathscr{G}'^{(x',x'')}_{y'} \,\square\, \mathscr{T}'^{(x')}_{x''} \,\square\, \mathscr{T}'_{x'} \right]\!\right]_{z\in\mathsf{Z}},$$

where $z \in \mathsf{Z}$ is shorthand for

$$\left(x', x'', y', y''\right) \in \mathsf{X}' \times \mathsf{X}'' \times \mathsf{Y}' \times \mathsf{Y}'',$$

which is exactly of the form (3.16).

Remark 25

In the procedure just described, there is a subtlety that has to be taken into account. For an instrument to be a well-defined operation in an OPT, it is necessary that it maps preparation-instruments to preparation-instruments of the theory, even in the presence of ancillary systems. The same requirement holds for observation-instruments. Hence, when we discuss the introduction of new operations—in this case, new conditional instruments—we are implicitly assuming that such additions do not violate this requirement, namely that the new operation still maps preparation-instruments [observation-instruments] to preparation-instruments [observation-instruments].
In order to abide to this requirement, we suppose that the theories to which Procedure 1 is applied already admit all the necessary preparation- and observation-instruments, so that no restriction is imposed on the conditional instruments that are added by the procedure.

We are now in a position to prove Theorem 21.

Proof. We need to show that $\tilde{\Theta}$, obtained through Procedure 1, is a well-defined OPT. This amounts to proving that, after the Cauchy completion, all conditional instruments are actual instruments of the theory, and that this enlarged collection of instruments is closed under sequential and parallel composition, as well as coarse-graining [134, 143, 159].

Closure under composition and coarse-graining follows directly from Theorem 9 and Theorem 10—and the analogous results for parallel composition—,together with Corollary 2. Therefore, the only remaining requirement is to verify that every conditional instrument belongs to $\mathsf{Instr}\left(\tilde{\Theta}\right)$.

By construction, we have added to $\tilde{\Theta}$ all instruments obtainable through a finite number of conditioning steps, and then completed the spaces of transformations and instruments with respect to the operational norm. We then need to show that for every instrument $\mathsf{T}_{\mathsf{X}} = [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and every labelled collection $\left\{\mathsf{G}_{\mathsf{Y}}^{(x)} = \left[\!\!\left[\mathscr{G}_y^{(x)}\right]\!\!\right]_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$, for any systems $\mathrm{A},\mathrm{B},\mathrm{C}\in\mathsf{Sys}\big(\tilde{\Theta}\big)$, the ensemble

$$\left\{\mathscr{G}_y^{(x)} \,\Box\, \mathscr{T}_x\right\}_{(x,y)\in\mathsf{X}\times\mathsf{Y}} \tag{3.17}$$

is an instrument of the theory.

For this verification it is sufficient to consider the case of instruments with a single conditioning step, since our goal is only to check that $\tilde{\Theta}$ satisfies Definition 21.

Suppose, for contradiction, that (3.17) is not an instrument. Since all conditional instruments obtainable from the tests of Θ have already been included in $\tilde{\Theta}$, the only remaining possibility is that at least one between T_{X} and $\left\{\mathsf{G}_{\mathsf{Y}}^{(x)}\right\}_{x\in\mathsf{X}}$ is an instrument obtained as a limit in the Cauchy completion step.

As any instrument can be seen as the limit of a constant sequence, we may treat both T_{X} and every element of $\left\{\mathsf{G}_{\mathsf{Y}}^{(x)}\right\}_{x\in\mathsf{X}}$ as limits. Thanks to Theorem 7 together with Theorem 9 and Theorem 10, we have that

$$\begin{aligned}\mathscr{G}_y^{(x)} \,\Box\, \mathscr{T}_x &= \left(\lim_{m\to\infty} \mathscr{G}_y^{(x)m}\right) \Box \left(\lim_{n\to\infty} \mathscr{T}_x^{n}\right)\\ &= \lim_{m\to\infty}\lim_{n\to\infty} \mathscr{G}_y^{(x)m} \,\Box\, \mathscr{T}_x^{n}.\end{aligned}$$

Therefore, $\mathscr{G}_y^{(x)} \,\Box\, \mathscr{T}_x$ is the limit of a sequence of transformations of Θ, for all $x\in\mathsf{X}$ and $y\in\mathsf{Y}$. Given that in Procedure 1 we add all the limits of instruments and transformations of Θ to $\tilde{\Theta}$, we obtain

$$\left[\!\!\left[\mathscr{G}_y^{(x)} \,\Box\, \mathscr{T}_x\right]\!\!\right]_{(x,y)\in\mathsf{X}\times\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})\,,$$

which is the desired result. Hence, $\tilde{\Theta}$ is a well-defined strongly causal OPT. □

3.7 No-restriction hypothesis

Alongside strong causality, the *no-restriction hypothesis* is another strong requirement that can be made regarding the spaces of instruments of an OPT.

In words, the no-restriction hypothesis states that

> All maps that satisfy all mathematical requirements for representing a transformation within the theory will be actual transformations of the theory [133],

which can be summarised in the motto:

> Everything that is not forbidden is allowed.

The idea is that, at least from a mathematical point of view, we should not impose artificial restrictions on the operations allowed within a theory. While natural, this is in fact a very strong assumption about the potential of a theory. Indeed, it essentially supposes that, in principle, the only reason why certain operations might not be achievable in a laboratory is due to technological limitations—until, at least, we find a reasonable principle for restricting the set.

Formally, the no-restriction hypothesis can be stated as follows.

Definition 22 (No-restriction hypothesis)

Let Θ be an OPT, $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, and $\mathsf{T}_{\mathsf{X}} \equiv \{\mathscr{T}_x\}_{x \in \mathsf{X}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ a collection of generalised transformations. Then, Θ is said to satisfy the *no-restriction hypothesis* if, for all $\mathrm{E} \in \mathsf{Sys}(\Theta)$, whenever $\mathsf{T}_{\mathsf{X}} \boxtimes \mathsf{I}_{\star}^{\mathrm{E} \to \mathrm{E}}$ maps preparation-instruments of AE to preparation-instruments of BE, it follows that $\mathsf{T}_{\mathsf{X}} \in \mathsf{Instr}(\Theta)$, i.e., T_{X} is an instrument of the theory [167].

An equivalent formulation is the following.

Definition 23 (No-restriction hypothesis)

Let Θ be an OPT, $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, and $\mathsf{T}_{\mathsf{X}} \equiv \{\mathscr{T}_x\}_{x \in \mathsf{X}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$ a collection of generalised transformations. Then, Θ is said to satisfy the *no-restriction hypothesis* if, for all $\mathrm{E} \in \mathsf{Sys}(\Theta)$, whenever $\mathsf{T}_{\mathsf{X}} \boxtimes \mathsf{I}_{\star}^{\mathrm{E} \to \mathrm{E}}$ maps observation-instruments of BE to observation-instruments of AE, it follows that $\mathsf{T}_{\mathsf{X}} \in \mathsf{Instr}(\Theta)$, i.e., T_{X} is an instrument of the theory.

Proof. We show that the sets of instruments described by Definition 22 and Definition 23 coincide.

First, let us prove that

$$\mathsf{Instr}(\text{Definition 22}) \subseteq \mathsf{Instr}(\text{Definition 23}),$$

where $\mathsf{Instr}(\text{Definition 22})$ denotes the set of generalised transformations satisfying Definition 22, i.e., those that are instruments of the theory under this property.

Let $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\text{Definition 22})$, with input system A and output system B. By hypothesis, $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \subset \mathsf{Transf}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ maps any preparation-instrument $[\![\rho_k]\!]_{k\in\mathsf{K}} \in \mathsf{Prep}(\mathrm{AEE}')$ into a preparation-instrument of BEE':

$$[\![(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{EE}'})\,|\rho_k)_{\mathrm{BEE}'}]\!]_{(x,k)\in\mathsf{X}\times\mathsf{K}} \in \mathsf{Prep}\big(\mathrm{BEE}'\big)\,.$$

Here the ancillary system has been split into two subsystems, E and E'. Given that they are arbitrary this can be done without loss of generality.

Now take a generic observation-instrument $[\![\mathrm{a}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{BE})$. By closure of OPTs under sequential composition, applying an observation-instrument to a preparation-instrument yields another valid preparation-instrument:

$$[\![(\mathrm{a}_y|_{\mathrm{BE}}\,(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{EE}'})\,|\rho_k)_{\mathrm{BEE}'}]\!]_{(y,x,k)} \in \mathsf{Prep}\big(\mathrm{E}'\big)\,.$$

Looking only at

$$\{(\mathrm{a}_y|\,(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}})\}_{(y,x)\in\mathsf{Y}\times\mathsf{X}} \subset \mathsf{Eff}_{\mathbb{R}}(\mathrm{BE})\,,$$

we see that this generalised effect maps any preparation-instrument of $\mathsf{Prep}(\mathrm{AEE}')$ into a preparation-instrument of E'. By the no-restriction hypothesis, this implies

$$[\![(\mathrm{a}_y|\,(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}})]\!]_{(y,x)\in\mathsf{Y}\times\mathsf{X}} \in \mathsf{Obs}(\mathrm{AE})\,.$$

Thus, any labelled collection of generalised transformations that maps preparation-instruments into preparation-instruments also maps observation-instruments into observation-instruments, showing that

$$\mathsf{Instr}(\text{Definition 22}) \subseteq \mathsf{Instr}(\text{Definition 23})\,.$$

An analogous argument shows that

$$\mathsf{Instr}(\text{Definition 23}) \subseteq \mathsf{Instr}(\text{Definition 22})\,.$$

Therefore, the two definitions characterise the same set of instruments. ☐

3.7.1 Relationship with strong causality

Intuitively, one might expect the no-restriction hypothesis to be a stronger requirement for an OPT than strong causality. However, this is not the case; the two notions are in general distinct.

To show this, let us first introduce a new class of OPTs.

Definition 24 (Dual OPTs)

The *dual* of an OPT is obtained by exchanging the spaces of states and effects, while keeping the same set of systems and the same composition rules.

An example of a dual OPT is the DCT, which was already introduced in section 3.5.

Given the equivalence between Definition 22 and Definition 23, the following lemma holds immediately.

Lemma 14

If an OPT satisfies the no-restriction hypothesis, then so does its dual.

This leads to the following results.

Lemma 15

No-restriction hypothesis $\not\Longrightarrow$ Strong causality.

Proof. A counterexample is given by DCT. □

Lemma 16

Strong causality $\not\Longrightarrow$ No-restriction hypothesis.

Proof. A counterexample is given by Minimal Strongly causal Bilocal Classical Theory (MSBCT). This theory, presented in section 6.4, satisfies strong causality but violates the no-restriction hypothesis, as it is strictly contained in BCT (section 3.4.2). □

Lemma 17

Classicality[a] + Strong causality $\not\Longrightarrow$ No-restriction hypothesis.

[a] For the precise definition of classicality we refer to Definition 27

Proof. A counterexample is again provided by MSBCT. □

3.8 Classicality

In conclusion to this chapter, we present the notion of classicality adopted in this thesis. As already observed in Remark 12, this notion of classicality is related to the geometrical structure of the state spaces of a theory.

In order to define a *classical* OPT, we first need some preliminary definitions.

Definition 25 (Simplicial OPTs)

A *simplicial* OPT Θ is a finite-dimensional OPT in which the extremal states of every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ are the vertices of a D_{A}-simplex[a] [124, 145, 159].

[a] A d-simplex is the convex hull of $d+1$ affinely independent vertices.

Definition 26 (Jointly perfect discriminability)

A set of deterministic states $\{\rho_n\}_{n\in N} \subset \mathsf{St}_1(\mathrm{A})$ is *jointly perfectly discriminable* if there exists an observation-instrument $[\![\mathrm{a}_x]\!]_{x\in N} \in \mathsf{Obs}(\mathrm{A})$ such that

$$(\mathrm{a}_x|\,\rho_{x'}) = \delta_{x,x'}$$

for all $x, x' \in N$.

We can now define *classical* OPTs as follows.

Definition 27 (Classicality)

A *classical* OPT is a simplicial OPT in which the pure states of every system are jointly perfectly discriminable [159]..

3.8.1 Properties of classical OPTs

For the sake of completeness, we collect here a series of results on simplicial and classical OPTs that are relevant for this thesis. All their proofs can be found in Ref. [159].

Theorem 22

Simplicial OPTs are causal [159].

Theorem 23

A simplicial OPT admits entangled states if and only if it does not satisfy local discriminability [159].

This last theorem is particularly interesting, as it guarantees that BCT admits entangled states. Therefore, it provides an example of a classical theory admitting entangled states.

The definition of entangled states that we adopt is the usual one. Given two systems A and B of a theory, the *separable states* of the bipartite system AB are those of the form

$$|\nu)_{\mathrm{AB}} = \sum_{i \in I} |\rho_i)_{\mathrm{A}} \, |\sigma_i)_{\mathrm{B}} \,, \tag{3.18}$$

with $\rho_i \in \mathsf{St}(\mathrm{A})$ and $\sigma_i \in \mathsf{St}(\mathrm{B})$, both different from the null state[2].

By negation, the *entangled states* are those that are *non-separable*.

We also highlight that a weaker version of Theorem 23 can be stated for generic OPTs.

Theorem 24

If an OPT does not satisfy local discriminability, then it admits entangled states [159].

The reason behind the latter theorem can be traced back to the fact that, for OPTs that do not satisfy local discriminability,

$$D_{\mathrm{AB}} > D_{\mathrm{A}} D_{\mathrm{B}}.$$

From this it follows that there are states that cannot be written as (3.18), since the linear space generated by separable states—of dimension $D_{\mathrm{A}} D_{\mathrm{B}}$—is strictly contained within $\mathsf{St}_{\mathbb{R}}(\mathrm{AB})$.

[2] The set of separable states defined in this way is generally larger than the one prepared using only Local Operations and Classical Communication (LOCC). We highlight, however, that these two sets coincide in the case of CT and QT [159].

3.8.2 The relationship with contextuality

Our definition of classicality via simplicial OPTs is closely related to that based on non-contextuality, under suitable assumptions.

Theorem 25 (Simpliciality and generalised non-contextuality)

Let an OPT satisfy:

- Causality,
- Convexity,
- Finite dimensionality (every system is finite dimensional),
- Local discriminability,

then it is generalised-non-contextual if and only if it is *simplex-embeddable*, i.e., it can be embedded into a simplicial theory [152, 153].

Operationally, simplex-embeddability means that there exists an embedding of the theory's state, effect, and transformation spaces into those of a simplicial theory, such that the probabilities predicted by the original theory are reproduced for all experiments by the embedded description. Equivalently, every experiment described by the original theory admits a simulation within a simplicial OPT.

This implies that simpliciality of the original OPT is a sufficient condition for generalised non-contextuality (under the above assumptions), since any simplicial OPT is trivially simplex-embeddable. However, simpliciality is not necessary: a OPT may be simplex-embeddable—and hence generalised-non-contextual—without itself being simplicial. An example is given by the *qutrit stabilizer subtheory* of QT [219, 220]. Further examples can be found, for instance, in Refs. [95, 221].

The relationship between generalised non-contextuality and simpliciality is not fully understood once one or more of the above assumptions is relaxed. This issue was highlighted by the case of BCT, for which the existence of an ontological model had long remained unclear, raising the possibility that, in the absence of local discriminability, simpliciality might no longer be sufficient to guarantee non-contextuality. Recent results, however, show that BCT is indeed non-contextual [222], thereby ruling it out as a counterexample. Nevertheless, beyond the setting defined by the above assumptions, a general implication from simpliciality to generalised non-contextuality has not yet been established.

3.8.3 Classical Theory (CT)

Having formally introduced the notion of classicality adopted in this work, we can now specify precisely what we mean by CT.

Definition 28 (Classical Theory (CT))

Classical Theory (CT) is the OPT Θ characterised by the following conditions [167]:

I) Θ is simplicial and convex,

II) Local discriminability holds,

III) For every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, the preparation-instruments are all collections of generalised states in $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ that sum to a point in the convex hull of $\mathsf{PurSt}(\mathrm{A})$,

IV) The no-restriction hypothesis holds.

3.8.4 What is quantum then?

We have now presented what we consider to be a good definition of classicality, at least in first approximation. This, therefore, opens the question: can we also give a definition of quantumness?

The problem with such a question is that it is not properly well-defined. Differently from the case of classicality—where we have a good notion of what it means for a process to be classical or not, grounded in the fact that the states of systems are well-defined (exactly the idea behind simplicial theories and ontological models)—, the same is not true for quantum processes.

As a first step toward answering this question, one could try to tackle a subproblem: does there exist a property that can solely set apart the quantum from the classical world?

Historically, many proposals have been made to fulfil this role. Schrödinger argued that entanglement was not "*one* but rather the characteristic trait of quantum theory, the one that enforces its entire departure from classical lines of thought" [223]. Heisenberg, on the other hand, believed that this privilege belonged to the existence of uncertainty relations [55]. Last but not least, there was Bohr, who instead argued that the principle of *complementarity* is what sets apart the quantum from the classical realm [100].

However, as of today, it is clear that the properties so far proposed as signatures of non-classicality are not satisfactory. For instance, one can devise theories that are classical yet admit entangled states, such as BCT [167]. Likewise, there exist classical theories that, in principle, could exhibit uncertainty relations, as in the case of MCT (section 5.5) [98]. The only property that has not yet been completely ruled out is complementarity. Although in this thesis we introduce an operational definition of complementarity (Definition 62), further work is required to clarify its relationship with classicality. In particular, theories such as MSBCT (section 6.4) [158]—a modified version of BCT—may provide a counterexample. A complete characterisation of this relationship, however, is left for future studies.

As the reader can notice, it was the advent of operational frameworks that allowed us to start closing in on this topic. In particular, major advancements were made thanks to the *axiomatisation program*—an effort aimed at identifying a minimal set of operational assumptions from which one can reconstruct standard Hilbert space quantum theory [124–126, 129–133].

Of particular relevance is the derivation by D'Ariano, Perinotti, and Chiribella [129, 130, 133], who—as already mentioned in Remark 13—were able to reconstruct quantum theory entirely from six axioms, which we report here again for convenience:

I.) Atomicity of composition,

II.) Perfect discriminability,

III.) Ideal compression,

IV.) Local discriminability,

V.) Causality,

VI.) Purification.

Among these, the principle of *purification*—the fact that whenever one is ignorant about the state of a physical system A, this ignorance can always be interpreted as arising from A being part of a larger system AB, about which one has full knowledge—stood out as the characterising feature of the quantum world. This principle is precisely what allows one to distinguish QT from CT [133]. The other five axioms are satisfied also by classical theory[3].

While there are strong indications that purification is indeed the property that sets apart QT from any other conceivable theory, this has yet to be formally proven.

[3] We highlight, however, that a complete reconstruction of CT from the first five axioms presented here—i.e., excluding purification—has yet to be fully carried out.

In fact, if we generalise our question and ask not only which property distinguishes the quantum from the classical world, but more broadly what sets quantum theory apart from any non-quantum behaviour, no definitive answer has yet been found.

The most intellectually honest answer we can currently give is that the violation of Bell-type inequalities provides a sufficient operational test for quantumness. In the paradigmatic CHSH scenario [40], classical local theories are bounded by the value 2, whereas quantum theory can reach correlation values up to $2\sqrt{2}$. No higher value can be achieved within quantum mechanics, as established by Tsirelson's bound [224, 225]. However, it is not yet understood why Nature enforces precisely such a bound, since, at least at a theoretical level, one can consistently propose theories that exceed it—such as *boxworld* [226], a GPT admitting PR-box correlations [227], i.e., reaching a value of 4 in the CHSH scenario. Many attempts have therefore been made to formulate principles that constrain admissible correlations and explain why quantum theory stops precisely at the Tsirelson bound; despite some of them coming remarkably close, none has so far fully succeeded. Without any claim of exhaustiveness, we recall:

- *Information causality* [228], which states that the total information a receiver can gain about data previously unknown to him, by using all his local resources and n classical bits communicated by a sender, is at most n bits.

- *Macroscopic locality* [229], which requires classical behaviour to be recovered in an appropriate macroscopic limit, and which can be used to constrain the set of correlations attainable between distant observers.

- *Local orthogonality* [230], an intrinsically multipartite principle stating that events corresponding to different outcomes of the same local measurement must be mutually exclusive.

- *Consistent exclusivity* [231], which requires that exclusivity constraints be applied consistently across different contexts, implying that the sum of probabilities of mutually exclusive events cannot exceed 1.

- *Almost quantum correlations* [232–234], it is a set of correlations that strictly contains the quantum set, while being compatible with many—though not all—of the principles proposed to characterize quantum correlations. This set is obtained by relaxing the standard quantum assumption that compound-local operations are described by tensor products of local operations, and by replacing this with a commutativity requirement for operations associated with different parties.

3.8.5 Quantum Theory (QT)

We can now state the definition of QT adopted in this thesis.

Definition 29 (Quantum Theory (QT))

Quantum Theory (QT) is the OPT characterised by the following postulates [130, 133]:

I) Atomicity of composition,

II) Perfect discriminability,

III) Ideal compression,

IV) Local discriminability,

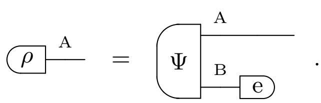
V) Causality,

VI) Purification.

The purification postulate We here report the formal definition of the purification principle. The definition of the last three assumptions that we have not yet presented are provided in appendix B.2.

Definition 30 (Purification)

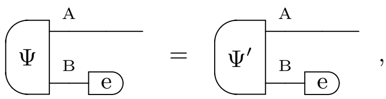
An OPT satisfies the property of *purification* if for every system A and for every state $\rho \in \mathsf{St}_1(\mathrm{A})$, there exists a system B and a pure state $\Psi \in \mathsf{PurSt}(\mathrm{AB})$ such that [133]



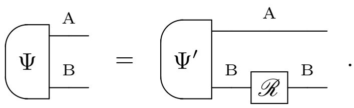
In the axiomatisation of D'Ariano, Perinotti, and Chiribella, the purification principle also includes a second assumption, which was later named *essential uniqueness of purification*. This property is indeed distinct from purification, as there exist theories that satisfy one but not the other. BCT is an example of a theory that does not satisfy purification, yet satisfies essential uniqueness of purification [167].

Definition 31 (Essential uniqueness of purification)

An OPT satisfies the property of *essential uniqueness of purification* if, whenever two pure states $\Psi, \Psi' \in \mathsf{PurSt}(\mathrm{AB})$ satisfy



then there exists a reversible transformation $\mathscr{R} \in \mathsf{RevTransf}(\mathrm{B} \to \mathrm{B})$ such that [133]



3.9 Summary of the assumptions

We here present a summary of all the assumptions introduced in this chapter that are adopted within this thesis. We furthermore collect some remarks that contain informations that we consider useful to be easily accessible.

Assumption 5 : Stabilisation of the outcome space cardinality in Cauchy sequences of instruments

For any Cauchy sequence of generalised instruments

$$\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}_{\mathbb{R}}(\mathrm{A} \to \mathrm{B})$$

in an OPT Θ, where A, B are two generic systems of the theory, the cardinality of the outcome spaces $\{\mathsf{X}^n\}_{n\in\mathbb{N}}$ stabilises as $n \to \infty$. Formally, there exists $N \in \mathbb{N}$ such that $|\mathsf{X}^n| \leq N$ for all $n \in \mathbb{N}$.

Remark 20

When proving results related to causality, a common strategy is to argue by contradiction, assuming the existence of more than one deterministic effect. This condition, however, is only meaningful for systems of dimension at least two. Indeed, a one-dimensional system can, by construction, admit only a single deterministic effect—and likewise a unique deterministic state. To visualise this, consider the trivial system: it has dimension one, and its state and effect spaces are subsets of the $[0, 1]$ segment.

Consequently, when proving results about causal OPTs, it is often implicitly assumed that the theory contains systems of dimension greater than one.
An OPT composed solely of one-dimensional systems[a] is trivially causal.

[a]OPTs of this kind are sometimes referred to as *trivial OPTs* (*Definition 49*).

4 Irreversibility, incompatibility & more

AFTER quite a ride in the development of the framework of OPTs and the encounter with some properties of interest, we are now ready to delve deeper into the main topic of this thesis. We can begin to tackle the problem delineated in section 1.3. To do so, we must start differentiating between the various phenomena encompassed by Heisenberg's *Gedankenexperiment*, or that have originated from it.

In particular, we identify three:

- The trade-off between collecting information about a physical system and disturbing it.
- *Compatibility*, namely the possibility of implementing multiple instruments simultaneously.
- *Irreversibility*, the existence of transformations that irreversibly disturb the system on which they are performed.

The first and third points are related, but they are not the same. For example, it is possible to irreversibly change the state of a system even without aiming at acquiring information about it. A randomisation procedure would be an instance of such an operation.

Given its relation to the other properties under discussion, we also introduce the notion of *broadcasting*.

4.1 Information and disturbance

One aspect of QT revealed by Heisenberg's *Gedankenexperiment* is the fact that retrieving information from a quantum system inevitably disturbs it. In the modern literature, this is known as the *No-Information Without Disturbance (NIWD) theorem*, which has been extensively studied with a variety of different quantifications of *information* and *disturbance* [70, 74, 133, 191, 235–238].

Here, we present the formulation of this property given in Ref. [191], specifically developed for the framework of OPTs.

4.1.1 No-information

Let us start by defining what it means for an instrument not to provide information about a system.

Definition 32 (No-information instrument)

An instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$, where A and B are arbitrary systems of the theory, is a *no-information instrument* if, for every choice of ancillary system E, deterministic state $\rho \in \mathsf{St}_1(\mathrm{AE})$, and deterministic effect $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{BE})$, there exist a deterministic state $\sigma \in \mathsf{St}_1(\mathrm{BE})$, and a deterministic effect $\mathrm{e}_m \in \mathsf{Eff}_1(\mathrm{AE})$ such that for all outcomes $x \in \mathsf{X}$:

$$(\mathrm{e}_k|_{\mathrm{BE}}\,(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) = \mathrm{p}_x\,(\mathrm{e}_m|_{\mathrm{AE}}\,, \tag{4.1a}$$

$$(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}})\,|\rho)_{\mathrm{AE}} = \mathrm{p}_x\,|\sigma)_{\mathrm{BE}}\,, \tag{4.1b}$$

where $[\![\mathrm{p}_x]\!]_{x\in\mathsf{X}}$ is a probability distribution [191].

We highlight that in the definition of no-information instrument just provided, we have already assumed that the probability distribution characterising the instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$ is independent of the particular deterministic states and effects on which the transformations of the instrument act. In the most general form, the definition would allow for such a dependence. The reason for imposing independence is simply that

$$(\mathrm{e}_k|_{\mathrm{BE}}\,(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}})\,|\rho)_{\mathrm{AE}} = \mathrm{q}_x(\rho) = \mathrm{r}_x(\mathrm{e}_k) = \mathrm{p}_x.$$

In words, a no-information instrument is an instrument that, when applied to a physical system, returns one of its outcomes at random, according to the probability distribution $[\![\mathrm{p}_x]\!]_{x\in\mathsf{X}}$, independently of the state (or effect) on which it acts. Therefore, it provides no information about the physical system on which it is performed.

Remark 26

The probability distribution that characterises a no-information instrument is also independent of the ancillary system that is being considered. On way to see this is to consider $\rho \boxtimes \sigma \in \mathsf{St}_1(\mathrm{AE})$, and $\mathrm{e}_k \boxtimes \mathrm{e}_m \in \mathsf{Eff}_1(\mathrm{BE})$. In this case

$$(\mathrm{e}_k|_{\mathrm{B}} (\mathrm{e}_m|_{\mathrm{E}} (\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) |\rho)_{\mathrm{A}} |\sigma)_{\mathrm{E}} = (\mathrm{e}_k|_{\mathrm{B}} \mathscr{T}_x |\rho)_{\mathrm{A}} = \mathrm{p}_x,$$

for all outcomes $x \in \mathsf{X}$. Therefore, the probabilities r_x do not depend on the ancillary system E.

4.1.2 No-disturbance

A non-disturbing instrument, instead, is an instrument that does not alter the system on which it acts when considered as a singleton instrument. In other words, it is a decomposition of the identity.

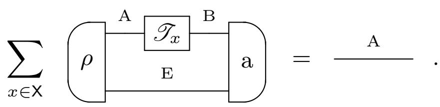
Definition 33 (Non-disturbing instrument)

An instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ is *non-disturbing* if

$$\sum_{x\in\mathsf{X}} \mathscr{T}_x = \mathscr{I}_{\mathrm{A}}. \tag{4.2}$$

Otherwise, we say that it is *disturbing* [191].

In light of (2.32), the condition (4.2) is satisfied if, for every system E of the theory and any state $\rho \in \mathsf{St}(\mathrm{AE})$ and effect $\mathrm{a} \in \mathsf{Eff}(\mathrm{BE})$, it holds that



Therefore, a non-disturbing instrument is one that, when applied to a physical system, acts as the identity on any possible state and effect of any composite system. This highlights the fact that, in order for an instrument to be non-disturbing, it must also preserve correlations with remote systems—it is not sufficient for it to act as the identity transformation only locally. The latter holds only for OPTs satisfying local discriminability (for more details, see the discussion surrounding (3.9)).

The difference between these two scenarios can be illustrated, for instance, by considering the decomposition of the identity in CT:

$$\left[\!\left[\ \overset{\mathrm{A}}{\text{---}}\boxed{\mathrm{i}}) \quad (\boxed{i}\overset{\mathrm{A}}{\text{---}} \ \right]\!\right]_{i\in\mathsf{I}}, \tag{4.3}$$

where $[\![\mathrm{i}]\!]_{i\in\mathsf{I}} \in \mathsf{Obs}(\mathrm{A})$ is the observation-instrument that perfectly discriminates the pure states of the system, $\{i\}_{i\in\mathsf{I}} = \mathsf{PurSt}(\mathrm{A})$. In the absence of local discriminability, it is immediate to verify that this instrument breaks correlations with remote systems. For instance, this is the case in BCT.

4.1.3 No-Information Without Disturbance (NIWD)

In conclusion, we can now define what it means for a theory to satisfy the NIWD property.

Definition 34 (No-Information Without Disturbance (NIWD))

An OPT Θ satisfies the *No-Information Without Disturbance (NIWD)* property if, for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ and any instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$, the following implication holds: if the instrument is non-disturbing, then it is also no-information [191].

Verifying whether an OPT satisfies the NIWD property according to Definition 34 is a demanding task, as it requires a complete characterisation of all the instruments in the theory. Fortunately, there exists an equivalent condition.

Theorem 26 (Atomicity of the identity and NIWD)

An OPT satisfies the property of NIWD if and only if the identity transformation is atomic for every system of the theory.

The proof provided in Ref. [191] for this last theorem relies on the existence of atomic states, which, as we have seen, is not always guaranteed due to the existence of amygdaloidal theories in the wide and diverse zoo of OPTs (Definition 12). We present here an alternative proof that overcomes this limitation.

Before proceeding with the proof, let us state a preliminary lemma.

Lemma 18

Let Θ be a generic OPT and A, B $\in \mathsf{Sys}(\Theta)$. Physical transformations preserve the closed cones generated by the state spaces: for any $\rho \in \overline{\mathsf{St}_+(\mathrm{A})}$ and any $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ it holds that

$$\mathscr{T}\,|\rho)_{\mathrm{A}} \in \overline{\mathsf{St}_+(\mathrm{B})}.$$

Proof. Since $\rho \in \overline{\mathsf{St}_+(\mathrm{A})}$, there exists a sequence $\left\{\rho^n\right\}_{n\in\mathbb{N}} \subset \mathsf{St}_+(\mathrm{A})$ such that

$$\lim_{n\to\infty} \rho^n = \rho.$$

By linearity $\mathscr{T}\,|\sigma)_{\mathrm{A}} \in \mathsf{St}_+(\mathrm{B})$ for all $\sigma \in \mathsf{St}_+(\mathrm{A})$ and, in particular,

$$\left\{\mathscr{T}\,|\rho^n)_{\mathrm{A}}\right\}_{n\in\mathbb{N}} \subset \mathsf{St}_+(\mathrm{B}).$$

Moreover, by (3.7), for all $n \in \mathbb{N}$ one has

$$\left\|\mathscr{T}\big(|\rho)_{\mathrm{A}} - |\rho^n)_{\mathrm{A}}\big)\right\|_{op} \leq \|\mathscr{T}\|_{op} \left\||\rho)_{\mathrm{A}} - |\rho^n)_{\mathrm{A}}\right\|_{op}. \tag{4.4}$$

Since $\|\mathscr{T}\|_{op} < \infty$ for physical transformations (Lemma 8) and $\rho^n \to \rho$, taking $n \to \infty$ in (4.4) gives

$$\lim_{n\to\infty} \mathscr{T}\,|\rho^n)_{\mathrm{A}} = \mathscr{T}\,|\rho)_{\mathrm{A}}.$$

As $\overline{\mathsf{St}_+(\mathrm{B})}$ is closed and contains all limits of sequences in $\mathsf{St}_+(\mathrm{B})$, it follows that $\mathscr{T}\,|\rho)_{\mathrm{A}} \in \overline{\mathsf{St}_+(\mathrm{B})}$, as claimed. □

We can now prove Theorem 26.

Proof. $\Longrightarrow$) Let A be a generic system of the OPT and let $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ be non-disturbing, hence no-information by NIWD. By definition, for every ancillary system E, deterministic state $\rho \in \mathsf{St}_1(\mathrm{AE})$, and deterministic effect $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{AE})$ there exist a probability distribution $[\![\mathrm{p}_x]\!]_{x\in\mathsf{X}}$, a deterministic state $\sigma \in \mathsf{St}_1(\mathrm{AE})$, and a deterministic effect $\mathrm{e}_m \in \mathsf{Eff}_1(\mathrm{AE})$ such that (4.1) holds.

Since

$$\sum_{x\in\mathsf{X}} \mathscr{T}_x = \mathscr{I}_{\mathrm{A}},$$

we must have $\rho = \sigma$ and $\mathrm{e}_k = \mathrm{e}_m$. Therefore,

$$\begin{aligned} (\mathrm{e}_k|_{\mathrm{AE}} \, (\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) &= \mathrm{p}_x \, (\mathrm{e}_k|_{\mathrm{AE}} \, , \\ (\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\rho)_{\mathrm{AE}} &= \mathrm{p}_x \, |\rho)_{\mathrm{AE}} \, . \end{aligned} \tag{4.5}$$

This relation holds for all deterministic states and effects.

By (2.32), to prove that $\mathscr{T}_x = \mathrm{p}_x \, \mathscr{I}_{\mathrm{A}}$ it suffices to show that (4.5) extends to arbitrary states in $\mathsf{St}(\mathrm{AE})$ or effects in $\mathsf{Eff}(\mathrm{AE})$.

Since the existence of atomic states is not assumed, we exploit Theorem 5 together with the procedure it supports, and work instead with generalised atomic states. Let Ψ be a generalised state spanning an extremal ray of $\overline{\mathsf{St}_+(\mathrm{AE})}$. Since such states are atomic within $\overline{\mathsf{St}_+(\mathrm{AE})}$, we must have

$$(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\Psi)_{\mathrm{AE}} = \lambda_x(\Psi) \, |\Psi)_{\mathrm{AE}} \, , \tag{4.6}$$

with $\lambda_x(\Psi) \geq 0$ and $\sum_x \lambda_x(\Psi) = 1$.

The fact that (4.6) holds is not entirely trivial: it requires that every component that can compose a decomposition of Ψ still belongs to the cone. Indeed, elements spanning extremal rays cannot be expressed as non-trivial sums of elements of the cone, although they could in principle be written as combinations of elements of the space in which the cone lives. In this case, the non-decomposability is guaranteed by Lemma 18, which ensures that $(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\Psi)_{\mathrm{AE}} \in \overline{\mathsf{St}_+(\mathrm{AE})}$. This also implies that $\lambda_x(\Psi) \geq 0$.

Lemma 18 applies also to effects, guaranteeing that the evaluation of an effect on a state returns a value in $\mathbb{R}_{\geq 0}$. Since we assume $\Psi \neq \varepsilon_{\mathrm{A}}$, there exists an effect—which may be chosen deterministic—$\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{AE})$ such that $(\mathrm{e}_k| \, \Psi)_{\mathrm{AE}} \neq 0$. Combining (4.5) and (4.6) yields

$$\mathrm{p}_x \, (\mathrm{e}_k| \, \Psi)_{\mathrm{AE}} = (\mathrm{e}_k|_{\mathrm{AE}} \, (\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\Psi)_{\mathrm{AE}} = \lambda_x(\Psi) \, (\mathrm{e}_k| \, \Psi)_{\mathrm{AE}} \, .$$

Hence $\lambda_x(\Psi) = \mathrm{p}_x$, independently of Ψ. Thus

$$(\mathscr{T}_x \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\Psi)_{\mathrm{AE}} = \mathrm{p}_x \, |\Psi)_{\mathrm{AE}} \, ,$$

for all states Ψ spanning extremal rays of $\overline{\mathsf{St}_+(\mathrm{AE})}$.

By Theorem 5, this extends to all $\rho \in \overline{\mathsf{St}_+(\mathrm{AE})}$, hence to all $\rho \in \mathsf{St}(\mathrm{AE})$. Finally, by (2.32), we conclude that $\mathscr{T}_x = \mathrm{p}_x \, \mathscr{I}_{\mathrm{A}}$ for all $x \in \mathsf{X}$.

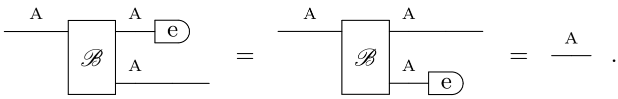
$\Longleftarrow$) Conversely, if the identity transformation is atomic, then any non-disturbing instrument is trivially no-information: the only admissible decomposition of the identity is the trivial one, where each element is proportional to the identity itself. □

4.2 Broadcasting

A direct consequence of the NIWD property is that, within an OPT satisfying it, broadcasting is impossible. Before proving this result, let us clarify what is meant by *broadcasting* in the OPT framework [144, 146, 158].

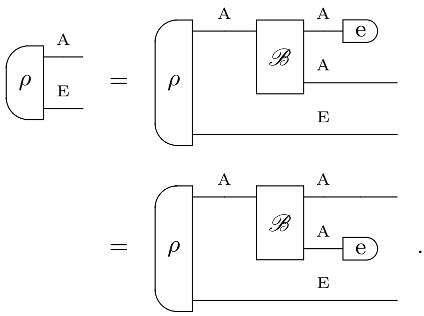
Definition 35 (Broadcasting transformation)

Let Θ be a causal OPT, and let A be a system of the theory. A deterministic transformation $\mathscr{B} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{AA})$ is called a *broadcasting* transformation if



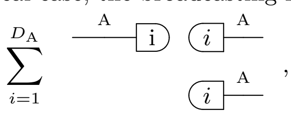
(4.7)

The condition (4.7) remains meaningful even in the presence of an ancillary system. More explicitly, a transformation $\mathscr{B} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{AA})$ is broadcasting if and only if, for every system E of the theory and every state $\rho \in \mathsf{St}(\mathrm{AE})$, the following holds:



(4.8)

It is well known that QT forbids broadcasting, whereas CT allows it [136, 144, 146, 239–245]. In the classical case, the broadcasting map takes the form



(4.9)

where $[\![\mathrm{i}]\!]_{i\in\mathsf{I}} \in \mathsf{Obs}(\mathrm{A})$ is the observation-instrument that perfectly discriminates the pure states of the system, $\{i\}_{i\in\mathsf{I}} = \mathsf{PurSt}(\mathrm{A})$.

We can now formalise what it means for an OPT to satisfy the property of broadcasting.

Definition 36 (Broadcasting)

A theory satisfies the property of *broadcasting* if every system admits a broadcasting transformation. Otherwise, we say that the theory satisfies the *no-broadcasting theorem*, or simply that it satisfies no-broadcasting.

We are now ready to prove the statement announced at the beginning of the section.

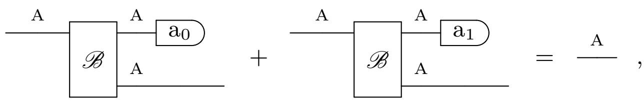
Theorem 27

Let Θ be a causal OPT, and let A be a system of the theory with $D_{\mathrm{A}} \geq 2$. If the identity transformation $\mathscr{I}_{\mathrm{A}}$ is atomic, then the system A does not admit a broadcasting channel $\mathscr{B} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{AA})$ [158].

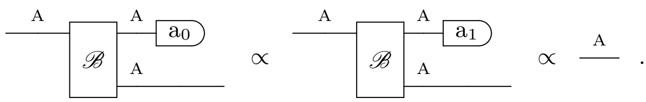
Proof. Suppose, for the sake of contradiction, that there exists a system A with $D_{\mathrm{A}} \geq 2$ such that $\mathscr{I}_{\mathrm{A}}$ is atomic and yet A admits a broadcasting channel $\mathscr{B} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{AA})$.

Consider a non-trivial decomposition $[\![\mathrm{a}_0, \mathrm{a}_1]\!] \neq [\![p_0\mathrm{e}, p_1\mathrm{e}]\!] \in \mathsf{Obs}(\mathrm{A})$ of the deterministic effect $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{A})$. Such an observation-test always exists; otherwise, the system A would have dimension $D_{\mathrm{A}} = 1$.

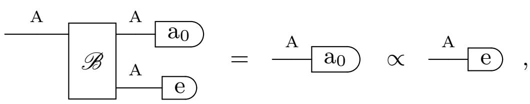
By (4.7) we then have



which, by the atomicity of $\mathscr{I}_{\mathrm{A}}$, implies



Using again (4.7), we further obtain



and analogously for a_1. This contradicts the assumption $D_{\mathrm{A}} \geq 2$. □

In other words, by Theorem 26, an OPT in which no information can be extracted without disturbing the system—that is, one satisfying NIWD—cannot admit a broadcasting transformation for any system. Conversely, the existence of a broadcasting channel would necessarily allow for information extraction without disturbance.

4.3 Compatibility

As discussed in the introduction, *compatibility* [84–86, 88, 89, 93–99] concerns the possibility of measuring physical quantities simultaneously, or more precisely, whether performing one measurement does not preclude the possibility of performing the other.

To define this notion for OPTs, we start from the well-established definition of compatibility for POVMs (1.1). Recall that two POVMs $[\![\mathrm{E}_x]\!]_{x\in\mathsf{X}}$ and $[\![\mathrm{F}_y]\!]_{y\in\mathsf{Y}}$ on a Hilbert space $\mathcal{H}_\mathrm{A}$ are said to be compatible if there exists a third POVM $[\![\mathrm{C}_{x,y}]\!]_{x,y\in\mathsf{X}\times\mathsf{Y}}$ on $\mathcal{H}_\mathrm{A}$ such that:

$$\begin{aligned} \mathrm{E}_x &= \sum_{y\in\mathsf{Y}} \mathrm{C}_{x,y} \quad \forall x\in\mathsf{X}, \\ \mathrm{F}_y &= \sum_{x\in\mathsf{X}} \mathrm{C}_{x,y} \quad \forall y\in\mathsf{Y}. \end{aligned} \tag{1.1}$$

In words, two POVMs are compatible if there exists a joint measurement whose outcomes can be classically post-processed to reproduce the statistics of both original measurements.

