Maximising the utility of Information Causality as an operational principle for bounding bipartite non-locality

Rothe, Thomas

Maximising the utility of Information Causality as an operational principle for bounding bipartite non-locality

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
PHYSICS

Author :	Thomas Rothe
Student ID :	1930443
Supervisor :	Dr. Jordi Tura i Brugués
Daily supervisor :	Jan Li, MSc
Second supervisor :	Prof. dr. Serge Fehr

Leiden, The Netherlands, August 28, 2024

Maximising the utility of Information Causality as an operational principle for bounding bipartite non-locality

Thomas Rothe

Instituut-Lorentz, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

August 28, 2024

Abstract

In quantum information theory, the presence of Bell non-local correlations is a key indicator of non-classical behavior in multipartite quantum systems. However, non-locality is not exclusive to quantum mechanics; more general theories with stronger non-local correlations than those achievable within the quantum formalism can be constructed. While distinguishing classical (local) correlations from non-local correlations can, in principle, be accomplished by a finite number of linear constraints called Bell inequalities, distinguishing between quantum and post-quantum correlations requires solving a hierarchy of SDP relaxations. To simplify the certification of quantum correlations, a whole research line has focused on searching for an operational principle that can explain the limited strength of quantum correlations. Among the proposed principles, information causality (IC) stands out as the most promising, though deriving general correlation bounds from it is also very complex.
We review the various attempts to formalise IC and their effectiveness in constraining bipartite non-locality, as well as the challenges encountered in studying this principle. In particular, we perform numerical experiments to showcase the insufficiency of all the currently proposed IC bounds to capture the full potential of the principle for correlations near the quantum boundary. Furthermore, we demonstrate the instability of two out of three bounds under non-locality distillation.

Contents

Chapter 1

Introduction

Classical theories of nature are built on the idea that spatially separated objects exist independently and influence each other only through local interactions. Quantum mechanics, a cornerstone theory for physical systems on the smallest scales, fundamentally challenges this classical view by featuring phenomena like entanglement, superposition, and contextuality —concepts that are widely known to contrast sharply with our everyday experiences on the macro-scale. Despite its counterintuitive aspects, quantum mechanics has proven remarkably successful in predicting and explaining experimental results. This success is noteworthy, given that its core formalism —involving state vectors, unitary evolution, and measurement postulates —is largely axiomatic. While these axioms were motivated by experimental observations, their abstract mathematical nature and the lack of an intuitive physical justification has led to a line of research that investigates whether quantum theory is the most general theory. More specifically, the question was raised whether any of the non-classical quantum effects are special to quantum physics or if alternative (hypothetical) theories can predict equivalent or even stronger phenomena. [1]

Central in this investigation has been the phenomenon of non-locality. Non-locality refers to the idea that systems, even when separated by vast distances, can exhibit correlations that cannot be explained by local actions and interactions alone. Importantly, this is unlike the more general concept of entanglement, which can sometimes be explained through classical correlations. In 1964, John Bell famously proved through his theorem [2] the occurrence of non-locality in multipartite quantum systems. This result ruled out an alternative explanation for non-local correlations, originally suggested in the EPR paper [3], which was based on the idea of "hidden"

variables.

Meanwhile, non-locality has found several applications in the certification of quantum devices [4, 5], cryptography [6–8] and the generation of randomness [9, 10], for example. Moreover, it turns out that non-locality is actually quite a fundamental phenomenon as it implies also some other non-classical features of quantum physics, like non-determinism [11]. Therefore, a seminal paper [12] by Popsecu and Rohrlich suggested to focus research on hypothetical theories which exhibit non-locality but still comply with relativistic causality. In the same work, they showed that this set of theories is much larger than the set of theories compatible with the quantum formalism, demonstrating that non-locality is clearly not a special feature of quantum mechanics. Since then, much effort was spent on finding physical principles that could serve as additional axioms, complementing the occurrence of non-locality and the adherence to relativistic causality. The hope is that these constraints on physical systems jointly identify quantum theory uniquely from the broader set of all non-local causal theories. While this was quickly achieved within the framework of so-called generalized probabilistic theories (GPTs) [13], GPTs still enforce a certain structure on (the fundamental description of) physical theories based on abstract entities like "states" and "effects". The ambition of the research line initiated by Popescu and Rohrlich, however, was rather to find a principle whose validity can be verified solely through the observation of certain non-local correlations in the experimental statistics of multipartite systems. That is, without assuming anything about how these statistics arise within the systems.

Many such *operational* principles [14] have been proposed in the last decades, but all of them were found to hold for non-local correlations that form a strictly larger sets than the set of valid quantum correlations. In fact, there is strong evidence that operational principles can ultimately only identify almost-quantum theory [15], which is a non-local theory with a formalism very similar to that of quantum theory but slightly more general with the respect to which measurements are allowed within a given multipartite system.

The principle of Information causality (IC), as originally proposed in [16] by Pawlowski et al, is the only exception for which it is still unknown to what extent it can characterise the set of non-local quantum correlations and whether it can rule out almost-quantum theory as a physical theory of nature. Though it was already shown that no operational principle,

including IC, can exactly describe the set of quantum correlations if its formulation is restricted to bipartite systems. [17]

Pragmatically speaking, IC says that the amount of information that can be gained by a receiver about a sender's dataset should *not* exceed the amount of information that may be transmitted to the receiver via some communication channel [16]. In the special case that nothing is communicated, this reduces to the principle of (relativistic) causality, implying that the receiver cannot obtain any information about the dataset. Basically, IC represents the idea that the existence of information in one system (e.g. the receiver) should *not* depend on the simultaneous existence of information in a different, spatially separate system (e.g. the sender) [18].

The first few works [16, 19–21] on IC already demonstrated its remarkable success and superior strength over other principles in approximating extremal quantum correlations. However, they also showcased some deficiencies and the challenges in making general claims about which non-local correlations satisfy the IC principle, limiting their results to very specific bipartite scenarios. Two other works, [22] and [23], independently made first attempts to generalise the study of IC by considering bipartite systems in more abstract frameworks, like causal structures [22], compared to the original IC paper [16]. However, even after a decade of research, there is still no convergent upper-bound on the set of non-local correlations that satisfy IC. So for many non-local correlations, it remains inconclusive whether they truly satisfy the IC principle or not.

Beside its use in bounding non-locality, IC was also shown to rule out very weak and very strong forms of compositing subsystems (i.e. certain tensor products) [24]. In addition, IC can be applied outside the field of quantum foundations. For example, IC might be useful to simplify procedures for the device-independent certification of quantum devices that exploit non-local correlations.
Moreover, as recently illustrated in [25], the security of certain quantum key distribution (QKD) protocols can be demonstrated under the reasonable assumption that both the communicating parties and the attacker adhere to the principle of information causality.

As very recent works [26, 27] made substantial progress in simplifying and systematizing the derivation of constraints that IC imposes on non-local correlations, we were motivated to use this for exploring various generalisations of the bipartite scenarios in which IC was previously studied. The

generalisation that we investigate in the main part of this thesis is accounting for non-locality distillation through so-called wirings, where multiple copies of a source of non-local correlations are combined to potentially create a single effective source of correlations with amplified non-locality. In the appendix, we also take a glimpse at generalisations through non-locality recycling and multi-bit channels.

We provide an overview of the different approaches to IC and expose their individual strengths and weaknesses, giving possible directions for further research on this topic. In particular, we aim to address in this thesis the following questions: **(1a)** What are (strongest) known bounds for IC ? **(1b)** To what extent are they quantum-tight ? **(1c)** If there is no universally strongest IC bound, what is the advantage of each bound and why ? **(2a)** To what extent can non-locality distillation through wirings violate current IC bounds ? **(2b)** Is there evidence for an IC bound that is stable under wirings ?