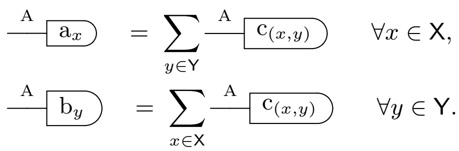
If no such joint measurement exists, the POVMs are said to be *incompatible*.

4.3.1 Observation-compatibility

The definition of compatibility for POVMs can be straightforwardly extended to observation-instruments (or, more generally, observation-tests) of an OPT.

Definition 37 (OPTs with full-compatibility of the observation-instruments)

A causal OPT Θ is said to satisfy *full-compatibility of the observation-instruments* if every pair of observation-instruments $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$, $[\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ of the theory, for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, are *compatible*. Namely, there exists a third test $[\![\mathrm{c}_{(x,y)}]\!]_{(x,y)\in\mathsf{X}\times\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ such that [94, 98]



(4.10)

We will often refer to the property of full-compatibility of the observation-instruments simply as *observation-compatibility* or *measurement-compatibility*.

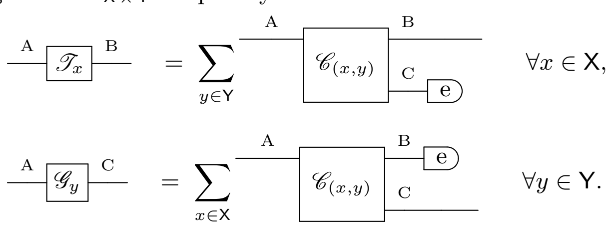
It is well known that CT satisfies observation-compatibility. For the sake of completeness—and perhaps a touch of pedantry—we nevertheless present a detailed proof in appendix D. By contrast, QT does not: for instance, two measurements along non-parallel directions on the state space of a qubit are incompatible.

4.3.2 Strong-compatibility

The compatibility condition expressed in (4.10) can be extended to the case of instruments with non-trivial output systems.

Definition 38 (Strong compatibility)

Let Θ be a causal OPT. Two instruments $\mathsf{T}_\mathsf{X} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and $\mathsf{G}_\mathsf{Y} \equiv [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})$ are said to be *strongly compatible*, denoted $\mathsf{T}_\mathsf{X} \bowtie \mathsf{G}_\mathsf{Y}$, if there exists a third instrument $\mathsf{C}_{\mathsf{X}\times\mathsf{Y}} \equiv [\![\mathscr{C}_{(x,y)}]\!]_{(x,y)\in\mathsf{X}\times\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{BC})$ such that both T_X and G_Y arise as marginals of $\mathsf{C}_{\mathsf{X}\times\mathsf{Y}}$. Explicitly:



(4.11)

Two instruments are said to be *weakly incompatible* if (4.11) does not hold, that is, if they cannot be obtained as marginals of any common joint instrument [94].

Remark 27

The definition provided above for strongly compatible instruments is not the most general possible. In the most general formulation, the outcome space of the instrument $\mathsf{C}_{\mathsf{X}\times\mathsf{Y}}$ need not coincide with the Cartesian product $X \times Y$. However, it is shown in Ref. [94] that the two formulations are equivalent.

Remark 28 (Causality and compatibility)

Following Ref. [94], we define compatibility in OPTs only for causal theories. Extending the definition to non-causal OPTs would require a careful reformulation of (4.11), since the existence of multiple deterministic effects would otherwise lead to ambiguities.
For the purposes of this thesis (and the research from which part of this material is drawn [98, 119, 158]), we do not pursue such an extension, since restricting to causal OPTs suffices to establish the main results.

4.3.3 Weak compatibility

We have referred to the definition of compatibility introduced in Definition 38 as *strong* because there also exists a weaker notion of compatibility.

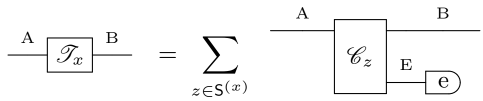
In Definition 38, two instruments are deemed compatible if they can be performed simultaneously, through the implementation of a third instrument. In the case of observation-instruments, where the physical system is discarded after the operation, compatibility can only be formulated in this way—as the existence of a joint measurement. However, when dealing with instruments with non-trivial output systems, an alternative form of compatibility can be defined, which may be regarded as a form of *"sequential" compatibility.*

The idea is to say that two instruments are compatible not only when they can be performed simultaneously (as in strong compatibility), but also when the execution of one still allows the recovery of the other, by means of a suitable generalised post-processing.

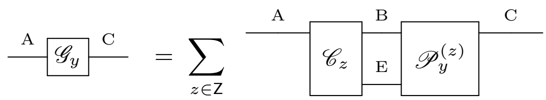
Definition 39 (Does not exclude)

Let Θ be a causal OPT. We say that an instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ *does not exclude* another instrument $[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in$

$\mathsf{Instr}(\mathrm{A} \to \mathrm{C})$ if there exist a test $[\![\mathscr{C}_z]\!]_{z\in\mathsf{Z}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{BE})$ and a post-processing, i.e., a labelled collection of instruments $\left\{ \left[\!\left[\mathscr{P}_y^{(z)} \right]\!\right]_{y\in\mathsf{Y}} \right\}_{z\in\mathsf{Z}} \subset \mathsf{Instr}(\mathrm{BE} \to \mathrm{C})$ such that



$\forall x \in \mathsf{X},$ (4.12a)



$\forall y \in \mathsf{Y},$ (4.12b)

where $\left\{\mathsf{S}^{(x)}\right\}_{x\in\mathsf{X}}$ is a suitable partition of Z [94]. If the above condition fails, we say that $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$ *excludes* $[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}}$.
We denote the fact that T_X does not exclude G_Y as

$$\mathsf{T}_\mathsf{X} \to \mathsf{G}_\mathsf{Y}.$$

In words, an instrument does not exclude another if there exists an implementation of the first—even possibly involving an ancillary environment—such that its output systems can be post-processed to reproduce the second instrument in its entirety. This is a form of *generalised post-processing*, where one is not restricted to classical data manipulation but may also exploit quantum (or, more generally, non-classical) systems.

We can now formulate the notion of *weak compatibility*.

Definition 40 (Weak compatibility)

Let Θ be a causal OPT. Two instruments T_X and G_Y are *weakly compatible* if each does not exclude the other [94]. In formulae:

$$\mathsf{T}_\mathsf{X} \leftrightarrow \mathsf{G}_\mathsf{Y}.$$

If one of the two instruments excludes the other, we say that they are *strongly incompatible*.

Remark 29

In general, $\mathsf{T_X} \to \mathsf{G_Y}$ does not imply $\mathsf{G_Y} \to \mathsf{T_X}$. Moreover, even when $\mathsf{T_X} \leftrightarrow \mathsf{G_Y}$ holds, the corresponding test $\mathsf{C_Z}$ and the post-processing implementing the two non-exclusion relations may differ.

As the name suggests, and as can be verified directly, weak compatibility is indeed weaker than strong compatibility. For the full proof of the following proposition, we refer the interested reader to Ref. [94].

Proposition 3

Let Θ be a causal OPT, and let $\mathsf{T_X} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{G_Y} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{C})$. If they are strongly compatible, $\mathsf{T_X} \bowtie \mathsf{G_Y}$, then they are weakly compatible, $\mathsf{T_X} \leftrightarrow \mathsf{G_Y}$ [94].

Conversely, the relationship between incompatibility notions is inverted: incompatibility defined via weak compatibility is a stronger condition than that defined via strong compatibility. Indeed, if an instrument excludes another, then it is impossible for them to be implemented simultaneously.

In the special case of observation-instruments, the notions of weak- and strong-compatibility turn out to coincide.

Proposition 4

Let Θ be a causal OPT, and let $\mathsf{a_X}, \mathsf{b_Y} \in \mathsf{Obs}(\mathrm{A})$ be two observation-instruments for a system $\mathrm{A} \in \Theta$. Then [94],

$$\mathsf{a_X} \bowtie \mathsf{b_Y} \quad \Longleftrightarrow \quad \mathsf{a_X} \leftrightarrow \mathsf{b_Y}.$$

The proof of this result can be found in Ref. [94].

4.3.4 Full-compatibility

Last, but not least, we can define what it means for an OPT to satisfy *full-compatibility*.

Definition 41 (Full-compatibility)

An OPT is said to satisfy *full-compatibility* if every pair of instruments $\mathsf{T} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ and $\mathsf{G} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{C})$ are weakly compatible, i.e. $\mathsf{T} \leftrightarrow \mathsf{G}$ [94].

Observation 10 (Clarification on the meaning of full compatibility)

A natural question concerns the relationship between the notion of full-compatibility introduced in this work (Definition 41, or Definition 37 for observation-instruments) and the standard notion of full-compatibility commonly adopted in the quantum-theory literature.

In the study of measurement compatibility in QT, two distinct notions are typically considered for a finite set of n measurements: *pairwise-compatibility* and *full-compatibility*. The former requires that every pair of measurements in the set be compatible. The latter, instead, requires the existence of a single measurement from which all n measurements of the set can be recovered as marginals.

These two notions are closely related, but not equivalent. Indeed, full-compatibility trivially implies pairwise-compatibility. However, the converse implication does not hold in general. There exist well-known examples of sets of POVMs that are pairwise-compatible while failing to be fully compatible. For a detailed discussion and explicit examples, we refer the interested reader to Refs. [233, 246–252].

For completeness, we recall that in QT the notions of pairwise-compatibility and full-compatibility coincide only for collections of Projection-Valued Measures (PVMs).

Let us now clarify the relation between our definition of full-compatibility and the notions introduced above.

Suppose that an OPT satisfies full-compatibility in the sense of Definition 37. We are considering observation-instruments in order to make contact with the existing literature, but the argument immediately extends to arbitrary instruments. This means that any two observation-instruments of the theory are compatible. Consider then a finite set of observation-instruments

$$\{\mathrm{a}_{\mathsf{X}}, \mathrm{b}_{\mathsf{Y}}, \ldots, \mathrm{c}_{\mathsf{Z}}\}\,.$$

By assumption, every pair of instruments in the set is compatible, and hence the set is pairwise-compatible in the standard quantum-theoretic sense.

However, our notion of full-compatibility is strictly stronger than pairwise-compatibility alone. Let $\mathsf{d}_{\mathsf{X}\times\mathsf{Y}}$ be an observation-instrument from which, for example, a_{X} and b_{Y} can be recovered as marginals. By construction, $\mathsf{d}_{\mathsf{X}\times\mathsf{Y}}$ remains compatible with any other observation-instrument in the set, such as c_{Z}. By iterating this procedure—which necessarily terminates because the set is finite—one eventually obtains an observation-instrument from which all the instruments in the set can be obtained as marginals. Therefore, any finite collection of observation-instruments also satisfies full-compatibility in the standard sense.
Summarising, although our notion of full-compatibility is defined in terms of pairwise-compatibility of instruments, it is in fact equivalent to the standard notion of full-compatibility, since pairwise-compatibility is required to hold between any possible instrument of the theory, including those from which other instruments can be recovered as marginals.

4.3.5 Full-Information Without Disturbance (FIWD)

Having introduced the property of non-exclusion (Definition 39), we can now present the dual notion to that of NIWD, namely Full-Information Without Disturbance (FIWD). In an OPT satisfying FIWD, it is always possible to extract information from a system without disturbing it. An example of such a theory is CT. In this case, the decomposition of the identity (4.3) provides the operation that makes this possible: one can recover information about the pure state of the classical system without altering it.

Definition 42 (Full-Information Without Disturbance (FIWD))

An OPT satisfies the property of Full-Information Without Disturbance (FIWD) if every instrument of the theory does not exclude the identity. Formally, for any instrument $\mathsf{T} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ it holds that $\mathsf{T} \to \mathscr{I}_{\mathrm{A}}$ [94].

4.3.6 The relationship between full-compatibility and FIWD

Interestingly, the two notions of full-compatibility and FIWD turn out to be equivalent [94].

Theorem 28

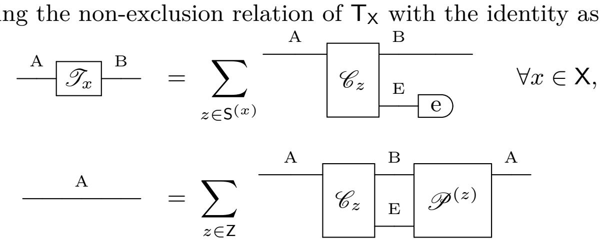
An OPT satisfies full-compatibility if and only if it satisfies FIWD [94].

The forward implication is immediate, while the converse follows from the next lemma. We highlight that the proof of Theorem 28 presented here is new, and differs from the original one given in Ref. [94].

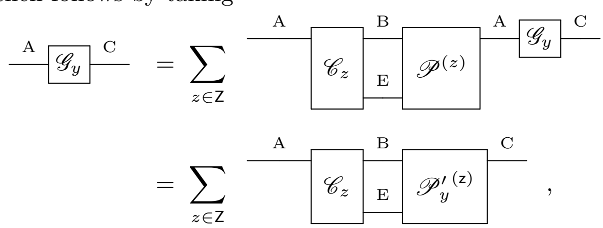
Lemma 19

For any two given tests $\mathsf{T}_\mathsf{X} = [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and $\mathsf{G}_\mathsf{Y} = [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})$, if T_X does not exclude the identity, then it also does not exclude G_Y.

Proof. Writing the non-exclusion relation of T_X with the identity as



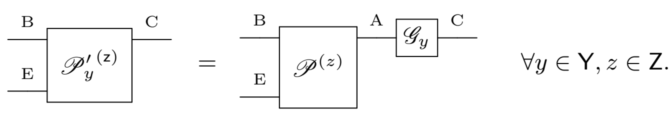
The result then follows by taking



for all outcomes $y \in \mathsf{Y}$. In the last equality we have defined the new post-processing

$$\left\{ \left[\!\!\left[\mathscr{P}'^{(z)}_y \right]\!\!\right]_{y\in\mathsf{Y}} \right\}_{z\in\mathsf{Z}} \subset \mathsf{Instr}(\mathrm{BE}\to\mathrm{C})\,,$$

where



□

Remark 30

Therefore, in an OPT where no information can be extracted without disturbing the system, there must exist incompatible instruments. Conversely, the mere presence of incompatible instruments does not guarantee that the theory satisfies NIWD. It only shows that some information cannot be extracted without disturbance, but not necessarily all.

Remark 31

Note that while NIWD does imply the existence of incompatible instruments, the same does not hold for observation-incompatibility. Indeed, MCT and MSBCT provide counterexamples: both satisfy NIWD and observation-compatibility.

4.4 Irreversibility

We can now finally introduce our notion of *irreversibility* [98, 158].

A notion that we have already encountered—the exclusion property—allows us to naturally introduce the notion of an instrument being *intrinsically irreversible*.

Definition 43 (Intrinsically irreversible instrument)

An instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ is *intrinsically irreversible* if it excludes the identity $\mathscr{I}_\mathrm{A}$ [98].

In words, this means that the action of the instrument disturbs the system A so strongly that it is not possible, not even in principle, to recover the original state of the system, independently of how $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$ is implemented. This captures the spirit of *irreversible disturbance* as intended by Heisenberg:

> [...] the interaction between observer and object causes uncontrollable and large changes in the system being observed, [...] [56].

This notion admits two equivalent characterisations.

Lemma 20

An instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ is intrinsically irreversible if and only if it is strongly incompatible with the identity transformation $\mathscr{I}_\mathrm{A}$ [98].

Lemma 21

An instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ is intrinsically irreversible if and only if it excludes some other instrument $[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{C})$ [98].

The first result follows directly from the definition: if an instrument excludes the identity, then it is strongly incompatible with it. This is in fact the only exclusion relation whose status can vary when studying compatibility with the identity, since the identity transformation does not exclude any instrument.

The second one is a direct application of Lemma 19.

We then say that an OPT has *irreversibility* if it admits intrinsically irreversible instruments.

Definition 44 (Irreversibility)

An OPT has *irreversibility* if it admits intrinsically irreversible instruments [98].

4.4.1 The relationship with compatibility

Both the definitions of intrinsic irreversibility and weak compatibility are based on the notion of exclusion. Therefore, it is natural to expect a close connection between them.

By definition, if an instrument is intrinsically irreversible, there exists another instrument with which it is strongly incompatible. Conversely, if two instruments are strongly incompatible, then at least one must exclude the other, and is therefore intrinsically irreversible.

From this observation it immediately follows that an OPT has irreversibility if and only if it admits instruments that are strongly incompatible.

Theorem 29

An OPT has irreversibility if and only if it admits instruments that are strongly incompatible.

This establishes that *irreversibility and compatibility are* in fact *equivalent notions*. Nevertheless, there is more that can be said between their relationship.

From a historical point of view, the primary objects of study in the context of compatibility were measurements, rather than generic operations of a theory. In the case of QT, this meant focusing on POVMs instead of channels. Therefore, if we really want to understand the relationship between compatibility and irreversibility, we should also analyse the relationship between *observation-compatibility* and *irreversibility*. This perspective also aligns more closely with Heisenberg's *Gedankenexperiment*, where no attention is paid to the fate of the system after its interaction with the photon. In the framework of OPTs, this situation is modelled by considering an operation with trivial output system, i.e., an effect.

In this case, the following result holds.

Theorem 30

Observation-incompatibility implies irreversibility, but the converse is not true [98, 158].

Proof. The implication immediately follows from Theorem 29. If two observation-instruments are incompatible, then at least one excludes the other. We recall that for observation-instruments the notions of weak and strong compatibility coincide ()Proposition 4).

For the converse implication, we provide two counterexamples: MCT (section 5.5) and MSBCT (section 6.4). Both of these OPTs exhibit irreversibility while satisfying observation-compatibility.

The reason for providing two counterexamples lies in the properties that characterise them. MCT is a minimal version of CT. The spaces of states and effects are left unchanged, but the set of transformations with non-trivial input and output systems is reduced to its bare minimum, to the point that the theory does not satisfy strong causality. MSBCT is then introduced to show that even when conditioning is allowed, irreversibility may still arise in the presence of observation-compatibility. This theory is a minimal version of BCT where conditioning is also allowed. □

Remark 32

Theorem 30 establishes, in a sense, a hierarchy among measurement methods. Extracting information from a system and discarding it afterwards can recover more information than performing a measurement while keeping track of the system's subsequent evolution. This fact is made explicit by the counterexamples of MCT and MSBCT, where every observation-instrument is compatible with any other, yet extracting information through the outcome of an instrument with a non-trivial output system can still lead to irreversible disturbance.

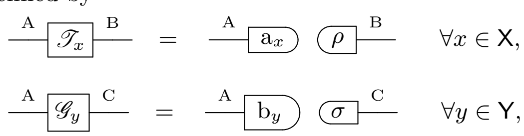
In the proof of Theorem 30 we relied on the fact that if two observation-instruments are incompatible, then at least one of them is intrinsically irreversible, since it excludes the other. However, we can prove a stronger statement.

Lemma 22

If an OPT does not satisfy observation-compatibility, then there exist at least two instruments with non-trivial input and output systems that are incompatible. Hence, they are intrinsically irreversible [98].

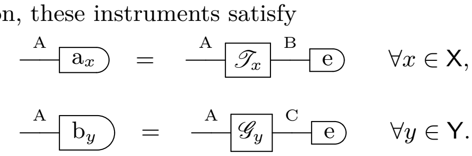
Proof. Let $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$ and $[\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ be two incompatible observation-instruments on the system A.

In any OPT, closure with respect to sequential composition allows one to construct two instruments $\mathsf{T}_\mathsf{X} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and $\mathsf{G}_\mathsf{Y} \equiv [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})$, defined by



where $\rho \in \mathsf{St}_1(\mathrm{B})$ and $\sigma \in \mathsf{St}_1(\mathrm{C})$ are arbitrary deterministic states of the systems B and C, respectively.

By construction, these instruments satisfy

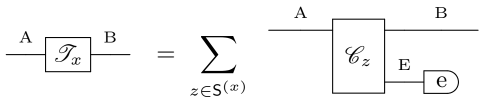


Suppose now, by contradiction, that $\mathsf{T}_\mathsf{X} \leftrightarrow \mathsf{G}_\mathsf{Y}$ holds. Then, neither instrument excludes the other, and consequently their compositions with the deterministic effect, $(e| \square \mathsf{T}_\mathsf{X} = a_\mathsf{X}$ and $(e| \square \mathsf{G}_\mathsf{Y} = b_\mathsf{Y}$, would define compatible observation-instruments. This contradicts the assumed incompatibility of the observation-instruments $[\![a_x]\!]_{x\in\mathsf{X}}$ and $[\![b_y]\!]_{y\in\mathsf{Y}}$.

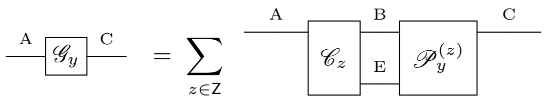
The implication

$$\mathsf{T}_\mathsf{X} \leftrightarrow \mathsf{G}_\mathsf{Y} \implies (e|\,\mathsf{T}_\mathsf{X} \leftrightarrow (e|\,\mathsf{G}_\mathsf{Y}$$

can be proven by direct calculation. If we express the non-exclusion relationship $\mathsf{T}_\mathsf{X} \to \mathsf{G}_\mathsf{Y}$ as

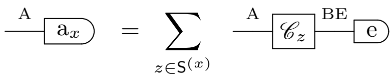


$\forall x \in \mathsf{X}$,

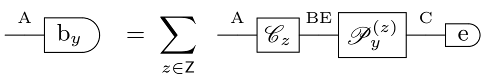


$\forall y \in \mathsf{Y}$,

then the same decomposition immediately characterises the non-exclusion relationship for the corresponding observation-instruments $a_\mathsf{X} \to b_\mathsf{Y}$,



$\forall x \in \mathsf{X}$,



$\forall y \in \mathsf{Y}$.

□

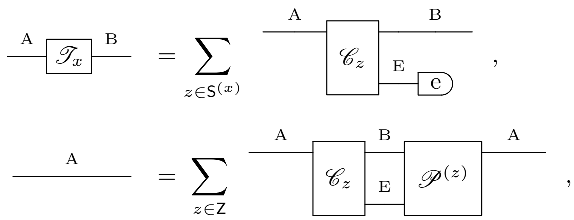
4.4.2 The relationship with NIWD

As we have seen with broadcasting, the atomicity of the identity transformation—equivalent to NIWD—is an extremely powerful structural requirement. We now show that it is also connected to irreversibility.

Theorem 31 (NIWD implies irreversibility)

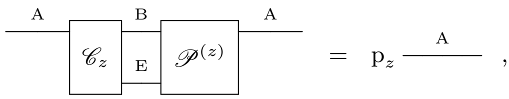
Let Θ be a causal OPT. Suppose there exists a system $A \in \mathsf{Sys}(\Theta)$ with $D_A \geq 2$, and that the identity transformation $\mathscr{I}_A$ is atomic. Then, there exists an instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(A \to B)$, for some system $B \in \mathsf{Sys}(\Theta)$, that is intrinsically irreversible. Hence, the theory exhibits irreversibility [98, 158].

Proof. Assume by contradiction that this is not the case. Namely, suppose there exists a system A with $D_{\mathrm{A}} \neq 1$ such that $\mathscr{I}_{\mathrm{A}}$ is atomic and no instrument with input system A is intrinsically irreversible. Then, for every instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ one can always find a suitable collection of conditional instruments such that

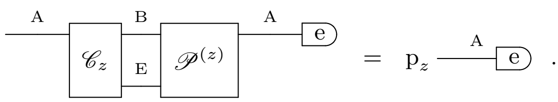


where we observe that, since the identity transformation is deterministic, the ensemble of conditioned instruments $\left\{\mathscr{P}^{(z)}\right\}_{z\in\mathsf{Z}} \subset \mathsf{Transf}_1(\mathrm{BE} \to \mathrm{A})$ are deterministic as well.

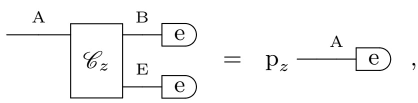
Since $\mathscr{I}_{\mathrm{A}}$ is atomic, it holds that:



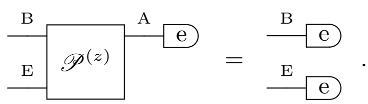
for some probability distribution $[\![\mathrm{p}_z]\!]_{z\in\mathsf{Z}}$. Applying the deterministic effect to both sides yields



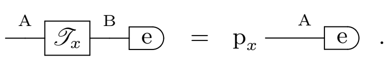
Since $\mathscr{P}^{(z)}$ is deterministic for all z, this reduces to



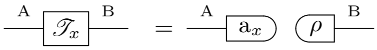
where we used that



Substituting back and summing over all outcomes, one obtains


(4.13)

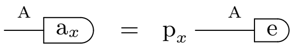
A particular family of instruments that exists in every OPT is given by those in which one performs a measurement followed by the preparation of a fixed deterministic state:



$\forall x \in \mathsf{X},$

where $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ can be any observation-instrument of the system and $\rho \in \mathsf{St}_1(\mathrm{A})$ can be any deterministic state.

In this case, (4.13) forces



$\forall x \in \mathsf{X}.$

Since $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$ is a generic observation-instrument of A, this equality implies that the only admissible observation-instruments are randomisations of the deterministic event $[\![\mathrm{p}_x\, \mathrm{e}]\!]_{x\in\mathsf{X}}$, which in turn implies that $\dim(\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})) = 1$. By duality, also $\dim(\mathsf{St}_{\mathbb{R}}(\mathrm{A})) = 1$, hence $D_{\mathrm{A}} = 1$, in contradiction with the hypothesis $D_{\mathrm{A}} \geq 2$. □

This theorem provides a direct way to see that NIWD implies irreversibility. Indeed, the same conclusion also follows from Theorem 28 and Theorem 29, together with Remark 30. Since NIWD entails incompatibility, and incompatibility is equivalent to irreversibility, it follows that NIWD necessarily implies irreversibility.

In conclusion, we observe that we can formulate Theorem 31 also in a different, but equivalent, way.

Theorem 32

An OPT has irreversibility (or equivalently, incompatibility) if and only if there exists some information that cannot be extracted without disturbance. Equivalently, irreversibility occurs when an OPT fails to satisfy FIWD.

4.4.3 The relationship with thermodynamic irreversibility

A noteworthy aspect of our definition of irreversibility is that it singles out a direction of time. In causal OPTs, there is already a preferred direction of time associated with the flow of information. Irreversibility, however, goes further: it establishes a preferred direction for the dynamical evolution of systems.

Take CT as an example. This OPT is causal but does not exhibit irreversibility[1]. In this case, while the flow of information is time-oriented, the evolution of

[1] One way to see this is that CT satisfies full-compatibility.

systems is not: whatever happens can always be undone, restoring the system to its original state. This is consistent with the standard formulation of classical theory, independently of the OPTs framework. From this perspective, irreversibility marks a sharp boundary between the quantum and classical worlds.

There is, of course, another form of irreversibility that plays a fundamental role in physics: that expressed by the second law of thermodynamics. Understanding how this thermodynamic notion relates to the operational one would be extremely interesting. Yet this constitutes a research program of its own, requiring a full reformulation of thermodynamics in operational terms—a task that remains open, though some promising steps have already been taken [253–255].

For the moment, we leave this question to future investigations.

4.5 Summary

This chapter was quite dense, full of new definitions and important results. We here present a summary of the most important ones that should be kept in mind.

We presented the different definitions of the three physical phenomena encompassed by Heisenberg's *Gedankenexperiment*:

- No-Information Without Disturbance (NIWD) (Definition 34),
- Observation-compatibility (Definition 37),
- Irreversibility (Definition 44).

In particular, we provided a novel definition of irreversibility, capturing the essence of Heisenberg's original proposal.

Furthermore, we characterised the relationships between them, showing that there are indeed different phenomena encompassed by the γ-ray microscope *Gedankenexperiment.*

We showed that there is an equivalence between the existence of information that cannot be extracted without disturbance, the existence of incompatible operations, and irreversibility (Theorem 28, Theorem 29). However, these implications fail if one restricts attention to the compatibility of measurements. In this case, a theory can satisfy NIWD—and therefore admit intrinsically irreversible operations—without exhibiting incompatible observation-instruments (Theorem 30). The reason is that ignoring the system after a measurement allows one to recover more information simultaneously. This is not contradictory: observation-instruments can never be non-disturbing, and, in fact, tracing out is arguably the most disturbing operation that can be performed. Hence, even at a logical level, there is no inconsistency in this last observation.

A Venn diagram representing the relationships established here is presented in figure 8.1.

5 Minimal OPTs

In many parts of this thesis we have mentioned MCT, describing it as a minimal version of CT, where the set of allowed transformations is reduced to the bare minimum. The scope of this chapter is to introduce the notion of *Minimal Operational Probabilistic Theory (MOPT)*, the class of theories to which MCT belongs, and to present a formal construction of this theory, explicitly showing that it provides a counterexample to the converse implication of Theorem 30.

This class of theories was conceived to answer the question of whether it is possible to coherently construct an OPT without assuming either strong causality or the no-restriction hypothesis. A MOPT is a theory in which the sets of states and effects can be freely chosen, while the set of allowed transformations is restricted to the minimal one compatible with the framework. Only states, effects, the braiding, and the identity transformation are included, together with their compositions and the limits of their Cauchy sequences.

This family of OPTs primarily serves as a source of counterexamples that reveal the relationships between different physical properties. Nevertheless, they can also be regarded as a more faithful model of a real-world laboratory, since experimenters typically do not have access to all theoretically implementable transformations, but only to certain subsets of them.

The study of frameworks where theories are subject to restrictions on the set of allowed operations has also been proposed in Ref. [216]. There, the authors introduce and investigate *accessible GPT fragments*, designed to capture scenarios in which states and effects are limited to those that are experimentally accessible in a given setting. Although the notion of accessible GPT fragments is close to that of MOPTs, an important difference must be highlighted. MOPTs, despite the restrictions, remain fully-fledged theories, whereas accessible GPT fragments are not, in general, GPTs themselves. One specific distinction is that in a MOPT the state and effect spaces must be separating, while this generally not hold for accessible GPT fragments.

5.1 Preliminary results

Before providing the formal definition of this class of theories, let us first introduce some preliminary definitions and results.

5.1.1 Elementary systems and minimal decomposition

We begin with a peculiarity that characterises the systems of MOPTs.

Through the notion of operational equivalence of systems (2.18), it is possible to reversibly transform a system into another, faithfully mapping the information contained in one onto the other. This kind of operation is extremely common in everyday life. For instance, a 16-dimensional system can be represented as the composition of four 2-dimensional systems. This is precisely the relation underlying the hexadecimal and binary representation of numbers, widely used in informatics. Similarly, the result of a six-faced die can be expressed as the joint outcome of a coin toss and a three-outcome random generator. The same reasoning applies to quantum systems as well.

However, this possibility can sometimes give rise to problematic situations for the coherent construction of an OPT, particularly when operational equivalence is induced by the identity transformation. More explicitly, it may happen that two composite systems turn out to be equivalent, even though their components are not. For example, in a generic OPT it could occur that four distinct systems $\mathrm{A}, \mathrm{B}, \mathrm{C}$, and D satisfy

$$\mathrm{AB} = \mathrm{CD}.$$

Such a situation leads to critical issues for the coherent construction of MOPTs, due to difficulties in handling permutations of these systems (Remark 36). Problems arise also in the definition of LQTs [172]. For this reason, the class of theories we are interested in is required to satisfy an additional property ruling out these cases.

Definition 45 (Elementary systems and minimal decomposition)

A system A is said to be *elementary* if $\mathrm{A} = \mathrm{BC}$ implies $\mathrm{B} = \mathrm{I}$ or $\mathrm{C} = \mathrm{I}$. Given a composite system S, we say that $\mathrm{S} = \mathrm{A}_1\mathrm{A}_2 \dots \mathrm{A}_k$ is a *minimal decomposition* of S into elementary systems if each A_i is elementary and non-trivial [158].

Definition 46 (Unique decomposition OPTs)

An OPT is said to have *unique decomposition* if for every system S, any two minimal decompositions into elementary systems,

$$\mathrm{S} = \mathrm{A}_1\mathrm{A}_2 \dots \mathrm{A}_k \qquad \text{and} \qquad \mathrm{S} = \mathrm{B}_1\mathrm{B}_2 \dots \mathrm{B}_l,$$

imply $k = l$ and $\mathrm{A}_i = \mathrm{B}_i$ for all $i = 1, \dots, k$.
In words, in an OPT with unique decomposition every composite system admits an unique decomposition in elementary systems [158].

Formally, in OPTs with unique decomposition, composite systems are completely specified by strings of characters denoting elementary systems. When two systems are composed, the resulting system is represented by the concatenation of the strings corresponding to the two subsystems.

Remark 33

When defining an OPT with unique decomposition, the set of elementary systems can be arbitrarily chosen, as long as it is consistent with the definition.
For example, in the case of classical systems, one could choose to regard a six-faced die as an elementary system, even though it can also be obtained as the composition of a coin and a three-outcome random generator. Systems can still be operationally equivalent, just not through the identity transformation. Denoting the three systems by A, B, and C, it is perfectly consistent to have

$$\mathrm{A} \cong \mathrm{BC}.$$

What is prohibited is that

$$\mathrm{A} = \mathrm{BC},$$

provided that all three systems are elementary.

Remark 34

A strong consequence of Remark 33 is that when we define an OPT with unique decomposition without explicitly specifying its set of elementary systems, we are in fact defining a *family* of theories.

In conclusion, we report a result that will be exploited in the following.

Lemma 23

Operationally equivalent systems have the same dimension.

This follows immediately from the definition, since operationally equivalent systems are connected by reversible transformations.

5.1.2 Braiding

We now introduce some notation and establish a few results concerning the braiding transformation.

Definition 47 (Set of braid transformations)

The *set of braid transformations*, whose elements are denoted by $\mathcal{S}$, is the collection of transformations generated by parallel and sequential composition of the braiding and identity transformations [98, 158].

In the special case of symmetric OPTs, braid transformations reduce to *permutations* of the systems on which they act.

Lemma 24

Consider a symmetric OPT with unique decomposition, and let $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{A} \to \mathrm{B})$ be a braid transformation that permutes the systems on which it acts as follows:

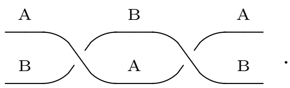
$$\mathrm{B}_j = \mathrm{A}_{\sigma(j)} ,$$

where $\mathrm{A} = \mathrm{A}_1 \dots \mathrm{A}_n$ and $\mathrm{B} = \mathrm{B}_1 \dots \mathrm{B}_n$ are the unique decompositions of A and B, respectively. Then the action of $\mathcal{S}$ is completely characterised by the permutation σ [158].

Proof. This is a direct consequence of the coherence theorem for symmetric monoidal categories (Theorem 43), of which symmetric OPTs are an instance. The theorem ensures that any transformation obtained by composing swap maps is uniquely determined by the permutation it induces on the ordering of the input systems [169, 256]. □

Remark 35

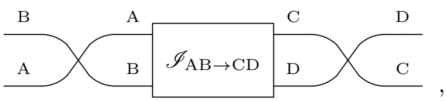
In non-symmetric OPTs, the reordering of subsystems is not sufficient to fully characterise the action of a braid transformation.
For example, compare the identity transformation $\mathscr{I}_{\mathrm{AB}}$ with the following circuit:



Although both reorder the systems in the same way, the resulting transformations are in general different [158].

Remark 36

When uniqueness of decomposition does not hold, a permutation cannot be completely characterised by how it permutes elementary systems.
Let S be a system that admits two distinct decompositions into elementary systems, $\mathrm{AB} = \mathrm{S} = \mathrm{CD}$, with $\mathrm{A} \neq \mathrm{C}, \mathrm{D}$ and $\mathrm{B} \neq \mathrm{C}, \mathrm{D}$. Consider the following permutation:



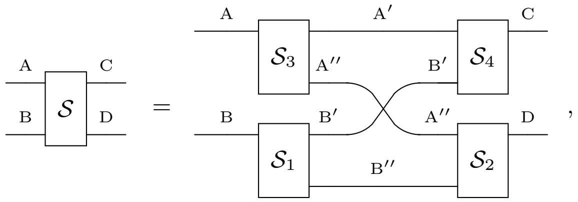
(5.1)

where $\mathscr{I}_{\mathrm{AB}\to\mathrm{CD}}$ denotes the identity test for the equivalence $\mathrm{AB} = \mathrm{CD}$. Since no relation is known between the systems BA and DC other than their operational equivalence, one cannot conclude that this permutation is completely characterised by the way it rearranges its input systems [158].

In the case of symmetric OPTs with unique decomposition, one can formulate a strong characterisation result for permutations.

Lemma 25 (Permutations on bipartite systems)

In every symmetric OPT with unique decomposition, for any permutation $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AB} \to \mathrm{CD})$ there exist systems A', B', A'', B'' and reversible transformations $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$ such that



(5.2)

where A, B are arbitrary systems of the theory, and C, D are systems such that CD has the same decomposition into elementary systems as AB. In general, any of $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}$ may be the trivial system, and the same holds for $\mathrm{A}', \mathrm{A}'', \mathrm{B}', \mathrm{B}''$ [158].

This result follows directly from Lemma 24. For completeness, a proof with all intermediate steps is provided in appendix E.

5.2 The definition

We now have all the elements to properly define MOPTs.

Definition 48 (Minimal Operational Probabilistic Theory (MOPT))

A *Minimal Operational Probabilistic Theory (MOPT)* is an OPT with unique decomposition in which the only allowed instruments are those obtainable by composing the elements of

$$\left\{ \mathsf{I}_{\star}^{\mathrm{A}\to\mathrm{A}} \ , \ \mathsf{S}_{\star}^{\mathrm{AB}\to\mathrm{BA}} \ , \ \left(\mathsf{S}^{-1}\right)_{\star}^{\mathrm{BA}\to\mathrm{AB}} \ , \ \rho_{\mathsf{X}} \ , \ \mathrm{a}_{\mathsf{X}} \right\}, \tag{5.3}$$

where $\rho_{\mathsf{X}} \in \mathsf{Prep}(\mathrm{A})$ and $\mathrm{a}_{\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ range over all possible preparation- and observation-instruments of the theory, together with the Cauchy completion of the above set.

Equivalently, the only allowed transformations are those obtainable by sequential and parallel composition of the elements of

$$\left\{ \mathscr{I}_{\mathrm{A}} \ , \ \mathscr{S}_{\mathrm{A,B}} \ , \ \mathscr{S}^{-1}_{\mathrm{A,B}} \ , \ \rho \ , \ \mathrm{a} \right\}, \tag{5.4}$$

for every $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, with $\rho \in \mathsf{Transf}(\mathrm{I} \to \mathrm{A})$ and $\mathrm{a} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{I})$, and the Cauchy completion of the corresponding spaces of transformations that belong to an instrument of the theory [158].

We now need to verify whether the theories defined in Definition 48 are indeed well-defined OPTs. In order to do so, one must check that the spaces of instruments and transformations are closed under sequential and parallel composition, as well as under coarse-graining. For the moment, we suspend the proof of this fact, since it requires a complete characterisation of all the instruments and transformations of the theory. We reassure the reader that this characterisation can be carried out, and that one can prove the theories to be well-defined in this sense. Moreover, thanks to Theorem 8, the theories remain well defined once these spaces are Cauchy completed.

5.3 Characterisation of instruments and transformations

The spaces of instruments and transformations of MOPTs were given only implicitly in Definition 48. However, it is possible to fully characterise them whenever they are not limits of Cauchy sequences. In this section we present a series of structural results that explicitly describe the instruments and transformations of this class of theories. We also address the case of limits: although a complete characterisation is available only for the deterministic transformations of causal symmetric MOPTs, it is still possible to establish strong structural results about them.

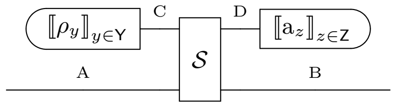
We provided preliminary versions of some of these results in the appendix of Ref. [98] and a more general formulation in Ref. [158].

5.3.1 Non-limit instruments and transformations

We begin with a complete characterisation of the instruments and transformations that can be obtained by sequential and parallel composition of the elements of (5.3) and (5.4), respectively.

Theorem 33 (Instruments of MOPTs)

In every MOPT, any instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ obtained by sequential and parallel composition of the elements of (5.3) is of the form



, (5.5)

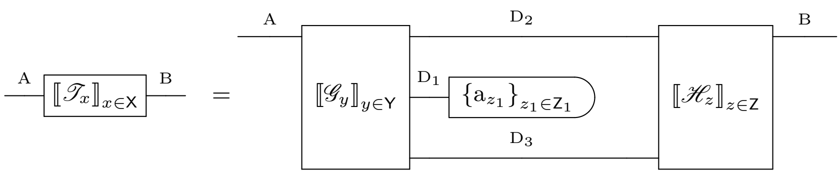
where $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AC}\to\mathrm{DB})$ is a suitable braid transformation, $[\![\rho_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Prep}(\mathrm{C})$, $[\![\mathrm{a}_z]\!]_{z\in\mathsf{Z}} \in \mathsf{Obs}(\mathrm{A})$, the outcome space is $\mathsf{X} = \mathsf{Y}\times\mathsf{Z}$, and $\mathrm{A},\mathrm{B},\mathrm{C},\mathrm{D} \in \mathsf{Sys}(\Theta)$ may also be the trivial system [98, 158].

Proof. The proof is constructive. We show that if we decompose any instrument of an MOPT into its constituent elements (5.3), then these elements can always be rearranged so as to obtain an instrument of the form (5.5).

Let $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ be an instrument satisfying the hypotheses of the theorem.

If its decomposition does not involve either a preparation-instrument or an observation-instrument, then necessarily $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} = \mathcal{S}$, that is, it has the form of (5.6) with $\mathrm{C}=\mathrm{D}=\mathrm{I}$.

Otherwise, suppose that the decomposition involves a total of n instruments, including both preparation- and observation-instruments, and consider one of them in isolation. Without loss of generality, we assume it to be an observation-instrument $\{\mathrm{a}_{z_1}\}_{z_1\in\mathsf{Z}_1} \in \mathsf{Obs}(\mathrm{D}_1)$; the procedure would be identical if carried out with a preparation-instrument. By isolating this observation-instrument, the overall instrument can then be rewritten as:



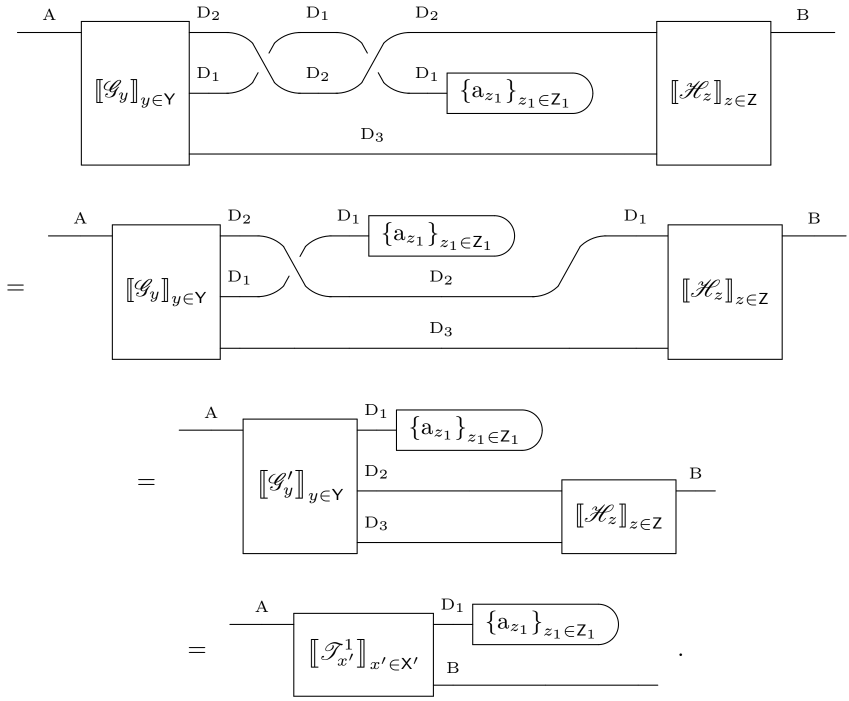
,

for suitable $[\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}}$ and $[\![\mathscr{H}_z]\!]_{z\in\mathsf{Z}}$, where

$$[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} = [\![\mathscr{H}_z]\!]_{z\in\mathsf{Z}} \,\Box\, \left(\mathscr{I}_{\mathrm{D}_2} \boxtimes \{\mathrm{a}_{z_1}\}_{z_1\in\mathsf{Z}_1} \boxtimes \mathscr{I}_{\mathrm{D}_3}\right) \,\Box\, [\![\mathscr{G}_y]\!]_{y\in\mathsf{Y}}\,,$$

and $\mathrm{D}_1,\mathrm{D}_2,\mathrm{D}_3 \in \mathsf{Sys}(\Theta)$.

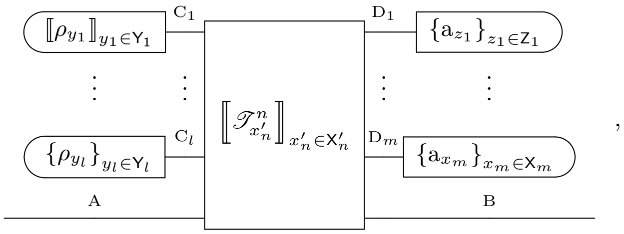
Using the reversibility of the braiding, one can equivalently rewrite the above circuit as



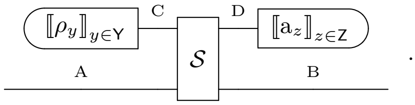
Iterating this procedure on $[\![\mathscr{T}^1_{x'}]\!]_{x'\in\mathsf{X}'}$ eventually yields, after n steps, a singleton instrument

$$\left[\!\!\left[\mathscr{T}^n_{x'_n}\right]\!\!\right]_{x'_n\in\mathsf{X}'_n} = \mathcal{S},$$

which is the desired result. Graphically,



which, by regrouping, is equal to

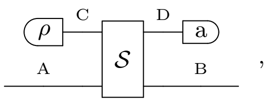


∎

The characterisation of the transformations follows immediately from that of the instruments.

Corollary 4 (Transformations of MOPTs)

In every MOPT, any transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ obtained by sequential and parallel composition of the elements of (5.4) is of the form

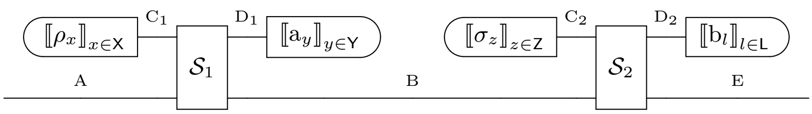


, (5.6)

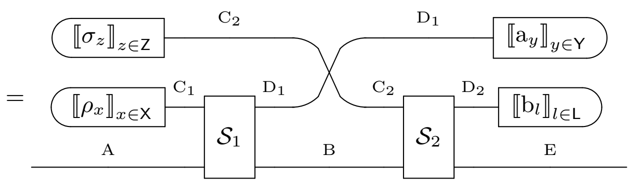
where $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AC} \to \mathrm{DB})$ is a suitable braid transformation, $\rho \in \mathsf{St}(\mathrm{C})$, $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$, and $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D} \in \mathsf{Sys}(\Theta)$ may also be the trivial system [98, 158].

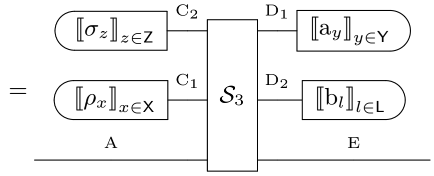
Now that we have a clear and complete characterisation of the instruments and transformations of MOPTs, we can show that they are closed under both sequential and parallel composition. The proof proceeds by direct calculation, verifying that composite instruments are still of the form (5.5).

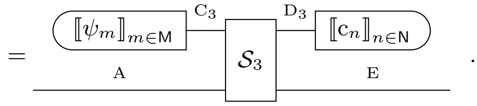
Let us begin with the case of sequential composition.



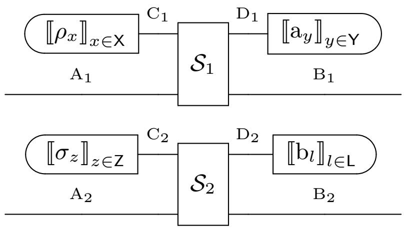
(5.7)



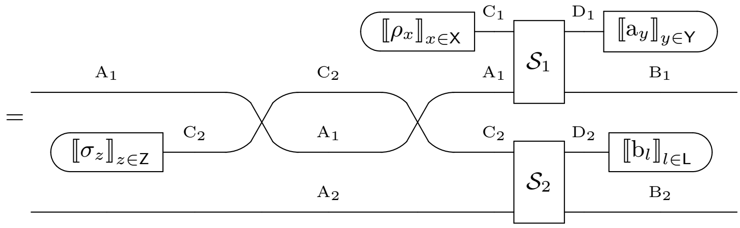


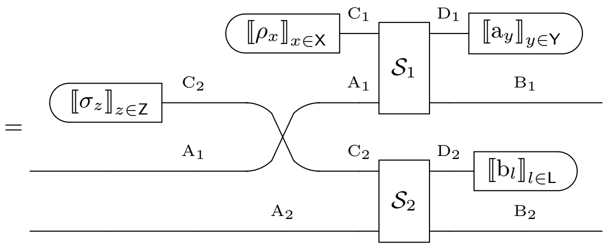


The case of parallel composition is analogous.

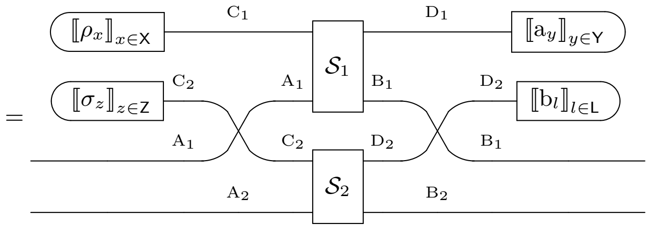


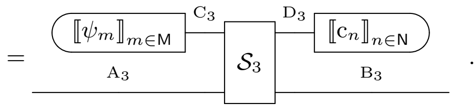
(5.8)





By following the same procedure on the right side of the circuit with $[\![\mathrm{b}_l]\!]_{l\in\mathsf{L}}$, the following expression is obtained:





In conclusion, we observe that (5.5) and (5.6) are preserved under coarse-graining, since this operation amounts to summing terms of the same form.

We have therefore shown that the ensembles of instruments and transformations obtained from sequential and parallel composition of the elements of (5.3) and (5.4), respectively, are closed under sequential composition, parallel composition, and coarse-graining. Well-definiteness after Cauchy completion is then ensured by Theorem 8. As anticipated, this establishes that MOPTs are well-defined OPTs.

Theorem 34 (MOPTs are well-defined)

MOPTs are well-defined OPTs.

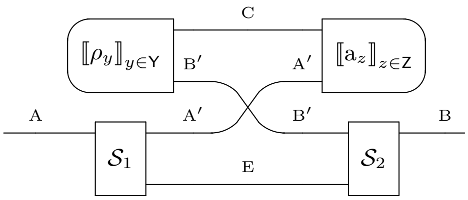
5.3.2 The case of symmetric MOPTs

In the case of symmetric MOPTs, Lemma 25 allows to provide a more refined characterisation of the set of instruments and transformations. By substituting (5.2) into (5.5) and (5.6), one obtains the following results.

Theorem 35 (Instruments of symmetric MOPTs)

In every symmetric MOPT, any instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ that can be obtained by parallel and sequential composition of the elements in

(5.3) is of the form:

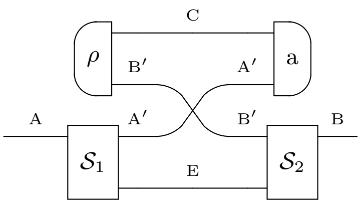


$$, \tag{5.9}$$

where $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{RevTransf}(\Theta)$ are suitable permutations, $[\![\rho_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Prep}(\mathrm{CB}')$, $[\![\mathrm{a}_z]\!]_{z\in\mathsf{Z}} \in \mathsf{Obs}(\mathrm{CA}')$, the outcome space is $\mathsf{X} = Y \times Z$, and $\mathrm{A}, \mathrm{B}, \mathrm{A}', \mathrm{B}', \mathrm{C}, \mathrm{E} \in \mathsf{Sys}(\Theta)$ may also be equal to the trivial system [98].

Corollary 5 (Transformations of symmetric MOPTs)

In every symmetric MOPT, any transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ that can be obtained by parallel and sequential composition of the elements in (5.4) is of the form:



$$, \tag{5.10}$$

where $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{RevTransf}(\Theta)$ are suitable permutations, $\rho \in \mathsf{St}(\mathrm{CB}')$, $\mathrm{a} \in \mathsf{Eff}(\mathrm{CA}')$, and $\mathrm{A}, \mathrm{B}, \mathrm{A}', \mathrm{B}', \mathrm{C}, \mathrm{E} \in \mathsf{Sys}(\Theta)$ may also be equal to the trivial system [98].

The diagrams in (5.9) and (5.10) are often referred to, informally, as *jellyfish* instruments and transformations, respectively.

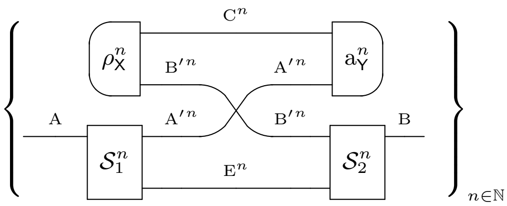
5.3.3 Limits of Cauchy sequences

Now that we have a clear characterisation of the class of instruments and transformations obtainable by composing the elements of (5.3) and (5.4), we can turn to the characterisation of the corresponding limit instruments and transformations. However, despite our efforts, the explicit form of generic instruments and transformations obtained as limits of Cauchy sequences of (5.3) and (5.4) remains an open question. Nonetheless, it is possible to prove structural results that characterise the behaviour of Cauchy sequences of instruments and transformations in MOPTs.

In particular, we begin by showing that the systems and permutations appearing in (5.9) and (5.10) must stabilise along a Cauchy sequence.

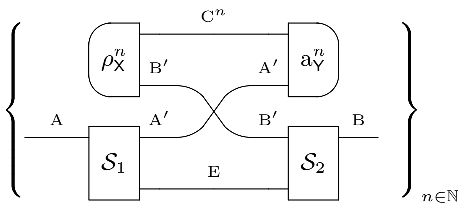
Theorem 36 (Stabilisation of Cauchy sequences of instruments in symmetric MOPTs)

In a symmetric MOPT, any Cauchy sequence of instruments obtained as parallel and sequential compositions of the elements in (5.3) [158],



$$\subset \mathsf{Instr}(\mathrm{A} \to \mathrm{B})\,, \quad (5.11)$$

admits a subsequence in which the systems E^n, A'^n, B'^n and the permutations $\mathcal{S}_1^n$, $\mathcal{S}_2^n$ stabilise, i.e., they can be taken to be fixed:



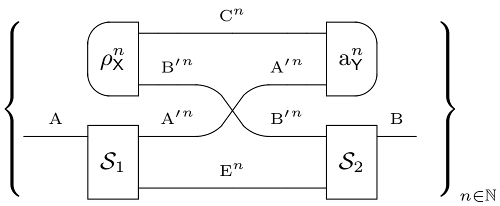
$$\subset \mathsf{Instr}(\mathrm{A} \to \mathrm{B})\,.$$

Proof. Let us start by considering a Cauchy sequence of the form given in (5.11). The proof can then be divided into two steps:

I) Consider the unique decompositions of A and B into elementary systems (Definition 46), which we know involves at most finitely many elementary systems[1]. Recall now that a permutation is completely determined by how it permutes the input wires (Lemma 24). Since there are only finitely many permutations of a finite set, there must exist a pair of permutations, say $\mathcal{S}^1$ and $\mathcal{S}^2$, that occur infinitely many times together in the sequence (5.11). We can therefore restrict our attention to the subsequence with this fixed pair of

[1] In the formulation of the OPT framework adopted here, the composition of an infinite number of systems is not supported.

permutations:



$\subset \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$.

Since (5.11) is a Cauchy sequence, all its subsequences are Cauchy and converge to the same limit.

II) We now turn our attention to the systems E^n. Because $\mathcal{S}^1$ and $\mathcal{S}^2$ are fixed and the decomposition into elementary systems is unique, the set of elementary systems making up the composite $\mathrm{A}'^n\mathrm{E}^n$ cannot change. The only possible variation with n is how these elementary systems are grouped together.

For instance, if $\mathrm{A}'^n = \mathrm{S}_1$ and $\mathrm{E}^n = \mathrm{S}_2\mathrm{S}_3\mathrm{S}_4$, then for some $n' \neq n$ it could be that $\mathrm{A}'^{n'} = \mathrm{S}_1\mathrm{S}_2$ and $\mathrm{E}^{n'} = \mathrm{S}_3\mathrm{S}_4$, or that $\mathrm{A}'^{n'} = \mathrm{S}_1\mathrm{S}_2\mathrm{S}_3$ and $\mathrm{E}^{n'} = \mathrm{S}_4$, or any other regrouping (including the original one), provided the order of the S_i is preserved.

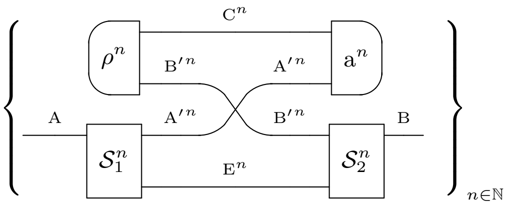
Since $\mathrm{A}'^n\mathrm{E}^n$ and $\mathrm{B}'^n\mathrm{E}^n$ are each composed of finitely many elementary systems, there must exist at least one system E that appears infinitely many times in the sequence. We restrict ourselves to considering the subsequence where E is fixed. Under this condition, the associated systems A′ and B′ are also automatically fixed. This concludes the proof.

□

Thanks to Theorem 7, the conclusion of Theorem 36 also applies to transformations.

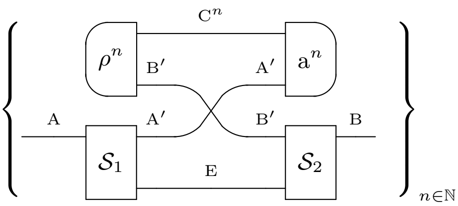
Corollary 6 (Stabilisation of Cauchy sequences of transformations in symmetric MOPTs)

In a symmetric MOPT, any Cauchy sequence of transformations obtained as parallel and sequential compositions of the elements in (5.3) [158],



$\subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B})\,,$

admits a subsequence in which the systems E^n, ${\mathrm{A}'}^n$, ${\mathrm{B}'}^n$ and the permutations $\mathcal{S}_1^n$, $\mathcal{S}_2^n$ stabilise, i.e., they can be taken to be fixed:



$\subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B})\,.$

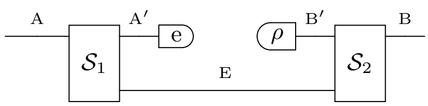
5.3.4 Limits of deterministic transformations

The only case in which it is possible to completely characterise the limits of Cauchy sequences is that of deterministic transformations in causal symmetric MOPTs.

To see why this is the case, let us first characterise the set of deterministic transformations of MOPTs.

Corollary 7 (Deterministic transformations)

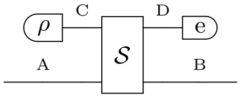
In a causal MOPT (also non-symmetric), every deterministic transformation obtained as a composition of the elements in (5.4) is of the form [98]:



. (5.12)

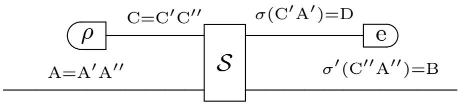
Proof. We recall that the most general transformation of a MOPT is given by (5.10).

In the particular case of deterministic transformations this becomes



with $\rho \in \mathsf{St}_1(\mathrm{C})$ and $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{D})$.

Even though in non-symmetric theories the action of a braid transformation cannot be fully characterised by the permutation it induces on the elementary systems, it still permutes them in a definite way. In the most general case, $\mathcal{S}$ maps part of the systems of A into B and the rest into D, which we denote by A′ and A″, respectively. Either A′ or A″ may coincide with the trivial system. The same holds for C. Consequently, one has

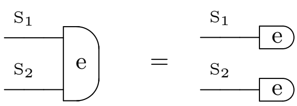


where, for a sequence of systems $\mathrm{S}_1 \ldots \mathrm{S}_n$, we write

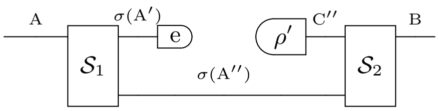
$$\sigma(\mathrm{S}_1 \ldots \mathrm{S}_n) \coloneqq \mathrm{S}_{\sigma(1)} \ldots \mathrm{S}_{\sigma(n)},$$

with σ and σ' suitable permutations.

Next, exploiting the uniqueness of the deterministic effect, which allows one to "split" it when acting on composite systems,



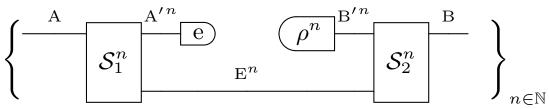
one can slide the deterministic effects towards the left, thanks to the naturality property of the braiding (2.19). The subset acting on the composite system C′ interacts with ρ, while the remainder stay "unresolved":



This is precisely (5.12). □

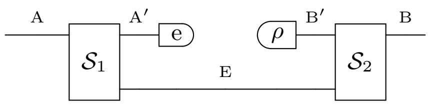
Putting together Corollary 7 and Corollary 6, we then arrive at the following theorem.

Theorem 37 (Limits of Cauchy sequences of deterministic transformations in symmetric MOPTs)
Consider a causal symmetric MOPT. Let



$\subset \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})$

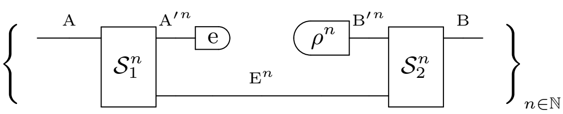
be a Cauchy sequence of deterministic transformations. Then its limit is given by



,

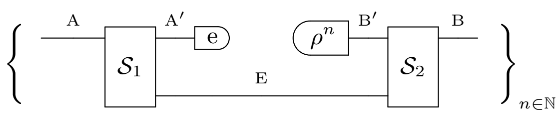
where $\rho = \lim_{n\to\infty}\rho^n \in \mathsf{St}_1(\mathrm{B}')$, with $\mathrm{B}' = \lim_{n\to\infty}\mathrm{B}'^n$. Analogous limits hold for $\mathrm{A}'^n \to \mathrm{A}'$, $\mathrm{E}^n \to \mathrm{E}$, and for the braid transformations $\mathcal{S}_1^n \to \mathcal{S}_1$ and $\mathcal{S}_2^n \to \mathcal{S}_2$ [98, 158].

Proof. The first step is to apply Corollary 6 to the Cauchy sequence



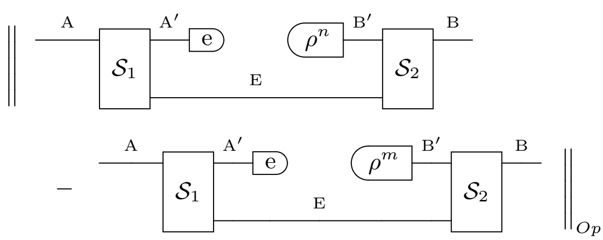
$\subset \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})\,,$

finding a subsequence where the permutations and the "inner" systems stabilise:



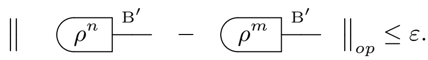
.

Since the original sequence is Cauchy, so is this one. Hence, for every $\varepsilon > 0$ there exists $\tilde{n} \in \mathbb{N}$ such that for all $n, m \geq \tilde{n}$,

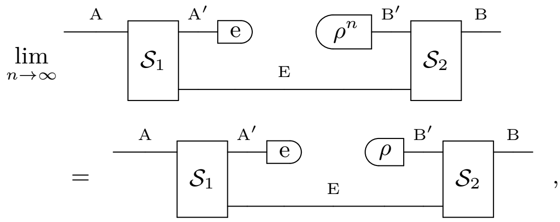


$\leq \varepsilon.$

By Lemma 6 and Lemma 7 we can discard the fixed pre/post-processing permutations and the ancillary system without increasing the norm, obtaining



Therefore $\{\rho^n\}_{n\in\mathbb{N}}$ is Cauchy in $\mathsf{St}_1(\mathrm{B}')$. By completeness there exists $\rho \in \mathsf{St}_1(\mathrm{B}')$ such that $\rho^n \to \rho$. It follows that



which is precisely the claimed limit. □

5.4 Properties of MOPTs

The characterisation of the instruments and transformations that we have carried out in the previous section, in particular Theorem 37, allows us to characterise the properties of MOPTs.

5.4.1 Atomicity of the identity

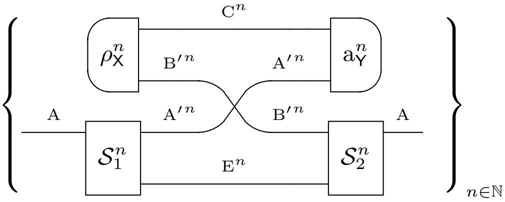
The first and most important property that we prove is that the identity transformation of every MOPT is atomic, independently of the set of states and effects that characterise them.

Theorem 38 (Identity of MOPTs is atomic)

In every causal symmetric MOPT Θ the identity transformation $\mathscr{I}_{\mathrm{A}}$ is atomic for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ [98, 158].

Proof. To prove that the identity is atomic, we have to show that the identity transformation does not admit any non-trivial decomposition, i.e., that the only instruments that can coarse-grain to the identity are of the form $[\![\mathrm{p}_x \mathscr{I}_{\mathrm{A}}]\!]_{x\in\mathsf{X}}$, where $[\![\mathrm{p}_x]\!]_{x\in\mathsf{X}}$ is a probability distribution. f There are two scenarios to consider. First, instruments obtained by sequential and parallel composition of the basic elements of the theory (5.3)—namely those of the form (5.9). Second, limits of such instruments. However, since any fixed instrument can be seen as the limit of a constant sequence, the first scenario is a special case of the second. We therefore work directly at the level of limits.

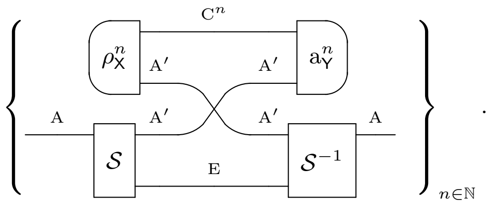
Consider a generic sequence of instruments



$\subset \mathsf{Instr}(\mathrm{A} \to \mathrm{A})\,,$

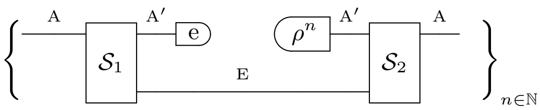
where we restrict to $\mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ because we are interested in instruments that could coarse-grain to the identity.

By Theorem 36 there exists a subsequence in which the permutations and some wires stabilise:



. (5.13)

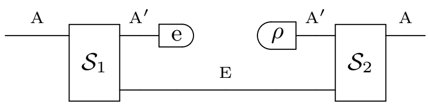
Now, take the full coarse-graining of each element of the sequence (5.13):



.

We recall that all deterministic transformations of MOPTs are of the form (5.12).

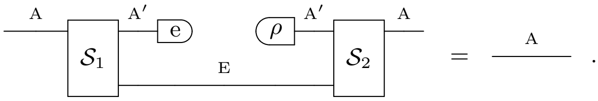
By Corollary 2, this is still a well-defined Cauchy sequence of deterministic transformations. Therefore, we can apply Theorem 37 to determine the limit of this sequence:



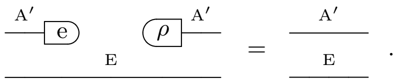
, (5.14)

with $\rho = \lim_{n\to\infty} \rho^n$.

We now determine when (5.14) equals the identity on A; explicitly, when:



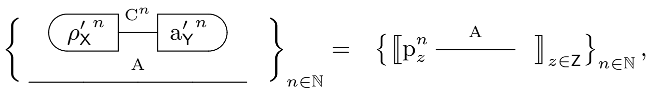
Applying the inverse of the permutations to the left and the right, this is equivalent to



This latter equation is satisfied only if

$$A' = I \qquad \text{and} \qquad E = A.$$

Substituting these constraints back into (5.13), one obtains that any sequence of instruments whose coarse-graining converges to the identity, together with the corresponding limit instrument, has the form



(5.15)

where $\left\{[\![p^n_z]\!]_{z\in Z}\right\}_{n\in\mathbb{N}}$ is a sequence of probability distributions. The limit is characterised by the probability distribution

$$\left\{[\![p^n_z]\!]_{z\in Z}\right\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} [\![p_z]\!]_{z\in Z}\,.$$

Summarising, we have shown that every admissible decomposition of the identity, even when limits are considered, has the form (5.15), namely all transformations composing it are proportional to the identity. Hence, the identity transformation is atomic. □

5.4.2 Irreversibility & More

We can now state the main consequences of how MOPTs manipulate information, given their structural constraints. The atomicity of the identity transformation plays a crucial role, as it entails a number of significant properties for OPTs, as discussed in chapter 4.

First, the atomicity of the identity transformation not only implies that every reversible transformation in a causal symmetric MOPT is atomic, but also that the only reversible transformations admissible in this class of theories are permutations—namely, braid transformations in symmetric theories.

Lemma 26

In any OPT, whenever the identity transformation is atomic for a system A, every reversible transformation with A as input or output must also be atomic [191].

This result follows from the fact that any reversible transformation can be composed sequentially with its inverse to obtain the identity. If a reversible transformation was not atomic—hence admitting a non-trivial decomposition—then, when composed with its inverse, it would yield a non-trivial decomposition of the identity. Therefore, the identity transformation could not be atomic. A fully rigorous proof of this implication can be found in Ref. [191].

Proposition 5

In any causal symmetric MOPT, every reversible transformation is atomic, and the set of reversible transformations coincides with the set of permutations [158].

Proof. The fact that all reversible transformations are atomic follows directly from Theorem 38 and Lemma 26.

We now show that the only reversible transformations are permutations. In MOPTs, the only reversible transformations obtainable from sequential and parallel composition of the elements in (5.4) are permutations. Hence, the only possible way for a different kind of reversible transformation to appear in a minimal theory would be as the limit of a sequence of transformations.

Since reversible transformations are deterministic (Lemma 1) and the set of deterministic transformations is Cauchy complete (Lemma 11), it suffices to consider Cauchy sequences of deterministic transformations. Let

$$\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})$$

be such a sequence converging to a reversible transformation:

$$\mathscr{R} = \lim_{n\to\infty} \mathscr{T}^n \in \mathsf{RevTransf}(\mathrm{A}\to\mathrm{B})\,.$$

Define the sequence

$$\{\mathscr{G}^n\}_{n\in\mathbb{N}} := \left\{\mathscr{R}^{-1}\,\mathscr{T}^n\right\}_{n\in\mathbb{N}}.$$

This sequence is still Cauchy and converges to $\mathscr{I}_{\mathrm{A}}$, by Lemma 6 (or equivalently, Lemma 10 for the sup norm).

By the same reasoning used in the proof of Theorem 38, any sequence of instruments whose coarse-graining converges to the identity must consist of transformations proportional to the identity itself. In this case, since the sequence is composed of deterministic maps, each transformation must coincide with the identity, implying that the sequence is constant. Hence,

$$\mathscr{I}_{\mathrm{A}} = \mathscr{R}^{-1} \square \mathscr{T}^{n} \quad \forall n \in \mathbb{N},$$

which implies

$$\mathscr{R} = \mathscr{T}^{n}, \quad \forall n \in \mathbb{N}.$$

Since $\{\mathscr{T}^{n}\}_{n\in\mathbb{N}}$ is a sequence of transformations obtained by composing the elements of (5.4), it follows that $\mathscr{R}$ is also a transformation composed of the elements of (5.4). Given that $\mathscr{R}$ is reversible, this means that it is a permutation. Therefore, the Cauchy completion of a causal symmetric MOPT does not introduce any new reversible transformations. In other words, the set of reversible transformations coincides with the set of permutations. □

Further consequences of the atomicity of the identity are the following.

Corollary 8

Every causal symmetric MOPT satisfies the property of NIWD [158].

This follows immediately from Theorem 26.

Moreover, from Proposition 5 and Theorem 31 we deduce the irreversibility of minimal theories.

Corollary 9

Every non-trivial causal symmetric MOPT exhibits irreversibility [158].

Where we define *trivial OPTs* in the following way.

Definition 49 (Trivial OPT)

A *trivial* OPT is an OPT Θ in which every system has dimension equal to one, i.e., $D_{\mathrm{S}} = 1$ for all $\mathrm{S} \in \mathsf{Sys}(\Theta)$.

Finally, in every minimal theory there can be no broadcasting channel, as it follows from Theorem 27.

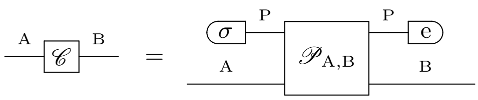
Corollary 10

Every causal symmetric MOPT does not admit a broadcasting channel [158].

As an addition, we also observe that no MOPTs can satisfy the programming theorem.

Definition 50 (Universal simulator)

Consider a causal OPT and a pair of systems A and B of the theory. A *universal simulator* for $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is a deterministic transformation $\mathscr{P}_{\mathrm{A},\mathrm{B}} \in \mathsf{Transf}_1(\mathrm{PA} \to \mathrm{PB})$ for some suitable system P, such that for every deterministic transformation $\mathscr{C} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$ there exists a program state $\sigma \in \mathsf{St}_1(\mathrm{P})$ with the property that [129]:



.

Definition 51 (No-programming)

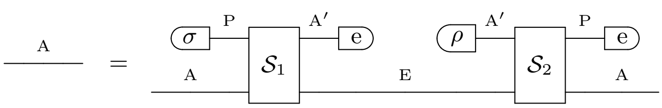
An OPT is said to satisfy *no-programming* if there exists some pair of systems A and B that does not admit a universal simulator [129].

Corollary 11

Every causal symmetric MOPT satisfies no-programming [158].

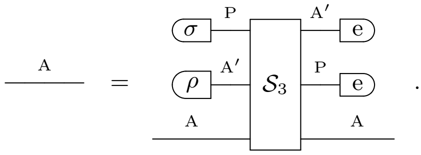
Proof. Assume by contradiction that a universal simulator $\mathscr{P}_{\mathrm{A},\mathrm{A}}$ for a system A exists, with σ the state that programmes the identity.

If we express $\mathscr{P}_{\mathrm{A},\mathrm{A}}$ in the most generic form a deterministic transformation can take in a MOPT (5.12), then the equality expressing the programming of the identity is

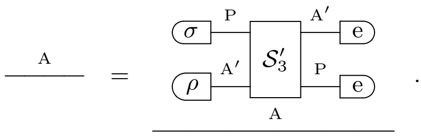


,

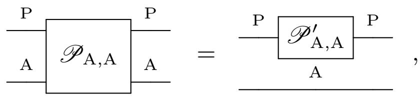
for a suitable state ρ and permutations $\mathcal{S}_1, \mathcal{S}_2$. Following the same steps as in (5.7), we can rewrite this as



For this equality to hold, system A must remain untouched when passing through the braid transformation $\mathcal{S}_3$. No part of A can be discarded or fixed by the states ρ and σ. Therefore, $\mathcal{S}_3$ must act trivially on A, and we obtain



Thus, the identity can be programmable only if $\mathscr{P}_{\mathrm{A,A}}$ acts trivially on system A, i.e., it is of the form



which cannot programme anything different than the identity, ruling out the existence of a universal simulator. □

We conclude with a remark discussing the relationship of MOPTs with strong causality and the no-restriction hypothesis.

Remark 37

The first observation is that causal MOPTs do not satisfy strong causality, as conditional instruments are, by construction, excluded in these theories[a].

The case of the no-restriction hypothesis is more delicate. The property (Definition 22) is highly sensitive to the choice of preparation- and observation-instruments that characterise a theory. Since the framework does not constrain how articulated or complex these may be, in general no statement can be made about whether a MOPT satisfies the hypothesis without additional assumptions. For instance, even a simple conditional preparation such as

$$\left\{\rho^{(x)}\right\}_{x\in\mathsf{X}} \triangleright \mathsf{a}_{\mathsf{X}} = \left[\!\left[\rho^{(x)} \,\square\, \mathrm{a}_x\right]\!\right]_{x\in\mathsf{X}}, \tag{5.16}$$

with $\mathrm{a}_\mathsf{X} \equiv \llbracket \mathrm{a}_x \rrbracket_{x \in \mathsf{X}} \in \mathsf{Obs}(\mathrm{S})$ and $\left\{\rho^{(x)}\right\}_{x \in \mathsf{X}} \subset \mathsf{St}_1(\mathrm{S})$, is not guaranteed to yield an admissible preparation-instrument. It may well be that no preparation-instrument can take the form

$$\left\llbracket \mathrm{p}\rho^{(x_1)},\ (1-\mathrm{p})\,\rho^{(x_2)} \right\rrbracket .$$

This is somewhat surprising, given that one of the original motivations for introducing MOPTs was precisely to construct theories that do not satisfy the no-restriction hypothesis. Research often unfolds in unexpected directions.
However, if no restriction is posed on what the preparation-instruments of a MOPT can be, then instruments of the form (5.16) suffice to rule out that the theory can satisfy the no-restriction hypothesis.

[a] A more detailed discussion on this point can be found at the beginning of the next chapter.

5.5 Minimal Classical Theory (MCT)

We now have all the elements to formalise MCT [98].

Definition 52 (Minimal Classical Theory (MCT))

Minimal Classical Theory (MCT) is the OPT Θ characterised by the following postulates:

I) Θ is simplicial and convex;

II) Local discriminability holds;

III) For every system $\mathrm{S} \in \mathsf{Sys}(\Theta)$, the preparation-instruments are all collections of generalised states in $\mathsf{St}_{\mathbb{R}}(\mathrm{S})$ summing to a point in the convex hull of $\mathsf{PurSt}(\mathrm{S})$;

IV) The no-restriction hypothesis holds for observation-instruments;

V) The theory is minimal.

In other words, MCT is the MOPT obtained by taking a minimal version of CT, as defined in Definition 28. The spaces of states and effects remain the same; the only difference lies in the spaces of transformations with non-trivial input and output systems.

Remark 38 (Classicality and non-contextuality for MCT)

By construction, MCT is classical. Indeed, the first postulate ensures that the theory is simplicial, while the no-restriction hypothesis for observation-instruments guarantees that the pure states of the simplex are perfectly discriminable. Equivalently, one can observe that MCT shares the same sets of preparation- and observation-instruments as CT. Since classicality is fully characterised by these structures, it follows that MCT is classical.
The theory is also Kochen–Specker non-contextual as well as generalised-non-contextual, by Theorem 25.

Remark 39

In accordance with Remark 34, Definition 52 does not define a single OPT, but rather a family of theories. To single out a specific theory, one has to specify the set of elementary systems. There are two natural choices:

- One may include an elementary system for each dimension. This allows the existence of high-dimensional systems that are not composite. For instance, a 4-dimensional system would not necessarily arise as the composition of two 2-dimensional ones.

- Alternatively, one may adopt the minimal choice of considering only systems of prime dimension as elementary, since these suffice to reconstruct, by composition, systems of arbitrary dimension.

In what follows we will speak of MCT as if it were a single theory, since the properties we establish hold irrespective of the particular choice of the set of elementary systems.

Observation 11

We observe that a theory analogous to MCT has also been proposed in Ref. [238]. In that work, the authors present a theory with instruments that are intrinsically irreversible, but whose associated observation-instruments are compatible. However, the theory does not satisfy observation-compatibility and thus cannot serve as a counterexample for decoupling irreversibility from observation-incompatibility in the way that MCT does.

5.5.1 Properties of MCT

We now summarise some additional properties of MCT, beyond those already established for all MOPTs.

Lemma 27

MCT satisfies observation-compatibility.

This result follows directly from the fact that MCT has the same observation-instruments as CT. Since CT satisfies observation-compatibility, so does MCT. The proof presented in appendix D applies unchanged in this case.

However, despite this property, MCT does not satisfy full-compatibility.

Lemma 28

MCT does not satisfy full-compatibility and features irreversibility.

This is a direct consequence of Corollary 8[2], together with Theorem 28 and Theorem 29.

Therefore, we have established that MCT indeed provides a valid counterexample to the converse implication of Theorem 30.

In conclusion, we observe that MCT fails to satisfy both strong causality and the no-restriction hypothesis, as captured by the following lemmas.

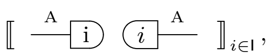
Lemma 29

MCT is not strongly causal.

Lemma 30

MCT does not satisfy the no-restriction hypothesis.

One way to see why these results hold is that MCT does not admit the decomposition of the identity:



(4.3)

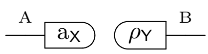
where $[\![\mathrm{i}]\!]_{i\in\mathsf{I}} \in \mathsf{Obs}(\mathrm{A})$ is the observation-instrument that perfectly discriminates the pure states $\{i\}_{i\in\mathsf{I}} = \mathsf{PurSt}(\mathrm{A})$.

[2] MCT satisfies the hypotheses of the corollary: it is minimal by construction, symmetric by Proposition 2, and causal by Theorem 22.

6 Minimal Strongly causal OPTs

THE class of OPTs introduced in the previous chapter, namely MOPTs, is quite interesting and provides a nice collection of counterexamples. Nonetheless, it is not entirely satisfactory. The original objective was to construct a class of theories where the set of allowed operations was restricted to the bare minimum. In retrospect, however, we might have taken this restriction a step too far.

Consider, for instance, MCT. From a purely mathematical perspective, it is a perfectly valid counterexample to Theorem 30. From a physical standpoint, however, it is less compelling. In MCT, the only operations with non-trivial input and output systems are permutations of subsystems between agents, together with instruments of the form



.

That is, experiments where a system is first measured, and then a new system is randomly prepared in one of its states, independently of the outcome of the previous measurement. In practice, this means that no action can ever depend on any previously obtained information: one can only act at random, completely oblivious to the past. A world without conditioning.

This naturally raises the question of whether we could construct a counterexample to Theorem 30 that does allow conditioning. Put differently: is conditioning, together with observation-compatibility, sufficient to guarantee full-compatibility? The answer we found is negative. There exists a strongly causal classical OPT that satisfies observation-compatibility while still exhibiting irreversibility. However, achieving this requires abandoning local discriminability. Indeed, if one adds all conditional instruments to MCT, the result is CT itself, which does satisfy full-compatibility.

The new counterexample we introduce here is called MSBCT. It is a particular instance of a broader class of OPTs, denoted Minimal Strongly causal Operational Probabilistic Theories (MSOPTs). In words, these are simply MOPTs supplemented with all conditioned instruments.

It turns out that even this enlarged class of theories satisfies a number of interesting no-go theorems, all ultimately stemming from the atomicity of the identity transformation. In this case, however, atomicity is guaranteed to hold only in the presence of entanglement.

6.1 The definition

Formally, MSOPTs are defined as follows.

Definition 53 (Minimal Strongly Causal Operational Probabilistic Theory)

A *Minimal Strongly causal Operational Probabilistic Theory (MSOPT)* is an OPT with unique decomposition in which the allowed tests are those of the corresponding MOPT (with the same systems), together with all their conditional instrument, and subsequently closed under Cauchy completion [158].

There is another equivalent way to construct a MSOPT: one can take an MOPT and then exploit Procedure 1.

Theorem 39 (Equivalence of the two construction procedures for MSOPTs)

Let Θ be an MOPT, and let I_0 denote the set of instruments generated from (5.3) by sequential and parallel composition. Define

$$\mathsf{Instr}_0 := \overline{\triangleright(I_0)},$$

$$\mathsf{Instr}_1 := \overline{\triangleright\left(\overline{I_0}\right)},$$

where $\triangleright(\cdot)$ denotes closure under well-defined conditioning operations, and $\overline{(\cdot)}$ denotes Cauchy completion. Then, $\mathsf{Instr}_0 = \mathsf{Instr}_1$. In particular, Instr_0 is the set of instruments obtained via the procedure in Definition 53, whereas Instr_1 is obtained by first constructing the MOPT and then applying Procedure 1.

Proof. Let us first unpack the two procedures.

Procedure I: start from I_0, close under conditioning, and then Cauchy-complete, obtaining $\mathsf{Instr}_0 = \overline{\triangleright(I_0)}$, as in Definition 53.

Procedure II: first Cauchy-complete I_0, then close under conditioning, and finally Cauchy-complete again, obtaining $\mathsf{Instr}_1 = \overline{\triangleright\left(\overline{I_0}\right)}$.

To prove the result—namely, $\mathsf{Instr}_0 = \mathsf{Instr}_1$—we show that each set of instruments is contained in the other.