Thesis outline

In chapter 2 we start by introducing the non-familiar reader to the device-independent paradigm, the topic of Bell non-locality, non-local games and the information causality (IC) principle. Also we give an overview on the different efforts that have been taken to extend the set of statistics for which we know that IC is violated with certainty. Thereby, we also introduce the idea of non-locality distillation by means of so-called "wirings".
We then present in chapter 3 the implementation of a modular numerical framework for studying device-independent non-locality in Julia, and we demonstrate this framework by using it to explore the current state of research on IC. Furthermore, we describe experiments on applying wirings to assess the strength of different IC constraints, and our attempts on finding new examples of wired distributions that violate the IC principle.
Finally, the thesis is concluded in chapter 4 with a discussion on the implications and limitations of our study of IC.
Appended are chapter A on the derivation and rationalisation of IC, chapter B on preliminary examinations of using non-locality recycling to further strengthen bounds on IC in non-local scenarios involving sequential queries of no-signaling boxes, and chapter C with a schematic of the setup for a non-local communication game involving a channel with multiple inputs and outputs.

Chapter 2

Background & Previous work

2.1 The device-independent characterisation of physical experiments

Since the beginning, quantum physics has been motivated by phenomena observed in physical experiments. As such, deriving its laws and mathematical formalism from a fully operational specification has been of special interest in theoretical research. Rather than describing the exact way that a physical experiment is realised in the laboratory, one thus only studies abstractly how the measurement outcomes $a, b, c, ...$ ("outputs") of some experiment are related to the values of a certain set of measurement settings $x, y, z, ...$ ("inputs"). In the famous Stern-Gerlach experiment, for example, one could think of x as a binary variable that determines the orientation of the magnetic field, while a indicates the observed direction in which the particles are deflected due to their spin. Another interesting example is the quantum double-slit experiment, where $x \in [2^2]$ could determine whether slit 1 and slit 2 are both blocked ($x = 0$), only slit 1 is blocked ($x = 1$), only slit 2 is blocked ($x = 2$), or both of the two slits are blocked ($x = 4$). Then a could indicate a detector "click" in a specific region of the screen behind the slits ($a \in \{0, 1\}$) or, if $a \in \mathbb{R}$, it might correspond to the exact position of the detection on the screen. [13] In principle, there is thus no restriction in the experimental description that require inputs $x, y, z, ...$ or outputs $a, b, c, ...$ to be discrete variables.

It is important to realise that the experiment itself is a blackbox and it only matters what the relation between inputs and outputs is, but not how it came about. Within the blackbox there could be classical physics,

quantum physics or something radically different going on. Furthermore, the relation between inputs and outputs is generally not specified by a deterministic algebraic expression, but in terms of some conditional probability distribution $P(a, b, c, ...|x, y, z, ...)$. While this is not necessary if the blackbox is restricted to classical physics, a probabilistic description is unavoidable in the case of a quantum blackbox due to the fundamentally indeterministic nature of quantum mechanics.
In the most simplistic view of physical experiments, we can thus focus on studying the structure of conditional probability distributions $P(a, b, c, ...|x, y, z, ...)$ while ignoring common abstract mathematical concepts from physical theories, like "states". Even the notion of a "measurement" is strictly speaking not defined since the outputs $a, b, c, ...$ and inputs $x, y, z, ...$ are just a set of values without a definite meaning. Nevertheless, we will mostly refer to $P(a, b, c, ...|x, y, z, ...)$ as an abstraction of performing measurements with settings $x, y, z, ...$ since measurements are the only physical operations within quantum mechanics from which interesting structures in the experimental statistics $P(a, b, c, ...|x, y, z, ...)$ emerge.

If one consistently holds on to this *device-independent* formalism and treats the emergence of any instance of such joint probability distributions as a blackbox, one enters the realm of what we call "boxworld". Concretely, we refer by boxworld[1] to the unconstrained collection of all valid probability distributions $P(a, b, c, ...|x, y, z, ...)$ for any fixed *alphabet* of inputs and outputs (e.g. a, b, x, y) for two inputs and two outputs). The main advantage of boxworld is a focus on the behavior of the physical system and the decoupling from any specific physical theory. That way, the number of assumptions is radically reduced and one can make more fundamental (and thus stronger) claims. This is of special importance in the field of cryptography, where one aims to guarantee security of a certain protocol independently of how it is physically implemented. [30, 31] Without the need to specify a Hilbert space dimension, for example, one can ensure that an adversary does not benefit from just scaling up his available resources.

2.1.1 Bell non-locality

When distinct input-output pairs (such as (x, a) and (y, b)) are associated with different space-like separated subsystems, one typically obtains only

[1]This "boxworld" is different but nearly equivalent to "boxworld" in the context of Generalised Probabilistic Theories (GPTs). In the end, both describe the same distributions $P(a, b, c, ...|x, y, z, ...)$ with the only difference being that the "boxworld" GPT enforces the existence of a certain state and effect space. [28, 29]

product distributions of the form $P(a,b,\ldots|x,y,\ldots) = P(a|x)P(b|y)\ldots$. These distributions preclude any correlation between output variables $a,b,\ldots$. However, by distributing additional information λ (potentially probabilistic in nature) across the subsystems prior to their separation, it becomes possible to generate correlated outputs. This is possible while still only requiring local operations, where each subsystem produces its output (e.g. a) based solely on its local input (e.g. x) and the shared information λ. In his famous paper, John Bell characterised distributions $P(a,b,...|x,y,...)$ for multipartite systems that incorporate such (hidden) variables λ as follows:

Definition 2.1.1. ***Bell locality*** *— A conditional distribution $P(a,b,\ldots|x,y,\ldots)$ is called (Bell) local if and only if*

$$P(a,b,\ldots|x,y,\ldots) = \int_\Lambda d\lambda Q(\lambda) P_\lambda(a|x) P_\lambda(b|y)\ldots \tag{2.1}$$

given some ensemble of marginal probability distributions $\{(Q(\lambda), P_\lambda(a|x), P_\lambda(b|y), ...)|\lambda \in \Lambda\}$, such that $Q(\lambda) \geq 0$ and $\sum_\lambda Q(\lambda) = 1$
Otherwise the distribution $P(a,b|x,y)$ is ***(Bell) non-local****.*

Within a multipartite system composed of subsystems A, B, and others, The marginal distributions $P_\lambda(a|x), P_\lambda(a|x), \ldots$ are called "processes" and denote the distributions from which each subsystem samples its respective output.
Further, λ is an abstract random variable that is distributed to and accessible from all of the subsystems. In this sense, we say that λ is globally shared. Abstractness of λ means that it can be represented by anything, from a fixed collection of numbers to a random number generator that can be sampled from within all subsystems.

From the standpoint of classical physics, these "local" distributions encompass everything that can be observed while adhering to the classical assumption of locality. This means that an outcome in one subsystem cannot be influenced by the input choice in any other subsystem, even when these input choices are made randomly and all precautions are taken to prevent any signal transmission between the subsystems. Consequently, *classical* distributions $P(a,b,\ldots|x,y,\ldots)$ are usually identified as the set of all distributions that take on a local form, as specified in definition 2.1.1.

Shifting our focus to quantum physics, one of the most striking and non-classical phenomena is the emergence of non-local correlations between

measurement outcomes, particularly when performing independent measurements on entangled multipartite quantum systems. In the EPR paper [3], a thought experiment was presented that famously demonstrated how non-locality as a strong form of quantum entanglement between space-like separated systems leads to an apparent contradiction between the principle of complementarity and local descriptions of reality.

While quantum entanglement in the sense of the inseparability of states is necessary to produce non-local correlations, entanglement is not a sufficient criterion for identifying non-locality. Within the more general framework of boxworld, definition 2.1.1 allows to pragmatically distinguish local from non-local (i.e. classical from non-classical) correlations between measurement outcomes $a, b, ...$ solely by means of distributions $P(a, b, ...|x, y, ...)$. In the realm of quantum, the specified measurements $x, y, ...$ with outcomes $a, b, ...$ are thereby performed in different (possibly space-like separated) subsystems A and B. We call such a setup a "Bell scenario" (or "Bell experiment") and specify it with a fixed number m_S of inputs and a fixed number of outputs o_S for subsystem S. [32] More briefly, for a bipartite Bell scenario, we refer to a (m_A, m_B, o_A, o_B) Bell scenario.

It is important to emphasise that there is generally no *unique* physical realisation (e.g. a quantum state and a set of quantum measurements) that relates the inputs $(x, y, \ldots)$ with the outputs $(a, b, \ldots)$ via a certain distribution $P(a, b, ...|x, y, ...)$. In some cases, for example, both classical and quantum systems might be able to produce the same distribution. In the following, we therefore refer collectively to all physical realisations of the distribution $P(a, b, ...|x, y, ...)$ as the *box* $P(a, b, ...|x, y, ...)$. Instead of stating that A and B share an entangled state, we can then say more generally that A and B share a box $P(a, b, ...|x, y, ...)$. Although we typically use the terms "box" and "distribution" interchangeably, it should be clear from context whether the physical realisation of the distribution $P(a, b, ...|x, y, ...)$ or the distribution $P(a, b, ...|x, y, ...)$ itself is meant.

From definition 2.1.1, it is clear that the number of valid Bell local distributions $P(a, b, ...|x, y, ...)$ is infinite, even for the simplest (2,2,2,2) Bell scenario. However, by the convexity of the decomposition in eq. 2.1 and the bounded nature of probabilities, it is possible to describe the set of Bell local distributions $\mathcal{L}$ compactly as the convex hull of a certain finite subset $\mathcal{P}_{ext}$. [32] This means that any $P(a, b, ...|x, y, ...)$ is a convex combination $\sum_\lambda q(\lambda) P^{(\lambda)}_{ext}(a, b, ...|x, y, ...)$ of so-called *extremal points* $P^{(\lambda)}_{ext}(a, b, ...|x, y, ...) \in$

$\mathcal{P}_{ext}$ such that $\sum_\lambda q(\lambda)$ and $q(\lambda) \geq 0$. The defining property of an extremal point hereby is that they are themselves not a (non-trivial) convex combination of other extremal points, so

$$P_{ext}^{(\lambda')}(a,b,...|x,y,...) = \sum_{\lambda \neq \lambda'} q(\lambda) P_{ext}^{(\lambda)}(a,b,...|x,y,...)$$

enforces $q(\lambda) = \delta_{\lambda=\lambda'}$.

According to a well-known theorem, due to Fine [32, 33], the extremal points for the local set $\mathcal{L}$ are exactly the deterministic distributions of the form $P_{det}(a,b,...|x,y,...) = \delta_{a=a(x)}\delta_{b=b(y)}\delta_{y=y'}\ldots$ for some discrete functionals $a(x) \in [o_A]$ and $b(y) \in [o_B]$. Note that the number of deterministic distributions P_{det} is finite in a fixed Bell scenario and that their convex hull (i.e. the local set $\mathcal{L}$) is a polytope. [32]

That deterministic distributions are extreme points is quite intuitive since a convex combination is nothing else than a weighted average of elements, where the weights are determined by probabilities. Sampling a' and a'' with probability $p < 1.0$ and $1-p$ respectively, for example, should never be equivalent to the situation of obtaining $\tilde{a}$ with certainty for any input x (i.e. $P(a|x) = \delta_{a=\tilde{a}}$) if $\tilde{a} \neq a'$ and $\tilde{a} \neq a''$. Put simply, determinism cannot be perfectly simulated by probabilistic sampling.

In this work, we will be fully focused on the simplest bipartite Bell scenario with input and output bits, i.e. $(2,2,2,2)$. In that case, we have the following convenient parameterisation for local deterministic distributions P_{LD} [19]

$$P_{LD}^{\alpha\gamma\beta\lambda} = \begin{cases} 1, & \text{if } a = \alpha x \oplus \gamma \,\&\, b = \beta y \oplus \lambda \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

with parameters $\alpha, \gamma, \beta, \lambda \in \{0,1\}$. Boxes corresponding to some $P(a,b|x,y) = P_{LD}^{\alpha\gamma\beta\lambda}$ are called Local-Deterministic boxes (LD boxes).

The significance of Bell non-locality in quantum information science is nowadays huge and goes far beyond the original aim in quantum foundations to refute local realism. Non-locality can be seen as a generic qualitative feature of quantum states and as a resource since it can be stored, converted and consumed. [1, 34]
One of the most important applications of non-locality in quantum information processing is self-testing, so the device-independent certification

of quantum devices. [7, 8] Additional applications include cryptography protocols [6–8] and the generation of randomness [9, 10].

Initially, the non-local property of the joint distributions $P(a, b|x, y)$ was actually thought to uniquely characterise quantum mechanics in multipartite systems. However, as it turns out, non-locality is in no way a special property of quantum physics but actually manifests as very strong correlations between random variables in more general (hypothetical) theories as well. [12]

2.1.2 No-signaling

Strictly speaking, not all (Bell) non-local distributions are the result of genuine non-classical behavior of physical systems. The input-output pairs in definition 2.1.1 (x, a) and (b, y), for instance, might belong to different but time-like separated subsystems A and B. In that case, any dependence of output b on input x (or a on y) is fully compatible with classical physics, since the subsystems can physically interact or communicate[2] with each other. Consequently, they can exhibit arbitrarily strong correlations, including non-local ones. This can occur, for instance, when the parties simply communicate their respective measurement settings x and y to each other.

In contrast, if one assumes that A and B are space-like separated and no communication takes place after fixing the inputs (x, y) and before determining the outputs (a, b), the observation of non-locality can no longer be justified from a classical perspective. In other words, what makes most types of quantum entanglement truly surprising and distinct from classical correlations is not merely the observation of non-local distributions. Rather it is the fact that these non-local distributions are observed in physical experiments even when the subsystems are space-like separated from each other and unable to communicate.

For the purpose of studying non-locality as a purely non-classical phenomenon, it is therefore assumed throughout this thesis that any distribution $P(a, b|x, y)$ satisfies the following no-signaling conditions with respect to subsystems A and B:

[2] Of course, they must still respect relativistic causality. That means waiting sufficiently long for the arrival of messages, which are transmitted at finite speed.

Definition 2.1.2. ***No-Signaling*** — *A conditional distribution $P(a,b|x,y)$ is called no-signaling (NS) if and only if the local marginals satisfy*

$$P(a|x,y) \equiv \sum_b P(a,b|x,y) = \sum_b P(a,b|x,y') \equiv P(a|x,y') \tag{2.3}$$

for $\forall y, y'$, even if $y \neq y'$, and

$$P(b|x,y) \equiv \sum_a P(a,b|x,y) = \sum_a P(a,b|x',y) \equiv P(b|x',y) \tag{2.4}$$

for $\forall x, x'$, even if $x \neq x'$.