We start with the inclusion $\mathsf{Instr}_0 \subseteq \mathsf{Instr}_1$. Since $I_0 \subseteq \overline{I_0}$, by monotonicity of the conditioning operation we have

$$\triangleright(I_0) \subseteq \triangleright\left(\overline{I_0}\right).$$

Here, monotonicity simply reflects the fact that enlarging the set of instruments provides more opportunities for conditioning. Then, taking the Cauchy completion of both sides gives

$$\overline{\triangleright(I_0)} \subseteq \overline{\triangleright\left(\overline{I_0}\right)}.$$

For the opposite inclusion, $\mathsf{Instr}_1 \subseteq \mathsf{Instr}_0$, observe that

$$\triangleright\left(\overline{I_0}\right) \subseteq \overline{\triangleright(I_0)},$$

which follows directly from Theorem 21. Indeed, as shown in the proof of that theorem, $\overline{\triangleright(I_0)}$ contains all conditional instruments obtained by conditioning between limits—that is, where both the conditioned instrument and the conditioning event may themselves be limits of sequences of instruments from the original set I_0. Taking the Cauchy completion of both sides then yields

$$\overline{\triangleright\left(\overline{I_0}\right)} \subseteq \overline{\triangleright(I_0)}.$$

We have thus proved both inclusions, and hence $\mathsf{Instr}_0 = \mathsf{Instr}_1$. Since the two constructions yield the same sets of instruments, they also yield the same sets of transformations, showing that the two resulting OPTs are equivalent. □

In what follows, we will use either construction—whichever best suits the argument at hand.

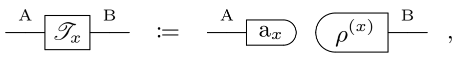
6.2 Characterisation of instruments and transformations

Since an MSOPT can be regarded as an extension of an MOPT—obtained by including conditional instruments—the sets of instruments and transformations of the former contain those of the latter. In particular, the set of instruments of an MSOPT strictly contains that of the corresponding MOPT:

$$\mathsf{Instr}(MOPT) \subset \mathsf{Instr}(MSOPT)\,, \tag{6.1}$$

where $\mathsf{Instr}(OPT)$ denotes the set of instruments of an OPT. The inclusion is strict precisely due to the presence of the conditional instruments.

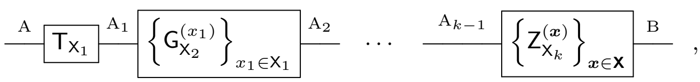
For example, while each transformation of the form



with A and B arbitrary systems of the theory, $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ a generic observation-instrument, and $\left\{\rho^{(x)}\right\}_{x\in\mathsf{X}} \subset \mathsf{St}_1(\mathrm{B})$ a collection of deterministic states, is a legitimate transformation in MOPTs, the corresponding collection $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$ forms an instrument only in MSOPTs.

The inclusion (6.1) guarantees that all the results we have proved to characterise the instruments of MOPTs still hold in the case of MSOPTs. For example, Theorem 33 still provides a characterisation of the instruments of these theories. However, in this case it does not characterise all instruments, but only the subset of non-conditional ones. A further explicit characterisation of the instruments of a generic MSOPT has not been possible. Even when restricting to deterministic transformations, the conditioning operation makes the task extremely challenging, if not impossible. Nevertheless, a full characterisation is not necessary in order to prove that MSOPTs satisfy some properties of interest.

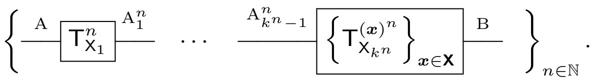
To characterise the instruments of MSOPTs, let us start to recall that the non-limit instruments (and consequently the transformations) added to an MOPT in order to obtain an MSOPT through closure under conditioning, as described in Procedure 1, are of the form



(6.2)

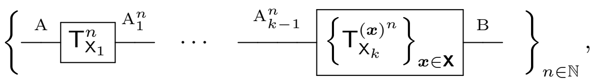
where $\boldsymbol{x} \coloneqq (x_1, x_2, \dots, x_{k-1})$, $\boldsymbol{\mathsf{X}} \coloneqq \mathsf{X}_1 \times \mathsf{X}_2 \times \cdots \times \mathsf{X}_{k-1}$, and each instrument appearing in the sequence of compositions belongs to the underlying MOPT, hence it is either of the form (5.5) (or (5.9) in the symmetric case) or a limit of a Cauchy sequence thereof. Moreover, according to Remark 23, the number of conditioning steps is arbitrary but finite (here equal to k).

One must then include the limits of Cauchy sequences of such conditional instruments, namely limits of sequences of the form



(6.3)

Notice that in (6.3) the systems on the "inside" wires A_i^n depend on the index of the sequence, while the input and output systems are fixed. Moreover, the number of conditioning steps k^n can also vary with the index of the sequence. However, for each n we can take the maximal value $k = \max_n k^n$ over the whole sequence, padding with conditioning by the identity at the extra steps $k - k^n$ if necessary. In this way, we can reduce to considering instruments that are limits of Cauchy sequences of the form



(6.4)

with a fixed number of conditioning steps.

Clearly, (6.2) is a special case of (6.4). Non-limit instrument can always be seen as the limits of constant sequences.

Summarising, the most general instrument belonging to a MSOPT is the limit of Cauchy sequences of the form

$$\left\{ \left\{ \mathsf{G}_{\mathsf{X}_k}^{(\boldsymbol{x})^n} \right\}_{\boldsymbol{x}\in\mathbf{X}} \triangleright \mathsf{T}_{\mathbf{X}}^n \right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{A} \to \mathrm{B}) \,, \tag{6.5}$$

where, for simplicity, we highlighted only the last conditioning step. Explicitly, one has

$$\left\{ \left\{ \mathsf{G}_{\mathsf{X}_k}^{(\boldsymbol{x})^n} \right\}_{\boldsymbol{x}\in\mathbf{X}} \right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{A}_{k-1} \to \mathrm{B}) \,,$$

and

$$\{\mathsf{T}_{\mathbf{X}}^n\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{A} \to \mathrm{A}_{k-1}) \,,$$

with $\boldsymbol{x} = (x_1, \dots, x_{k-1}) \in \mathbf{X} = \mathsf{X}_1 \times \cdots \times \mathsf{X}_{k-1}$. Consequently, the most general transformation of a MSOPT is given by the limit of Cauchy sequences of the form

$$\left\{ \mathscr{G}_{x_k}^{(\boldsymbol{x})^n} \,\square\, \mathscr{T}_{\boldsymbol{x}}^n \right\}_{n\in\mathbb{N}} \subset \mathsf{Transf}(\mathrm{A} \to \mathrm{B}) \,.$$

6.3 Properties of MSOPTs

We now have sufficient elements to establish a series of results concerning the properties satisfied by MSOPTs.

Before starting, we remark that, by construction, MSOPTs are strongly causal OPTs.

Theorem 40

MSOPTs are strongly causal.

6.3.1 Atomicity of the identity

Our first—and most relevant—result shows that, in the presence of entanglement, the identity transformation is guaranteed to be atomic in symmetric MSOPTs.

Theorem 41 (Identity of MSOPTs is atomic)

In every symmetric MSOPT that admits a spanning set of entangled states for every composite system, the identity transformation is atomic for every system [158].

Proof. To prove that the identity transformation is atomic, we have to show that no instrument of the MSOPT can coarse-grain to it—under these hypothesis—unless it is a trivial decomposition of it, i.e., the instrument is a randomisation of the identity: $[\![\mathrm{p}_x \mathscr{I}_{\mathrm{A}}]\!]_{x \in \mathsf{X}}$, for a probability distribution $[\![\mathrm{p}_x]\!]_{x \in \mathsf{X}}$.

Following the same line of reasoning as in the proof of Theorem 38, we analyse under which conditions the limit instrument of a Cauchy sequence of instruments can coarse-grain to the identity. We then show that, whenever the coarse-graining is the identity, every element of the sequence—and hence its limit—must be a randomisation of the identity.

This strategy is sufficient to cover all instruments of the theory, since any instrument can be regarded as the limit of the corresponding constant sequence.

Strictly speaking, it would be enough to establish the statement for conditional instruments and their limits: by Theorem 38, no other kind of decomposition of the identity is possible. Nevertheless, since non-conditional instruments are a special case of conditional ones (corresponding to trivial conditioning), the argument below applies to arbitrary instruments of a MSOPT.

Formally, consider a generic sequence of instruments in a MSOPT, of the form given in (6.5):

$$\left\{ \left\{ \mathsf{G}^{(\boldsymbol{x})\,n}_{\mathsf{X}_k} \right\}_{\boldsymbol{x}\in\mathbf{X}} \triangleright \mathsf{T}^n_{\mathbf{X}} \right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}(\mathrm{A}\to\mathrm{A})\,, \tag{6.5}$$

where we assume that the overall input and output systems coincide, since we are interested in decompositions of the identity.

The proof then proceeds in two steps:

I) First, we consider the sequence of deterministic transformations obtained by completely coarse-graining over the outcome space $\mathbf{X}\times\mathsf{X}_k$ the instruments in (6.5), and study under which conditions it converges to the identity. In other words, we characterise the conditions under which the following equality holds:

$$\lim_{n\to\infty}\sum_{\boldsymbol{x}\in\mathbf{X}} \xrightarrow{\mathrm{A}} \boxed{\mathscr{T}^n_{\boldsymbol{x}}} \xrightarrow{\mathrm{A}^n_{k-1}} \boxed{\mathscr{G}^{(\boldsymbol{x})\,n}_{\mathsf{X}_k}} \xrightarrow{\mathrm{A}} \quad = \quad \xrightarrow{\mathrm{A}} \quad , \tag{6.6}$$

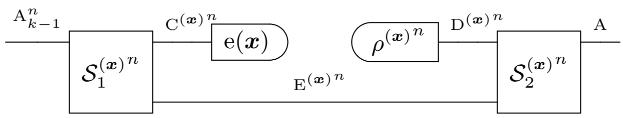
where $\mathscr{G}^{(\boldsymbol{x})\,n}_{\mathsf{X}_k}$ denotes the deterministic transformation obtained by fully coarse-graining the instrument $\left[\!\left[\mathscr{G}^{(\boldsymbol{x})}_{x_k}\right]\!\right]^n_{x_k\in\mathsf{X}_k}$ over X_k for a fixed $\boldsymbol{x}\in\mathbf{X}$, for each $n\in\mathbb{N}$.

The main difficulty lies in the fact that the transformations appearing in the sequence above exhibit a complicated dependence on the index n: at an arbitrary conditioning step k' both the input system $\mathrm{A}^n_{k'-1}$ and the output system $\mathrm{A}^n_{k'}$ depend on n, thus preventing us from directly applying the stabilisation result of Theorem 36. Nevertheless, this problem can be circumvented. Since the overall input and output systems are fixed, we can start from them and make statements about the transformations they are connected to. In particular, we proceed from the last conditioning step, analysing what can be said about the instruments $\left[\!\left[\mathscr{G}^{(\boldsymbol{x})}_{x_k}\right]\!\right]^n_{x_k\in\mathsf{X}_k}$.

II) Using the constraints obtained in the previous step, we then examine the structure of the transformations inside the sequence of conditional instruments. The details of this second step are more involved and are presented in full in appendix F.

Let us then dive into the proof.

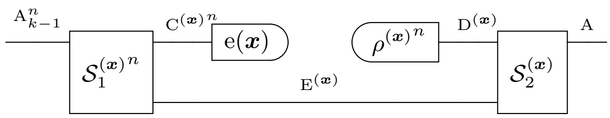
Item I): By Corollary 7 and Theorem 37, the deterministic map $\mathscr{G}^{(\boldsymbol{x})\,n}_{\mathsf{X}_k}$ takes the form



,

for all $\boldsymbol{x} \in \mathbf{X}$ and every $n \in \mathbb{N}$, where $\rho^{(\boldsymbol{x})\,n} \in \mathsf{St}_1\big(\mathrm{D}^{(\boldsymbol{x})\,n}\big)$ are suitable deterministic states. Although the dependence on $\boldsymbol{x}$ in the deterministic effect could be omitted—since the deterministic effect is unique in causal theories—we retain it in the notation since, in the subsequent analysis of the transformations that coarse-grain to this one (carried out in Item II), the effects associated with the observation-instruments summing to the deterministic effect will in fact depend on the outcome $\boldsymbol{x}$.

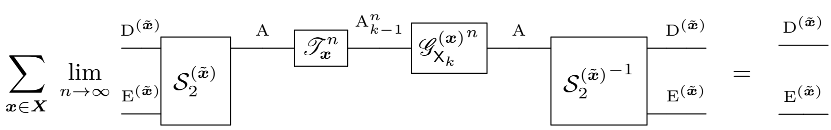
As anticipated at the beginning of the proof, to proceed in our argument we cannot apply Theorem 36 directly, since the input system A^n_{k-1} is not fixed along the sequence. Nevertheless, the same line of reasoning used in the proof of that theorem can still be applied to the fixed system A. Given that the number of elementary systems composing A is finite, the number of ways in which these subsystems can be permuted and grouped is also finite. Therefore, there exists a subsequence where one of these combinations of permutation and grouping is realised infinitely many times. We restrict to such a subsequence, in which $\mathscr{G}^{(\boldsymbol{x})\,n}_{\mathsf{X}_k}$ takes the form



, (6.7)

where $\mathrm{D}^{(\boldsymbol{x})\,n}$, $\mathrm{E}^{(\boldsymbol{x})\,n}$, and $\mathcal{S}_2^{(\boldsymbol{x})\,n}$ are now constant.

If we now consider an arbitrary outcome $\tilde{\boldsymbol{x}} \in \mathbf{X}$ and consider the permutation $\mathcal{S}_2^{(\tilde{\boldsymbol{x}})}$ associated with it, we can rewrite the condition (6.6) as

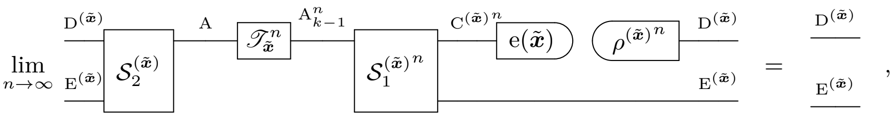


, (6.8)

where exploited Corollary 2 to exchange the operations of limit and coarse-graining and we applied the permutation $\mathcal{S}_2^{(\tilde{\boldsymbol{x}})}$ [and its inverse] to the left [right] on both sides of the equality.

In particular, the term of the coarse-graining associated to the outcome $\tilde{\boldsymbol{x}} \in \mathbf{X}$

takes the form



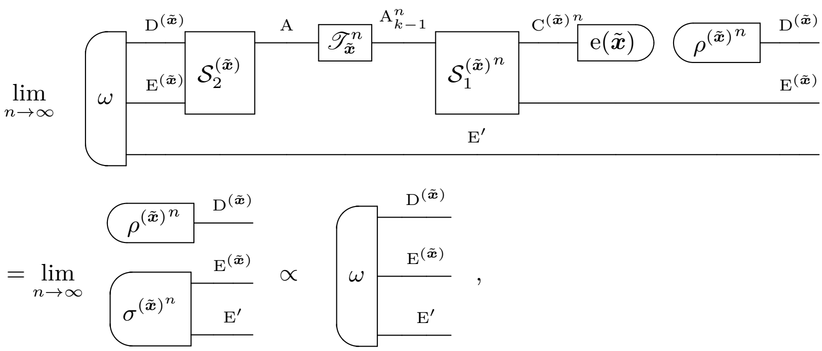
(6.9)

which is obtained by explicitly substituting (6.7) into (6.8) and cancelling the permutation with its inverse.

Let now $\omega \in \mathsf{St}_1\left(\mathrm{D}^{(\tilde{\boldsymbol{x}})}\mathrm{E}^{(\tilde{\boldsymbol{x}})}\mathrm{E}'\right)$ be a deterministic atomic entangled state, where E' is an arbitrary ancillary system. The existence of such a state follows from the assumption that entangled states are spanning, together with Corollary 3, which guarantees the presence of atomic states for every system in a strongly causal OPT. One way to see how these results imply the existence of an atomic entangled state is the following. Start from an arbitrary entangled state and consider its refinement set, namely the set of states that coarse-grain to the state of interest. This set either contains a single element—in which case the state is atomic—or contains more than one element, in which case at least one of them must be entangled. This is guaranteed by the assumption that entangled states are spanning and by the fact that entangled states cannot be obtained as sums of separable states. By iterating this procedure, one eventually obtains an atomic entangled state. This construction, albeit somewhat unusual, is realised in many theories, such as QT and BCT. Finally, since in strongly causal OPTs every state is proportional to a deterministic one (Theorem 20), we can always take ω to be deterministic.

If we now evaluate both sides of (6.8) on ω, atomicity implies that each term of the coarse-graining on the left-hand side must be proportional to ω itself.

In particular, the term associated to the outcome $\boldsymbol{x} = \tilde{\boldsymbol{x}}$ takes the form

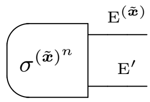


(6.10)

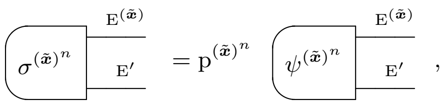
where we used (6.9) to write the left-hand side of the equation.

We observe that, thanks to the monotonicity of the operational norm (Lemma 6), the sequence obtained by composing (6.8) with a deterministic state is still Cauchy.

Since $\mathscr{T}_{\tilde{\boldsymbol{x}}}^{n}$ is, in general, not deterministic, also the state



is not deterministic. However, thanks to Theorem 20, we can rewrite it as



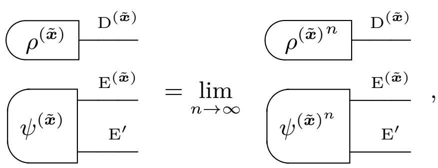
with $\psi^{(\tilde{\boldsymbol{x}})^n} \in \mathsf{St}_1\left(\mathrm{D}^{(\tilde{\boldsymbol{x}})}\right)$ a deterministic state and $\mathrm{p}^{(\tilde{\boldsymbol{x}})^n} \in [0,1]$ an appropriate rescaling factor. This can be done for any $n \in \mathbb{N}$.

Lemma 13 and Lemma 12 then allow us to characterise the limit appearing on the right-hand side of (6.10). The former characterises the limits of deterministic transformations rescaled by a factor in $[0,1]$, stating that the limit is given by the composition of the limits of the sequences of rescaling factors and of the sequence of deterministic states taken separately. The latter guarantees that the limit of deterministic compound-local transformations is again deterministic and compound-local; in particular, limits of deterministic separable states remain separable and deterministic.

Explicitly, Lemma 13 allows us to evaluate the two limits separately:

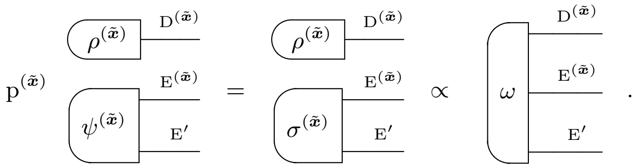
$$\mathrm{p}^{(\tilde{\boldsymbol{x}})} = \lim_{n \to \infty} \mathrm{p}^{(\tilde{\boldsymbol{x}})^n},$$

and



where Lemma 12 guarantees that the limit is still a deterministic separable state.

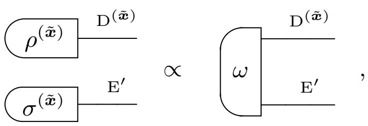
Putting everything together, we obtain that the limit is given by



What we have found is an impossible condition: the proportionality between a separable and an entangled state.

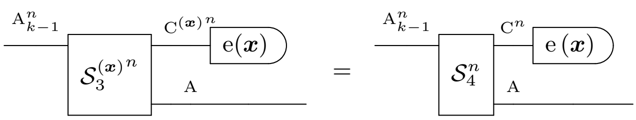
A condition that would guarantee that this does not occur is the triviality of $\mathrm{D}^{(\tilde{\boldsymbol{x}})}$. Indeed, the condition $\mathrm{D}^{(\tilde{\boldsymbol{x}})} = \mathrm{I}$ is necessary for (6.10) to hold, since it is a necessary condition for the transformation on the left-hand side of (6.9) not to be entanglement-breaking.

We observe that the condition $\mathrm{E}^{(\tilde{\boldsymbol{x}})} = \mathrm{I}$ would not solve the problem. Indeed, in this case (6.10) would still lead to an expression of the form



which is again problematic. This also highlights the importance of considering the presence of an ancillary system, a necessary condition to completely characterise the action of instruments and transformations in the OPT framework.

Then, since entangled states are spanning and $\tilde{\boldsymbol{x}}$ was chosen arbitrarily, it follows that $\mathrm{D}^{(\tilde{\boldsymbol{x}})} = \mathrm{I}$ for every $\tilde{\boldsymbol{x}} \in \mathbf{X}$. This implies that (6.7) must actually be of the form



$, \quad \forall \boldsymbol{x} \in \mathbf{X}, \qquad (6.11)$

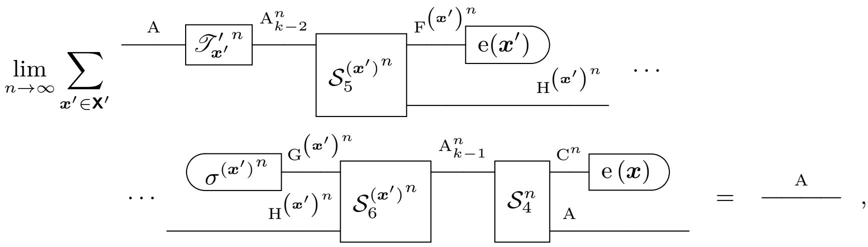
where

$$\mathcal{S}_3^{(\boldsymbol{x})\,n} := \left(\mathscr{I}_{\mathrm{C}^{(\boldsymbol{x})\,n}} \boxtimes \mathcal{S}_2^{(\boldsymbol{x})} \right) \,\Box\, \mathcal{S}_1^{(\boldsymbol{x})\,n}.$$

The equality between the left- and right-hand side of (6.11) follows from the fact that both A^n_{k-1} and A do not depend on $\boldsymbol{x}$, together with the observation that a permutation is completely specified by how it permutes the elementary subsystems composing the system on which it acts (Lemma 24). Therefore, the $\boldsymbol{x}$-dependent part of the action of $\mathcal{S}_3^{(\boldsymbol{x})\,n}$ reduces to a local permutation acting on $\mathrm{C}^{(\boldsymbol{x})\,n}$, which can be absorbed into the deterministic effect.

We observe that it may occur that the resolution of the limit of the left-hand side of (6.8), when applied to every atomic entangled state, yields the null state. Since entangled states form a spanning set—and in particular the atomic ones do, as the non-atomic states arise as their sums—this implies that the transformation under consideration must be the null one.

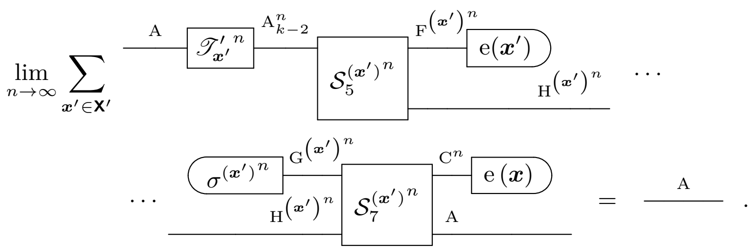
We can now proceed to make explicit the $(k-1)$-th conditioning step in (6.6). In particular, we express the coarse-graining of the instrument at position $k-1$ as a generic deterministic transformation (see Theorem 37), followed by the deterministic transformation at position k derived in (6.11). The resulting expression is



where $\boldsymbol{x}' = (x_1, \ldots, x_{k-2}) \in \mathbf{X}' = \mathsf{X}_1 \times \cdots \times \mathsf{X}_{k-2}$ denotes the outcomes of the first $k-2$ instruments. By grouping together the permutations $\mathcal{S}_6^{(\boldsymbol{x}')^n}$ and $\mathcal{S}_4^n$ as

$$\mathcal{S}_7^{(\boldsymbol{x}')^n} := \mathcal{S}_4^n \,\Box\, \mathcal{S}_6^{(\boldsymbol{x}')^n},$$

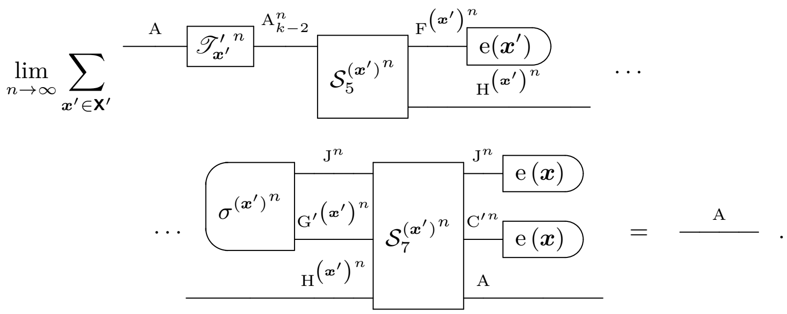
the equality can be rewritten as



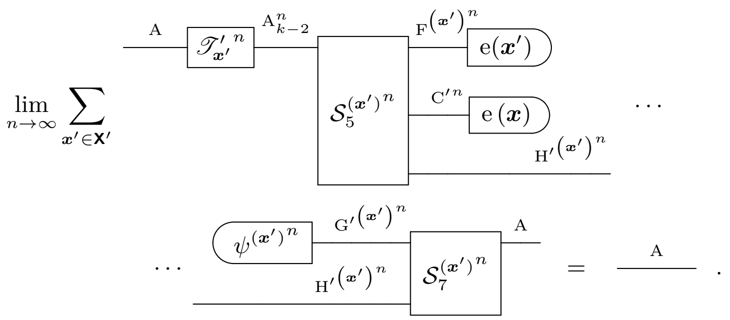
As in the case of $\mathrm{e}\,(\boldsymbol{x})$, the dependence on $\boldsymbol{x}'$ could be omitted from the deterministic effect $\mathrm{e}\,(\boldsymbol{x}')$. Nevertheless, we retain it in the notation, since the individual events of the corresponding observation-instrument generally depend on $\boldsymbol{x}'$.

We cannot now directly apply the same reasoning used to conclude the triviality of the system $\mathrm{D}^{(\boldsymbol{x})}$ to infer the triviality of the sequence of systems $\mathrm{G}^{(\boldsymbol{x}')^n}$, since $\mathrm{G}^{(\boldsymbol{x}')^n}$ and C^n may share a collection of elementary subsystems, whose composition we denote by J^n[1].

By factorising the deterministic effect over the composite system $J^n C'^n$, the previous expression can be rewritten as



Now, exploiting the naturality property of the braiding (2.19), we can slide the deterministic effects to the left, obtaining



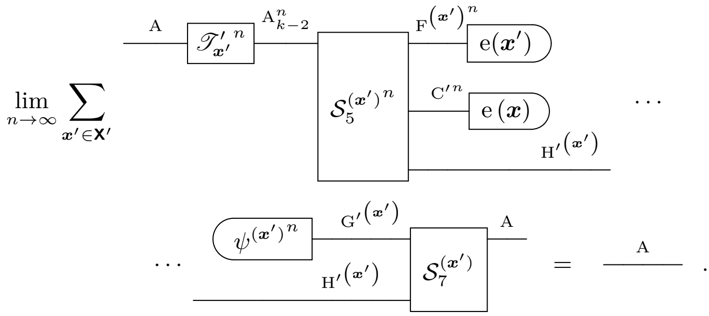
(6.12)

The deterministic effect acting on J^n passes through the permutation and acts on $\sigma^{(\boldsymbol{x}')^n}$, returning $\psi^{(\boldsymbol{x}')^n}$. By contrast, the deterministic effect acting on C'^n does not interact with the state—since we assumed that all systems shared between C^n and $G^{(\boldsymbol{x}')^n}$ are collected into J^n—and is therefore carried back up to the permutation $\mathcal{S}_5^{(\boldsymbol{x}')^n}$.

We can now find a subsequence where the systems $G'^{(\boldsymbol{x}')}$ and $H'^{(\boldsymbol{x}')}$, and the

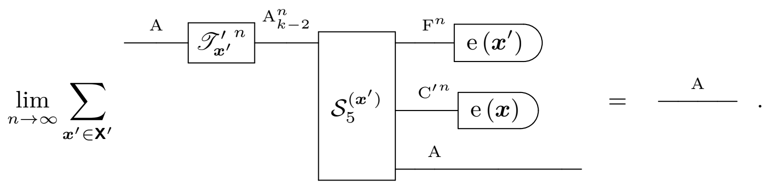
[1] Strictly speaking, J^n is defined only up to a local permutation. This is, however, irrelevant for the present discussion, as such a local permutation can always be absorbed into the deterministic effect.

permutation $\mathcal{S}_7^{(\boldsymbol{x}')}$ are fixed:



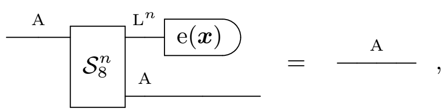
Then, we can apply the permutation $\mathcal{S}_7^{(\boldsymbol{x}')}$ [and its inverse] to the left-hand [right-hand] side of both circuits and repeat the argument used following (6.8). This shows that the sequence of systems $\mathrm{G}'^{(\boldsymbol{x}')\,n}$ must stabilise to the trivial system, otherwise the equality with the identity could not be satisfied.

Summarising all the conditions derived so far, the relation (6.6) reduces to



(6.13)

Iterating this reasoning for all the k conditioning steps, one finds that the condition in (6.6) reduces to



(6.14)

for some sequence of permutations $\mathcal{S}_8^n$ and systems L^n.

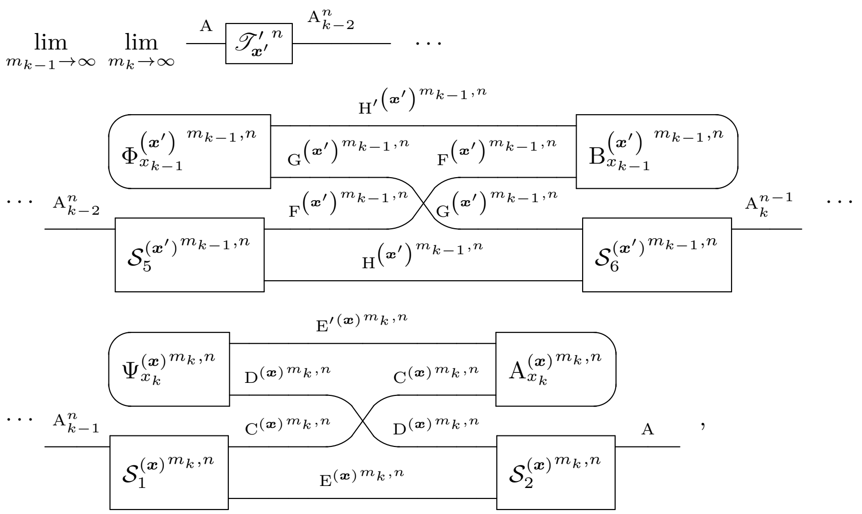
However, the left-hand side of the equality is well defined only if $\mathrm{L}^n = \mathrm{I}$ and $\mathcal{S}_8^n = \mathscr{I}_\mathrm{A}$ for all $n \in \mathbb{N}$. Indeed, since permutations act by reordering systems, it is impossible for a permutation to map between input and output systems that differ by more than a reordering of the elementary systems composing them.

Item II): Having derived a series of conditions on the systems and transformations appearing in a generic sequence of instruments whose coarse-graining converges to the identity by analysing the behaviour of the coarse-graining in the limit, we now trace these constraints back to the transformations composing the instruments of the sequence. The goal is to show that all such transformations must be proportional to the identity whenever (6.6) is satisfied. Explicitly, the strategy is to repeat the reasoning above, this time applied to the individual transformations composing the initial sequence (of the form (6.5)), rather than to the sequence of its full coarse-grainings.

Since we are now interested in the single transformations within the instruments of the sequence (6.5), in the following we fix an arbitrary outcome

$$\left(\boldsymbol{x}', x_{k-1}, x_k\right) \in \mathsf{X}' \times \mathsf{X}_{k-1} \times \mathsf{X}_k.$$

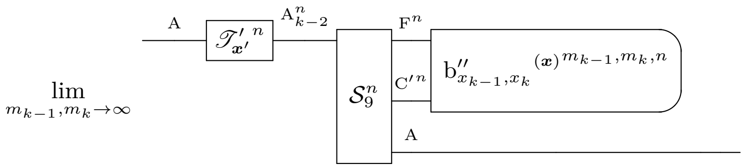
Recalling that each transformation appearing in the conditioning steps of (6.5) is the limit of transformations of the form (5.10), we can write them explicitly as



(6.15)

where we have made explicit only the last two transformations, namely the k-th and $(k-1)$-th ones, in the sequential composition.

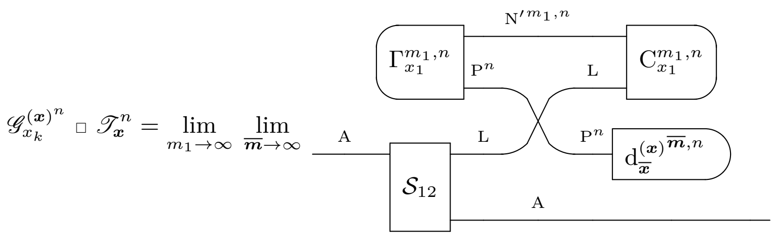
By following the same reasoning that led to (6.13) but now for transformations (see appendix F for details), one finds that (6.15) reduces to



, (F.2)

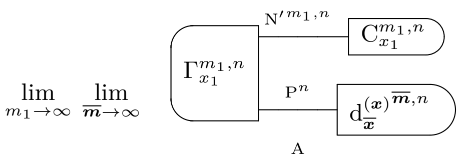
where a suitable observation-instruments appear in place of the deterministic effect in (6.13).

Further iterations lead, in analogy with (6.14), to the following expression for all $n \in \mathbb{N}$:



, (F.3)

where $\overline{\boldsymbol{m}} = (m_2, \dots, m_k)$ and $\overline{\boldsymbol{x}} = (x_2, \dots, x_k)$. Recalling the triviality of L, we conclude that any transformation in the instrument (6.5) must be of the form



, (6.16)

namely a probability composed with the identity on A.

Summarising, we started from a generic Cauchy sequence of instruments of an MSOPT, built from elementary processes, of the form (6.5), and showed that if the limit of its coarse-grainings is the identity, then the sequence must be of the form

$$\left\{ \left[\!\left[\lim_{m\to\infty} p^{m,n}_x \mathscr{I}_A \right]\!\right]_{x\in \mathsf{X}} \right\}_{n\in\mathbb{N}},$$

where $m = (m_1, \dots, m_k)$, $x = (x_1, \dots, x_k)$, and $\mathsf{X} = \mathsf{X}_1 \times \cdots \times \mathsf{X}_k$.

By Lemma 13, resolving the limit over m, we can conclude that in order for (6.5) to converge to the identity, it must be of the form

$$\left\{ [\![\mathrm{p}_x^n \mathscr{I}_{\mathrm{A}}]\!]_{x \in \mathsf{X}} \right\}_{n \in \mathbb{N}},$$

with $[\![\mathrm{p}_x^n]\!]_{x \in \mathsf{X}}$ a probability distribution for all $n \in \mathbb{N}$.

The same lemma also allows us to conclude that the limit instrument—obtained by taking the limit over n—must be of the form

$$\left[\!\!\left[\mathrm{p}'_x \mathscr{I}_{\mathrm{A}} \right]\!\!\right]_{x \in \mathsf{X}},$$

with $[\![\mathrm{p}'_x]\!]_{x \in \mathsf{X}}$ a probability distribution.

This concludes the proof, as we have shown that the only instruments that can decompose the identity are randomisations of it. □

Remark 40

We remark that Theorem 41 relies on the existence of entangled states. Indeed, entangled states are fundamental in order to characterise the implications of the equality (6.9), as done in (6.10).
The use of a separable state would not have proven effective, as, differently from the case of MOPTs, conditioning could allow one to implement the identity transformation on separable states. Any state given as input on $\mathrm{D}^{(\tilde{\boldsymbol{x}})}$ could potentially be conditionally prepared again through $\rho^{(\tilde{\boldsymbol{x}})^{(n)}}$ in (6.9).
If we lift the assumption of the existence of a spanning set of entangled states, counterexamples can be found. A striking one is provided by CT. In fact, completing MCT to an MSOPT reproduces CT in its entirety, whose identity transformation is not atomic (4.3).The reason why the strong completion of MCT leads to CT, together with its implications, is discussed in detail in section 6.5.
Therefore, entanglement proves to be a necessary condition for MSOPTs to exhibit an atomic identity transformation. However, whether this is also a sufficient condition remains an open question and is left for future research.

6.3.2 Irreversibility & more

As in the case of MOPTs, thanks to Theorem 41 we can conclude that MSOPTs satisfy a number of information-theoretic properties. Importantly, these properties persist even after Cauchy completion under strong causality.

First, we observe that MSOPTs satisfying the hypothesis of Theorem 41 do not admit reversible transformations other than permutations:

Corollary 12

In a MSOPT satisfying the hypothesis of Theorem 41, all reversible transformations are atomic and coincide with permutations [158].

The first part of the result follows from Lemma 26. While, the second part can be proved following the same argument as in Proposition 5; also in this case the proof of Theorem 41 shows that the only sequence of deterministic transformations that can converge to the identity is the constant one.

Proceeding further, from Theorem 41 one directly recovers NIWD:

Corollary 13

Every MSOPT satisfying the hypothesis of Theorem 41 has NIWD [158].

Irreversibility is also inherited, as an immediate consequence of Theorem 31:

Corollary 14

Every non-trivial MSOPT satisfying the hypothesis of Theorem 41 has irreversibility [158].

Finally, from Theorem 27 one obtains the no-broadcasting property:

Corollary 15

Every MSOPT satisfying the hypothesis of Theorem 41 does not admit a broadcasting channel [158].

The only property of minimal theories that has not yet been extended to the case with conditional instruments is the no-programming theorem (Definition 51). Establishing this result would require a complete characterisation of deterministic transformations, which is still lacking.

Remark 41

We conclude this part by analysing the relationship between MSOPTs and the no-restriction hypothesis. In this regard, we find counterexamples in both directions. A minimal strongly causal version of CT, denoted Minimal Strongly causal Classical Theory (MSCT), satisfies the no-restriction hypothesis because, as we will discuss in a moment, it coincides with CT. By contrast, MSBCT does not. Hence, restricting the set of instruments and then allowing for conditioning neither implies nor forbids the satisfaction of the no-restriction hypothesis.
This is in line with the observations on the no-restriction hypothesis made in Remark 37, where we discuss the relationship between this property and MOPTs. Given the high degree of arbitrariness with which the set of instruments characterising an OPT can be defined, making general statements about the conditions leading to the satisfaction of the no-restriction hypothesis is highly non-trivial. In fact, no general statement can be made about the relationship between a MSOPT and whether it satisfies the no-restriction hypothesis.

6.4 Minimal Strongly causal Bilocal Classical Theory (MSBCT)

We now have all the ingredients to formalise MSBCT [158]. This theory is the MSOPT counterpart of BCT, a locally classical theory with entangled states. MSBCT is a toy-theory that is locally classical and satisfies strong causality, yet it still exhibits irreversibility and does not admit a broadcasting channel.

6.4.1 Postulates of the theory

Formally, Minimal Strongly causal Bilocal Classical Theory (MSBCT) is the OPT Θ characterised by the following postulates [158].

The first postulate is on the nature of the systems of the theory.

Postulate 1 (Classicality, convexity and system types)

The theory Θ is classical and its state spaces are convex. In addition to the trivial system, for every integer $D \geq 1$, $\mathsf{Sys}(\Theta)$ contains a type of system of dimension D.

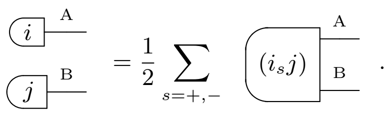
The second postulates establishes the parallel composition rule for systems and states.

Postulate 2 (Parallel composition of systems and states)

For every pair of systems A, B $\in \mathsf{Sys}(\Theta)$, the dimension of the composite system AB is given by the following rule:

$$D_{\mathrm{AB}} = D_{\mathrm{BA}} = \begin{cases} 2D_{\mathrm{A}}D_{\mathrm{B}}, & \textit{if } \mathrm{A}, \mathrm{B} \neq \mathrm{I}, \\ D_{\mathrm{A}}, & \textit{if } \mathrm{B} = \mathrm{I}. \end{cases}$$

Let $\mathrm{I} \neq \mathrm{A}, \mathrm{B}, \mathrm{C} \in \mathsf{Sys}(\Theta)$. Denoting the pure states of the composite system AB as $\mathsf{PurSt}(\mathrm{AB}) = \{(i_- j), (i_+ j) \mid 1 \leq i \leq D_{\mathrm{A}}, 1 \leq j \leq D_{\mathrm{B}}\}$, for all state $i \in \mathsf{PurSt}(\mathrm{A})$, $\mathrm{j} \in \mathsf{PurSt}(\mathrm{B})$ the following parallel composition rule holds:



The relation between states of the same composite system defined via compositions in different orders is given by

$$\left((i_{s_1} j)_{s_2} l\right) = \left(i_{s_1} \left(j_{s_1 s_2} l\right)\right),$$

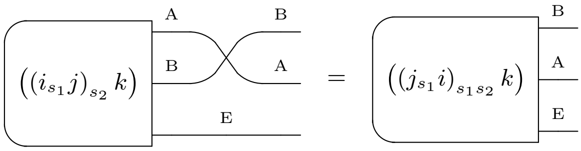
for all indices i, j, k and signs s_1, s_2.

The previous postulate also implies that MSBCT is a bilocal theory (Definition 15) by virtue of Theorem 13. We recall that satisfying the property of bilocal discriminability implies that in this theory two distinct states of a composite system cannot be discriminated by local measurements (even when assisted by classical communication): in theories satisfying bilocal discriminability also global measurements of pairs of subsystems are needed. In the case of MSBCT, this is evident if one considers the state $(i_s j) \in \mathsf{St}(\mathrm{AB})$ with $s = \pm$. Measuring locally the states $i \in \mathsf{St}(\mathrm{A})$ and $j \in \mathsf{St}(\mathrm{B})$ provides no information about the sign s. Such information can be recovered only through a measurement on the composite system AB.

We can then define the braiding transformation for MSBCT.

Postulate 3 (Swap transformation)

The theory is symmetric. Considering $\mathrm{I} \neq \mathrm{A}, \mathrm{B}, \mathrm{E} \in \mathsf{Sys}(\Theta)$, the swap $\mathscr{S}_{\mathrm{A,B}}$ is defined as follows



. (6.17)

And, in conclusion, its spaces of states and effects.

Postulate 4 (Preparation- and observation-instruments)

For any system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, a collection $[\![\rho_x]\!]_{x \in \mathsf{X}} \subset \mathsf{St}(\mathrm{A})$ is a preparation-instrument if and only if it is normalised, namely

$$\sum_{x \in \mathsf{X}} (\mathrm{e}|\ \rho_x)_{\mathrm{A}} = 1.$$

An observation-instrument on A is any collection $[\![\mathrm{a}_y]\!]_{y \in \mathsf{Y}} \subset \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ of generalised effects such that, for every $\mathrm{E} \in \mathsf{Sys}(\Theta)$, the set $\{\mathrm{a}_y \boxtimes \mathscr{I}_{\mathrm{E}}\}_{y \in \mathsf{Y}}$ maps preparation-instruments of AE to preparation-instruments of E.

The effect sets $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ for each A are determined by Postulate 1 through the property of *joint perfect discriminability* [159].

In other words, this postulate is stating that Θ satisfies the no-restriction hypothesis (Definition 22) for observation-instruments.

The postulates introduced so far are in common with those of BCT. The essential distinction between the two theories lies in their spaces of instruments and transformations.

Postulate 5 (Minimality and strong causality)

The theory Θ is minimal and strongly causal, in the sense of Definition 53.

We can then summarise everything in the following definition.

Definition 54 (Minimal Strongly causal Bilocal Classical Theory (MSBCT))

Minimal Strongly causal Bilocal Classical Theory (MSBCT) [158] is the OPT Θ characterised by the following postulates:

I) Classicality, convexity, and system types (Postulate 1).

II) Parallel composition of systems and states (Postulate 2).

III) Swap transformation (Postulate 3).

IV) Preparation- and observation-instruments (Postulate 4).

V) Minimality and strong causality (Postulate 5).

In other words, MSBCT is the MSOPT obtained from BCT [167]. The spaces of states and effects are the same in both theories; the essential difference lies in the spaces of transformations with non-trivial input and output systems.

Remark 42 (Classicality and non-contextuality for MSBCT)

Unlike the case of MCT, one cannot conclude that MSBCT is Kochen-Specker non-contextual and generalised-non-contextual from Theorem 25, since the theory satisfies bilocal discriminability. Nevertheless, it has been recently proven that BCT admits a generalised-non-contextual ontological model [222]. Hence, MSBCT does as well, being a sub-theory of BCT. Its generalised-non-contextual ontological model is simply obtained by restricting that of BCT to the set of instruments and transformations that characterise MSBCT.

Remark 43

In line with Remark 34, Definition 54 does not specify a single OPT, but rather a family of theories. To single out one theory from this family, it is necessary to specify the set of elementary systems. There are two natural options:

- One may adopt the most conservative choice, postulating that for every dimension there exists a corresponding elementary system. This allows for the possibility of high-dimensional systems that are not composite. For instance, in the case of MSBCT, there is no reason to exclude that an 8-dimensional system might be elementary rather than the composition of two 2-dimensional ones.

- Alternatively, one may adopt a more minimalist approach, according to which high-dimensional systems are always regarded as composites of lower-dimensional ones. In particular, for MSBCT, this amounts to declaring that the elementary systems are precisely those of odd dimension. Indeed, from these, the compositional rule of Postulate 2 suffices to reconstruct systems of arbitrary dimension. The presence of a 1-dimensional system not operationally equivalent to the trivial one guarantees the reconstruction of all even-dimensional systems. The fact that this assumption is minimal then follows from the fact that the only way to obtain an odd number as a product is for all factors to be odd.

In what follows we will speak of MSBCT as if it were a single theory, since the properties we establish hold irrespective of the particular choice of the set of elementary systems.

6.4.2 Properties of MSBCT

Having introduced MSBCT, we now turn to the main properties that it satisfies in addition to those already established for all MSOPTs.

First of all, the theory is *locally classical*. By this we mean that, when restricted to single systems or to single-system operations on composite systems—that is, to compound-local instruments—the theory is indistinguishable from CT. The differences emerge only when comparing local measurements on the subsystems of a composite system with global measurements on the same composite system, the latter granting access to additional information—specifically, the sign charactering the pure states of composite systems.

This leads to a peculiarity. Although MSBCT does not admit broadcasting channels Corollary 15, it nevertheless contains an instrument that broadcasts single-system states, namely the classical broadcasting channel of CT (4.9). The reason why this map fails to broadcast generic states is that it does not preserve correlations in the presence of an environment (4.8). Indeed, information encoded in entangled states cannot be faithfully broadcast by (4.9). For example, employing this channel one would lose the information corresponding to the sign s of the pure states $(i_s j)$.

Implicit in the discussion of the absence of broadcasting channels in MSBCT is the fact that the theory satisfies the hypothesis of Theorem 41. Indeed, the identity transformation of MSBCT is atomic. Consequently, MSBCT is an OPT that satisfies NIWD (Corollary 13) and exhibits irreversibility (Corollary 14).

Lemma 31

MSBCT does not satisfy full-compatibility and features irreversibility.

Thus, MSBCT provides then an evidence that strong causality and classicality do not prevent no-broadcasting and NIWD. Even in presence of strong causality, NIWD and no-broadcasting cannot be taken *per se* as signatures of non-classicality.

However, since MSBCT shares the same set of observation-instruments as BCT, it inherits observation-compatibility from the latter. This can be verified directly by exploiting the characterisation of observation-instruments in Postulate 4. The proof presented in appendix D applies unchanged also in this case.

Lemma 32

MSBCT satisfies observation-compatibility.

Through the last two results we have established that MSBCT provides a counterexample to the converse implication of Theorem 30. In particular, we see that even the joint presence of strong causality and classicality does not forbid the existence of "incompatible" dynamics (irreversibility), despite the fact that all observation-tests of MSBCT are compatible.

As a final remark, we observe that in MSBCT no mixed state admits a purification—hence, MSBCT does not satisfy *purification* (Definition 30)—, and the operational superposition principle [130, 133] is not satisfied. These features follow from the no-go results established in Ref. [159] for general simplicial theories that are not locally tomographic.

Furthermore, the characterisation of reversible transformations also shows that MSBCT fails to satisfy the property of *essential uniqueness of purification* (Definition 31). Indeed, no permutation in $\mathsf{RevTransf}(\mathrm{B} \to \mathrm{B})$ can map a pure state of the form $(i_s j) \in \mathsf{PurSt}(\mathrm{AB})$ into one of the form $(i'_{s'} j') \in \mathsf{PurSt}(\mathrm{AB})$ whenever, for example, $i \neq i'$. This follows directly from the definition of the swap operation in these theories (Postulate 3), which cannot modify the state of a system on which it does not act directly.

In conclusion of this section, we note that the atomicity of the identity transformation for every system in MSBCT sets it apart from BCT. In the latter theory, there always exists a non-trivial instrument that decomposes the identity [159]. Consequently, the transformation spaces of MSBCT are *strictly* contained within those of BCT:

$$\mathsf{Instr}(MSBCT) \subset \mathsf{Instr}(BCT)\,.$$

This result also follows from Corollary 12, since BCT admits reversible transformations other than permutations.

6.4.3 Minimal Bilocal Classical Theory (MBCT)

To conclude this chapter, we note that in the process of constructing MSBCT, we have also implicitly defined another OPT.

Definition 55 (Minimal Bilocal Classical Theory (MBCT))

Minimal Bilocal Classical Theory (MBCT) [158] is the OPT Θ characterised by the following postulates:

I) Classicality, convexity, and system types (Postulate 1);

II) Parallel composition of systems and states (Postulate 2);

III) Swap transformation (Postulate 3);

IV) Preparation- and observation-instruments (Postulate 4);

V) Minimality (Definition 48).

In other words, MBCT is the minimal version of BCT [167].

MBCT satisfies all the properties that we have discussed so far in connection with MSBCT. However, the fact that these properties also hold in a minimal classical theory *without* strong causality is not particularly noteworthy, since they had already been established for MCT, even in the absence of bilocal discriminability.

Let us recall that, due to the presence of conditional instruments, the set of instruments of MBCT is strictly contained in that of MSBCT (6.1). Hence, these two theories are indeed distinct OPTs.

Summarising, we have shown that it is possible to construct three different theories, characterised by three distinct sets of instruments, each strictly contained in the next:

$$\mathsf{Instr}(MBCT) \subset \mathsf{Instr}(MSBCT) \subset \mathsf{Instr}(BCT)\,. \tag{6.18}$$

6.5 Relationship between an OPT and the corresponding MSOPT

At the end of the previous section, after introducing both the minimal and the minimal strongly causal versions of BCT—namely MBCT and MSBCT—we have shown that these three theories are distinct, with their sets of instruments related by

$$\mathsf{Instr}(MBCT) \subset \mathsf{Instr}(MSBCT) \subset \mathsf{Instr}(BCT)\,. \tag{6.18}$$

In this section we investigate more closely the general relationship between a theory and its minimal strongly causal counterpart.

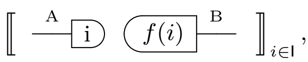
By construction, an MOPT and the corresponding MSOPT are always distinct, since the latter necessarily includes conditional instruments, and therefore

$$\mathsf{Instr}(MOPT) \subset \mathsf{Instr}(MSOPT)\,. \tag{6.1}$$

However, the same does not always hold when considering an arbitrary OPT and its minimal strongly causal version.

As already anticipated, CT provides a counterexample.

The reason why this happens is that the addition of all conditional instruments to MCT also includes, in particular, all instruments of the form

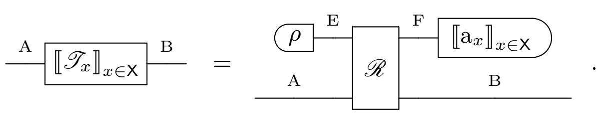


(6.19)

for an arbitrary function $f : \mathsf{I} \to \mathsf{J}$, where I and J denote the sets of pure states of A and B, respectively. By coarse-graining these instruments one can recover any reversible transformation of CT. The conclusion then follows from the fact that every instrument of CT admits a reversible dilation [133, 159, 167].

Definition 56 (Reversible dilation)

We say that an instrument $[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{B})$ admits a *reversible dilation* if there exist two ancillary systems E and F, a reversible transformation $\mathscr{R} \in \mathsf{RevTransf}(\mathrm{EA} \to \mathrm{FB})$, a deterministic state $\rho \in \mathsf{St}_1(\mathrm{E})$, and an observation-instrument $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$ such that



Analogously, for a single transformation one considers an effect instead of an observation-instrument.

Summarising, in the case of CT, it holds that

$$\mathsf{Instr}(MCT) \subset \mathsf{Instr}(MSCT) \equiv \mathsf{Instr}(CT)\,, \tag{6.20}$$

where Minimal Strongly causal Classical Theory (MSCT) is the MSOPT obtained from CT.

Remark 44

An interesting consequence of the relationship just established is that, once a set of elementary systems is fixed for CT, the whole classical world and its dynamics can be fully described in terms of measure-and-reprepare instruments (6.19).

An analogous statement to Remark 44 does not hold in the case of QT.

Indeed, since the minimal strongly causal version of QT, denoted Minimal Strongly causal Quantum Theory (MSQT), satisfies the hypothesis of Theorem 41, it also fulfils Corollary 12. However, QT possesses a rich set of reversible operations, namely all unitary evolutions, which can describes more general and interesting dynamics that the mere permutation of the systems on which the unitarity is acting. Therefore, MSQT is a genuinely different theory from QT.

Summarising, in the case of QT we have

$$\mathsf{Instr}(MQT) \subset \mathsf{Instr}(MSQT) \subset \mathsf{Instr}(QT)\,, \tag{6.21}$$

where MQT denotes the minimal version of QT.

An interesting question that naturally arises concerns the relationship between our result and the universality of *measurement-based quantum computation* [257–263]. A more in depth discussion of this matter is carried out in the conclusions. Nevertheless, at present we have only a limited intuition on how these aspects are related, and we therefore leave this question open for future research.

Remark 45

When comparing inclusions such as (6.18) for BCT, (6.20) for CT, or (6.21) for QT, it is important to stress that the three theories must be defined over the same set of systems. In particular, they must share the same choice of elementary systems.
What differences may arise between formulations of an OPT defined by

different sets of elementary systems, or between such formulations and the traditional one without this specification, remains an open question for future research.

6.6 Local discriminability, strong causality and simpliciality

The theories developed in this work also allow us to establish the logical independence of three central properties of OPTs: *local discriminability*, *strong causality*, and *simpliciality*.

First, none of these properties implies either of the others. Theories can be exhibited that enjoy exactly one out of the three: FQT is strongly causal but neither simplicial nor locally discriminable; MBCT is simplicial but neither locally discriminable nor strongly causal; finally, MQT is locally discriminable but neither simplicial nor strongly causal.

Furthermore, even any two of them together do not imply the third. QT is both strongly causal and locally discriminable, yet non-simplicial; MCT is locally discriminable and simplicial, yet not strongly causal; finally, MSBCT is strongly causal and simplicial, yet not locally discriminable.

Taken together, these examples show that the three notions are mutually independent structural features of OPTs. A diagrammatic summary of these relationships is provided in figure 8.3.

7 Complementarity

AFTER having analysed and set apart the different physical phenomena that are captured by Heisenberg's *Gedankenexperiment*, we now turn to the notion of *complementarity*. This was introduced by Bohr as a principle from which Heisenberg's irreversibility arises, but in the more recent literature on the operational approach to quantum foundations it has often been taken to be equivalent to incompatibility.

In this chapter we aim to straighten that path, proposing a new definition of complementarity that is distinct from compatibility and more closely aligned with Bohr's original idea of the principle. We then proceed to compare complementarity with compatibility and irreversibility, showing that they are indeed distinct notions, with complementarity being the strongest. It remains true, however, that in the specific case of QT, complementarity and incompatibility coincide.

7.1 Operational definition

Bohr first introduced his idea of complementarity in the lecture he delivered in Como in 1927 [100]. His central claim is that, contrary to the classical case, when dealing with quantum objects it is not possible to adopt a single point of view from which the system can be fully described. Instead, several complementary descriptions, each capturing a different aspect of the system, are necessary. For example, he expressed this idea in relation to the dual nature of light:

> The two views of the nature of light are rather to be considered as different attempts at an interpretation of experimental evidence in which the limitation of the classical concepts is expressed in complementary ways. [100]

Similarly, referring to the wave-particle duality of material particles, he stated:

> [...] here again we are not dealing with contradictory but with complementary pictures of the phenomena, which only together offer a natural generalisation of the classical mode of description. [100]

Finally, in connection with Heisenberg's relation $\Delta x \Delta p_x \approx h$, he observed:

> [...] he has stressed the peculiar reciprocal uncertainty which affects all measurements of atomic quantities. Before we enter upon his results it will be advantageous to show how the complementary nature of the description appearing in this uncertainty is unavoidable already in an analysis of the most elementary concepts employed in interpreting experience. [100]

As these examples show—and several others in the same lecture confirm—Bohr's notion of complementarity is not restricted to the wave-particle duality of light, but extends to all quantum phenomena.

So, how can one capture this idea and translate it into an operational language?

Our definition relies on the notion of *property* of a system, meant to capture a specific aspect of the system under consideration. Two properties are said to be complementary if knowing one does not guarantee any knowledge of the other. In particular, if a system is measured to be in a definite state with respect to one property, then we obtain no information about its state with respect to a complementary property. In this sense, each property captures only one facet of the system under study—exactly as in Bohr's original proposal [119].

To formally define the notion of property, we first introduce some preliminary definitions.

Definition 57 (Repeatable instruments)

Let $\mathsf{T}_\mathsf{X} = [\![\mathscr{T}_x]\!]_{x \in \mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$. We call T_X *repeatable* if

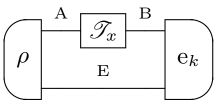
$$\mathscr{T}_x \mathscr{T}_{x'} = \delta_{x,x'} \mathscr{T}_x \quad \forall x, x' \in \mathsf{X}.$$

Definition 58 (Atomic instruments)

Let $\mathsf{T}_{\mathsf{X}} = [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{A})$. We say that T_{X} is an *atomic instrument* if $\mathscr{T}_x$ is atomic for all $x \in \mathsf{X}$.

Definition 59 (Verifier state)

Consider a generic OPT and let $\mathsf{T}_{\mathsf{X}} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ be an instrument of the theory. We say that a state $\rho \in \mathsf{St}_1(\mathrm{AE})$, where E is a generic system of the theory, is a *verifier state* for the instrument, denoted $\rho \in \Upsilon(\mathsf{T}_{\mathsf{X}})$, if there exists $x \in \mathsf{X}$ such that



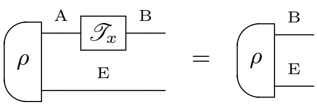
$= 1$ (7.1)

for all deterministic effects $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{BE})$.
Since the outcome $x \in \mathsf{X}$ satisfying (7.1) is unique, we also say that ρ is a verifier state for the particular transformation $\mathscr{T}_x$, denoted $\rho \in \Upsilon(\mathscr{T}_x)$.

One can also introduce a stronger notion of verifier state.

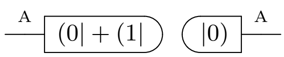
Definition 60 (Strong verifier state)

Given an instrument $\mathsf{T}_{\mathsf{X}} \equiv [\![\mathscr{T}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$, a state $\rho \in \mathsf{St}_1(\mathrm{AE})$, with E a generic system of the theory, is called a *strong verifier* if there exists $x \in \mathsf{X}$ such that



. (7.2)

It can be immediately checked that if a state satisfies (7.2), then it also satisfies (7.1). The converse is in general not true. Consider, for example, the states $|0)$ and $(|0) + |1))/2$ for the transformation



.

While both states are verifiers for this transformation, only the former is a strong verifier, since the latter does not reproduce itself at the output.

We can now use these definitions to formally introduce the notion of *elementary property*.

Definition 61 (Elementary property)

Let $\mathsf{P}_\mathsf{X} \equiv [\![\mathscr{P}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{A})$ be a repeatable atomic instrument such that every transformation composing it admits a verifier state, i.e. $\Upsilon(\mathscr{P}_x) \neq \emptyset$ for all $x \in \mathsf{X}$. We call an *elementary property (of system* A*)* any such instrument P_X [119].

The notion of elementary property encodes the idea that a system—and here we emphasise that we are concerned with systems themselves, not with the particular states they may be prepared in—has some aspects that can be precisely defined. Our focus is on fundamental properties of systems, namely those that do not descend from other properties. This is why we require the transformations composing a property to be atomic. For example, if one considers a die, an elementary property would be the instrument that verifies which specific face is facing up, with its precise value. We are not interested in coarse-grainings, such as verifying only whether the outcome is even or odd. Similarly, in QT, an elementary property would be an instrument that checks whether a qubit is in one of two orthogonal pure states.

In principle, one could also define a more general notion of *property*, as instruments that are coarse-grainings of elementary properties. However, this could give rise to situations in which properties are assigned even to noisy systems, because coarse-graininied operations would lack sufficient sensitivity to detect the noise.

Another crucial aspect of the definition is repeatability. This reflects the *desideratum* that multiple agents should be able to check whether a system has a certain property, passing the system between them. For this reason, we require that the system is not destroyed by the operation. A property is something that can be verified repeatedly, not merely once.

Finally, we introduce the notion of *complementarity* for elementary properties.

Definition 62 (Complementary elementary properties)

Let P and Q be two elementary properties. They are *complementary* if there exists $\rho \in \Upsilon(\mathsf{P}) \cup \Upsilon(\mathsf{Q})$ such that $\rho \notin \Upsilon(\mathsf{P}) \cap \Upsilon(\mathsf{Q})$.
In other words, two elementary properties are complementary whenever there exists a verifier state of one that is not a verifier state of the other. If an OPT admits complementary elementary properties, we say that it has *complementarity* [119].

Following the idea introduced at the beginning of this section, *complementarity* is defined as the existence of systems for which multiple points of view are necessary in order to fully characterise them. If two properties do not share a verifier state, this means that the information encoded in such a state about the system does not provide any further information about some other aspect of it.

That *elementary properties*—rather than general properties—are the relevant quantities of interest becomes clearer when analysing situations in which properties are complementary. Consider, for example, the following two non-elementary properties for a qutrit in QT:

$$\{\Pi_0 \cdot \Pi_0 + \Pi_1 \cdot \Pi_1,\ \Pi_2 \cdot \Pi_2\},$$
$$\{\Pi_+ \cdot \Pi_+ + \Pi_- \cdot \Pi_-,\ \Pi_2 \cdot \Pi_2\},$$

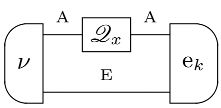
where $\Pi_i = |i\rangle\langle i|$, $\{|i\rangle\}_{i=0,1,2}$ is the canonical basis, and $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. It is immediate to see that these two properties share all verifier states. However, we are not interested in such cases: complementarity must be unrelated to the noise introduced in the measurement by ignoring features through coarse-graining transformations.

7.2 Different degrees of complementarity

When considering two complementary properties, the extent to which their verifier states fail to overlap can be used to quantify different degrees of complementarity: the more the definition of one property undermines knowledge of the other, the stronger their complementarity. Entropy provides a natural tool to capture this loss of information.

Definition 63 (Degrees of complementarity - State-wise)

Let P and $\mathsf{Q} \equiv [\![\mathscr{Q}_x]\!]_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ be two complementary elementary properties, and let $\nu \in \Upsilon(\mathsf{P})$ be a verifying state of P. For every $\mathrm{e}_k \in \mathsf{Eff}_1(\mathrm{AE})$, define the conditional probabilities



$$:= \mathrm{p}_\nu^{\mathsf{P},\mathsf{Q}}(x \,|\, \mathrm{e}_k)\,,$$

and the corresponding Shannon entropy

$$H_\nu^{\mathsf{P},\mathsf{Q}}(\mathrm{e}_k) := -\sum_{x\in\mathsf{X}} \mathrm{p}_\nu(x \,|\, \mathrm{e}_k) \log(\mathrm{p}_\nu(x \,|\, \mathrm{e}_k))\,. \tag{7.3}$$

We say that P and Q are complementary to the following degrees [119]:

I) **Strong:**

$$\min_{e_k \in \mathsf{Eff}_1(\mathrm{AE})} H_\nu^{\mathsf{P},\mathsf{Q}}(e_k) = \log(|\mathsf{X}|) .$$

In words, the second property is completely undermined: every outcome occurs with equal probability, independently of the effect chosen.

II) **Mild:**

$$\max_{e_k \in \mathsf{Eff}_1(\mathrm{AE})} H_\nu^{\mathsf{P},\mathsf{Q}}(e_k) \leq \log(|\mathsf{X}|) ,$$

with $p_\nu^{\mathsf{P},\mathsf{Q}}(x \mid e_k) \in (0,1)$ for all $x \in \mathsf{X}$ and for all e_k. Here one has some limited information, but every outcome remains possible.

III) **Weak:**

$$\max_{e_k \in \mathsf{Eff}_1(\mathrm{AE})} H_\nu^{\mathsf{P},\mathsf{Q}}(e_k) \leq \log(|\mathsf{X}|) ,$$

with $p_\nu^{\mathsf{P},\mathsf{Q}}(x \mid e_k) \in [0,1)$ for all $x \in \mathsf{X}$ and for all e_k. In this case, some outcomes are excluded entirely, hence one has partial knowledge about the second property.

We stress that the distinction between mild and weak complementarity lies in whether all outcomes remain possible, or some are strictly forbidden.

Furthermore, using the definition of $H_\nu^{\mathsf{P},\mathsf{Q}}(e_k)$ introduced in Definition 63, one immediately obtains the following result.

Lemma 33

Let P and $\mathsf{Q} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ be two elementary properties, and let $\nu \in \Upsilon(\mathsf{P})$. Then ν is a verifier for Q if and only if

$$H_\nu^{\mathsf{P},\mathsf{Q}}(e_k) = 0 \quad \forall e_k \in \mathsf{Eff}_1(\mathrm{AE}) ,$$

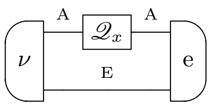
where the entropy $H_\nu^{\mathsf{P},\mathsf{Q}}(e_k)$ is defined in (7.3) [119].

The relations characterising the different degrees of complementarity presented in Definition 63 take a particularly simple form in the case of causal OPTs, where the deterministic effect is unique.

Lemma 34 (Degrees of complementarity - State-wise, causal case)

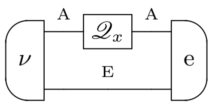
Consider a causal OPT and let P and $\mathsf{Q} \equiv [\![\mathscr{Q}_x]\!]_{x \in \mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ be two complementary elementary properties. Let $\nu \in \Upsilon(\mathsf{P})$, with $\nu \notin \Upsilon(\mathsf{Q})$. Then, with respect to the unique deterministic effect $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{AE})$, the state ν generates complementarity between the two properties in one of the following three degrees [119]:

I) **Strong:**



$$= \frac{1}{|\mathsf{X}|} \quad \forall x \in \mathsf{X},$$

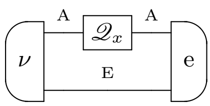
II) **Mild:**



$$\in (0,1) \quad \forall x \in \mathsf{X},$$

but not equal to $\frac{1}{|\mathsf{X}|}$ for all $x \in \mathsf{X}$,

III) **Weak:**



$$\in [0,1) \quad \forall x \in \mathsf{X},$$

but not equal to $\frac{1}{|\mathsf{X}|}$ for all $x \in \mathsf{X}$.

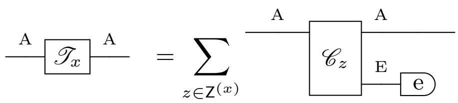
7.3 Complementarity implies incompatibility

We now turn to the first of the two main results established in this chapter: in causal OPTs, *complementarity implies incompatibility*. This statement follows as a corollary of the following theorem.

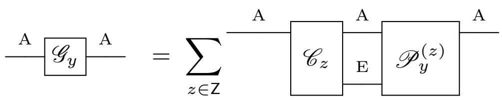
Theorem 42

Let Θ be a causal OPT. Consider two instruments $\mathsf{T}_\mathsf{X} \equiv [\![\mathscr{T}_x]\!]_{x \in \mathsf{X}}$ and $\mathsf{G}_\mathsf{Y} \equiv [\![\mathscr{G}_y]\!]_{y \in \mathsf{Y}}$ in $\mathsf{Instr}(\mathrm{A} \to \mathrm{A})$, with T_X repeatable and G_Y both atomic and such that $\mathscr{G}_y \neq \varepsilon_{A\to A}$ for all outcomes $y \in \mathsf{Y}$. If T_X does not exclude G_Y ($\mathsf{T}_\mathsf{X} \to \mathsf{G}_\mathsf{Y}$), then the set of verifier states of G_Y is included in that of T_X, namely $\Upsilon(\mathsf{G}_\mathsf{Y}) \subseteq \Upsilon(\mathsf{T}_\mathsf{X})$ [119].

Proof. Let us start by explicitly recalling the instruments that characterise the non-exclusion relationship $\mathsf{T}_\mathsf{X} \to \mathsf{G}_\mathsf{Y}$ (4.12):

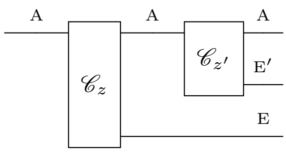


$\forall x \in \mathsf{X}$,



$\forall y \in \mathsf{Y}$.

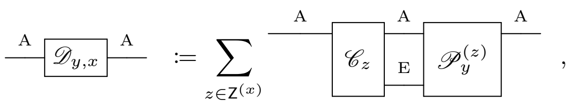
Since T_X is repeatable, it holds that $\mathscr{T}_x \mathscr{T}_{x'} = \mathscr{T}_{x'} \mathscr{T}_x = \varepsilon_{A\to A}$ whenever $x \neq x'$, which implies that



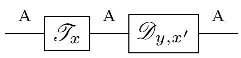
$= \varepsilon_{A\to A}$, (7.4)

whenever $z \in \mathsf{Z}^{(x)}$, $z' \in \mathsf{Z}^{(x')}$ with $x \neq x'$, as can be checked by direct calculation.

Let us now then define

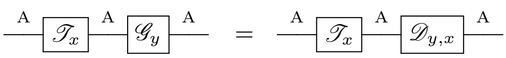


for all $(x, y) \in \mathsf{X} \times \mathsf{Y}$. From (7.4) one immediately obtains



$= \varepsilon_{A\to A}$, (7.5)

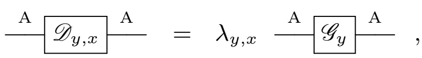
whenever $x \neq x'$ for all $y \in \mathsf{Y}$. Consequently



(7.6)

since $\mathscr{G}_y = \sum_{x\in\mathsf{X}} \mathscr{D}_{y,x}$,

Next, because G_Y is atomic, each $\mathscr{D}_{y,x}$ must be proportional to $\mathscr{G}_y$:



for some $\lambda_{y,x} \in [0,1]$. Substituting into (7.6), we obtain

$$\xrightarrow{\mathrm{A}}\boxed{\mathscr{T}_x}\xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \quad = \quad \lambda_{y,x} \xrightarrow{\mathrm{A}}\boxed{\mathscr{T}_x}\xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \quad . \tag{7.7}$$

This is possible only if $\lambda_{y,x} \in \{0,1\}$. Moreover, for each $y \in \mathsf{Y}$, there exists a unique $x \in \mathsf{X}$ with $\lambda_{y,x} = 1$.

To establish both existence and uniqueness, we proceed by contradiction.

Existence. Suppose that for some $y \in \mathsf{Y}$ one has $\lambda_{y,x} = 0$ for all $x \in \mathsf{X}$. Then

$$\begin{aligned} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \quad &= \sum_{x\in\mathsf{X}} \xrightarrow{\mathrm{A}}\boxed{\mathscr{D}_{y,x}}\xrightarrow{\mathrm{A}} \\ &= \sum_{x\in\mathsf{X}} \lambda_{y,x} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \\ &= \varepsilon_{\mathrm{A}\to\mathrm{A}}, \end{aligned}$$

which contradicts the assumption that $\mathscr{G}_y \neq \varepsilon_{\mathrm{A}\to\mathrm{A}}$. Hence, for every y there exists at least one x with $\lambda_{y,x} = 1$.

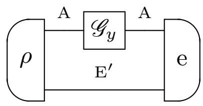
Uniqueness. Suppose instead that for some $y \in \mathsf{Y}$ there exists a subset $X' \subseteq X$ with $\lambda_{y,x} = 1$ for all $x \in \mathsf{X}'$ and $|X'| > 1$. Then

$$\begin{aligned} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \quad &= \sum_{x\in\mathsf{X}} \xrightarrow{\mathrm{A}}\boxed{\mathscr{D}_{y,x}}\xrightarrow{\mathrm{A}} \\ &= \sum_{x\in\mathsf{X}} \lambda_{y,x} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \\ &= \sum_{x\in\mathsf{X}'} \lambda_{y,x} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \\ &= \sum_{x\in\mathsf{X}'} \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \\ &= |\mathsf{X}'| \xrightarrow{\mathrm{A}}\boxed{\mathscr{G}_y}\xrightarrow{\mathrm{A}} \quad , \end{aligned}$$

which is possible only if $|X'| = 1$. Therefore, the outcome x such that $\lambda_{y,x} = 1$ is unique.

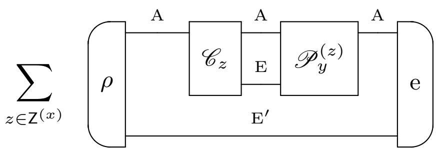
To conclude the proof, let $\rho \in \mathsf{St}_1(\mathrm{AE}')$ be a verifier state for G_Y, where E' is a generic system of the theory. By definition, there exists $y \in \mathsf{Y}$ such that

$\rho \in \Upsilon(\mathscr{G}_y)$, that is,



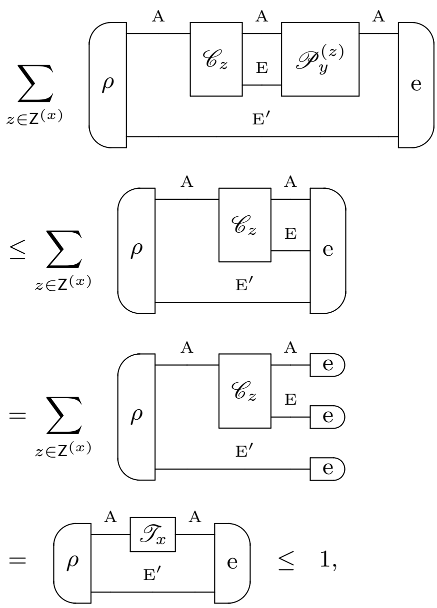
$= 1.$

Since for each y there exists a unique x with $\mathscr{G}_y = \mathscr{D}_{y,x}$, the above condition can be rewritten as



$= 1.$

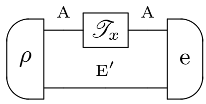
Now, since in general $\mathscr{P}_y^{(z)}$ is not deterministic, we must have



where the last inequality follows from the fact that $\mathscr{T}_x$ is not, in general, deterministic.

Combining this with the condition that the probability of verifying $\mathscr{G}_y$ is exactly

1, we deduce



$= 1.$

Hence $\rho \in \Upsilon(\mathscr{T}_x) \subseteq \Upsilon(\mathsf{T}_\mathsf{X})$. Since ρ was arbitrary in $\Upsilon(\mathsf{G}_\mathsf{Y})$, it follows that

$$\Upsilon(\mathsf{G}_\mathsf{Y}) \subseteq \Upsilon(\mathsf{T}_\mathsf{X}).$$

□

Corollary 16 (Complementarity implies incompatibility)

In causal OPTs, complementary elementary properties are strongly incompatible. Equivalently, weakly compatible elementary properties are non-complementary [119].

Proof. Let $\mathsf{P}_\mathsf{X}, \mathsf{Q}_\mathsf{Y} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$ be two elementary properties. If they are weakly compatible ($\mathsf{P}_\mathsf{X} \leftrightarrow \mathsf{Q}_\mathsf{Y}$), then by Theorem 42 one has

$$\Upsilon(\mathsf{P}_\mathsf{X}) = \Upsilon(\mathsf{Q}_\mathsf{Y}).$$

This condition means precisely that the two properties are non-complementary. Hence, complementary elementary properties cannot be weakly compatible, i.e. they are strongly incompatible. □

Corollary 16 allows us to conclude our analysis of the relationship among the different physical properties under consideration. As Bohr had already intuited, *complementarity lies at the core*: the very limitations in our ability to provide a comprehensive description of a physical system imply the existence of operations that are not compatible, and therefore, by Theorem 29, the existence of operations that irreversibly disturb the system on which they act.

What about the converse of Corollary 16? Technically, a counterexample exists. Indeed, MCT features strongly incompatible instruments, yet it does not exhibit complementarity, since the only elementary property it admits is the identity.

The more intriguing question is whether a theory admitting more than a single property, none of which are complementary, can still feature strongly incompatible instruments. This problem remains open: we have neither found a counterexample nor proved that the existence of properties forces strong incompatibility to imply complementarity. We have the intuition that the characterisation of the elementary properties of MSBCT could provide further insight. The peculiarity of this theory might indeed allow for elementary properties beyond the identity. However, such a characterisation has not yet been carried out.

We also highlight that the relationship between complementarity and observation-incompatibility, as well as that with simpliciality, is currently unknown. In particular, the latter would be of great interest, since it could clarify the role of complementarity as a useful property to distinguish between the quantum and the classical worlds.

7.4 The case of quantum theory

In the particular case of QT—and, by restriction, of CT—one can show that complementarity and incompatibility are in fact equivalent notions.

The key to proving this result lies in the possibility of explicitly characterising the properties of QT.

Let A and B be two quantum systems, with associated Hilbert spaces $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$. Consider then a quantum instrument $\mathsf{T}_\mathsf{X} = \{\mathscr{T}_x\}_{x\in\mathsf{X}}$, where each $\mathscr{T}_x$ is a quantum operation. We recall that quantum operations are CPTNI maps.

Every quantum operation admits a *Kraus decomposition*, namely a family of operators $\{T_y\}_{y\in\mathsf{Y}^{(x)}}$ such that

$$\mathscr{T}_x(\rho) = \sum_{y\in\mathsf{Y}^{(x)}} T_y \rho T_y^\dagger,$$

for any state ρ, where the set $\left\{\mathsf{Y}^{(x)}\right\}_{x\in\mathsf{X}}$ forms a partition of X.

It is immediate to see that a quantum operation is *atomic* if and only if it is *single Kraus*, meaning that its Kraus decomposition contains a single operator, i.e., $\mathsf{Y}^{(x)} = \{x\}$. In other words, atomic operations generate the extremal rays of the cone of completely positive maps.

Following the definition, elementary properties in QT are repeatable quantum instruments whose quantum operations are atomic, i.e., single Kraus. Repeatability together with atomicity then implies

$$T_x^2 = \mathrm{e}^{i\phi}\, T_x.$$

This result can be obtained by a direct calculation. Indeed, by repeatability one has

$$\mathscr{T}_x(\mathscr{T}_x(\rho)) = \mathscr{T}_x(\rho)\,,$$

for every transformation $\mathscr{T}_x$ of the instrument. Expressing the transformation through the single Kraus operator associated with the operation, this condition reads

$$T_x T_x \, \rho \, T_x^\dagger T_x^\dagger = T_x \, \rho \, T_x^\dagger,$$

which implies the relation above.

Applying the Moore-Penrose pseudo-inverse of T_x to both sides yields

$$T_x = e^{i\phi} \Pi_x,$$

meaning that each Kraus operator T_x is proportional (up to a phase) to the projector Π_x onto its image. The global phase factor plays no role, since Kraus operators appear only in expressions of the form $T_x \rho T_x^\dagger$. We therefore ignore it.

Hence, in QT the Kraus operators of the quantum operations composing an elementary property are orthogonal projectors. In this sense, elementary properties correspond precisely to unnormalised projective measurements of the von Neumann-Lüders type.

Thanks to Corollary 16, we already know that weak compatibility implies no-complementarity. We now prove that the converse also holds.

Since in QT the elementary properties are composed of projectors, all verifier states are strong verifiers (7.2). Formally, given an elementary property $\mathsf{P}_\mathsf{X} \equiv \llbracket \Pi_x^\mathsf{P} \rrbracket_{x \in \mathsf{X}} \in \mathsf{Instr}(\mathrm{A} \to \mathrm{A})$, for every $\rho \in \Upsilon\left(\mathsf{P}_\mathsf{X}\right)$ there exists an outcome $x \in \mathsf{X}$ such that

$$\Pi_x^\mathsf{P} \left|\rho\right)_\mathrm{A} = \left|\rho\right)_\mathrm{A}.$$

To see this, let $\rho \in \Upsilon\left(\Pi_x^\mathsf{P}\right)$. Then

$$\mathsf{Tr}\left[\Pi_{x'}^\mathsf{P} \rho\right] = 0,$$

for all $x' \in \mathsf{X} \setminus \{x\}$. Hence,

$$\bigoplus_{x' \in \mathsf{X} \setminus \{x\}} \mathsf{Supp}\left(\Pi_{x'}^\mathsf{P}\right) \subseteq \mathsf{Ker}(\rho),$$

which implies

$$\mathsf{Supp}(\rho) \subseteq \left(\bigoplus_{x' \in \mathsf{X} \setminus \{x\}} \mathsf{Supp}\left(\Pi_{x'}^\mathsf{P}\right) \right)^\perp .$$

Since the projectors form a complete resolution of the identity,

$$\sum_{x \in \mathsf{X}} \Pi_x^{\mathsf{P}} = \mathbb{I}, \tag{7.8}$$

—which follows from the fact that we are considering all the Kraus operators of the quantum operations composing the instrument—, it follows that

$$\left(\bigoplus_{x' \in \mathsf{X} \setminus \{x\}} \mathsf{Supp}\left(\Pi_{x'}^{\mathsf{P}}\right) \right)^{\perp} = \mathsf{Supp}\left(\Pi_x^{\mathsf{P}}\right).$$

Therefore,

$$\mathsf{Supp}(\rho) \subseteq \mathsf{Supp}\left(\Pi_x^{\mathsf{P}}\right),$$

which gives

$$\Pi_x^{\mathsf{P}} \rho \Pi_x^{\mathsf{P}} = \rho.$$

In other words, ρ is a strong verifier for Π_x^{P}.

Now, consider two non-complementary elementary properties $\mathsf{P}_\mathsf{X} \equiv \llbracket \Pi_x^{\mathsf{P}} \rrbracket_{x \in \mathsf{X}}$ and $\mathsf{Q}_\mathsf{Y} \equiv \llbracket \Pi_y^{\mathsf{Q}} \rrbracket_{y \in \mathsf{Y}}$. By definition of non-complementarity, their sets of verifier states must coincide:

$$\Upsilon\left(\mathsf{P}_\mathsf{X}\right) = \Upsilon\left(\mathsf{Q}_\mathsf{Y}\right).$$

In particular, for every $x \in \mathsf{X}$ there must exist $y \in \mathsf{Y}$ such that

$$\Upsilon\left(\Pi_x^{\mathsf{P}}\right) = \Upsilon\left(\Pi_y^{\mathsf{Q}}\right).$$

Otherwise, non-complementarity would be violated. Indeed, suppose two states $\rho, \sigma \in \Upsilon\left(\Pi_x^{\mathsf{P}}\right)$, with $\rho \in \Upsilon\left(\Pi_y^{\mathsf{Q}}\right)$ and $\sigma \in \Upsilon\left(\Pi_{y'}^{\mathsf{Q}}\right)$ for distinct outcomes $y \neq y'$. Then any convex mixture of ρ and σ would be a verifier state for P_X but not for Q_Y, contradicting the assumption.

Since in QT an orthogonal projector is uniquely determined by its support, we obtain

$$\Pi_x^{\mathsf{P}} = \Pi_{\varphi(x)}^{\mathsf{Q}},$$

for a bijection $\varphi : \mathsf{X} \to \mathsf{Y}$. Hence the two instruments coincide up to relabelling of outcomes. As an instrument is trivially compatible with itself, we conclude that in QT non-complementarity implies compatibility.

8 Conclusion

THANK YOU for having followed us in this journey in the universe of OPTs. It all began with a simple question: can we gain a deeper understanding of what is truly encompassed by Heisenberg's *Gedankenexperiment* of the γ-ray microscope, and of the relationships between the physical properties it involves? This question has guided us to the resolution of a debate that has lasted for almost a century. We have shown that *complementarity*, *measurement incompatibility*, and *irreversibility* are distinct—though closely related—physical properties.

As we have seen, in order to establish this result it was necessary to rely on the framework of OPTs. This necessity arose from the need to consider a broader class of theories than just QT and CT. Since the former satisfies all three properties, while the latter satisfies none, neither could serve to investigate the relationships among complementarity, observation-incompatibility, and irreversibility.

The framework of OPTs provides the ideal ground for this purpose, as it was devised precisely to allow for the characterisation of the interplay between different physical properties from a theory-agnostic perspective.

The first necessary step towards our result was the formalisation of the properties of interest:

- **Observation-incompatibility**: Definition 37. Defined as the existence of incompatible observation-instruments, namely instruments that cannot be performed simultaneously.

- **(Strong) Incompatibility**: Definition 40. Defined as the existence of strongly incompatible instruments, i.e., cases where the implementation of one excludes the other.