Moreover, we define the (bipartite) no-signaling set $\mathcal{NS}$, which contains all distributions $P(a,b|x,y)$ compatible with conditions 2.3 and 2.4. In words, the no-signaling condition eq. 2.3 (eq. 2.4) thus ensures that the marginal distribution $P(a|x)$ ($P(b|y)$) over the local output a (b) for one subsystem is independent of the input y (x) in the other subsystem. Note hereby that no-signaling, in contrast to Bell locality, still allows the final value of a (b) to depend on y (x), but only in such a way that y (x) does not effect the locally observed output statistics in subsystem A (B).
From a more high-level physics perspective, the no-signaling conditions enforce special relativity on space-like separated subsystems. Specifically, they address relativistic causality, which dictates that no signal or influence can propagate faster than the speed of light.

2.1.3 Facets & testing membership in compact convex sets

Classifying a probability distribution by its (mathematical) structure, like in definition 2.1.1, allows for a very basic but universal interpretation of what type of resources (e.g. quantum state & measurements) are required to produce those particular statistics in physical experiments. This manifests concretely in the study of causal relations between random variables [22] and the aforementioned self-testing of quantum devices [4, 5, 30], for example.

It would be infeasible to list all distributions with a certain structure and to compare each with a given distribution $P(a,b|x,y)$. However, if a given set $\mathcal{S}$ is generated from a finite subset of distributions, like the local polytope $\mathcal{L}$, testing membership of $P(a,b|x,y)$ in $\mathcal{S}$ is then equivalent to determining whether any convex combination of those extremal points of $\mathcal{S}$ recovers $P(a,b|x,y)$.

Instead of representing a polytope $\mathcal{S}$ (i.e. a certain type of compact convex set) by its extremal points, it is also sufficient to specify the facets of $\mathcal{S}$. To understand this, it is easiest to embed the distribution $P(a,b|x,y)$ as a vector in some vector space, where each dimension corresponds to a specific tuple (a,b,x,y). The value of each dimension represents the probability associated with that tuple. For example, if a, b, x, and y have binary values, the vector would have 16 dimensions, each corresponding to one combination of $(a,b,x,y) \in \{0,1\}^4$.
Facets of $\mathcal{S}$ can be thought of as the "faces" or "sides" of the polytope in this vector space. Mathematically, each facet is defined by a hyperplane that contains a maximal set of linearly independent extremal points, with all other extremal points lying on one side of this hyperplane. These facets allow for a compact representation of the polytope's boundaries as an alternative to the collection of extremal points. [32] Facets of $\mathcal{S}$ are then those hyperplanes in the vector space which contain a maximal linearly independent subset of extremal points, while the remaining extremal points need to be on the same side of the hyperplane. [32]
That there always exist some extremal points that are linearly dependent is hereby a consequence of the normalisation and no-signaling constraints on $P(a,b|x,y)$. Concretely, they reduce the dimension of the space of bipartite no-signaling distributions from $m_A o_A m_B o_B$ to $m_A(o_A-1)m_B(o_B-1)+m_A m_B(o_A-1)(o_A-1)$, while the number of extremal points is unaffected and stays $m_A o_A m_B o_B$. So technically not every tuple (a,b,x,y) is an (independent) component of the vector $P(a,b|x,y)$, since some $P(\tilde{a},\tilde{b}|\tilde{x},\tilde{y})$ can be derived from other the probabilities at other values of a, b, x and y.

Like any hyperplane, a facet $F^{(i)}$ can be expressed as a linear combination $\sum_{a,b,x,y} v^{(i)}_{a,b,x,y} P(a,b|x,y) = f^{(i)}$, with $\overrightarrow{v^{(i)}} \in \mathbb{R}^{dim(P(a,b|x,y))}$ the vector normal and $f \in \mathbb{R}$ a constant offset, such that $\sum_{a,b,x,y} v_{a,b,x,y} P_{ext}(a,b|x,y) \leq f$ for all extremal distributions $P_{ext}(a,b|x,y)$ of $\mathcal{S}$. [32]
To test whether $P(a,b|x,y) \in \mathcal{S}$, one has to check for every facet of $\mathcal{S}$ on which side $P(a,b|x,y)$ resides. Specifically, if

$$\sum_{a,b,x,y} v^{(i)}_{a,b,x,y} P(a,b|x,y) \leq f^{(i)}$$

for all facets $F^{(i)}$ then $P(a,b|x,y) \in \mathcal{S}$.

2.1.4 Bell functionals & Bell inequalities

Even independently from a specific set and its facets, hyperplanes (and their associated inequalities) are a powerful tool to characterise the correlations in distributions $P(a,b|x,y)$. A very important set of tools in the field of Bell non-locality are Bell functionals and Bell inequalites:

Definition 2.1.3. ***Bell functional*** *— A Bell functional is a linear form* $P(a,b|x,y) \mapsto I\left(P(a,b|x,y)\right) \in \mathbb{R}$ *where*

$$I(P) \equiv \sum_{(i,j,k,l)\in\{0,1\}^4} v_{ijkl} P(a=i,b=j|x=k,y=l)$$

for some coefficients $v_{ijkl} \in \mathbb{R}$*.*

So any linear combination of the different components of $P(a,b|x,y)$ is a Bell functional.

Bell inequalities are then linear criteria that divide the space of no-signaling distributions into half-spaces, one of which is strictly non-local. Concretely:

Definition 2.1.4. ***Bell inequality*** *— A Bell inequality is an inequality of the form* $I(P) \leq I_L$ *with a Bell functional* $I(P)$ *and some threshold* $I_L \in \mathbb{R}$ *such that if* $P(a,b|x,y)$ *violates the inequality* ($I(P) > I_L$) *then* $P(a,b|x,y)$ *is Bell non-local* ($P(a,b|x,y) \notin \mathcal{L}$)*.*

The converse (i.e. Bell non-locality implies violation of an (arbitrary) Bell inequality) does generally not hold. Also, satisfying a Bell inequality does not necessarily certify that $P(a,b|x,y)$ is a member of the local set $\mathcal{L}$. If a Bell inequality $I(P) \leq I_L$ is an inequality that corresponds to a facet of the local set $\mathcal{L}$, then the Bell inequality is called *tight*.

Note that not all Bell functionals are associated to a well-defined Bell inequality. A simple counter-example is the average probability of equal outputs in a (m_A, m_B, o_A, o_A) scenario:

$$\frac{1}{m_A m_B o_A} \sum_{x,y} P(a=b|x,y) = \frac{1}{m_A m_B o_A} \sum_{x,y} \sum_{k} P(a=k,b=k|x,y).$$

Indeed, it is a linear combination of different components of $P(a,b|x,y)$, which makes it a Bell functional with coefficients $v_{abxy} = \delta[a=b]$. However, it reaches its maximum value of 1.0 already for certain local deterministic distributions $P(a,b|x,y) = \delta[a=k] \cdot \delta[b=k]$ where $k \in \{0, o_A - 1\}$.

So no threshold I_L can exist such that the set of distributions which satisfy $\frac{1}{m_A m_B o_A} \sum_{x,y} \sum_k P(a = k, b = k|x, y) > I_L$ is non-empty and does not contain any Bell local distribution $P(a, b|x, y) \in \mathcal{L}$. Therefore, definition 2.1.4 is not applicable to $I(P) = \frac{1}{m_A m_B o_A} \sum_{x,y} \sum_k P(a = k, b = k|x, y)$, although it is a Bell functional according to definition 2.1.3.