- **Irreversibility**: Definition 44. Defined as the existence of operations that irreversibly alter the system on which they are performed.

- **Complementarity**: Definition 62. Defined as the existence of elementary properties of a physical system that cannot be simultaneously well defined.

Once formal notions of these properties had been established, it became possible to characterise the relationships among them.

A first consequence of the definitions is that

$$\text{Observation-incompatibility} \implies \text{(Strong) incompatibility} \iff \text{Irreversibility},$$

which represents one of the main results of this thesis (Theorem 29, Theorem 30):

> *The existence of quantities that cannot be simultaneously measured implies the existence of operations that irreversibly alter the physical system on which they act.*

This clarifies the relationship between two of the distinct phenomena encompassed by Heisenberg's *Gedankenexperiment*.

We were also able to show that the implication is strict. Much of the work, in fact, consisted in constructing suitable counterexamples to the converse, namely that irreversibility does not imply observation-incompatibility.

The two counterexamples—*MCT* (section 5.5) and *MSBCT* (section 6.4)—are specific instances of two broader classes of OPTs introduced in this thesis, namely *MOPTs* (chapter 5) and *MSOPTs* (chapter 6). These classes were designed to explore the consequences of restricting the allowed dynamics in a general theory of information processing. In particular,

- **MCT** is the MOPT obtained from CT, i.e., a classical theory where the only allowed operations are preparation-instruments, observation-instruments, swaps, the identity, and their compositions and limits;

- **MSBCT** is the MSOPT obtained from BCT, i.e., the minimal version of the theory that also includes all conditional instruments.

All MOPTs and all MSOPTs admitting a spanning set of entangled states share an important structural feature: they do not allow a non-trivial decomposition of the identity, that is, the identity transformation is atomic (Theorem 38, Theorem 41). Atomicity of the identity is a powerful property, as it implies several other important properties:

- **NIWD**, namely that extracting information from a system is only possible at the cost of disturbing it (Theorem 26);
- **No-broadcasting**, prohibiting the "copying" of states, including their remote correlations (Theorem 27);
- **Irreversibility** (Theorem 31).

This is precisely what allowed us to conclude that irreversibility does not imply observation-incompatibility, since both MCT and MSBCT exhibit full compatibility of observation-instruments, inherited from the theories from which they derive, despite featuring irreversibility.

Moreover, MCT and MSBCT also show that NIWD and no-broadcasting cannot be taken *per se* as signatures of non-classicality, not even in the presence of conditional instruments. Indeed, MSBCT further highlights that strong causality alone is insufficient to guarantee a non-trivial decomposition of the identity, even under the assumption of classicality.

The notion of classicality we adopt in this thesis is related to the geometry of the state spaces (Definition 27), namely, simpliciality. It is worth noting, however, that both MCT and MSBCT are also Kochen-Specker non-contextual and generalised-non-contextual [152, 153, 222].

We stress that the results obtained here are not in contrast with those of Refs. [144, 146]. Their conclusion—that a set of states is broadcastable if and only if it is contained in a simplex whose vertices are jointly perfectly discriminable—relies on assuming both local discriminability and the possibility of classical conditioning on measurement outcomes. The absence of local discriminability rules out MSBCT, while the absence of classical conditioning rules out MCT, from satisfying the premises of their result.

MCT and MSBCT, being non-contextual, could also serve as toy models for testing the implications established in Ref. [264], where the relation between no-broadcasting and contextuality in probabilistic theories is extensively analysed.

Beyond atomicity, another striking feature of all MOPTs and all MSOPTs admitting a spanning set of entangled states is that these theories do not admit reversible transformations other than permutations (Proposition 5, Corollary 12). At first sight this may appear to be a mere curiosity, but it actually provides a clear criterion to distinguish a theory from its minimal and minimal strongly causal versions. In the case of QT, this implies that

$$\mathsf{Instr}(MQT) \subset \mathsf{Instr}(MSQT) \subset \mathsf{Instr}(QT)\,. \tag{6.21}$$

By contrast, an analogous result does not hold for CT, since this theory does not satisfy the hypothesis of Corollary 12. Indeed, because every instrument of CT admits a reversible dilation, one has

$$\mathsf{Instr}(MCT) \subset \mathsf{Instr}(MSCT) \equiv \mathsf{Instr}(CT)\,. \tag{6.20}$$

This result is remarkable, as it shows that in the classical world all dynamics can be reduced to an appropriate sequence of measurements and preparations conditioned on measurement outcomes. In the quantum world, however, this is not the case, thereby highlighting a sharp distinction in the way the two theories process information.

A natural question then arises: which additional instruments must be added to MSQT to fully recover QT, and, more generally, to an MSOPT to recover the OPT from which it is derived?

A plausible route towards resolving the quantum case is suggested by the paradigm of *measurement-based quantum computation* [257–263]. This is a an universal computational paradigm—which means that is able to approximately reconstruct all unitary dynamics—driven by sequences of conditional measurements and Pauli transformations on a suitable initial entangled state. This setting is remarkably close to that of MSQT, except for the explicit availability of Pauli transformations. Might these suffice to recover the full QT?

Resolving this could motivate a parallel line of inquiry for generic OPTs. Can any computation achievable in a given OPT be reduced to an operation in the associated MSOPT with the addition of at most a small set of further instruments? Addressing such questions could provide further insight into the computational properties of generic operational theories [145, 150].

Beyond their implications for theory reconstruction and computation, MOPTs and MSOPTs might also prove useful in furthering our understanding of cryptographic protocols. Since all these theories satisfy the NIWD property, they provide a natural proving ground for studying under which conditions a theory can support information-theoretically secure cryptographic protocols for key distribution, in analogy with QT [38, 39, 138, 265–270]. The strong restrictions on the set of transformations would allow greater control over the variables involved in a communication process, allowing the problem to be investigated in a simplified setting. In particular, it is worth asking whether an OPT such as MSBCT could allow for one, despite being classical. We conjecture that this might be indeed feasible, by exploiting the fact that an eavesdropper's intervention would necessarily be detectable, since any attempt to acquire information about the system would irreversibly disturb it.

Last, but not least, we observe that MOPTs allow one to construct OPTs by focusing exclusively on state and measurement spaces. This mirrors the approach adopted in prepare-and-measure scenarios that are widely studied in the GPT literature. Since MOPTs constitute minimal working examples of OPTs that can be extracted from GPTs in the prepare-and-measure scenario, they provide a natural bridge between these two areas of research.

Having concluded the characterisation of the properties of MOPTs and MSOPTs, we then turned to Bohr's notion of *complementarity*. We proposed a new operational definition of this concept, based on the existence of elementary properties of a physical system that cannot be simultaneously well defined (Definition 62). This definition retraces Bohr's original idea of complementarity as the impossibility of fully describing a physical system from a single point of view, thereby distinguishing it from incompatibility and hence from much of the current literature on the subject [71, 76, 112–118].

In particular, we showed that for generic OPTs (Corollary 16),

$$\text{Complementarity} \implies \text{(Strong) Incompatibility},$$

which in turn implies

$$\text{Complementarity} \implies \text{Irreversibility}.$$

Thus, we settle the long-standing debate on whether complementarity can indeed, as Bohr proposed, be regarded as a principle from which Heisenberg's irreversibility follows [100].

We know that, in general, incompatibility does not imply complementarity. MCT provides a counterexample: it is a theory without complementarity—since the identity is the only elementary property—yet it still admits incompatible instruments. However, in the presence of multiple elementary properties, then, the relationship is unknown. In this respect, MSBCT could be an interesting case to analyse. Here, the identity is an elementary property, but the theory may also admit an instrument of the form (4.3) as an elementary property. Among the conditions that must be checked, atomicity appears the most challenging, since a full characterisation of the instruments of the theory is still lacking.

The only case where a complete picture is available is QT, where we proved that

$$\text{Complementarity} \iff \text{(Strong) Incompatibility}.$$

Finally, we observe that the relationship between complementarity and observation-incompatibility remains unsettled. At present, we lack intuition about what the connection might be, and its characterisation is left for future research. Similarly, the link between complementarity and classicality has yet to be fully clarified. We believe that the study of MSBCT could provide valuable insights also in this latter case.

In conclusion, we propose a series of diagrams summarising all the relationships between different properties that we have presented in this thesis:

- In figure 8.1 we summarise the relationships between complementarity, observation-incompatibility, strong-incompatibility, and irreversibility.

  The case of MSBCT is highlighted with a question mark, as it is not yet clear whether this theory exhibits complementarity.

- In figure 8.2 we summarise the relationship between strong-causality, causality, and the no-restriction hypothesis.

- In figure 8.3, we summarise the relationships among strong causality, causality, local discriminability, and simpliciality. In this context, we also introduce three new MOPTs: the minimal versions of RQT (Minimal Real Quantum Theory (MRQT)), of FQT (Minimal Fermionic Quantum Theory (MFQT)), and of the LQTs (Minimal Latent Quantum Theories (MLQTs)).

Figures such as Bohr, Heisenberg, Einstein, and Schrödinger were extraordinary. Their intuition and their ability to foresee a whole new physical world, and to bring it to life, remain astonishing—we truly stand on the shoulders of giants. Yet, looking back, one cannot fail to notice that the conceptual and mathematical instruments available to them were far less powerful than those at our disposal today, which inevitably limited what they could achieve. It has been a privilege to take part in this great tradition and to give a small contribution in bringing to light new notions that enrich our understanding of this beautiful, chaotic, and sometimes counter-intuitive quantum world. There is still much we do not understand about quantum theory, but step by step we are getting there.

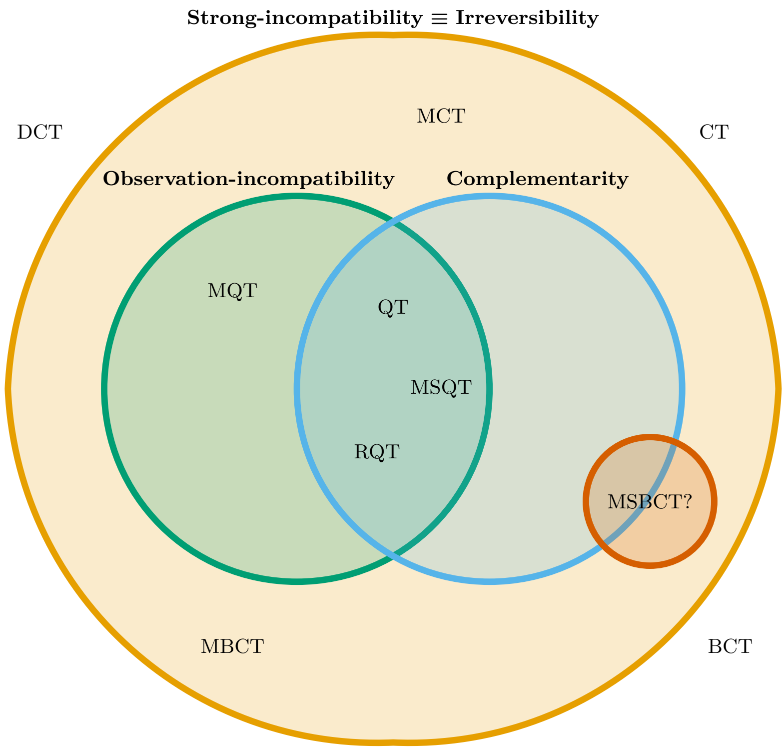


Figure 8.1: Venn diagram summarising the relationships between complementarity, observation-incompatibility, strong-incompatibility, and irreversibility. The case of MSBCT is highlighted with a question mark, as it is not yet clear whether this theory exhibits complementarity.

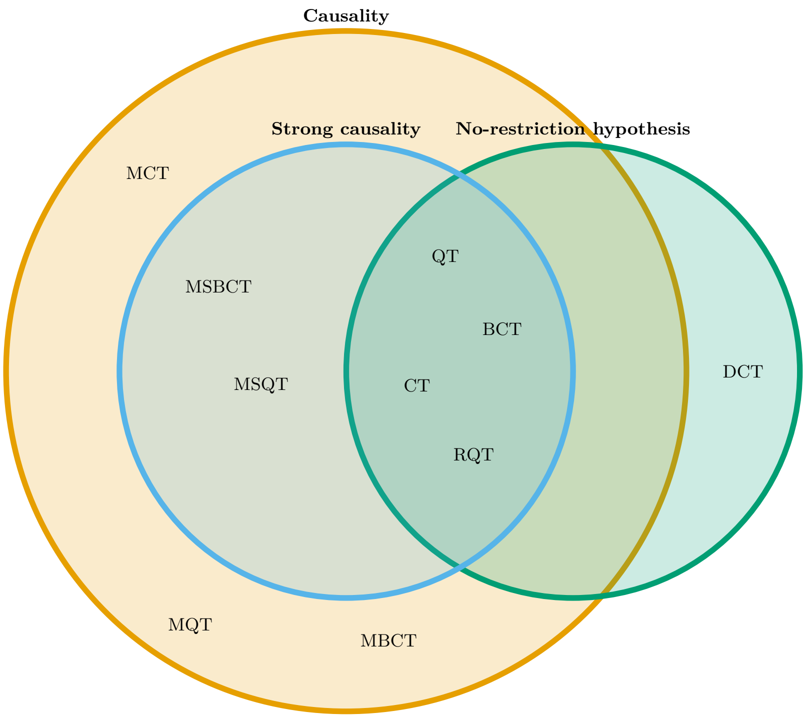


Figure 8.2: Venn diagram summarising the relationship between strong-causality, causality ,and the no-restriction hypothesis.

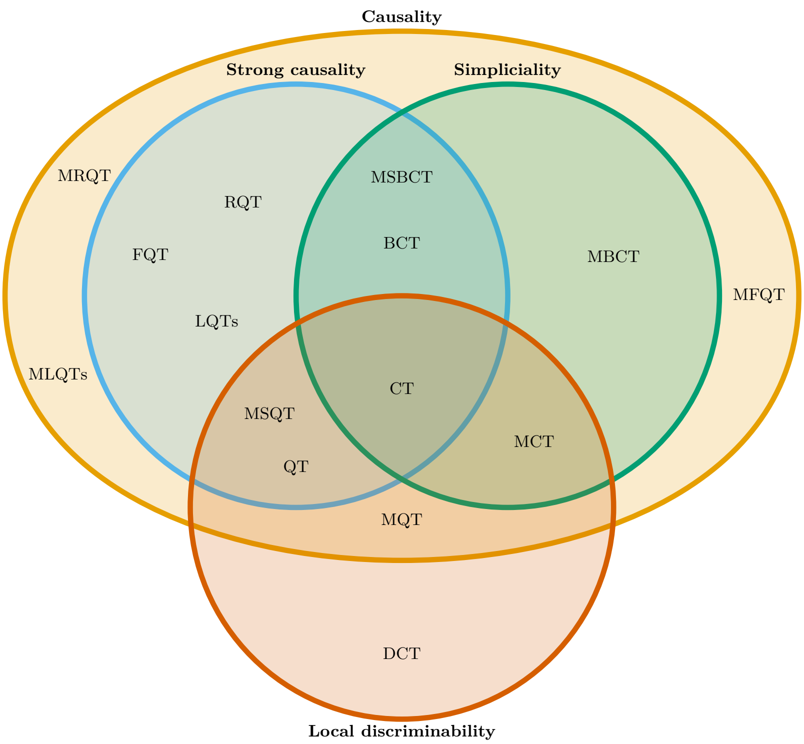


Figure 8.3: Venn diagram summarising the relationships between strong-causality, causality, local discriminability, and simpliciality.

A Category Theory

CATEGORY THEORY is a general mathematical framework, introduced in the middle of the 20th century by Eilenberg and Mac Lane in their seminal work [271]. Originally motivated by problems in algebraic topology, it soon became a cornerstone of modern mathematics. The central idea of category theory is to formalise the relationships between abstract objects by encoding them into *morphisms*, that is, “maps” between objects.

In what follows we provide a concise overview of the definitions and results from category theory that are relevant for this thesis. For a more in-depth discussion of category theory and its relationship with quantum theory, we refer the interested reader to Refs. [169–171].

Let us begin by recalling the formal definition of a *category*:

Definition 64 (Category)

A *category* C consists of:

- A collection of *objects*: A, B, ... $\in \mathsf{Obj}(C)$.
- A collection of *morphisms* (or *arrows*): f, g, ... $\in \mathsf{Mor}(C)$.
- For each morphism f, there are given objects

$$\mathsf{dom}(f)\,, \quad \mathsf{cod}(f)\,,$$

called the *domain* (or *source*) and *codomain* (or *target*) of f. We write:

$$f : \mathrm{A} \to \mathrm{B}$$

to indicate that $\mathrm{A} = \mathsf{dom}(f)$ and $\mathrm{B} = \mathsf{cod}(f)$ and we say that f is a morphism form object A to object B. The collection of all such morphisms is denoted with $\mathsf{Mor}(\mathrm{A}, \mathrm{B})$.

- A binary operation $\circ$, called *composition of morphisms*, such that for any three objects A, B, $\mathrm{C} \in C$ defines:

$$\circ : \mathsf{Mor}(\mathrm{B}, \mathrm{C}) \times \mathsf{Mor}(\mathrm{A}, \mathrm{B}) \to \mathsf{Mor}(\mathrm{A}, \mathrm{C}) .$$

The composition of $f \in \mathsf{Mor}(\mathrm{A}, \mathrm{B})$ and $g \in \mathsf{Mor}(\mathrm{B}, \mathrm{C})$ is written as $g \circ f$ or simply gf. The composition rule is furthermore required to satisfy:

  - **Associativity**: for any $f \in \mathsf{Mor}(\mathrm{A}, \mathrm{B})$, $g \in \mathsf{Mor}(\mathrm{B}, \mathrm{C})$, and $h \in \mathsf{Mor}(\mathrm{C}, \mathrm{D})$, it holds that

$$h \circ (g \circ f) = (h \circ g) \circ f,$$

  - **Existence of the identity element**: for any object $\mathrm{A} \in \mathsf{Obj}(C)$ there exists a morphism $1_{\mathrm{A}} \in \mathsf{Mor}(\mathrm{A}, \mathrm{A})$ called the *identity morphism* for A, such that for any morphism $f \in \mathsf{Mor}(\mathrm{A}, \mathrm{B})$ we have

$$1_{\mathrm{B}} \circ f = f = f \circ 1_{\mathrm{A}}.$$

A.1 Monoidal categories

As we have already seen in the introduction to the framework of OPTs, it is sometimes necessary to represent relationships between multiple objects simultaneously—“in parallel”—rather than only through sequential composition. In category theory, this can be achieved by introducing a notion of parallel composition, formally captured by the notion of a *monoidal category*.

Definition 65 (Monoidal category)

A *monoidal category* C is a category equipped with the following structures:

- A bifunctor $\otimes : C \times C \to C$, called the *monoidal product*[a].
- An object I, called the *monoidal unit*, *unit object*, or *identity object*.
- A natural isomorphism

  $$\alpha : ((\cdot \otimes \cdot) \otimes \cdot) \xrightarrow{\simeq} (\cdot \otimes (\cdot \otimes \cdot))$$

  with components of the form

  $$\alpha_{\mathrm{A,B,C}} : (\mathrm{A} \otimes \mathrm{B}) \otimes \mathrm{C} \to \mathrm{A} \otimes (\mathrm{B} \otimes \mathrm{C})\,,$$

  called the *associator*.
- A natural isomorphism

  $$\lambda : \mathrm{I} \otimes \cdot \xrightarrow{\simeq} \cdot$$

  with components of the form

  $$\lambda_{\mathrm{A}} : \mathrm{I} \otimes \mathrm{A} \to \mathrm{A},$$

  called the *left-unitor*.
- A natural isomorphism

  $$\rho : \cdot \otimes \mathrm{I} \xrightarrow{\simeq} \cdot$$

  with components of the form

  $$\rho_{\mathrm{A}} : \mathrm{A} \otimes \mathrm{I} \to \mathrm{A},$$

  called the *right-unitor*.

These structures must satisfy the following coherence conditions, expressed through commutative diagrams that must hold for all involved objects:

- The *triangle identity*:

$$\begin{array}{ccc} (A \otimes 1) \otimes B & \xrightarrow{\alpha_{A,I,B}} & A \otimes (1 \otimes B) \\ & {\scriptstyle \rho_A \otimes 1_B} \searrow \qquad \swarrow {\scriptstyle 1_A \otimes \lambda_B} & \\ & A \otimes B & \end{array} \quad . \tag{A.1}$$

- The *pentagon identity*:

$$\begin{array}{ccc} & (D \otimes A) \otimes (B \otimes C) & \\ {\scriptstyle \alpha_{D \otimes A,B,C}} \nearrow & & \searrow {\scriptstyle \alpha_{D,A,B \otimes C}} \\ ((D \otimes A) \otimes B) \otimes C & & D \otimes (A \otimes (B \otimes C)) \\ {\scriptstyle \alpha_{D,A,B} \otimes 1_C} \downarrow & & \uparrow {\scriptstyle 1_D \otimes \alpha_{A,B,C}} \\ (D \otimes (A \otimes B)) \otimes C & \xrightarrow[\alpha_{D,A \otimes B,C}]{} & D \otimes ((A \otimes B) \otimes C) \end{array} \quad . \tag{A.2}$$

[a] When discussing categories, the symbol $\otimes$ does not denote the (algebraic) tensor product, but simply the bifunctor defining the structure of the monoidal category.

When discussing the relationship between OPTs and category theory in section 2.2.1, we highlighted that Event_Θ is a *strict* monoidal category. Formally, this is defined as follows:

Definition 66 (Strict monoidal category)

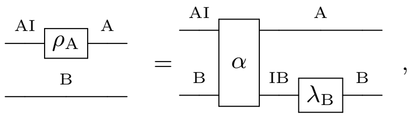
A *strict monoidal category* C is a monoidal category whose associator, left-unitor, and right-unitor are all identity morphisms.

In the case of OPTs, the left-unitor and right-unitor were implicitly defined in (2.13), while the associator was implicitly defined in (2.12). As can be seen, all these morphisms were chosen to be equal to the identity event. Therefore, Event_Θ is indeed *strict*.

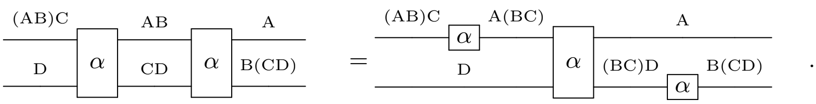
An important feature of strict monoidal categories is that they automatically satisfy the triangle (A.1) and pentagon (A.2) identities.

In the OPTs formalism, these two identities are represented as follows:

- The triangle identity:



- The pentagon identity:



For completeness, we now recall the definitions of a *functor* and a *natural transformation*.

Definition 67 (Functor)

A *functor* F from a category C to a category D is a map sending each object $\mathrm{A} \in \mathsf{Obj}(C)$ to an object $F(\mathrm{A}) \in \mathsf{Obj}(D)$, and associating each morphism $f \in \mathsf{Mor}(\mathrm{A},\mathrm{B})$ in C with a morphism $F(f) \in \mathsf{Mor}(F(\mathrm{A}) \to F(\mathrm{B}))$ in D, such that the following two conditions hold:

- F preserves composition:

$$F(g \circ f) = F(g) \circ F(f),$$

  for any $f \in \mathsf{Mor}(\mathrm{A},\mathrm{B})$ and $g \in \mathsf{Mor}(\mathrm{B},\mathrm{C})$,

- F preserves identity morphisms:

$$F(1_{\mathrm{A}}) = 1_{F(\mathrm{A})},$$

  for every object $\mathrm{A} \in \mathsf{Obj}(C)$.

Definition 68 (Natural transformation)

If F and G are two functors between categories C and D, a *natural transformation* η from F to G is a family of morphisms satisfying the following conditions:

- To every object $\mathrm{A} \in \mathsf{Obj}(C)$, the natural transformation associates a morphism

$$\eta_{\mathrm{A}} : F(\mathrm{A}) \to G(\mathrm{A}),$$

called the *component of* η *at* A, which is a morphism in D.

- The components must satisfy the following naturality condition: for every morphism $f : \mathrm{A} \to \mathrm{B}$ in C,

$$\eta_{\mathrm{B}} \circ F(f) = G(f) \circ \eta_{\mathrm{A}}.$$

This equality can also be illustrated by the following commutative diagram:

$$\begin{array}{ccccc}
\mathrm{A} & \quad & F(\mathrm{A}) & \xrightarrow{\eta_{\mathrm{A}}} & G(\mathrm{A}) \\
{\scriptstyle f}\big\downarrow & & {\scriptstyle F(f)}\big\downarrow & & \big\downarrow{\scriptstyle G(f)} \\
\mathrm{B} & & F(\mathrm{B}) & \xrightarrow[\eta_{\mathrm{B}}]{} & G(\mathrm{B})
\end{array}\quad.$$

A.2 Braided monoidal categories

We are now left with one final piece of the puzzle needed to fully reconstruct the compositional structure underlying Event_Θ: the definition of *braided monoidal categories*.

Definition 69 (Braided monoidal category)

A *braided monoidal category* is a monoidal category C equipped with a natural isomorphism

$$\mathscr{S}_{\mathrm{A,B}} : \mathrm{A} \otimes \mathrm{B} \to \mathrm{B} \otimes \mathrm{A},$$

called the *braiding*, such that the *hexagon identities* are satisfied for all objects involved. These identities express the compatibility of the braiding with the associator of the monoidal structure:

$$
\begin{array}{ccccc}
(\mathrm{A} \otimes \mathrm{B}) \otimes \mathrm{C} & \xrightarrow{\alpha_{\mathrm{A,B,C}}} & \mathrm{A} \otimes (\mathrm{B} \otimes \mathrm{C}) & \xrightarrow{\mathscr{S}_{\mathrm{A,B}\otimes\mathrm{C}}} & (\mathrm{B} \otimes \mathrm{C}) \otimes \mathrm{A} \\
\Big\downarrow {\scriptstyle \mathscr{S}_{\mathrm{A,B}} \otimes 1_{\mathrm{C}}} & & & & \Big\downarrow {\scriptstyle \alpha_{\mathrm{B,C,A}}} \\
(\mathrm{B} \otimes \mathrm{A}) \otimes \mathrm{C} & \xrightarrow[\alpha_{\mathrm{B,A,C}}]{} & \mathrm{B} \otimes (\mathrm{A} \otimes \mathrm{C}) & \xrightarrow[1_{\mathrm{B}} \otimes \mathscr{S}_{\mathrm{A,C}}]{} & \mathrm{B} \otimes (\mathrm{C} \otimes \mathrm{A})
\end{array} ,
$$

and

$$
\begin{array}{ccccc}
\mathrm{A} \otimes (\mathrm{B} \otimes \mathrm{C}) & \xrightarrow{\alpha^{-1}_{\mathrm{A,B,C}}} & (\mathrm{A} \otimes \mathrm{B}) \otimes \mathrm{C} & \xrightarrow{\mathscr{S}_{\mathrm{A}\otimes\mathrm{B,C}}} & \mathrm{C} \otimes (\mathrm{A} \otimes \mathrm{B}) \\
\Big\downarrow {\scriptstyle 1_{\mathrm{A}} \otimes \mathscr{S}_{\mathrm{B,C}}} & & & & \Big\downarrow {\scriptstyle \alpha^{-1}_{\mathrm{C,A,B}}} \\
\mathrm{A} \otimes (\mathrm{C} \otimes \mathrm{B}) & \xrightarrow[\alpha^{-1}_{\mathrm{A,C,B}}]{} & (\mathrm{A} \otimes \mathrm{C}) \otimes \mathrm{B} & \xrightarrow[\mathscr{S}_{\mathrm{A,C}} \otimes 1_{\mathrm{B}}]{} & (\mathrm{C} \otimes \mathrm{A}) \otimes \mathrm{B}
\end{array} .
$$

We recall that in the case of OPTs, the hexagon identities can be graphically represented as in (2.20), where the associator is not made explicit, since it is equal to the identity.

If the braiding operation "squares" to the identity $\mathscr{S}_{\mathrm{A,B}} \circ \mathscr{S}_{\mathrm{B,A}} = 1_{\mathrm{AB}}$, then the braided monoidal category is called *symmetric monoidal category*.

Symmetric monoidal categories satisfy the following useful result.

Theorem 43 (Coherence for symmetric monoidal categories)

Every diagram in a symmetric monoidal category made up of associators, unitors, swaps, and in which both sides have the same underlying permutation, commutes [169].

A.3 Graphical calculus for category theory

We have seen that for OPTs, it is possible to define a sound and complete graphical calculus. The well-definiteness of this graphical reasoning method relies on a more general result, which provides a foundation for graphical calculation in generic categories. This procedure is described in detail in Ref. [171].

For our purposes, the most important result is the following:

Theorem 44 (Correctness of graphical calculus for braided monoidal categories)

A well-typed equation between morphisms in a braided monoidal category follows from the axioms if and only if it holds in the graphical language up to spatial isotopy [171].

We define *spatial isotopy* as follows:

Definition 70 (Spatial isotopy)

Two diagrams in the graphical language of a braided monoidal category are said to be *spatially isotopic* if one can be continuously deformed into the other within a three-dimensional space (typically imagined inside a rectangular box), without breaking or creating wires, and while keeping fixed the boundary positions where wires intersect the sides of the box. That is, the deformation respects the input and output positions of the diagram, and only involves internal movement of boxes and wires.

B Cones

In this appendix, we delve deeper into the results presented at the end of section 3.1. Specifically, we show why these results hold by providing their proofs together with the mathematical theorems on which they rely.

B.1 Proof of Theorem 4

We begin with the proof of Theorem 4.

Theorem 4

The cone generated by a bounded, closed set B with non-zero distance from the origin is closed.

Proof. The set B acts as the base of the cone.

Let $\{q_n\}_{n\in\mathbb{N}} \subset \mathsf{ConicHull}(B)$ be a sequence in the conic hull of B with limit

$$p = \lim_{n\to\infty} q_n.$$

Since the sequence converges, for every $\varepsilon > 0$ there exists $\tilde{n} \in \mathbb{N}$ such that for all $n > \tilde{n}$,

$$\|q_n - p\| \leq \varepsilon.$$

We consider two cases:

1. $p = 0$. Then $p \in \textsf{ConicHull}(B)$ by definition of conic hull.

2. $p \neq 0$. Since $q_n \in \textsf{ConicHull}(B)$, we can write $q_n = k_n b_n$ with $b_n \in B$ and $k_n \geq 0$ for all n.

   Since $\|p\| > 0$, and the sequence $\{q_n\}$ converges to p, we have for sufficiently small ε:

   $$\|q_n\| \geq \|p\| - \varepsilon.$$

   Assuming without loss of generality that $\|b_n\| = 1$, it follows that

   $$k_n = \|q_n\| \geq \|p\| - \varepsilon.$$

   (More generally, since B is bounded away from zero, there exist constants $0 < c \leq \|b_n\| \leq C$ for all n, and the argument can be suitably rescaled.)

   Since $\{q_n\}$ converges, so does $\{\|q_n\|\}_{n\in\mathbb{N}}$ to $\|p\|$, and we obtain:

   $$\left\| b_n - \frac{p}{k_n} \right\| \leq \frac{\varepsilon}{k_n} \leq \frac{\varepsilon}{\|p\| - \varepsilon}.$$

   We now estimate:

   $$\begin{aligned} \left\| b_n - \frac{p}{\|p\|} \right\| &\leq \left\| b_n - \frac{p}{k_n} \right\| + \left\| \frac{p}{k_n} - \frac{p}{\|p\|} \right\| \\ &= \left\| b_n - \frac{p}{k_n} \right\| + \left\| p \left(\frac{1}{k_n} - \frac{1}{\|p\|} \right) \right\| \\ &\leq \frac{\varepsilon}{\|p\| - \varepsilon} + \frac{\varepsilon}{\|p\| - \varepsilon} = \frac{2\varepsilon}{\|p\| - \varepsilon}. \end{aligned}$$

   Up to taking ε sufficiently small, this yields:

   $$\left\| b_n - \frac{p}{\|p\|} \right\| \leq \frac{4\varepsilon}{\|p\|}.$$

   We have therefore shown that $b_n \to \frac{p}{\|p\|} \in B$, since B is closed. It follows that $p = \|p\| \cdot \frac{p}{\|p\|} \in \textsf{ConicHull}(B)$.

$\square$

B.2 Proof of Theorem 5

The case of Theorem 5 requires a more careful discussion. While the statement is, *per se*, a direct application of Klee's theorem, the delicate point is to verify that the cones arising in OPTs satisfy its hypotheses.

Theorem 5

In an OPT, the set $\overline{\mathsf{St}_+(\mathrm{A})}$ is the convex hull of its extremal rays together with the null state.

For reference, we recall Klee's theorem.

Theorem 45 (Klee's theorem)

If C is a closed, convex, and locally compact subset of a locally convex space, and if C contains no straight line, then C is the closed convex hull of its extreme points and extreme rays [272].

A proof of the theorem can be found in Ref. [273]. For completeness, we also recall the notion of extreme ray.

Definition 71 (Extreme ray)

An *extreme ray* of C is a closed half-line $R \subseteq C$ such that every open interval in C intersecting R is entirely contained in R [272].

This definition of extremal ray is equivalent to the standard indecomposability condition: any element of an extremal ray can only be written as a sum of elements proportional to it. More precisely:

Lemma 35

Let C be a convex cone in a real vector space[a] and let $R \subseteq C$ be a closed half-line, i.e., an object of the form

$$R \equiv \{\lambda x \,|\, \lambda \in \mathbb{R}_{\geq 0}\}$$

with $x \in C \setminus \{0\}$. The following are equivalent:

I) Every open interval in C intersecting R is entirely contained in R.

II) R is extremal in the usual indecomposable sense: if $y, z \in C$ and $y + z \in R$, then $y, z \in R$.

[a]The result can also be shown to hold in the case of complex vector spaces.

Proof. *(I $\implies$ II)* Suppose $y, z \in C$ with $y + z \in R$. Consider the open interval

$$I := \{ty + (1-t)z \mid t \in (0,1)\} \subseteq C.$$

Its midpoint is $(y+z)/2 \in R$, hence I intersects R. By (I), $I \subseteq R$, in particular the endpoints y and z belong to R. Thus R is indecomposable.

(II $\implies$ I) Let $I = \{ty + (1-t)z \mid t \in (0,1)\} \subseteq C$ be an open interval with a point $w \in I \cap R$, say

$$w = t_0 y + (1-t_0)z, \qquad t_0 \in (0,1).$$

Since C is a cone, for any $\lambda > 0$ we have

$$\lambda w = (\lambda t_0)y + \lambda(1-t_0)z \in C.$$

By indecomposability (II), this implies $y, z \in R$. Hence every element of I is a convex combination of points in R, so $I \subseteq R$. This proves (I). $\square$

In the language of OPTs, the elements of extremal rays of a cone are precisely the atomic ones (Definition 7).

To apply Theorem 45 in our setting, the relevant object is $\overline{\mathsf{St}_+(\mathrm{A})}$, with A any system of an OPT. By construction, this set is closed and convex (2.53). Under the assumptions of this thesis (Assumption 3), $\mathsf{St}_{\mathbb{R}}(\mathrm{A})$ is a finite-dimensional vector space and hence locally convex. Moreover, every closed subset of a locally compact space is locally compact. Since $\mathbb{R}^d$ is locally compact by the Heine-Borel theorem, it follows that $\overline{\mathsf{St}_+(\mathrm{A})}$ satisfies the local compactness hypothesis as well.

The remaining hypothesis to check is that $\overline{\mathsf{St}_+(\mathrm{A})}$ does not contain any straight line, namely that the cone is *pointed*.

Definition 72 (Pointed convex cone)

A convex cone C is *pointed* if does not contain any straight line.

This verification is immediate by the following result.

Theorem 46

If a closed cone $C \subseteq V$, where V is a real vector space[a], containing 0, contains $x_0 + \mathsf{Span}_{\mathbb{R}}(y)$ for some $x_0 \in C$ and $y \in V$, then it contains $\mathsf{Span}_{\mathbb{R}}(y)$.

[a]The result can also be shown to hold in the case of complex vector spaces.

Remark 46

The assumption of closedness in Theorem 46 is essential. For example,

$$C = \{(x, y) \mid y > 0\} \cup \{0\}$$

is a non-closed cone that contains any full line parallel to the x-axis (namely $\{(x, y) \mid y = c\}$, with $c \in \mathbb{R}_{\geq 0}$), but not the x-axis.

Proof. To prove the result we show that $x_0 + \mathsf{Span}_{\mathbb{R}}(y) \subseteq C$ implies $y, -y \in C$. A sufficient condition to show that $\mathsf{Span}_{\mathbb{R}}(y) \subseteq C$, since cones are closed under multiplication by positive scalars.

For y, consider

$$\frac{1}{n}(x_0 + ny) = y + \frac{x_0}{n} \in C, \quad \forall n \in \mathbb{N}.$$

By closedness, $\lim_{n\to\infty}\left(y + \frac{x_0}{n}\right) = y \in C$.

Analogously, for $-y$,

$$\frac{1}{n}(x_0 - ny) = -y + \frac{x_0}{n} \in C, \forall n \in \mathbb{N},$$

and closedness gives $\lim_{n\to\infty}\left(-y + \frac{x_0}{n}\right) = -y \in C$. □

We can now verify pointedness for the cones of interest in OPTs.

Theorem 47

Let Θ be a generic OPT. Then $\overline{\mathsf{St}_+(\mathrm{A})}$ is a pointed convex cone for any $\mathrm{A} \in \mathsf{Sys}(\Theta)$.

Proof. By construction, $\overline{\mathsf{St}_+(\mathrm{A})}$ is a closed convex cone. We now show that it is also pointed.

By Lemma 18, for every state $\rho \in \overline{\mathsf{St}_+(\mathrm{A})}$ and every effect $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$ one has

$$(\mathrm{a}|\, \rho)_{\mathrm{A}} \geq 0.$$

This holds since Lemma 18 applies also to effects, so evaluating an effect on an element of $\overline{\mathsf{St}_+(\mathrm{A})}$ yields a value in $\mathbb{R}_{\geq 0}$, i.e., in $\overline{\mathsf{St}_+(\mathrm{I})}$.

Suppose, by contradiction, that $\overline{\mathsf{St}_+(\mathrm{A})}$ contains a line. Then, by Theorem 46, there exists a non-zero $\rho \in \overline{\mathsf{St}_+(\mathrm{A})}$ with $\mathsf{Span}_{\mathbb{R}}(\rho) \subseteq \overline{\mathsf{St}_+(\mathrm{A})}$. In particular, for any $\lambda < 0$, the vector $\sigma := \lambda\rho$ also belongs to $\overline{\mathsf{St}_+(\mathrm{A})}$, and hence

$$0 \leq (\mathrm{a}|\, \sigma)_{\mathrm{A}} = \lambda\, (\mathrm{a}|\, \rho)_{\mathrm{A}} \leq 0, \quad \forall\, \mathrm{a} \in \mathsf{Eff}(\mathrm{A})\,,$$

where the second inequality comes from the fact that $\lambda < 0$. This implies

$$(\mathrm{a}|\, \rho)_{\mathrm{A}} = 0 \quad \forall\, \mathrm{a} \in \mathsf{Eff}(\mathrm{A})\,.$$

By separation of states and effects, this forces $\rho = \varepsilon_{\mathrm{A}}$, contradicting the assumption that ρ was non-zero. Thus $\overline{\mathsf{St}_+(\mathrm{A})}$ contains no non-trivial line, i.e., it is pointed. □

We have thus verified that $\overline{\mathsf{St}_+(\mathrm{A})}$ satisfy all the hypotheses of Klee's theorem (Theorem 45) for any system A of an OPT. Consequently, Theorem 5 holds.

Observation 12

The results presented in this appendix are proven in the finite-dimensional case. The possible extension to infinite dimensions remains to be investigated.

C

The remaining postulates of QT

In this appendix, we present the formal definition of the last three postulates of QT (Definition 29) that were not discussed elsewhere in this thesis.

Definition 73 (Atomicity of composition)

The sequential composition of two atomic transformations is itself atomic [133].

Definition 74 (Perfect discriminability)

Every deterministic state that is not maximally mixed is perfectly discriminable from some other state [133]. Formally, for every system A and every state $\rho_0 \in \mathsf{St}_1(\mathrm{A})$ that is not maximally mixed, there exists another state $\rho_1 \in \mathsf{St}_1(\mathrm{A})$ and an observation-instrument $[\![\mathrm{a}_0, \mathrm{a}_1]\!] \in \mathsf{Obs}(\mathrm{A})$ such that

$$(\mathrm{a}_i|\, \rho_j)_{\mathrm{A}} = \delta_{i,j}.$$

Definition 75 (Ideal compression)

Every state can be compressed in a lossless and efficient way, i.e., every state admits an ideal compression protocol [133].

Ideal compression In order to define what an ideal compression protocol is, let us start by defining *compression* protocols.

Definition 76 (Compression protocol)

A *compression protocol* for a state $\rho \in \mathsf{St}(\mathrm{A})$ is a triple $(\mathscr{E}, \mathscr{D}, \mathrm{B})$, where B is another system of the theory (usually of smaller dimension), and $\mathscr{E} \in \mathsf{Transf}_1(\mathrm{A} \to \mathrm{B})$ and $\mathscr{D} \in \mathsf{Transf}_1(\mathrm{B} \to \mathrm{A})$ are two deterministic transformations, respectively denoted the *encoding* and the *decoding*.

The intuition behind a compression protocol is to store the information contained in a state of some system into another system (usually of smaller dimension) via an encoding operation $\mathscr{E}$, in such a way that the original information can later be fully recovered by a decoding operation $\mathscr{D}$.

We can then define the following.

Definition 77 (Lossless compression protocol)

We say that a compression protocol $(\mathscr{E}, \mathscr{D}, \mathrm{B})$ for a state $\rho \in \mathsf{St}(\mathrm{A})$ is *lossless* if it holds that [133]:

$$\mathscr{D}\mathscr{E} \left|\alpha\right)_{\mathrm{A}} = \left|\alpha\right)_{\mathrm{A}}, \qquad \forall \alpha \in \mathsf{Ref}(\rho).$$

Definition 78 (Efficient compression protocol)

We say that a compression protocol $(\mathscr{E}, \mathscr{D}, \mathrm{B})$ for a state $\rho \in \mathsf{St}(\mathrm{A})$ is *efficient* if every state of system B arises as a codeword for some state in $\mathsf{Ref}(\rho)$, that is [133],

$$\forall \beta \in \mathsf{St}(\mathrm{B}), \qquad \exists \alpha \in \mathsf{Ref}(\rho) : \left|\beta\right)_{\mathrm{B}} = \mathscr{E} \left|\alpha\right)_{\mathrm{A}}.$$

Finally, we define ideal compression.

Definition 79 (Ideal compression)

A compression protocol is *ideal* if and only if it is both lossless and efficient [133].