A particularly important instance of a Bell inequality in a bipartite $(2, 2, 2, 2)$ Bell scenerio is the Clauser-Horne-Shimony-Holt (CHSH) inequality. [35] First, let us define the CHSH (Bell) functional with coefficients $v_{ijkl} = (-1)^{kl}\delta_{i=j}$ in definition 2.1.3, such that

$$\begin{aligned} I_{CHSH} &= \sum_{a,b,x,y} (-1)^{xy}\delta_{a=b}P(a, b|x, y) \\ &= \sum_{x,y}(-1)^{xy}\left(\sum_{k=0}^{1} P(k, k|x, y)\right) = \sum_{x,y}(-1)^{xy}P(a = b|x, y) \end{aligned} \tag{2.5}$$

The CHSH functional is usually written more compactly in terms of the so-called (probabilistic) Bell correlators $E_{kl} \equiv \sum_{a=b} P(a, b|x = k, y = l) - \sum_{a\neq b} P(a, b|x = k, y = l) = 2P(a = b|x = k, y = l) - 1$, by simply rescaling and translating the terms of I_{CHSH}. This yields the equivalent linear functional

$$S_{CHSH} = E_{00} + E_{01} + E_{10} - E_{11} \tag{2.6}$$

which is what we will from now on call the CHSH functional. Furthermore, we will refer to the value of S_{CHSH} as the *CHSH value*.

2.1.5 The CHSH inequality

The Bell inequality associated to S_{CHSH} is then the *CHSH inequality* and reads

$$S_{CHSH} \leq 2.$$

If one prefers an expression in terms of probabilities, rather than Bell correlators E_{xy}, the equivalent (probabilistic) Bell inequality reads $I_{CHSH} \leq 3$.

The CHSH inequality is a tight Bell inequality but, in a $(2, 2, 2, 2)$ scenario, satisfying the CHSH inequality $S_{CHSH} \leq 2$ alone is not conclusive evidence that a probability distribution $P(a, b|x, y)$ is local. This is also visually illustrated in figure 2.1. Distributions satisfying the probabilistic Bell inequality $I_{CHSH} \leq 3$ fall within the entire region below the horizontal dashed line

that aligns with the blue boundary of the local set $\mathcal{L}$. Any distribution above this line is definitely non-local since the local polytope $\mathcal{L}$ is entirely situated below the line. However, also below the line there are non-local regions, i.e. regions which fall outside of the blue-bordered polytope $\mathcal{L}$.

Even if a distribution $P(a,b|x,y)$ satisfies $S_{CHSH} \leq 2$, it might still violate a different Bell inequality and thus turn out to be non-local. To determine with certainty whether $P(a,b|x,y)$ is local, one needs to evaluate other "versions" of the CHSH inequality as well. This is similar to the previously mentioned need to consider all facets of $\mathcal{L}$, and their associated inequalites, for testing membership in $\mathcal{L}$. In figure 2.1, two other versions of the (probabilistic) CHSH inequality are illustrated by the vertical dashed lines to the left and right of the blue-bordered polytope $\mathcal{L}$.

Mathematically, the different versions of the CHSH inequality only differ in the choice of the Bell functional, while the threshold value of 2 stays the same. The complete family of CHSH functionals can be obtained by simply changing the signs of the four coefficients $s_i = \pm 1$ in 2.6:

$$S_{CHSH}^{(s_1,s_2,s_3,s_4)} = s_1 E_{00} + s_2 E_{01} + s_3 E_{10} - s_4 E_{11} \tag{2.7}$$

Any choice of an odd number of negative signs gives a tight Bell inequality $S_{CHSH}^{(s_1,s_2,s_3,s_4)} \leq 2$, such that the parity $\prod_i s_i = -1$. [19, 32] The different CHSH functionals are equivalent to the canonical one in 2.6 up to permutations of the labels $\{0,1\}$ for the inputs x or y, and interchanging subsystems A and B by re-labelling x as y and y as x. The CHSH inequality $S_{CHSH} \leq 2$ that corresponds to the CHSH functional $S_{CHSH} \equiv S_{CHSH}^{(1,1,1,-1)}$ in 2.6 is what we call the *canonical CHSH inequality*, or simply *the* CHSH inequality.

As part of the CHSH inequalities $S_{CHSH}^{(s_1,s_2,s_3,s_4)} \leq 2$, the CHSH functionals $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ play a key role in the qualitative detection of non-locality. However, the CHSH functionals are by themselves also a quantifier of the correlation strength between outputs (a,b) with respect to inputs (x,y) for any given distribution $P(a,b|x,y)$. Concretely, the correlation strength of $P(a,b|x,y)$ is defined as the largest CHSH value within the family of CHSH functionals, i.e.

$$\max_{(s_1,s_2,s_3,s_4)} S_{CHSH}^{(s_1,s_2,s_3,s_4)}\left(P(a,b|x,y)\right).$$

In that regard, the threshold value of 2 in 2.1.5 is the maximum correlation strength among all local distributions $P(a,b|x,y) \in \mathcal{L}$.

For non-local distributions $P(a,b|x,y) \notin \mathcal{L}$, the degree to which the CHSH functional $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ exceeds the threshold of 2 quantifies how strongly non-local the observed correlations are. That is, the "amount" of non-locality in $P(a,b|x,y) \notin \mathcal{L}$.

Conveniently, checking the violation of inequalities in the CHSH family is sufficient to detect any type of non-locality in the simplest bipartite Bell scenario. In context of physical experiments, this makes CHSH inequalities a popular choice for testing whether an observed distribution $P(a,b|x,y)$ is indeed non-classical or not. [30, 32] Curiously, generalisations of the CHSH inequality to multipartite or more complex bipartite Bell scenarios are less effective. Some types of non-locality do not violate these generalized CHSH inequalities, requiring other (families of) inequalities to classify a distribution as non-local with certainty. [31]

2.1.6 The CHSH non-local game

An alternative approach to identify non-local distributions through explicit criteria is based on non-local games.
We focus here on bipartite non-local games which describe a scenario involving two space-like separated parties, A and B. They share some no-signaling box which matches the distribution $P(a,b|x,y) \in \mathcal{NS}$, whereby no-signaling is implied by space-like separation.

First, each party S is queried with a random dit-string $Q^{(\mathrm{S})} \in \mathbb{Z}_d^{n_S}$, according to some an ensemble $\{P(Q^{(\mathrm{S})}), Q^{(\mathrm{S})}\}$.
Subsequently, each party uses their local marginal of $P(a,b|x,y)$ and their query $Q^{(\mathrm{S})}$ to map them to some response $R^{(\mathrm{S})} \in \mathbb{Z}_d^{m_S}$. The maps

$$(Q^{(\mathrm{A})}, P(a|x)) \mapsto R^{(\mathrm{A})} \text{ and } (Q^{(\mathrm{B})}, P(b|y)) \mapsto R^{(\mathrm{B})}$$

are what we refer to as the *strategy* of party A and B respectively. Together the strategies of both parties $(R^{(\mathrm{A})}, R^{(\mathrm{B})})$ form a *protocol*.
Finally, a *winning condition* $V : (R^{(\mathrm{A})}, R^{(\mathrm{B})}) \mapsto \{0,1\}$ is evaluated on the responses to decide whether the game is won ($V \mapsto 1$) or lost ($V \mapsto 0$). For some fixed queries $Q^{(\mathrm{A})}$ and $Q^{(\mathrm{B})}$, the *success probability* is then defined as

$$P_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} \equiv P(V(R^{(\mathrm{A})}, R^{(\mathrm{B})}) = 1 \,|Q^{(\mathrm{A})}, Q^{(\mathrm{B})}).$$

The ultimate goal for both parties in a specific non-local game is to maximise some objective function, which is a functional of the success probabilities $P_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}}$. To simplify the interpretation of the objective value, the functional is often chosen to be linear:

$$\sum_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} v_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} P_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}}$$

with coefficients $v_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} \in \mathbb{R}$. A popular choice for such a linear objective function is the average success probability:

$$\left(D_Q^{(\mathrm{A})} \cdot D_Q^{(\mathrm{B})}\right)^{-1} \cdot \sum_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} P_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}}$$

A simple but very important example of a non-local game is the CHSH game. The queries and responses are single bits (i.e. $d = 2$ and $n_A = n_B = m_A = m_B = 1$) and the winning condition is defined as $V(R^{(\mathrm{A})}, R^{(\mathrm{B})}) = \delta_{R^{(\mathrm{A})} \oplus R^{(\mathrm{B})} = Q^{(\mathrm{A})} \cdot Q^{(\mathrm{B})}}$. This means that party A and B are only successful in winning the game if they produce equal responses $R^{(\mathrm{A})} = R^{(\mathrm{B})}$ for any queries $Q^{(\mathrm{A})}$ and $Q^{(\mathrm{B})}$, except if $Q^{(\mathrm{A})} = Q^{(\mathrm{B})} = 1$.
Furthermore, the overall objective of the CHSH game is to maximize the sum of success probabilities,

$$\begin{aligned}\sum_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} P(V(R^{(\mathrm{A})}, R^{(\mathrm{B})}) = 1 \,|Q^{(\mathrm{A})}, Q^{(\mathrm{B})}) &= \sum_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} P(R^{(\mathrm{A})} \oplus R^{(\mathrm{B})} = Q^{(\mathrm{A})} \cdot Q^{(\mathrm{B})} \,|Q^{(\mathrm{A})}, Q^{(\mathrm{B})}) \\ &= P(R^{(\mathrm{A})} = R^{(\mathrm{B})} \,|0,0) + P(R^{(\mathrm{A})} = R^{(\mathrm{B})} \,|0,1) + P(R^{(\mathrm{A})} \neq R^{(\mathrm{B})} \,|1,1).\end{aligned}$$

Since $P(R^{(\mathrm{A})} \neq R^{(\mathrm{B})} \,|1,1) = 1 - P(R^{(\mathrm{A})} = R^{(\mathrm{B})} \,|1,1)$ the objective function becomes

$$-1 + \sum_{Q^{(\mathrm{A})},Q^{(\mathrm{B})}} (-1)^{Q^{(\mathrm{A})} \cdot Q^{(\mathrm{B})}} \cdot P(R^{(\mathrm{A})} = R^{(\mathrm{B})} \,|Q^{(\mathrm{A})}, Q^{(\mathrm{B})})$$

However, we can drop the constant "-1" since it is irrelevant for maximising the sum.

Note that the above is still independent of the strategies of party A and B in the CHSH game. To proceed, we now need to specify the protocol.
Let party A choose the query as local input to the no-signaling box $P(a.b|x,y)$, such that $Q^{(\mathrm{A})} = x$, and let A give the local output as response $R^{(\mathrm{A})} = a$.

Equivalently, let $Q^{(\mathrm{B})} = y$ and $R^{(\mathrm{B})} = b$ for party B.
For this specific protocol, the above objective then becomes $\sum_{x,y}(-1)^{xy} \cdot P(a = b\,|x, y)$. When comparing this to eq. 2.5, we can see that the objective for a given protocol is just a Bell functional. Consequently, maximising the objective function over distributions $P(a, b|x, y) \in \mathcal{NS}$ for the CHSH game is identical to maximising the (probabilistic) CHSH functional I_{CHSH}.

The equivalence between Bell functionals and the success probability of non-local games generalises beyond the special case demonstrated here. So for every non-local game with a given protocol, there is a corresponding Bell functional. Which of the two formulations is more convenient depends on the context of application. For some physical experiments, for example, one would like to detect non-locality in an observed distribution $P(a, b|x, y)$ of measurements labelled by inputs (x, y). In that case, one can simply substitute $P(a, b|x, y)$ into a Bell functional $I(P)$ and check for the violation of the corresponding Bell inequality $I(P) \leq I_L$. However, in the field of cryptography, it is usually more natural to think in terms of non-local games and success probabilities. In that case, an objective function value that exceeds a certain threshold can also certify whether the parties shared a non-local resource or not.
This is because from the perspective of non-local games, a Bell functional like S_{CHSH} is nothing more than a linear combination of success probabilities between different (independent) sets of inputs (x, y) of the no-signaling box $P(a, b|x, y)$. In fact, any non-local game objective that only depends on $P(a, b|x, y)$ does correspond to some Bell functional. However, the converse does not hold true, which means that not all Bell functionals can be re-formulated as the objective function of some non-local game.

2.1.7 PR-boxes & the no-signaling polytope

Although the locality threshold $I_L = 2$ for the CHSH functionals $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ in 2.7 is sufficient for distinguishing between local and non-local distributions, it does not yet saturate their maximum. Indeed, since $0 \leq P(a, b|x, y) \leq 1$ and the Bell correlators $-1 \leq E_{xy} \leq 1$, we see that $S_{CHSH}^{(s_1,s_2,s_3,s_4)} \leq 4$ for any variant of the CHSH functional with parameters (s_1, s_2, s_3, s_4).
The game-based formulation of the canonical CHSH functional S_{CHSH} helps to identify when the maximum value of 4 is reached, since optimal success in the CHSH game corresponds to a maximised S_{CHSH}. With the previously specified protocol, the winning condition for the CHSH game reads $x \cdot y = a \oplus b$. The distribution $P(a, b|x, y) = \delta_{x\cdot y=\tilde{a}\oplus\tilde{b}}\delta_{a=\tilde{a}}\delta_{b=\tilde{b}}$

does satisfy this condition for any fixed pair of outputs $(\tilde{a}, \tilde{b}) \in \{0,1\} \times \{0,1\}$ with certainty, i.e. $P(x \cdot y = \tilde{a} \oplus \tilde{b}|x,y) = 1$ for all input pairs (x,y). These distributions, however, turn out to be signaling ($P(a,b|x,y) \notin \mathcal{NS}$) for any output pair $(\tilde{a}, \tilde{b})$. The unique distribution $P(a,b|x,y)$ that is no-signaling and for which any sample perfectly satisfies $x \cdot y = a \oplus b$, is the uniform probabilistic mixture of distributions $P_{PR}(a,b|x,y) = \frac{1}{4}\sum_{\tilde{a},\tilde{b}} \delta_{x\cdot y=\tilde{a}\oplus\tilde{b}}\delta_{a=\tilde{a}}\delta_{b=\tilde{b}}$. [32] Evaluating the canonical CHSH functional S_{CHSH} on P_{PR} then indeed gives the optimum of 4
More generally, each CHSH functional in the family $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ corresponds to a CHSH game with a different winning condition and consequently yield other, unique distributions $P(a,b|x,y) \in \mathcal{NS}$ for which $S_{CHSH}^{(s_1,s_2,s_3,s_4)} = 4$.