For a more detailed discussion of compression protocols in generic OPTs—including, for instance, a generalisation of the notion of informational *entropy* to general theories of information processing—we refer the reader to Refs. [274–276].

D Observation-compatibility in CT

We here show a formal proof of the fact that CT satisfies observation-compatibility. This is a refined argument of the one presented in Ref. [98].

Consider two observation-instruments $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}\,, [\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ for a generic system A of the theory. The most generic form they can take is

$$[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}} = [\![\mathrm{p}_0^0\,(0| + \cdots + \mathrm{p}_{D_\mathrm{A}}^0\,(\,D_\mathrm{A}\,|, \ldots, \mathrm{p}_0^m\,(0| + \cdots + \mathrm{p}_{D_\mathrm{A}}^m\,(\,D_\mathrm{A}\,|]\!] \in \mathsf{Obs}(\mathrm{A})\,,$$
$$[\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}} = \left[\!\!\left[\mathrm{q}_0^0\,(0| + \cdots + \mathrm{q}_{D_\mathrm{A}}^0\,(\,D_\mathrm{A}\,|, \ldots, \mathrm{q}_0^k\,(0| + \cdots + \mathrm{q}_{D_\mathrm{A}}^k\,(\,D_\mathrm{A}\,|\right]\!\!\right] \in \mathsf{Obs}(\mathrm{A})\,,$$

where $\left\{\mathrm{p}_j^i\right\}_{i=0}^m$ and $\left\{\mathrm{q}_j^i\right\}_{i=0}^k$ are probability distribution for all $j = 0, \ldots, D_\mathrm{A}$, and $[\![\mathrm{i}]\!]_{i=0}^{D_\mathrm{A}} \in \mathsf{Obs}(\mathrm{A})$ denotes the observation-instrument that jointly perfectly discriminates the pure states of the system.

To prove compatibility, let us construct the observation-instrument

$$\left[\!\!\left[\mathrm{p}_j^i\,\mathrm{q}_j^{i'}\,\,(\mathrm{j}|\right]\!\!\right]_{i=0,\ldots,m;\; i'=0,\ldots,k;\; j=0,\ldots,D_\mathrm{A}} \in \mathsf{Obs}(\mathrm{A})\,.$$

One can verify directly that this is a valid observation-instrument of CT. Moreover, coarse-graining over the outcomes (i', j) yields $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$, while coarse-graining over the outcomes (i, j) yields $[\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}}$. Hence, $[\![\mathrm{a}_x]\!]_{x\in\mathsf{X}}$ and $[\![\mathrm{b}_y]\!]_{y\in\mathsf{Y}}$ are compatible.

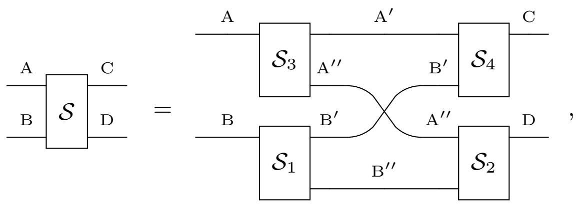
E

Proof of Lemma 25

We report in this appendix the full proof of Lemma 25, detailing all intermediate steps. The argument builds on successive applications of Lemma 24.

Lemma 25 (Permutations on bipartite systems)

In every symmetric OPT with unique decomposition, for any permutation $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AB} \to \mathrm{CD})$ there exist systems A', B', A'', B'' and reversible transformations $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$ such that



(5.2)

where A, B are arbitrary systems of the theory, and C, D are systems such that CD has the same decomposition into elementary systems as AB. In general, any of A, B, C, D may be the trivial system, and the same holds for $\mathrm{A}', \mathrm{A}'', \mathrm{B}', \mathrm{B}''$ [158].

Proof. Let us start by considering the decomposition of AB into its elementary subsystems:

$$\mathrm{A}_1, \ldots, \mathrm{A}_n, \mathrm{B}_1, \ldots, \mathrm{B}_m.$$

The action of $\mathcal{S}$ is to permute its input systems (Lemma 24):

$$\mathrm{A}_1, \ldots, \mathrm{A}_n, \mathrm{B}_1, \ldots, \mathrm{B}_m$$

$$\downarrow \mathcal{S}$$

$$\sigma(\mathrm{A}_1), \ldots, \sigma(\mathrm{A}_n), \sigma(\mathrm{B}_1), \ldots, \sigma(\mathrm{B}_m) = \mathrm{C}_1, \ldots, \mathrm{C}_l, \mathrm{D}_1, \ldots, \mathrm{D}_k,$$

where, thanks to the hypothesis of uniqueness of decomposition, it holds that $\sigma(\mathrm{A}_1) = \mathrm{C}_1$, and so on. We highlight that the permutation is acting on all the elementary subsystems of AB. Therefore, it is possible that $\sigma(\mathrm{B_i}) = \mathrm{C_j}$ for some i, j.

If we now define $N = \{1, \ldots, n\}$, $M = \{1, \ldots, m\}$, $L = \{1, \ldots, l\}$, $K = \{1, \ldots, k\}$, the most general action of $\mathcal{S}$ can be described as

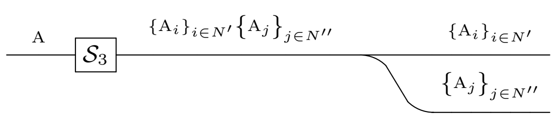
$$\{\mathrm{A}_i\}_{i \in N'} \xrightarrow{\mathcal{S}} \{\mathrm{C}_i\}_{i \in L'},$$
$$\{\mathrm{A}_j\}_{j \in N''} \xrightarrow{\mathcal{S}} \{\mathrm{D}_j\}_{j \in K'},$$

where $N = N' \cup N''$, $|N'| = |L'|$, $|N''| = |K'|$. Analogously, for B we have

$$\{\mathrm{B}_i\}_{i \in M'} \xrightarrow{\mathcal{S}} \{\mathrm{C}_i\}_{i \in L''},$$
$$\{\mathrm{B}_j\}_{j \in M''} \xrightarrow{\mathcal{S}} \{\mathrm{D}_j\}_{j \in K''},$$

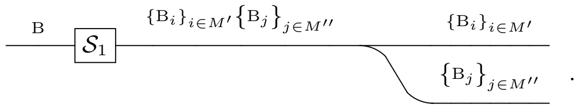
where $M = M' \cup M''$, $|M'| = |L''|$, $|M''| = |K''|$, and $L = L' \cup L''$, $K = K' \cup K''$. Here $|S|$ denotes the cardinality of the set S.

We now show that this permutation can always be realised by a transformation of the same form as in (5.2). First, observe that for system A one can always find a permutation that reorganises the subsystems so that those mapped into states of C appear on top, and those mapped into D appear on the bottom:

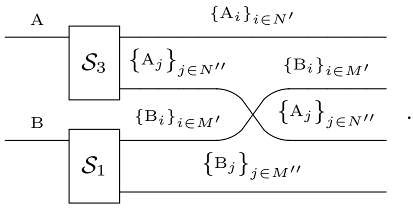


,

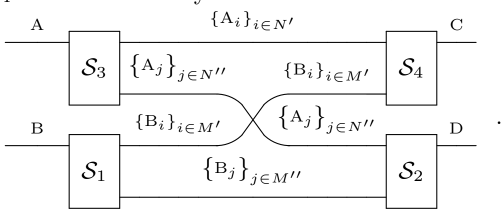
where for the moment the ordering within the subsets $\{\mathrm{A}_i\}_{i \in N'}$ and $\{\mathrm{A}_j\}_{j \in N''}$ is not important. The same reasoning applies to B:



Next, we must exchange the subsystems of A that are mapped into D with those of B that are mapped into C. This is achieved by swapping $\{A_j\}_{j\in N''}$ with $\{B_i\}_{i\in M'}$:



Finally, $\mathcal{S}_2$ and $\mathcal{S}_4$ reorder the subsystems to obtain C and D:



Therefore, we have shown that for any permutation $\mathcal{S}$, it is always possible to find a decomposition of the form (5.2) that achieves the same permutation of elementary systems. Since permutations are completely characterised by how they permute their input systems, the equality between the two transformations follows (Lemma 24). □

F Details of the proof of Theorem 41

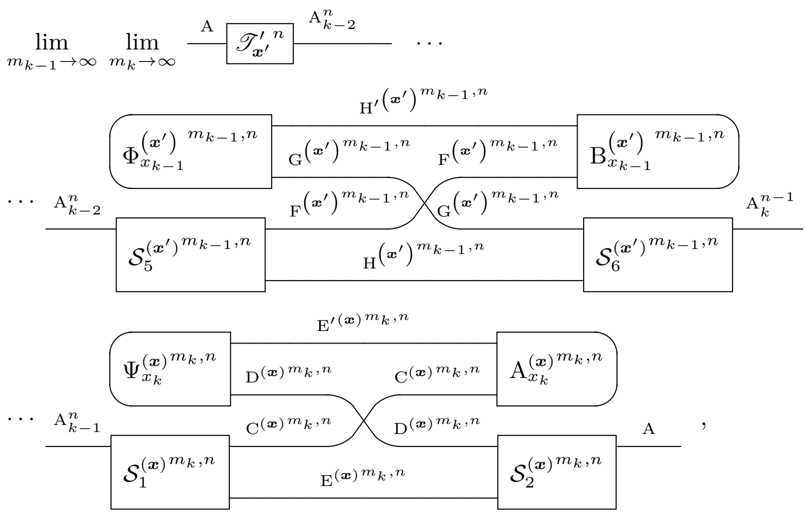
In this appendix we provide a detailed analysis of the structure of the single transformations composing the instruments in the sequence (6.5). Our goal is to prove that whenever the coarse-graining of such transformations yields the identity, as in (6.6), each transformation must necessarily be proportional to the identity itself. This establishes that the identity is an atomic map.

Since we are interested here in the study of transformations, in the following we fix an arbitrary outcome of the conditional instrument, namely

$$\left(\boldsymbol{x}', x_{k-1}, x_k\right) \in \boldsymbol{X}' \times \mathsf{X}_{k-1} \times \mathsf{X}_k.$$

We highlighted in the list of outcomes the last two, x_{k-1} and x_k, since these correspond to the final two transformations in the sequential composition that will be the focus of our analysis.

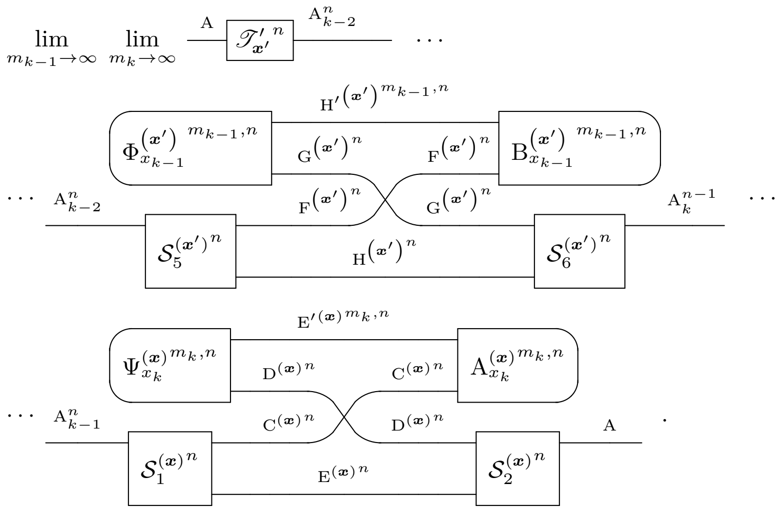
The general form of these transformations is



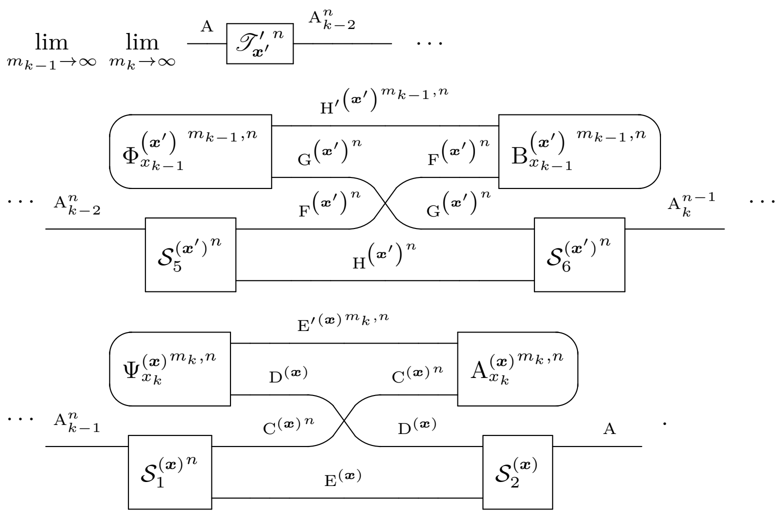
(6.15)

where the limits with respect to m_{k-1} and m_k appear because the transformations composing the instruments of (6.5) are, by construction, those of MOPTs. Consequently, they may themselves be limits of sequences of transformations of the form (5.10).

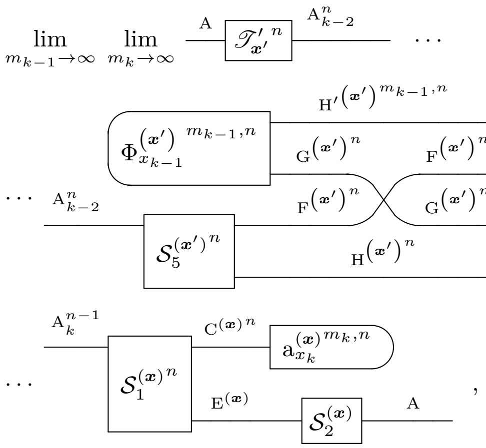
Given Corollary 6, one can already eliminate the dependence of many of the components on the indices m_{k-1} and m_k. The resulting structure takes the form



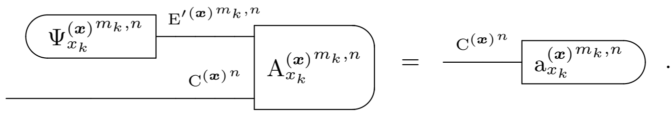
Next, by following the same reasoning used in the derivation of (6.7), one can further eliminate the dependence on the index n from the final transformation in the composition. Explicitly, the structure reduces to



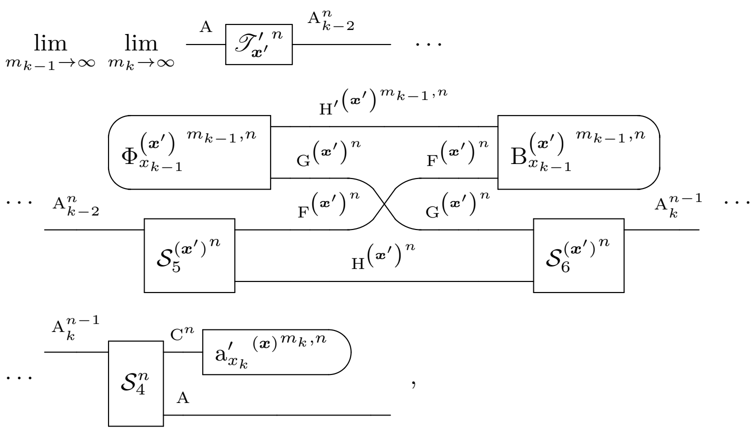
But we have already proved that $\mathrm{D}^{(\boldsymbol{x})} = \mathrm{I}$. Consequently, the overall transformation simplifies to



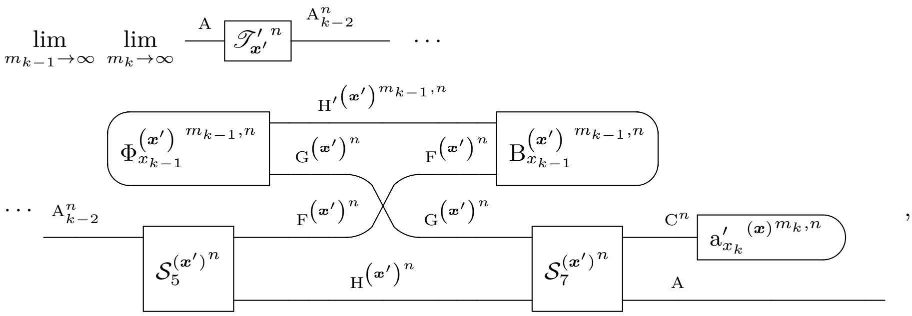
where



The latter equation, following the same steps that lead to (6.11), takes the form



or equivalently



(F.1)

where, as in (6.11), the local permutation on the system C^n has been absorbed into the effect $a_{x_k}^{(\boldsymbol{x})^{m_k,n}}$.

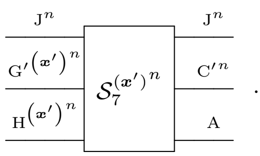
At this stage it is also clear why we chose to preserve the explicit outcome dependence in the deterministic effect $e(\boldsymbol{x})$ acting on C^n. While the deterministic effect itself is independent of the outcome, the associated observation-instrument

$$\left[\!\left[a'^{(\boldsymbol{x})^{m_k,n}}_{x_k} \right]\!\right]_{x_k \in \mathsf{X}_k} \in \mathsf{Obs}(C^n)$$

does retain such dependence.

We have now reached a delicate point, the one which, in the case of the coarse-grained transformation, ultimately leads to (6.12). Since we cannot decompose the observation-instrument in the same way as we did previously for the deterministic effect, the situation is slightly more involved here. Nevertheless, with some care the problem can still be addressed.

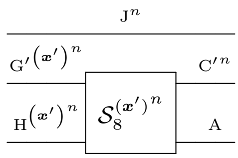
Let us therefore start again by splitting the systems, $G^{(\boldsymbol{x}')^n} = J^n\, G'^{(\boldsymbol{x}')^n}$ and $C^n = J^n\, C'^n$[1], and then study the corresponding permutation



Given that permutations are completely characterised by how they permute their input and output systems (Lemma 24), if two different permutations act

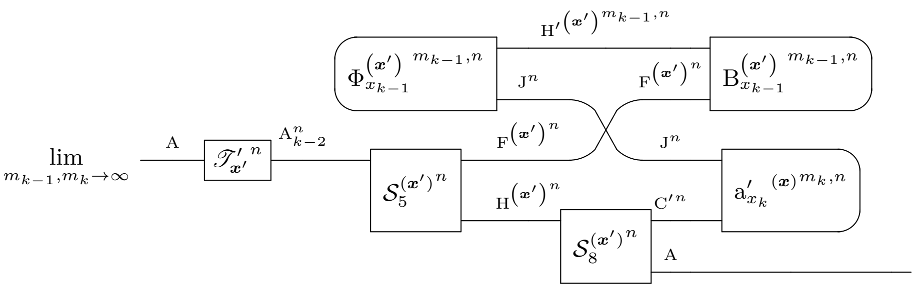
[1] We recall that the system J^n is defined only up to a local permutation, which can always be absorbed into the observation-instrument.

in the same way on their input and output wires, then they represent the same transformation. Hence, an equivalent transformation to $\mathcal{S}_7^{(\boldsymbol{x}')^n}$ is



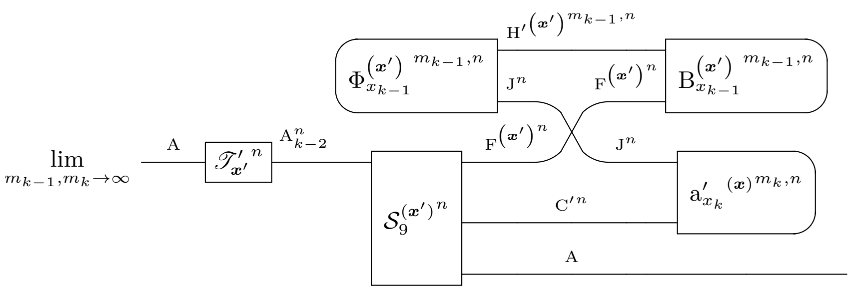
,

which, when substituted into (F.1), yields



,

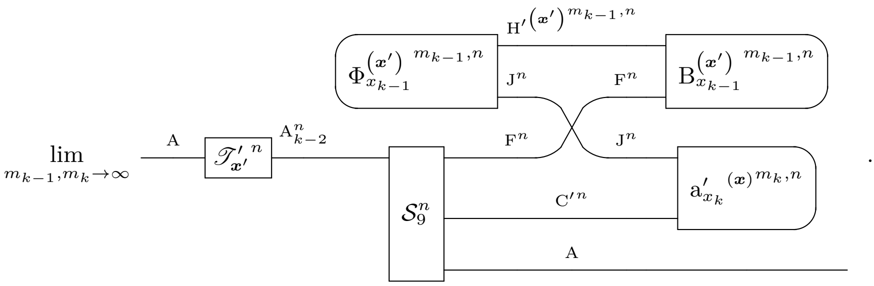
or, equivalently,



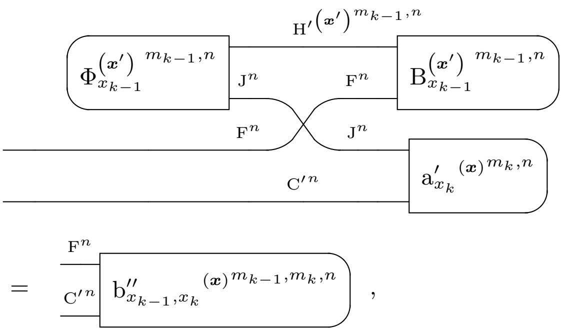
,

where, in addition, we have exploited the fact that one can always find a subsequence for which $G'^{(\boldsymbol{x}')^n} = I$ (6.12).

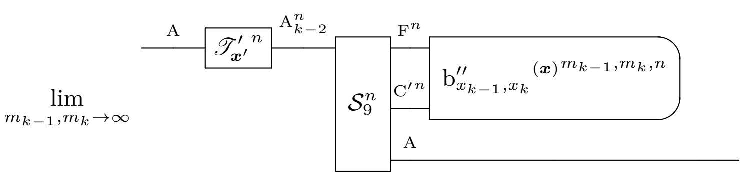
Following the same procedure as before, one can now remove the dependence on the outcomes $\boldsymbol{x}'$ both from the permutation and from the system $F^{(\boldsymbol{x}')^n}$. The transformation thus becomes



Exploiting then the fact that

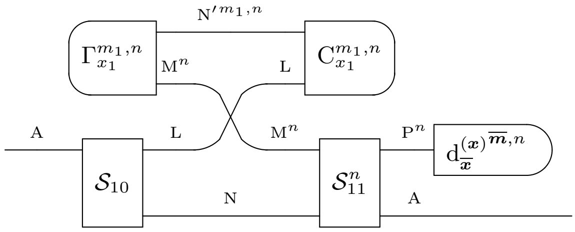


for a suitable observation-instrument, one arrives at

 , (F.2)

which is still of the form (F.1), as can be seen by expanding $\mathscr{T}'^{\,n}_{\boldsymbol{x}'}$ and isolating the $(k-2)$-th conditional step. Hence, the procedure followed from (F.1) can be iterated.

In particular, by stopping just before the final step, one obtains

$\lim_{m_1 \to \infty} \lim_{\overline{\boldsymbol{m}} \to \infty}$ 

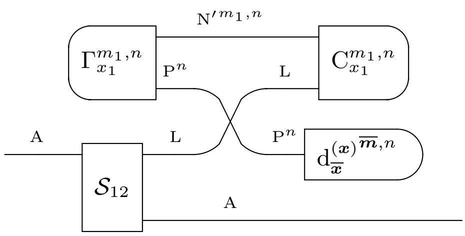
,

where $\overline{\boldsymbol{m}} = (m_2, \ldots, m_k)$, $\overline{\boldsymbol{x}} = (x_2, \ldots, x_k)$, $\mathcal{S}_{11}^n$ is a suitable permutation, P^n is a suitable system, and

$$\left[\!\!\left[\left\{ \mathrm{d}_{\overline{\boldsymbol{x}}}^{(\boldsymbol{x})\,\overline{\boldsymbol{m}},n} \right\}_{\overline{\boldsymbol{x}} \in \overline{\mathsf{X}} = \mathsf{X}_2 \times \cdots \times \mathsf{X}_k} \right]\!\!\right] \in \mathsf{Obs}(\mathrm{P}^n)$$

is a suitable observation-instrument. Here the dependence on m_1 has already been removed wherever possible, and we assume we are working within the subsequence where the leftmost permutation (together with its systems) is fixed (6.14).

Since the effect $\mathrm{d}_{\overline{\boldsymbol{x}}}^{(\boldsymbol{x})\,\overline{\boldsymbol{m}},n}$ must be completely absorbed by the state $\Gamma_{x_1}^{m_1,n}$ (otherwise it would be impossible to decompose the identity), it follows that $\mathrm{M}^n = \mathrm{P}^n$. Then, exploiting again the fact that permutations are completely characterised by their action on the system wires, and regrouping all possible components, the transformation reduces to

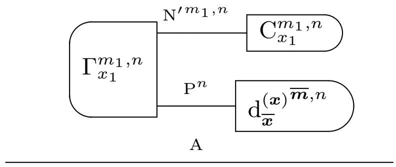
$\lim_{m_1 \to \infty} \lim_{\overline{\boldsymbol{m}} \to \infty}$ 

. (F.3)

We highlight that $\mathcal{S}_{12}$ does not depend on n, since it was obtained by composing $\mathcal{S}_{10}$ with a permutation $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{N} \to \mathrm{A})$, and neither N nor A depend on the index n of the sequence.

Recalling now that a necessary condition to decompose the identity is that $\mathrm{L} = \mathrm{I}$ (6.14), we obtain that any transformation composing the instrument (6.5) must be of the form

$$\lim_{m_1\to\infty}\ \lim_{\overline{\boldsymbol{m}}\to\infty}$$



, (6.16)

which is precisely the desired result, namely an instrument given by the identity rescaled by a probability distribution. This is sufficient to conclude the proof, as done in the main text.

Acknowledgements

THERE are many people without whom this thesis would never have been possible, and I am profoundly grateful for their support and encouragement throughout these past three years.

My deepest thanks go to Prof. Paolo Perinotti and Dr. Alessandro Tosini, whose supervision has been invaluable. Under your guidance, I was able to grow as a researcher, far beyond what I could have ever hoped for. Thank you for the trust you placed in me and for the opportunities you offered, enabling me to work on stimulating projects, to travel, and to collaborate with colleagues around the world.

I am also sincerely indebted to Prof. Robert W. Spekkens and Dr. Elie Wolfe for welcoming me during my six-month visit at the Perimeter Institute for Theoretical Physics. Your hospitality and mentorship made me feel part of your group from the very beginning. You helped me to develop into a more rounded and complete researcher.

I owe particular gratitude to Dr. Marco Erba, whose guidance and collaboration have been fundamental during these years. His introduction to the framework of Operational Probabilistic Theories changed my research path. I will always remember the moment during his lectures in which I decided that this would be what I want to work on in the years to come.

I am thankful to Prof. Teiko Heinosaari and Dr. Ravi Kunjwal for the time and attention dedicated to evaluating my thesis. Their observations have helped improve the clarity and rigour of the work.

I am further grateful to Prof. Daniela Rebuzzi, coordinator of the doctoral school, for her tireless dedication to the PhD program and for the support and guidance she has always extended to me. As a student representative, I had the privilege of witnessing first-hand her commitment to the program, and I will never be grateful enough. I also warmly thank Prof. Claudio Dappiaggi, vice-coordinator of the PhD program, whose thoughtful advice and suggestions have been invaluable to my growth and development throughout this journey.

I would also like to thank Prof. Giulia Mancini for all the support and for having taught me how to navigate better the academic environment. I would also like to thank her in conjunction with Prof. Andrea Negri for the trust they put in me, making myself their teaching assistant.

My journey would not have been the same without the support of some incredible researcher that I had the luck to meet during my Ph.D., Dr. Francesca Brero, Dr. Jacopo Surace, Dr. Giacomo Guarnieri, Dr. Leonardo Vaglini, and Dr. Giovanni Chesi, who has also been my philosophy and literature teacher for the past three years. Thank you for all the beauty you made me discover.

I would also like to thank the Blanceflor Boncompagni Ludovisi Foundation, née Bildt, for its support of my research period at the Perimeter Institute, alongside the Erasmus+ Traineeship programme.

I wish to express my gratitude to the research groups that became my academic homes over the years. In particular, I thank the QUIT group in Pavia, together with members of the biomedical physics group, as well as the Foundations group at the Perimeter Institute, whose support and companionship made this period especially meaningful. I am also grateful to the Foundations group of ICTQT in Gdańsk, who consistently made me feel welcome whenever I visited.

Last, but not least, I would like to thank my family, friends, and those close to me, who have followed me throughout this journey, supporting and encouraging me at every step.

List of publications

Published

- D. Rolino, M. Erba, A. Tosini, and P. Perinotti. “Minimal operational theories: Classical theories with quantum features”. New Journal of Physics (2025)
- M. Erba, P. Perinotti, D. Rolino, and A. Tosini. “Measurement incompatibility is strictly stronger than disturbance”. Physical Review A **109**, 022239 (2024)

Pre-print

- D. Rolino, P. Perinotti, and A. Tosini. “Quantum complementarity” (2025). arXiv:2510.10800

Bibliography

[1] M. Planck. "On an Improvement of Wien's Equation for the Spectrum". In D. TER Haar, editor, The Old Quantum Theory. Pages 79–81. The Commonwealth and International Library: Selected Readings in Physics. Pergamon (1967). Originally published in Verh. Dtsch. Phys. Ges. Berlin 2, 202 (1900).

[2] M. Planck. "On the Theory of the Energy Distribution Law of the Normal Spectrum". In D. TER Haar, editor, The Old Quantum Theory. Pages 82–90. The Commonwealth and International Library: Selected Readings in Physics. Pergamon (1967). Originally published in Verh. Dtsch. Phys. Ges. Berlin 2, 237 (1900).

[3] N. E. Bomark and R. Renstrøm. "The ultraviolet myth". PoS **EPS-HEP2023**, 660 (2024).

[4] A. Einstein. "über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt". Annalen der Physik **322**, 132–148 (1905).

[5] N. Bohr. "On the constitution of atoms and molecules, Part I". The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **26**, 1–25 (1913).

[6] A. H. Compton. “A Quantum Theory of the Scattering of X-rays by Light Elements”. Physical Review **21**, 483–502 (1923).

[7] W. Gerlach and O. Stern. “Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld”. Zeitschrift für Physik **9**, 349–352 (1922).

[8] W. Heisenberg. “Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen.”. Zeitschrift für Physik **33**, 879–893 (1925).

[9] M. Born and P. Jordan. “Zur Quantenmechanik”. Zeitschrift für Physik **34**, 858–888 (1925).

[10] M. Born, W. Heisenberg, and P. Jordan. “Zur Quantenmechanik. II.”. Zeitschrift für Physik **35**, 557–615 (1926).

[11] E. Schrödinger. “An Undulatory Theory of the Mechanics of Atoms and Molecules”. Physical Review **28**, 1049–1070 (1926).

[12] A. Einstein, B. Podolsky, and N. Rosen. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?”. Physical Review **47**, 777–780 (1935).

[13] W. Heisenberg. “Physics and philosophy: The revolution in modern science”. Prometheus Books. (1958).

[14] H. Everett. “On the foundations of quantum mechanics”. PhD thesis. Princeton University. (1957).

[15] H. Everett. “"Relative State" Formulation of Quantum Mechanics”. Reviews of Modern Physics **29**, 454–462 (1957).

[16] G. C. Ghirardi, A. Rimini, and T. Weber. “Unified dynamics for microscopic and macroscopic systems”. Physical Review D **34**, 470–491 (1986).

[17] D. Bohm. “A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I”. Physical Review **85**, 166–179 (1952).

[18] E. P. Wigner. “Remarks on the mind-body question”. In J. Mehra, editor, Philosophical Reflections and Syntheses. Pages 247–260. Springer Berlin Heidelberg, Berlin, Heidelberg (1995).

[19] M. Pusey. “Is QBism 80% complete, or 20%?” (2016). YouTube video.

[20] Č. Brukner. “A No-Go Theorem for Observer-Independent Facts”. Entropy **20**, 350 (2018).

[21] D. Frauchiger and R. Renner. "Quantum theory cannot consistently describe the use of itself". Nature Communications **9**, 3711 (2018).

[22] S. Gao. "Quantum theory is incompatible with relativity: A new proof beyond Bell's theorem and a test of unitary quantum theories" (2019). PhilSci-Archive:16155.

[23] M. Leifer. "What Are Copenhagenish Interpretations and Should They Be Perspectival?" (2019). YouTube video.

[24] K.-W. Bong, A. Utreras-Alarcón, F. Ghafari, Y.-C. Liang, N. Tischler, E. G. Cavalcanti, G. J. Pryde, and H. M. Wiseman. "A strong no-go theorem on the Wigner's friend paradox". Nature Physics **16**, 1199–1205 (2020).

[25] P. Allard Guérin, V. Baumann, F. Del Santo, and Č. Brukner. "A no-go theorem for the persistent reality of Wigner's friend's perception". Communications Physics **4**, 1–7 (2021).

[26] D. Schmid, Y. Yīng, and M. Leifer. "A review and analysis of six extended Wigner's friend arguments" (2024). arXiv:2308.16220.

[27] J. S. Bell. "On the Einstein Podolsky Rosen paradox". Physics Physique Fizika **1**, 195–200 (1964).

[28] J. S. Bell. "On the Problem of Hidden Variables in Quantum Mechanics". Reviews of Modern Physics **38**, 447–452 (1966).

[29] J. S. Bell. "Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy". Cambridge University Press. Cambridge (2004). 2 edition.

[30] D. Bohm. "Quantum theory". Dover Books on Physics Series. Dover Publications. (1951).

[31] R. P. Feynman. "Simulating physics with computers". International Journal of Theoretical Physics **21**, 467–488 (1982).

[32] D. Deutsch. "Quantum theory, the Church–Turing principle and the universal quantum computer". Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **400**, 97–117 (1985).

[33] P. W. Shor. "Algorithms for quantum computation: Discrete logarithms and factoring". In Proceedings 35th Annual Symposium on Foundations of Computer Science. Pages 124–134. (1994).

[34] L. K. Grover. “A fast quantum mechanical algorithm for database search”. In Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing. Pages 212–219. Stoc ’96Philadelphia, Pennsylvania, USA and New York, NY, USA (1996). Association for Computing Machinery.

[35] A. Acín, N. Gisin, and L. Masanes. “From Bell’s Theorem to Secure Quantum Key Distribution”. Physical Review Letters **97**, 120405 (2006).

[36] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. “Quantum entanglement”. Reviews of Modern Physics **81**, 865–942 (2009).

[37] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. “Random numbers certified by Bell’s theorem”. Nature **464**, 1021–1024 (2010).

[38] M. A. Nielsen and I. L. Chuang. “Quantum Computation and Quantum Information: 10th Anniversary Edition”. Cambridge University Press. Cambridge (2010). 10th anniversary edition edition.

[39] C. H. Bennett and G. Brassard. “Quantum cryptography: Public key distribution and coin tossing”. Theoretical Computer Science **560**, 7–11 (2014).

[40] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. “Proposed Experiment to Test Local Hidden-Variable Theories”. Physical Review Letters **23**, 880–884 (1969).

[41] S. Kochen and E. P. Specker. “The problem of hidden variables in quantum mechanics”. In C. A. Hooker, editor, The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution. Pages 293–328. Springer Netherlands, Dordrecht (1975). Originally published in Journal of Mathematics and Mechanics 17, 59–87 (1967). (Available at: `http://www.jstor.org/stable/24902153`).

[42] D. M. Greenberger, M. A. Horne, and A. Zeilinger. “Going beyond bell’s theorem”. In M. Kafatos, editor, Bell’s Theorem, Quantum Theory and Conceptions of the Universe. Pages 69–72. Springer Netherlands, Dordrecht (1989).

[43] N. D. Mermin. “Simple unified form for the major no-hidden-variables theorems”. Physical Review Letters **65**, 3373–3376 (1990).

[44] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger. “Bell’s theorem without inequalities”. American Journal of Physics **58**, 1131–1143 (1990).

[45] L. Hardy. "Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories". Physical Review Letters **68**, 2981–2984 (1992).

[46] N. D. Mermin. "Hidden variables and the two theorems of John Bell". Reviews of Modern Physics **65**, 803–815 (1993).

[47] L. Hardy. "Nonlocality for two particles without inequalities for almost all entangled states". Physical Review Letters **71**, 1665–1668 (1993).

[48] R. W. Spekkens. "Contextuality for preparations, transformations, and unsharp measurements". Physical Review A **71**, 052108 (2005).

[49] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. "Bell nonlocality". Reviews of Modern Physics **86**, 419–478 (2014).

[50] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres". Nature **526**, 682–686 (2015).

[51] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger. "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons". Physical Review Letters **115**, 250401 (2015).

[52] C. J. Wood and R. W. Spekkens. "The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning". New Journal of Physics **17**, 033002 (2015).

[53] R. W. Spekkens. "Noncontextuality: How we should define it, why it is natural, and what to do about its failure" (2017). PIRSA:17070035.

[54] M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín, and M. Navascués. "Quantum theory based on real numbers can be experimentally falsified". Nature **600**, 625–629 (2021).

[55] W. Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". Zeitschrift für Physik **43**, 172–198 (1927).

[56] W. Heisenberg. "The physical principles of the quantum theory.". University of Chicago Press. Chicago (1930).

[57] A. E. Ruark. “The Limits of Accuracy in Physical Measurements”. Proceedings of the National Academy of Sciences **14**, 322–328 (1928).

[58] J. von Neumann. “Mathematical foundations of quantum mechanics”. Goldstine Printed Materials. Princeton University Press. (1955).

[59] A. Messiah. “Quantum mechanics”. Number v. 2 in Quantum Mechanics. Elsevier Science. (1961).

[60] W. E. Lamb, Jr. “An operational interpretation of nonrelativistic quantum mechanics”. Physics Today **22**, 23–28 (1969).

[61] M. Jammer. “The philosophy of quantum mechanics”. Wiley. New York, (1974).

[62] C. Roychoudhuri. “Heisenberg’s microscope—A misleading illustration”. Foundations of Physics **8**, 845–849 (1978).

[63] M. Jammer. “The conceptual development of quantum mechanics”. History of Modern Physics and Astronomy Series. Tomash Publishers. (1989).

[64] L. D. Landau and E. M. Lifshits. “Quantum Mechanics: Non-Relativistic Theory”. Volume v.3 of Course of Theoretical Physics. Butterworth-Heinemann. Oxford (1991).

[65] V. B. Braginsky, F. Y. Khalili, and K. S. Thorne. “Quantum Measurement”. Cambridge University Press. Cambridge (1992).

[66] M. J. W. Hall. “Noise-dependent uncertainty relations for the harmonic oscillator”. Physical Review A **49**, 42–47 (1994).

[67] M. Beller. “Quantum dialogue: The making of a revolution”. University of Chicago Press. (1999).

[68] W. M. de Muynck. “Preparation and Measurement: Two Independent Sources of Uncertainty in Quantum Mechanics”. Foundations of Physics **30**, 205–225 (2000).

[69] A. Peres, editor. “Quantum Theory: Concepts and Methods”. Springer Netherlands. Dordrecht (2002).

[70] G. M. D’Ariano. “On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement”. Fortschritte der Physik **51**, 318–330 (2003).

[71] D. Saha, M. Oszmaniec, L. Czekaj, M. Horodecki, and R. Horodecki. "Operational foundations for complementarity and uncertainty relations". Physical Review A **101**, 052104 (2020).

[72] M. Ozawa. "Controlling quantum state reductions". Physics Letters A **282**, 336–342 (2001).

[73] M. Ozawa. "Uncertainty relations for joint measurements of noncommuting observables". Physics Letters A **320**, 367–374 (2004).

[74] M. Ozawa. "Uncertainty relations for noise and disturbance in generalized quantum measurements". Annals of Physics **311**, 350–416 (2004).

[75] P. Busch and C. Shilladay. "Complementarity and uncertainty in Mach–Zehnder interferometry and beyond". Physics Reports **435**, 1–31 (2006).

[76] P. Busch, T. Heinonen, and P. Lahti. "Heisenberg's uncertainty principle". Physics Reports **452**, 155–176 (2007).

[77] P. Busch, P. Lahti, and R. F. Werner. "Proof of Heisenberg's Error-Disturbance Relation". Physical Review Letters **111**, 160405 (2013).

[78] H. P. Robertson. "The Uncertainty Principle". Physical Review **34**, 163–164 (1929).

[79] E. U. Condon. "Remarks on Uncertainty Principles". In A. O. Barut, H. Odabasi, and A. van der Merwe, editors, Selected Scientific Papers of E.U. Condon. Pages 93–96. Springer, New York, NY (1991). Originnaly published in Science LXIX, 573 (1929).

[80] E. Schrödinger. "Zum heisenbergschen unschärfeprinzip". Physikalische Zeitschrift **26**, 301–303 (1930).

[81] L. Maccone and A. K. Pati. "Stronger Uncertainty Relations for All Incompatible Observables". Physical Review Letters **113**, 260401 (2014).

[82] D. Deutsch. "Uncertainty in Quantum Measurements". Physical Review Letters **50**, 631–633 (1983).

[83] H. Maassen and J. B. M. Uffink. "Generalized entropic uncertainty relations". Physical Review Letters **60**, 1103–1106 (1988).

[84] P. Busch, P. J. Lahti, and P. Mittelstaedt. "The Quantum Theory of Measurement". Volume 2 of Lecture Notes in Physics Monographs. Springer Berlin Heidelberg. Berlin, Heidelberg (1996).

[85] T. Heinosaari, D. Reitzner, and P. Stano. "Notes on Joint Measurability of Quantum Observables". Foundations of Physics **38**, 1133–1147 (2008).

[86] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens. "Comparing the degrees of incompatibility inherent in probabilistic physical theories". Europhysics Letters **103**, 10002 (2013).

[87] C. Addis, T. Heinosaari, J. Kiukas, E.-M. Laine, and S. Maniscalco. "Dynamics of incompatibility of quantum measurements in open systems". Physical Review A **93**, 022114 (2016).

[88] T. Heinosaari, T. Miyadera, and M. Ziman. "An invitation to quantum incompatibility". Journal of Physics A: Mathematical and Theoretical **49**, 123001 (2016).

[89] T. Heinosaari. "Simultaneous measurement of two quantum observables: Compatibility, broadcasting, and in-between". Physical Review A **93**, 042118 (2016).

[90] C. Carmeli, T. Heinosaari, D. Reitzner, J. Schultz, and A. Toigo. "Quantum Incompatibility in Collective Measurements". Mathematics **4**, 54 (2016).

[91] T. Heinosaari and T. Miyadera. "Incompatibility of quantum channels". Journal of Physics A: Mathematical and Theoretical **50**, 135302 (2017).

[92] S. N. Filippov, T. Heinosaari, and L. Leppäjärvi. "Necessary condition for incompatibility of observables in general probabilistic theories". Physical Review A **95**, 032127 (2017).