We can parameterise the family of all maximally non-local, no-signaling distributions in the simplest bipartite $(2,2,2,2)$ scenario conveniently by $\mu, \nu, \sigma \in \{0,1\}$: [19]

$$P_{PR}^{\mu\nu\sigma}(a,b|x,y) = \begin{cases} \frac{1}{2} & \text{if } a \oplus b = xy \oplus \mu x \oplus \nu y \oplus \sigma \\ 0, & \text{otherwise} \end{cases} \tag{2.8}$$

The boxes which correspond to distributions of this form are called Popescu-Rohrlich (PR) boxes. Of particular importance is $P_{PR}^{000}(a,b|x,y)$, which we refer to as the canonical PR-box. [32] Beyond the simplest Bell scenario that we consider in this work, generalisations of PR-boxes to more complex scenario's do exist as well. [36]

PR-boxes are extreme examples of non-local no-signaling resources that even go beyond what quantum physics can realise. This makes PR-boxes evidently unphysical. [1, 32] Being extreme no-signaling distributions, however, makes them characteristic for the $\mathcal{NS}$ set. In fact, in the $(2,2,2,2)$ Bell scenario, the 8 non-local PR distributions (eq. 2.8) and 16 local deterministic boxes (eq. 2.2) together generate $\mathcal{NS}$. [32] Because of the finite number of extremal points, the no-signaling set $\mathcal{NS}$ is thus also a polytope, just like the local set $\mathcal{L}$. Notice that all extremal points of $\mathcal{L}$ are also vertices of $\mathcal{NS}$. This implies that also some of the edges and facets of $\mathcal{NS}$ and $\mathcal{L}$ match up exactly.

Although the 16 extremal distributions of $\mathcal{NS}$ already provide valuable insights about the no-signaling set $\mathcal{NS}$ as a whole, it is sometimes also useful to consider convex combinations of the extremal points. To simplify calculations, as well as the interpretation and visualisation of our results,

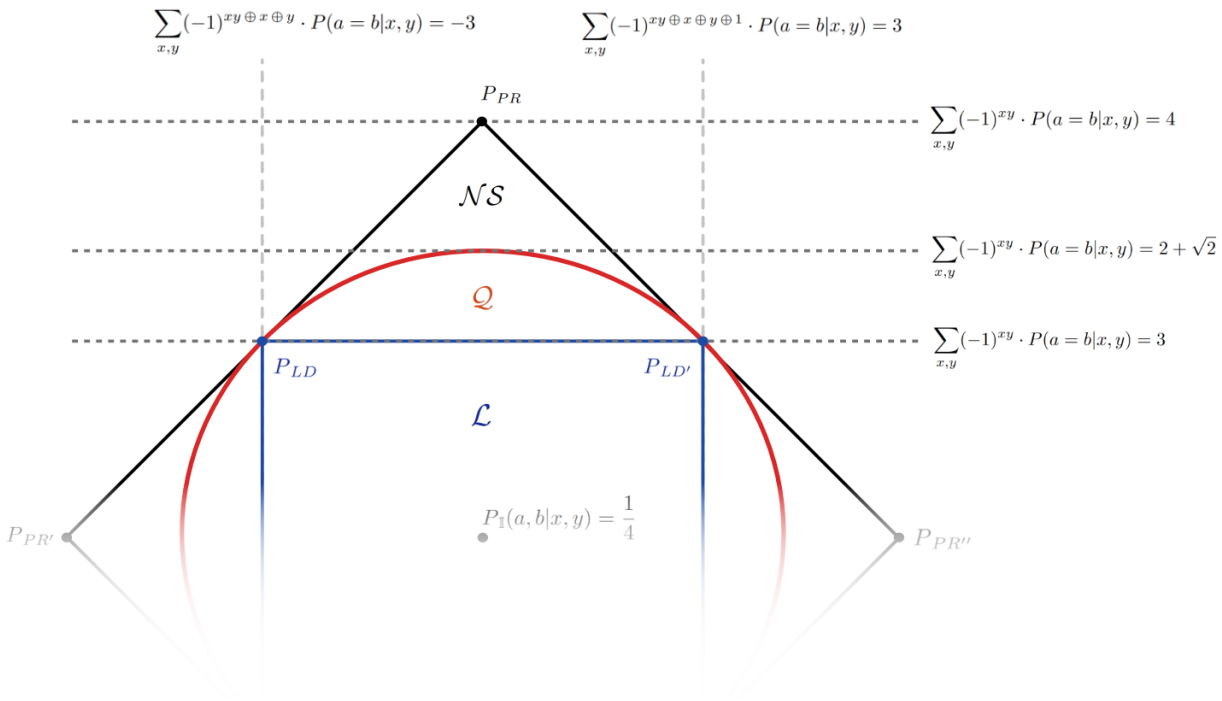


Figure 2.1: *Schematic illustration of an arbitrary 2D slice of the $\mathcal{NS}$-polytope in comparison with the corresponding projections of the strictly contained quantum set $\mathcal{Q}$ (red) and local set $\mathcal{L}$ (blue). The value of the canonical CHSH functional S_{CHSH} varies along the vertical as indicated by the horizontal dashed lines. The slice gives an extreme example of the non-trivial boundary of the quantum set $\mathcal{Q}$ since it is completely curved. Illustration based on figure 2 in [37]*

we will usually not study all mixtures in $\mathcal{NS}$ at once.
Rather one focuses on mixtures of only three distributions, i.e.

$$\eta_1 P(a,b|x,y) + \eta_2 Q(a,b|x,y) + (1-\eta_1-\eta_2)R(a,b|x,y)$$

with $P, Q, R \in \mathcal{NS}$, $\eta_1 + \eta_2 \leq 1.0$ and $0.0 \leq \eta_1, \eta_2 \leq 1.0$. We will often call such a family of mixtures a *slice* of the no-signaling set $\mathcal{NS}$. Hereby, "slice" refers to the geometric interpretation of selecting the unique 2D cross-section of $\mathcal{NS}$, embedded in the high-dimensional vector space of distributions $P(a,b|x,y)$, which contains the three points P, Q and R.

The relation between the no-signaling set $\mathcal{NS}$, the quantum set $\mathcal{Q}'$ (defined below) and the local set $\mathcal{L}$ is illustrated in figure 2.1 for a specific family of mixtures of three PR-boxes $\eta_1 P_{PR}^{000} + (1-\eta_1-\eta_2)P_{PR}^{111} + \eta_2 P_{PR}^{001}$. Hereby, the coefficient η_1 varies along the vertical of the figure, while the horizontal corresponds to the value of η_2. For this slice, η_1 is equal to the

value of the canonical CHSH functional S_{CHSH} up to some scaling factor. Accordingly, the dashed horizontal lines intersect the boundary of each of the sets at their respective maximum value of I_{CHSH} from eq. 2.5. Observe furthermore that $\mathcal{L} \subseteq \mathcal{Q}' \subseteq \mathcal{NS}$. This makes sense since quantum entangled states indeed feature non-classical behavior ($\mathcal{L} \subseteq \mathcal{Q}'$), while they should still obey special relativity ($\mathcal{Q}' \subseteq \mathcal{NS}$).

2.1.8 Non-locality as a resource

Quantifying non-locality by the value of CHSH functionals $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$, also begs the question of how this quantity changes when performing certain operations on the physical system. In a typical bipartite Bell scenario, the two space-like separated parties are restricted to local operations in their individual subsystems and possibly some classical communication. For this set of operations, often briefly denoted as LOCC (i.e. Local Operations & Classical Communication), any pre-established non-locality is either preserved or "consumed" over time, but it can never increase. [31, 32] Only if the subsystems are brought together and a certain global operation is applied to the joint system, more non-locality might be created. In that sense, non-locality in entanglement theory is often interpreted as a type of limited resource, which can be used to share information (or randomness) between multiple parties.