[93] F. Buscemi, E. Chitambar, and W. Zhou. "Complete Resource Theory of Quantum Incompatibility as Quantum Programmability". Physical Review Letters **124**, 120401 (2020).

[94] G. M. D'Ariano, P. Perinotti, and A. Tosini. "Incompatibility of observables, channels and instruments in information theories". Journal of Physics A: Mathematical and Theoretical **55**, 394006 (2022).

[95] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens. "Contextuality without Incompatibility". Physical Review Letters **130**, 230201 (2023).

[96] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini. "Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication". Quantum **7**, 1035 (2023).

[97] M. Yamada and T. Miyadera. “Incompatibility of quantum channels in general probabilistic theories”. Physical Review A: Atomic, Molecular, and Optical Physics **110**, 062210 (2024).

[98] M. Erba, P. Perinotti, D. Rolino, and A. Tosini. “Measurement incompatibility is strictly stronger than disturbance”. Physical Review A **109**, 022239 (2024).

[99] M. Plávala, O. Gühne, and M. T. Quintino. “All incompatible measurements on qubits lead to multiparticle bell nonlocality”. Physical Review Letters **134**, 200201 (2025).

[100] N. Bohr. “The Quantum Postulate and the Recent Development of Atomic Theory”. Nature **121**, 580–590 (1928).

[101] N. Bohr. “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?”. Physical Review **48**, 696–702 (1935).

[102] W. Pauli. “General Principles of Quantum Mechanics”. Springer Berlin Heidelberg. Berlin, Heidelberg (1980).

[103] B. Englert. “Fringe Visibility and Which-Way Information: An Inequality”. Physical Review Letters **77**, 2154–2157 (1996).

[104] M. O. Scully, B.-G. Englert, and H. Walther. “Quantum optical tests of complementarity”. Nature **351**, 111–116 (1991).

[105] G. Jaeger, A. Shimony, and L. Vaidman. “Two interferometric complementarities”. Physical Review A **51**, 54–67 (1995).

[106] S. Dürr, T. Nonn, and G. Rempe. “Origin of quantum-mechanical complementarity probed by a ‘which-way’ experiment in an atom interferometer”. Nature **395**, 33–37 (1998).

[107] G. Bimonte and R. Musto. “On interferometric duality in multibeam experiments”. Journal of Physics A: Mathematical and General **36**, 11481 (2003).

[108] J. Martinez-Linares and D. A. Harmin. ““Quality” of a which-way detector”. Physical Review A **69**, 062109 (2004).

[109] U. Schilling and J. von Zanthier. “Phase-dependent which-way information”. Physics Letters A **376**, 3479–3484 (2012).

[110] F. De Zela. “Comment on "Fringe Visibility and Which-Way Information: An Inequality"” (2013). arXiv:1305.0861.

[111] J. Prabhu Tej, A. R. Usha Devi, H. S. Karthik, Sudha, and A. K. Rajagopal. "Quantum which-way information and fringe visibility when the detector is entangled with an ancilla". Physical Review A **89**, 062116 (2014).

[112] K. Kraus. "Complementary observables and uncertainty relations". Physical Review D **35**, 3070–3075 (1987).

[113] P. Busch, M. Grabowski, and P. J. Lahti. "Operational Quantum Physics". Volume 31 of Lecture Notes in Physics Monographs. Springer Berlin Heidelberg. Berlin, Heidelberg (1995).

[114] P. Busch and P. J. Lahti. "The complementarity of quantum observables: Theory and experiments". La Rivista del Nuovo Cimento **18**, 1–27 (1995).

[115] D. Petz. "Complementarity in quantum systems". Reports on Mathematical Physics **59**, 209–224 (2007).

[116] A. Plotnitsky. "What is complementarity?: Niels Bohr and the architecture of quantum theory". Physica Scripta **T163**, 014002 (2014).

[117] C.-Y. Hsieh, R. Uola, and P. Skrzypczyk. "Quantum complementarity: A novel resource for unambiguous exclusion and encryption" (2023). arXiv:2309.11968.

[118] L. Serino, G. Chesi, B. Brecht, L. Maccone, C. Macchiavello, and C. Silberhorn. "Complementarity-based complementarity" (2024). arXiv:2406.11395.

[119] D. Rolino, P. Perinotti, and A. Tosini. "Quantum complementarity" (2025). arXiv:2510.10800.

[120] G. Birkhoff and J. von Neumann. "The Logic of Quantum Mechanics". Annals of Mathematics **37**, 823–843 (1936). arXiv:1968621.

[121] P. J. Lahti. "Characterization of quantum logics". International Journal of Theoretical Physics **19**, 905–923 (1980).

[122] G. Ludwig. "Foundations of Quantum Mechanics". Springer Berlin Heidelberg. Berlin, Heidelberg (1985).

[123] A. Wilce. "Test Spaces and Orthoalgebras". In B. Coecke, D. Moore, and A. Wilce, editors, Current Research in Operational Quantum Logic: Algebras, Categories, Languages. Pages 81–114. Fundamental Theories of Physics. Springer Netherlands, Dordrecht (2000).

[124] L. Hardy. "Quantum Theory From Five Reasonable Axioms" (2001). arXiv:quant-ph/0101012.

[125] C. A. Fuchs. "Quantum Mechanics as Quantum Information (and only a little more)" (2002). arXiv:quant-ph/0205039.

[126] G. Brassard. "Is information the key?". Nature Physics **1**, 2–4 (2005).

[127] G. M. D'Ariano. "Probabilistic theories: What is special about Quantum Mechanics?". In A. Bokulich and G. Jaeger, editors, Philosophy of Quantum Information and Entanglement. Pages 85–126. Cambridge University Press (2010). 1 edition.

[128] R. W. Spekkens. "Evidence for the epistemic view of quantum states: A toy theory". Physical Review A **75**, 032110 (2007).

[129] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Probabilistic theories with purification". Physical Review A **81**, 062348 (2010).

[130] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Informational derivation of quantum theory". Physical Review A **84**, 012311 (2011).

[131] L. Masanes and M. P. Müller. "A derivation of quantum theory from physical requirements". New Journal of Physics **13**, 063001 (2011).

[132] B. Dakić and Č. Brukner. "Quantum Theory and Beyond: Is Entanglement Special?". In H. Halvorson, editor, Deep Beauty: Understanding the Quantum World through Mathematical Innovation. Pages 365–392. Cambridge University Press, Cambridge (2011).

[133] G. M. D'Ariano, G. Chiribella, and P. Perinotti. "Quantum Theory from First Principles: An Informational Approach". Cambridge University Press. (2016). 1 edition.

[134] P. Perinotti. "Cellular automata in operational probabilistic theories". Quantum **4**, 294 (2020).

[135] A. S. Holevo. "Bounds for the quantity of information transmitted by a quantum communication channel". Rossiiskaya Akademiya Nauk. Problemy Peredachi Informatsii **9**, 3–11 (1973).

[136] W. K. Wootters and W. H. Zurek. "A single quantum cannot be cloned". Nature **299**, 802–803 (1982).

[137] S. Wiesner. "Conjugate coding". SIGACT News **15**, 78–88 (1983).

[138] C. H. Bennett. "Quantum cryptography using any two nonorthogonal states". Physical Review Letters **68**, 3121–3124 (1992).

[139] C. H. Bennett and S. J. Wiesner. “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states”. Physical Review Letters **69**, 2881–2884 (1992).

[140] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”. Physical Review Letters **70**, 1895–1899 (1993).

[141] L. Hardy. “Disentangling Nonlocality and Teleportation” (1999). arXiv:quant-ph/9906123.

[142] G. Chiribella, G. M. D’Ariano, and P. Perinotti. “Quantum from Principles”. In G. Chiribella and R. W. Spekkens, editors, Quantum Theory: Informational Foundations and Foils. Volume 181, pages 171–221. Springer Netherlands, Dordrecht (2016).

[143] M. Erba, P. Perinotti, and G. M. D’Ariano. “Categorical Physics I – Review of a constructive framework for Operational Probabilistic Theories”. Forthcoming (2026).

[144] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. “Cloning and Broadcasting in Generic Probabilistic Theories” (2006). arXiv:quant-ph/0611295.

[145] J. Barrett. “Information processing in generalized probabilistic theories”. Physical Review A **75**, 032304 (2007).

[146] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. “Generalized No-Broadcasting Theorem”. Physical Review Letters **99**, 240501 (2007).

[147] L. Hardy and W. K. Wootters. “Limited Holism and Real-Vector-Space Quantum Theory”. Foundations of Physics **42**, 454–473 (2012).

[148] P. Janotta and R. Lal. “Generalized probabilistic theories without the no-restriction hypothesis”. Physical Review A **87**, 052131 (2013).

[149] L. Lami. “Non-classical correlations in quantum mechanics and beyond”. PhD thesis. Universitat Autònoma de Barcelona. (2018).

[150] J. Barrett, N. De Beaudrap, M. J. Hoban, and C. M. Lee. “The computational landscape of general physical theories”. npj Quantum Information **5**, 41 (2019).

[151] M. Plávala. “General probabilistic theories: An introduction”. Physics Reports **1033**, 1–64 (2023).

[152] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens. "Characterization of Noncontextuality in the Framework of Generalized Probabilistic Theories". PRX Quantum **2**, 010331 (2021).

[153] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens. "A structure theorem for generalized-noncontextual ontological models". Quantum **8**, 1283 (2024).

[154] S. Abramsky and B. Coecke. "A categorical semantics of quantum protocols". In Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, 2004. Pages 415–425. (2004).

[155] B. Coecke. "Kindergarten Quantum Mechanics: Lecture Notes". AIP Conference Proceedings **810**, 81–98 (2006).

[156] B. Coecke and A. Kissinger. "Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning". Cambridge University Press. Cambridge (2017).

[157] A. Kissinger and J. van de Wetering. "Picturing quantum software: An introduction to the ZX-calculus and quantum compilation". Preprint. (2024). url: `https://github.com/zxcalc/book`.

[158] D. Rolino, M. Erba, A. Tosini, and P. Perinotti. "Minimal operational theories: Classical theories with quantum features". New Journal of Physics (2025).

[159] G. M. D'Ariano, M. Erba, and P. Perinotti. "Classical theories with entanglement". Physical Review A **101**, 042118 (2020).

[160] R. W. Spekkens. "The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein" (2019). arXiv:1909.04628,.

[161] P. W. Bridgman. "The logic of modern physics". Macmillan Paperbacks. Macmillan. (1927).

[162] J. M. Keynes. "A treatise on probability". Dover Publications. London, England (1921).

[163] R. Carnap. "Logical foundations of probability". Chicago University of Chicago Press. Chicago] (1950).

[164] R. T. Cox. "Algebra of probable inference". The Johns Hopkins Press. (1961).

[165] E. T. Jaynes. "Probability Theory: The Logic of Science". Cambridge University Press. Cambridge (2003).

[166] J. Pearl. "Causality". Cambridge University Press. Cambridge (2009). 2 edition.

[167] G. M. D'Ariano, M. Erba, and P. Perinotti. "Classicality without local discriminability: Decoupling entanglement and complementarity". Physical Review A **102**, 052216 (2020).

[168] P. Perinotti. "Private comunication".

[169] S. Mac Lane. "Categories for the Working Mathematician". Volume 5 of Graduate Texts in Mathematics. Springer New York. New York, NY (1978).

[170] S. Awodey. "Category Theory". Oxford University Press. (2006).

[171] C. Heunen and J. Vicary. "Categories for Quantum Theory: An Introduction". Oxford University PressOxford. (2019). 1 edition.

[172] M. Erba and P. Perinotti. "The composition rule for quantum systems is not the only possible one" (2024). arXiv:2411.15964.

[173] S. Abramsky and B. Coecke. "Categorical Quantum Mechanics". In K. Engesser, D. M. Gabbay, and D. Lehmann, editors, Handbook of Quantum Logic and Quantum Structures. Pages 261–323. Elsevier, Amsterdam (2009).

[174] "Facts and figures about the LHC". Data retrieved on the official CERN website. `https://home.cern/resources/faqs/facts-and-figures-about-lhc`.

[175] S. Soltani, M. Erba, D. Schmid, and J. H. Selby. "Noncontextual ontological models of operational probabilistic theories" (2025). arXiv:2502.11842,.

[176] D. Centeno, M. Erba, T. D. Galley, D. Schmid, J. H. Selby, R. W. Spekkens, S. Soltani, J. Surace, A. Wilce, and Y. Yīng. "Symmetry-induced failures of tomographic locality: Constructing foil theories by twirling". Physical Review A **112**, L030202 (2025).

[177] D. Schmid, J. H. Selby, and R. W.. Spekkens. "Unscrambling the omelette of causation and inference: The framework of causal-inferential theories" (2021). arXiv:2009.03297.

[178] N. Gisin, J. Bancal, Y. Cai, P. Remy, A. Tavakoli, E. Zambrini Cruzeiro, S. Popescu, and N. Brunner. "Constraints on nonlocality in networks from no-signaling and independence". Nature Communications **11**, 2378 (2020).

[179] J. Bancal and N. Gisin. “Nonlocal boxes for networks”. Physical Review A **104**, 052212 (2021).

[180] G. Ludwig. “An Axiomatic Basis for Quantum Mechanics”. Springer Berlin Heidelberg. Berlin, Heidelberg (1985).

[181] J. Watrous. “On one-dimensional quantum cellular automata”. In Proceedings of IEEE 36th Annual Foundations of Computer Science. Pages 528–537. (1995).

[182] W. van Dam. “Quantum Cellular Automata”. PhD thesis. University of Nijmegen. (1996).

[183] B. Schumacher and R. F. Werner. “Reversible quantum cellular automata” (2004). arXiv:quant-ph/0405174.

[184] J. Gruska. “Quantum computing”. In Wiley Encyclopedia of Computer Science and Engineering. Pages 2297–2310. John Wiley & Sons, Ltd (2009).

[185] P. Arrighi, V. Nesme, and R. Werner. “Unitarity plus causality implies localizability”. Journal of Computer and System Sciences **77**, 372–378 (2011).

[186] D. Gross, V. Nesme, H. Vogts, and R. F. Werner. “Index Theory of One Dimensional Quantum Walks and Cellular Automata”. Communications in Mathematical Physics **310**, 419–454 (2012).

[187] A. Bisio, G. M. D’Ariano, and A. Tosini. “Dirac quantum cellular automaton in one dimension: $\mathit{Zitterbewegung}$ and scattering from potential”. Physical Review A **88**, 032301 (2013).

[188] G. M. D’Ariano and P. Perinotti. “Derivation of the Dirac equation from principles of information processing”. Physical Review A **90**, 062106 (2014).

[189] A. Bisio, G. M. D’Ariano, and A. Tosini. “Quantum field as a quantum cellular automaton: The Dirac free evolution in one dimension”. Annals of Physics **354**, 244–264 (2015).

[190] A. Bisio, G. M. D’Ariano, P. Perinotti, and A. Tosini. “Weyl, Dirac and Maxwell Quantum Cellular Automata”. Foundations of Physics **45**, 1203–1221 (2015).

[191] G. M. D’Ariano, P. Perinotti, and A. Tosini. “Information and disturbance in operational probabilistic theories”. Quantum **4**, 363 (2020).

[192] M. Erba. “Private comunication”.

[193] Z.-D. Li, Y.-L. Mao, M. Weilenmann, A. Tavakoli, H. Chen, L. Feng, S.-J. Yang, M.-O. Renou, D. Trillo, T. P. Le, N. Gisin, A. Acín, M. Navascués, Z. Wang, and J. Fan. "Testing Real Quantum Theory in an Optical Quantum Network". Physical Review Letters **128**, 040402 (2022).

[194] M.-C. Chen, C. Wang, F.-M. Liu, J.-W. Wang, C. Ying, Z.-X. Shang, Y. Wu, M. Gong, H. Deng, F.-T. Liang, Q. Zhang, C.-Z. Peng, X. Zhu, A. Cabello, C.-Y Lu, and J.-W Pan. "Ruling Out Real-Valued Standard Formalism of Quantum Theory". Physical Review Letters **128**, 040403 (2022).

[195] D. Wu, Y.-F. Jiang, X.-M. Gu, L. Huang, B. Bai, Q.-C. Sun, X. Zhang, S.-Q. Gong, Y. Mao, H.-S. Zhong, M.-C. Chen, J. Zhang, Q. Zhang, C.-Y. Lu, and J.-W. Pan. "Experimental Refutation of Real-Valued Quantum Mechanics under Strict Locality Conditions". Physical Review Letters **129**, 140401 (2022).

[196] S. Sarkar, D. Trillo, M. O. Renou, and R. Augusiak. "Gap between quantum theory based on real and complex numbers is arbitrarily large" (2025). arXiv:2503.09724.

[197] Y. Yīng, M. Ciudad Alañón, D. Centeno, J. Surace, M. Maciel Ansanelli, R. Liu, D. Schmid, and R. W. Spekkens. "On whether quantum theory needs complex numbers: The foil theories perspective" (2025). arXiv:2506.08091.

[198] S. B. Bravyi and A. Y. Kitaev. "Fermionic Quantum Computation". Annals of Physics **298**, 210–226 (2002).

[199] G. M. D'Ariano, F. Manessi, P. Perinotti, and A. Tosini. "The Feynman problem and fermionic entanglement: Fermionic theory versus qubit theory". International Journal of Modern Physics A **29**, 1430025 (2014).

[200] G. M. D'Ariano and A. Tosini. "Emergence of space–time from topologically homogeneous causal networks". Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **44**, 294–299 (2013).

[201] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron. "Quantum computations without definite causal structure". Physical Review A **88**, 022318 (2013).

[202] G. Chiribella, G. M. D'Ariano, and P. Perinotti. "Theoretical framework for quantum networks". Physical Review A **80**, 022339 (2009).

[203] A. Bisio and P. Perinotti. "Theoretical framework for higher-order quantum theory". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **475**, 20180706 (2019).

[204] L. Apadula, A. Bisio, and P. Perinotti. "No-signalling constrains quantum computation with indefinite causal structure". Quantum **8**, 1241 (2024).

[205] O. Oreshkov, F. Costa, and Č Brukner. "Quantum correlations with no causal order". Nature Communications **3**, 1092 (2012).

[206] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č Brukner. "Witnessing causal nonseparability". New Journal of Physics **17**, 102001 (2015).

[207] O. Oreshkov and C. Giarmatzi. "Causal and causally separable processes". New Journal of Physics **18**, 093020 (2016).

[208] Ä. Baumeler, A. Feix, and S. Wolf. "Maximal incompatibility of locally classical behavior and global causal order in multiparty scenarios". Physical Review A **90**, 042106 (2014).

[209] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner. "The simplest causal inequalities and their violation". New Journal of Physics **18**, 013008 (2015).

[210] A. Feix, M. Araújo, and Č. Brukner. "Causally nonseparable processes admitting a causal model". New Journal of Physics **18**, 083040 (2016).

[211] A. A. Abbott, C. Giarmatzi, F. Costa, and C. Branciard. "Multipartite causal correlations: Polytopes and inequalities". Physical Review A **94**, 032131 (2016).

[212] T. van der Lugt, J. Barrett, and G. Chiribella. "Device-independent certification of indefinite causal order in the quantum switch" (2023). arXiv:2208.00719.

[213] J. Wechs, C. Branciard, and O. Oreshkov. "Existence of processes violating causal inequalities on time-delocalised subsystems". Nature Communications **14**, 1471 (2023).

[214] M. Planck. "Der kausalbegriff in der physik". Barth. Leipzig (1948).

[215] G. M. D'Ariano, F. Manessi, and P. Perinotti. "Determinism without causality". Physica Scripta **2014**, 014013 (2014).

[216] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens. "Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality". Physical Review A **107**, 062203 (2023).

[217] A. Bluhm, L. Leppäjärvi, and I. Nechita. “On the simulation of quantum multimeters” (2024). arXiv:2402.18333.

[218] L. Leppäjärvi and M. Sedlák. “Postprocessing of quantum instruments”. Physical Review A **103**, 022615 (2021).

[219] R. W. Spekkens. “Quasi-Quantization: Classical Statistical Theories with an Epistemic Restriction”. In G. Chiribella and R. W. Spekkens, editors, Quantum Theory: Informational Foundations and Foils. Pages 83–135. Fundamental Theories of Physics. Springer Netherlands, Dordrecht (2016).

[220] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey. “Uniqueness of Noncontextual Models for Stabilizer Subtheories”. Physical Review Letters **129**, 120403 (2022).

[221] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens. “What is Nonclassical about Uncertainty Relations?”. Physical Review Letters **129**, 240401 (2022).

[222] S. Soltani, M. Erba, D. Schmid, and J. H. Selby. “Decoupling local classicality from classical explainability: A noncontextual model for bilocal classical theory and a locally-classical but contextual theory” (2025). arXiv:2511.19266.

[223] E. Schrödinger. “Discussion of Probability Relations between Separated Systems”. Mathematical Proceedings of the Cambridge Philosophical Society **31**, 555–563 (1935).

[224] B. S. Tsirel’son. “Quantum generalizations of Bell’s inequality”. Letters in Mathematical Physics **4**, 93–100 (1980).

[225] B. S. Tsirel’son. “Quantum analogues of the Bell inequalities. The case of two spatially separated domains”. Journal of Soviet Mathematics **36**, 557–570 (1987).

[226] D. Gross, M. Müller, R. Colbeck, and O. C. O. Dahlsten. “All Reversible Dynamics in Maximally Nonlocal Theories are Trivial”. Physical Review Letters **104**, 080402 (2010).

[227] S. Popescu and D. Rohrlich. “Quantum nonlocality as an axiom”. Foundations of Physics **24**, 379–385 (1994).

[228] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski. “Information causality as a physical principle”. Nature **461**, 1101–1104 (2009).

[229] M. Navascués and H. Wunderlich. "A glance beyond the quantum model". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **466**, 881–890 (2009).

[230] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín. "Local orthogonality as a multipartite principle for quantum correlations". Nature Communications **4**, 2263 (2013).

[231] A. Cabello. "Simple explanation of the quantum violation of a fundamental inequality". Physical Review Letters **110**, 060402 (2013).

[232] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín. "Almost quantum correlations". Nature Communications **6**, 6288 (2015).

[233] T. Gonda, R. Kunjwal, D. Schmid, E. Wolfe, and A. B. Sainz. "Almost Quantum Correlations are Inconsistent with Specker's Principle". Quantum **2**, 87 (2018).

[234] A. B. Sainz, Y. Guryanova, A. Acín, and M. Navascués. "Almost-Quantum Correlations Violate the No-Restriction Hypothesis". Physical Review Letters **120**, 200402 (2018).

[235] C. A. Fuchs and A. Peres. "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information". Physical Review A **53**, 2038–2045 (1996).

[236] L. Maccone. "Information-disturbance tradeoff in quantum measurements". Physical Review A **73**, 042307 (2006).

[237] P. Busch. ""No information without disturbance": Quantum limitations of measurement". In Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony. Pages 229–256. Springer Netherlands, Dordrecht (2009).

[238] T. Heinosaari, L. Leppäjärvi, and M. Plávala. "No-free-information principle in general probabilistic theories". Quantum **3**, 157 (2019).

[239] D. Dieks. "Communication by EPR devices". Physics Letters A **92**, 271–272 (1982).

[240] H. P. Yuen. "Amplification of quantum states and noiseless photon amplifiers". Physics Letters A **113**, 405–407 (1986).

[241] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. "Noncommuting Mixed States Cannot Be Broadcast". Physical Review Letters **76**, 2818–2821 (1996).

[242] A. Daffertshofer, A. R. Plastino, and A. Plastino. "Classical No-Cloning Theorem". Physical Review Letters **88**, 210601 (2002).

[243] T. A. Walker and S. L. Braunstein. "Classical Broadcasting is Possible with Arbitrarily High Fidelity and Resolution". Physical Review Letters **98**, 080501 (2007).

[244] M. Piani, P. Horodecki, and R. Horodecki. "No-Local-Broadcasting Theorem for Multipartite Quantum Correlations". Physical Review Letters **100**, 090502 (2008).

[245] S. Luo. "On Quantum No-Broadcasting". Letters in Mathematical Physics **92**, 143–153 (2010).

[246] K. Kraus, A. Böhm, J. D. Dollard, and W. H. Wootters, editors. "States, Effects, and Operations Fundamental Notions of Quantum Theory". Volume 190 of Lecture Notes in Physics. Springer Berlin Heidelberg. Berlin, Heidelberg (1983).

[247] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman. "Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity". Physics Reports **506**, 1–39 (2011).

[248] Yeong-Cherng Liang, Robert W. Spekkens, and Howard M. Wiseman. "Erratum to "Specker's parable of the over-protective seer: A road to contextuality, nonlocality and complementarity" [Phys. Rep. 506 (2011) 1–39]". Physics Reports **666**, 110–111 (2017).

[249] C. Heunen, T. Fritz, and M. L. Reyes. "Quantum theory realizes all joint measurability graphs". Physical Review A **89**, 032121 (2014).

[250] R. Kunjwal, C. Heunen, and T. Fritz. "Quantum realization of arbitrary joint measurability structures". Physical Review A: Atomic, Molecular, and Optical Physics **89**, 052126 (2014).

[251] R. Kunjwal and S. Ghosh. "Minimal state-dependent proof of measurement contextuality for a qubit". Physical Review A: Atomic, Molecular, and Optical Physics **89**, 042118 (2014).

[252] S. A. Yadavalli, N. Andrejic, and R. Kunjwal. "Qualitative equivalence between incompatibility and Bell nonlocality" (2020). arXiv:2008.10100.

[253] A. Hulse, B. Schumacher, and M. D. Westmoreland. "Axiomatic information thermodynamics". Entropy. An International and Interdisciplinary Journal of Entropy and Information Studies**20** (2018).

[254] M. Lostaglio. "Quantum Fluctuation Theorems, Contextuality, and Work Quasiprobabilities". Physical Review Letters **120**, 040602 (2018).

[255] L. Giannelli and G. Chiribella. "Information-theoretic derivation of energy and speed bounds" (2024). arXiv:2403.13223.

[256] A. Joyal and R. Street. "Braided Tensor Categories". Advances in Mathematics **102**, 20–78 (1993).

[257] D. Gottesman and I. L. Chuang. "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations". Nature **402**, 390–393 (1999).

[258] R. Raussendorf and H. J. Briegel. "A One-Way Quantum Computer". Physical Review Letters **86**, 5188–5191 (2001).

[259] D. W. Leung. "Two-qubit Projective Measurements are Universal for Quantum Computation" (2002). arXiv:quant-ph/0111122.

[260] M. A. Nielsen. "Quantum computation by measurement and quantum memory". Physics Letters A **308**, 96–100 (2003).

[261] R. Raussendorf, D. E. Browne, and H. J. Briegel. "Measurement-based quantum computation on cluster states". Physical Review A **68**, 022312 (2003).

[262] D. W. Leung. "Quantum computation by measurements". International Journal of Quantum Information **02**, 33–43 (2004).

[263] R. Jozsa. "An introduction to measurement based quantum computation" (2005). arXiv:quant-ph/0508124.

[264] P. Jokinen, M. Weilenmann, M. Plávala, J. Pellonpää, J. Kiukas, and R. Uola. "No-broadcasting characterizes operational contextuality" (2024). arXiv:2406.07305.

[265] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin. "Experimental quantum cryptography". Journal of Cryptology **5**, 3–28 (1992).

[266] S. M. Barnett and S. J. D. Phoenix. "Information-theoretic limits to quantum cryptography". Physical Review A **48**, R5–R8 (1993).

[267] A. K. Ekert, B. Huttner, G. M. Palma, and A. Peres. "Eavesdropping on quantum-cryptographical systems". Physical Review A **50**, 1047–1056 (1994).

[268] A. Kent. “Unconditionally Secure Bit Commitment”. Physical Review Letters **83**, 1447–1450 (1999).

[269] H. Lo and H. F. Chau. “Unconditional Security of Quantum Key Distribution over Arbitrarily Long Distances”. Science **283**, 2050–2056 (1999).

[270] D. Mayers. “Unconditional security in quantum cryptography”. Journal of the ACM **48**, 351–406 (2001).

[271] S. Eilenberg and S. Mac Lane. “General theory of natural equivalences”. Transactions of the American Mathematical Society **58**, 231–294 (1945).

[272] N. M. Roy. “Extreme Points of Convex Sets in Infinite Dimensional Spaces”. The American Mathematical Monthly **94**, 409–422 (1987).

[273] G. Köthe. “Topological Vector Spaces I”. Volume 159 of Grundlehren Der Mathematischen Wissenschaften. Springer Berlin Heidelberg. Berlin, Heidelberg (1983).

[274] P. Perinotti, A. Tosini, and L. Vaglini. “Shannon theory beyond quantum: Information content of a source”. Physical Review A **105**, 052222 (2022).

[275] P. Perinotti, A. Tosini, and L. Vaglini. “Shannon theory for quantum systems and beyond: Information compression for fermions”. In A. Plotnitsky and E. Haven, editors, The Quantum-like Revolution: A Festschrift for Andrei Khrennikov. Pages 135–156. Springer International Publishing, Cham (2023).

[276] P. Perinotti, A. Tosini, and L. Vaglini. “Which entropy for general physical theories?”. Quantum **8**, 1384 (2024).

Symbols

$\mathcal{S}$, 164
$\mathscr{S}_{\mathrm{A,B}}$, 31
$\mathsf{S}_{\star}^{\mathrm{AB}\to\mathrm{BA}}$, 31

$|\cdot|$, 19
$\mathrm{G}(\cdot)$, 35
Υ, 44
$\mathfrak{C}_{\mathcal{K}(\mathsf{X})}$, 40
$\mathsf{cod}(\cdot)$, 242
$\triangleright$, 114
$\triangleright(\cdot)$, 190
$\succcurlyeq$, 65
$\mathsf{ConicHull}(\cdot)$, 66
$\mathsf{ConvRef}(\cdot)$, 73

D_{A}, 54
$\mathsf{dom}(\cdot)$, 242
$\to$, 146

$\mathsf{Eff}_{\mathbb{R}}(\cdot)$, 53
$\mathsf{Eff}(\cdot)$, 46
$\mathsf{Eff}_1(\cdot)$, 46
$\sim$, 45
Event_{Θ}, 34
$\mathscr{T}_x$, 17
$\mathscr{T}_{\mathsf{X}^{(k)}}$, 40
$[\![\mathscr{T}_{\mathsf{X}^{(k)}}]\!]_{k\in\mathsf{K}}$, 40
$\mathscr{T}$, 17
$\mathsf{Event}(\cdot)$, 17, 18
$[\![\mathscr{T}_x]\!]_{x\in\mathsf{X}}$, 17

$\mathscr{I}_{\mathrm{A}}$, 23
1_{A}, 242
$\mathsf{I}_{\star}$, 23
$\mathsf{Instr}(\cdot)$, 46
$\mathsf{Instr}(OPT)$, 192
$\mathsf{Instr}^{(N)}(\mathrm{A}\to\mathrm{B})$, 68
$\mathsf{Instr}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$, 59
$\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A}\to\mathrm{B})$, 59
$\mathscr{R}^{-1}$, 30
R^{-1}, 30

$\mathsf{Mor}(\cdot)$, 34, 241

$\|\cdot\|_{gen}$, 67
$\varepsilon_{\mathrm{A}\to\mathrm{B}}$, 49

$\mathsf{Obj}(\cdot)$, 34, 241
a_x, 25
$\mathsf{Obs}(\cdot)$, 46
$\mathrm{a}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{I}}$, 25
$d_{op}(\cdot,\cdot)$, 63
$\|\cdot\|_{op}$, 62, 64, 67
$\mathsf{X}^{(k)}$, 39
$\mathcal{K}(\mathsf{X})$, 39
$\left\{\mathsf{X}^{(k)}\right\}_{k\in\mathsf{K}}$, 39
$\mathsf{Part}(\mathsf{X})$, 40
$\mathsf{Out}(\cdot)$, 23

$\boxtimes$, 24
ρ_x, 25
$\mathsf{Prep}(\cdot)$, 46
$\rho_{\mathsf{X}}^{\mathrm{I}\to\mathrm{A}}$, 25
$\mathsf{PurSt}(\cdot)$, 75

$\mathsf{Ref}(\cdot)$, 72
$(\cdot|$, 25
$|\cdot)$, 25

$\bullet$, 37
$\square$, 21
$\circ$, 242
$\star$, 18

$\mathsf{St}_{\mathbb{R}}(\cdot)$, 53
$\mathsf{St}(\cdot)$, 46
$\mathsf{St}_1(\cdot)$, 46
$\mathsf{St}_+(\cdot)$, 66
$\bowtie$, 144
$\|\cdot\|_{sup}$, 65
$J(\cdot)$, 65
$\cong$, 31
$\Delta^{\{n\}}_{\mathrm{A}_1\ldots\mathrm{A}_n}$, 102
$\mathsf{Sys}(\cdot)$, 17

$\otimes$, 56
T_{X}, 17
$\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}$, 17
T, 17
$\mathsf{Test}(\cdot)$, 17
$\mathsf{Test}(\mathrm{A}\to\mathrm{B})$, 17
$\mathsf{Transf}(\mathrm{A}\to\mathrm{B})$, 46
$\mathsf{Transf}_1(\mathrm{A}\to\mathrm{B})$, 46
$\mathsf{Transf}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$, 55
$\left\{\widetilde{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}}\right\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)}$, 48
$\left\{\widehat{\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}}\right\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)}$, 48
$\mathsf{Transf}(\cdot)$, 46
$\mathsf{Transf}_1(\cdot)$, 46
I, 24

$\Upsilon(\cdot)$, 219

$\leftrightarrow$, 146

Index

γ-ray microscope
Gedankenexperiment, 5

Amygdaloidal OPT, 75
Associator, 243
Atomic, 72
Atomicity of composition, 255

Bilocal Classical Theory (BCT), 104
Bilocal discriminability, 101
Bilocal tomography, *see* Bilocal discriminability
Braided monoidal category, 246
Braiding, 31, 246
Hexagon identities, *see* Hexagon identities
Naturality, *see* Naturality
Broadcasting, 142
Broadcasting transformation, 141

Category, 241
Arrow, *see* Category - Morphism
Braided strict monoidal, 34
Functor, *see* Functor
Monoidal, *see* Monoidal Category
Braided, *see* Braided monoidal category
Strict, *see* Monoidal category - Strict
Symmetric, *see* Symmetric monoidal category
Morphism, 34, 241
Codomain, 242
Composition, 242
Domain, 242
Identity, 242
Source, 242
Target, 242
Objects, 34, 241
Causal OPT, 105, 107, 108

Causality, 105, 107, 108
  Condition, 106
Classical conditioning, *see* Conditioning operation
Classical control on outcomes, *see* Strong causality
Classical OPT, 126
Classical post-processing, 39
Classical Theory (CT), 12, 129
Closed circuit, 34
Coarse-graining, 40
  Deterministic, 41
  Event, *see* Event - Coarse-grained
  Test, *see* Test - Coarse-grained
Compatibility
  Full, *see* Full-compatibility
  Measurement, *see* Observation-compatibility
  Observation, *see* Observation compatibility
  Observation-instrument, *see* Observation-compatibility
  POVM, 9
  Strong, *see* Strong compatibility
  Weak, *see* Weak compatibility
Complementarity, 10, 220
  Fringe visibility, 10
  Mild, 222, 223
  Strong, 222, 223
  Weak, 222, 223
  Which way, 10
Compound-local-events, 26
Compound-local-test, 26
Compression protocol, 256
  Decoding, *see* Decoding
  Efficient, 256
  Encoding, *see* Encoding
  Ideal, 257
  Lossless, 256
Conditioning operation, 114

Conic hull, 66
Context, 49
Contextuality, 49
Convex OPT, 73
Convex-refinement, 73
  Set, 73

Decoding, 256
Determinism, *see* Deterministic OPT
Deterministic, *see* Deterministic event
Deterministic OPT, 113
Does not exclude, 145
Dual Classical Theory (DCT), 106, 125
Dual OPT, 125

Effect, 46
  Generalised, 53
  Null, *see* Null-effect
Elementary property, 220
  Complementary, 220
Encoding, 256
  Lossless, 256
Entangled state, 127
Epistemic state, 50
Equivalent norms, 65
Essential uniqueness of purification, 133
Event, 17
  Braiding, 31
  Coarse-grained, 40
  Deterministic, 18
  Inverse, 30
  Reversible, 30
Event$_\Theta$, 34
Extremal, 73
Extreme ray, 251

Fermionic Quantum Theory (FQT), 104

Foundations of quantum theory, 4
Full-compatibility, 148
Full-Information Without Disturbance (FIWD), 149
Functor, 245

Generalised Probabilistic Theory (GPT), 12
Generalised-non-contextuality, 51

Heisenberg's
- Gedankenexperiment, *see* γ-ray microscope Gedankenexperiment
- microscope, *see* γ-ray microscope Gedankenexperiment

Hexagon identities, 32, 246

Ideal compression, *see* Compression protocol - Ideal
Identity, 23
- Event, 23
- Test, 23

Identity object, *see* Monoidal unit
Incompatibility, 143
- POVM, 9
- Strong, *see* Strong incompatibility
- Weak, *see* Weak incompatibility

Instrument, 46
- Atomic, 219
- Conditional, 114
- Disturbing, 137
- Does not exclude, *see* Does not exclude
- Follows, 108
- Generalised, 59
- Intrinsically irreversible, 151
- Jellyfish, 173
- No-information, 136
- Non-disturbing, 137
- Precedes, 108
- Repeatable, 218
- Reversible dilation, *see* Reversible dilation
- Strong compatibility, *see* Strong compatibility
- Strong incompatibility, *see* Strong incompatibility
- Weak compatibility, *see* Weak compatibility
- Weak incompatibility, *see* Weak incompatibility

Inverse, 30
Irreversibility, 152
Irreversible disturbance, 5, 9

Jointly perfect discriminable, 126

Klee's theorem, 251
Kraus decomposition, 228
- Single Kraus, 228

Latent Quantum Theories (LQTs), 56
Left-unitor, 243
Local discriminability, 99
Local tomography, *see* Local discriminability
Local-test of system A from system AE, 30

Measure-and-re-prepare, 26
Measurement compatibility, *see* Observation-compatibility
Minimal Bilocal Classical Theory (MBCT), 213
Minimal Classical Theory (MCT), 186
Minimal Operational Probabilistic Theory (MOPT), 166
Minimal Quantum Theory (MQT), 215

Minimal Strongly causal Bilocal Classical Theory (MSBCT), 125, 207, 210
Minimal Strongly causal Classical Theory (MSCT), 207, 215
Minimal Strongly causal Operational Probabilistic Theory (MSOPT), 190
Minimal Strongly causal Quantum Theory (MSQT), 215
Miracle of scalars, 36
Mixed, 74
Monoidal category, 243
  Associator, *see* Associator
  Braided, *see* Braided monoidal category
  Left-unitor, *see* Left-unitor
  Natural transformation, *see* Natural transformation
  Pentagon identity, *see* Pentagon identity
  Right-unitor, *see* Right-unitor
  Strict, 244
  Symmetric, *see* Symmetric monoidal category
  Triangle identity, *see* Triangle identity
Monoidal Product, 243
Monoidal unit, 243
Multimeter, *see* Conditioning operation

N-local discriminability, 98
N-local tomography, *see* N-local discriminability
Natural transformation, 245
Naturality, 32
No-broadcasting theorem, 142
No-Information Without Disturbance (NIWD), 136, 138
No-programming, 184
No-restriction hypothesis, 122, 123
No-signalling, 105
Norm
  Instrument, 67
Null-effect, 53
Null-state, 53
Null-transformation, 49

Observation-compatibility, 144
Observation-instrument, 46
  Compatibility, *see* Observation-compatibility
Observation-test, 25
Observations, 25
Ontic state, 50
Ontological model, 50
  Generalised-non-contextual, 50
Operational completeness, 68, 70
Operational distance, 63
  Effect, 67
  Instruments, 68
  States, 63
  Transformations, 64
Operational equivalence
  System, 31
Operational norm
  Effect, 67
  Instrument, 67
  State, 62
  Transformation, 64
Operational Probabilistic Theory (OPT), 17
  Amygdaloidal, *see* Amygdaloidal OPT
  Convex, *see* Convex OPT
  Symmetric, *see* Symmetric OPT
  Unquotiented, 46
Operationalism, 16
Operationally indistinguishable, 45
Outcome, 17
Outcome space, 17

Partition, 40

Parallel composition, 26
Pentagon identity, 244
Perfect discriminability, 255
Post-processing, *see* Strong causality
Generalised, 146
Pre-OPT, 96
Preparation-instrument, 46
Preparation-tests, 25
Preparations, 25
Prepare-and-measure, 26
Probabilistic equivalence, 45
Projection-Valued Measures (PVMs), 148
Property, 220
Elementary, *see* Elementary property
Pure, 74
Purification, 132

Quantum foundations, *see* Foundations of quantum theory
Quantum logic, 11
Quantum Theory (QT), 17, 132
Quantum-switch, 111

Real Quantum Theory (RQT), 102
Rebit, 103
Refinement, 72
Set, 72
Reversible dilation, 214
Right-unitor, 243

Scalar, 26
Scalar multiplication, 37
Scalar-tests, 26
Separable state, 127
Sequential composition, 21
Simplex, 126
Simplicial OPT, 126
Singleton set, 18
Spatial isotopy, 33, 248
State, 46
Entangled, *see* Entangled state
Generalised, 53
Null, *see* Null-state
Pure, 75
Separable, *see* Separable state
Verifier, *see* Verifier
Strong, *see* Strong verifier
Strong causality, 115
Strong compatibility, 144
Strong incompatibility, 146
Strong Verifier, 219
Strongly causal OPT, 115
Sup norm, 65
Swap, 32
Symmetric monoidal category, 247
Symmetric OPT, 32
System, 17
Composite, 24
Minimal decomposition, 163
Dimension, 54
Elementary, 163
Operatinal equivalence, *see* Operational equivalence - System
Size, *see* System - Dimension
Trivial, 24
Type, 20

Test, 17
Braiding, 31
Coarse-grained, 40
Inverse, 30
Reversible, 30
Singleton, 18
Tomographically complete, 47
Transformation, 46
Atomic, *see* Atomic
Braid, 164

Broadcasting, *see* Broadcasting transformation
Extremal, *see* Extremal
Follows, 108
Generalised, 55
Jellyfish, 173
Mixed, *see* Mixed
Null, *see* Null-transformation
Permutation, 164
Precedes, 108
Pure, *see* Pure
Transformations, 46
Triangle identity, 244
Trivial OPT, 183

Uncertainty relation
Entropic, 9
Heisenberg, 6
Robertson, 8
Unique decomposition OPT, 163
Unit object, *see* Monoidal unit
Universal simulator, 184

Verifier, 219
Strong, *see* Strong verifier

Weak compatibility, 146
Weak incompatibility, 144