2.1.9 The quantum set

In our quest to identify which distributions $P(a,b|x,y) \in \mathcal{NS}$ can be realised with the quantum formalism of quantum states and general POVM measurements, it is of special interest what value any CHSH functionals $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ can reach when restricting ourselves to quantum mechanics.
The most pragmatic way to define the set of quantum distributions $\mathcal{Q}$ is the following: [32]

Definition 2.1.5. ***Quantum set*** $\mathcal{Q}$ *— In a bipartite* $(2,2,2,2)$ *Bell scenario, a distribution* $P(a,b|x,y)$ *is called quantum iff*

$$P(a,b|x,y) = Tr\left(\rho(\Pi_a^x \otimes \Pi_b^y)\right)$$

for some positive-semidefinite operator $\rho \in \mathcal{L}\left(\mathcal{H}_A \otimes \mathcal{H}_B\right)$ *on the product between two Hilbert spaces* $\mathcal{H}_A$ *and* $\mathcal{H}_B$*, such that* $\mathrm{Tr}\,\rho = 1$*. Further,* $\Pi_a^x \in \mathcal{L}(\mathcal{H}_A)$ *(*$\Pi_b^y \in \mathcal{L}(\mathcal{H}_B)$*) are positive-semidefinite operators such that* $\sum_a \Pi_a^x = \mathbb{I}$ *for any* $x \in [m_A]$ *(*$\sum_b \Pi_b^y = \mathbb{I}$ *for any* $y \in [m_B]$*). These operators correspond to* m_A *(*m_B*)*

measurements with each o_A (o_B) possible outcomes.
Operator "$\otimes$" represents the usual tensor product of real linear vector spaces.

For an "if" instead of "iff", one can shorten def. 2.1.5 by dropping $\mathrm{Tr}\,\rho = 1$ and the normalisation of density operators $\sum_a \Pi_a^x = \mathbb{I}$ ($\sum_b \Pi_b^y = \mathbb{I}$). For any given $P(a,b|x,y)$, these conditions are already implied from the normalisation of the distribution.

The above definition is a natural choice from the perspective of quantum information theory, as it inherently constructs multipartite systems from individual subsystems using the tensor product formalism. This approach mirrors the way qubits, the fundamental units of quantum information, are integrated in the field of quantum information processing.
However, the resulting set $\mathcal{Q}$ is not closed and, therefore, has no sharp boundary that one could use as reference to compare $\mathcal{Q}$ to other subsets of $\mathcal{NS}$. Moreover, the use of tensor product operations in the definition for distributions in $\mathcal{Q}$ complicates the analytical characterisation of $\mathcal{Q}$. This means that only few of the conjectured mathemtical properties of $\mathcal{Q}$ have actually been proven yet, while much more is known about $\mathcal{Q}'$, which corresponds to a weaker definition of the quantum set: [32]

Definition 2.1.6. *Quantum set $\mathcal{Q}'$ In a bipartite $(2,2,2,2)$ Bell scenario, a distribution $P(a,b|x,y)$ is called quantum if*

$$P(a,b|x,y) = Tr(\rho \Pi_a^x \Pi_b^y)$$

for some positive-semidefinite operator $\rho \in \mathcal{L}(\mathcal{H})$ on some Hilbert space $\mathcal{H}$, such that $\mathrm{Tr}\,\rho = 1$. Further, $\Pi_a^x \in \mathcal{L}(\mathcal{H})$ ($\Pi_b^y \in \mathcal{L}(\mathcal{H})$) are positive-semidefinite operators such that $\sum_a \Pi_a^x = \mathbb{I}$ for any $x \in [m_A]$ ($\sum_b \Pi_b^y = \mathbb{I}$ for any $y \in [m_B]$). These operators correspond correspond to m_A (m_B) measurements with each o_A (o_B) possible outcomes.
Additionally, measurement operators on different subsystems must commute, i.e.

$$[\Pi_a^x, \Pi_b^y] = 0, \ \forall a,b,x,y$$

The set $\mathcal{Q}'$ defined by 2.1.6 is actually a strict superset of the actual quantum set $\mathcal{Q}$ from definition 2.1.5. The two sets differ only in what structure is demanded on the measurement operators. In quantum information, we use more often the tensor product structure for composite systems in 2.1.5 to implicitly enforce commutativity of operators on different subsystems. Defining the relaxation $\mathcal{Q}'$ as the quantum set, however, is usually

preferred in quantum foundations since $\mathcal{Q}'$ is closed by construction and a more common choice in adjacent fields of research, such as quantum field theory. We are not further concerned with the minimal (though non-zero [38]) difference between $\mathcal{Q}$ and $\mathcal{Q}'$. Therefore, we will keep referring to $\mathcal{Q}'$ from definition 2.1.6 as *the* quantum set.

The well-known Tsirelson bound of $S_{CHSH} \leq 2\sqrt{2}$ then gives the maximal quantum violation of the CHSH inequality. [39] See section 3.2.2 of [32] for an instructive derivation of the Tsirelson bound.
Remarkably, the Tsirelson bound is achievable within the quantum framework only by preparing a singlet state, up to local isometries, and performing a suitable projective measurement. [30, 32] This unique property of the singlet state is of interest for applications in self-testing and re-emphasizes the relevance of the singlet in quantum physics more generally.

However, beyond the few quantum distributions $P(a, b|x, y) \in \mathcal{Q}'$ which reach the Tsirelson bound, it is very difficult to characterise the boundary of the set $\mathcal{Q}'$. While the quantum set $\mathcal{Q}'$ is convex, like $\mathcal{L}$ and $\mathcal{NS}$, it has an infinite number of extremal points. Therefore, it is not a polytope and does not underlie any other compact mathematical description.
Only a few exceptions are known for which an exact expression for the boundary of $\mathcal{Q}'$ exist. Those only apply to well-studied regions of the no-signaling polytope, like quantum voids [40] and a subset of distributions arising from dichomotic quantum measurements [41]. However, to recover the full boundary of the quantum set $\mathcal{Q}'$, we usually need to fall back on solving semidefinite programs (SDPs). The most famous and widely used set of SDPs is the so-called Navascues-Pironio-Acin (NPA) hierachy. [42]

When visualising the quantum set within the vector space of no-signaling distributions, some parts of the boundary actually turn out to be smooth, non-linear curves. This is in contrast to the flat edges and sharp corners of the polytopes corresponding to $\mathcal{L}$ and $\mathcal{NS}$. [32, 37] An extreme case of the non-linear boundary of $\mathcal{Q}'$, in direct comparison with the other two sets, is illustrated in figure 2.1.

2.2 Bounding quantum correlations operationally

The problem of describing the quantum set of distributions by purely operational principles was first raised in [1]. It was based on the finding that special relativity, in the form of the no-signaling conditions 2.3 and 2.4,

actually allow for much stronger correlations than quantum mechanics, such as those of PR-boxes. The extent of this gap between quantum and no-signaling correlations was recently shown to become more significant when considering Bell scenario's with higher number of inputs or outputs. [43]

Thus clearly the no-signaling principle can not be a characteristic principle of quantum theory, at least as long as no further assumptions are made about the underlying physical system.
If we, for example, assume that the two space-like separated subsystems can individually be described within the quantum formalism, then the no-signaling principle is actually able to constrain any shared bipartite box $P(a,b|x,y)$ to the quantum set, i.e. $P(a,b|x,y) \in \mathcal{Q}'$. [44] That is, whenever the no-signaling principle holds, the box $P(a,b|x,y)$ can always be *simulated* with bipartite quantum states and local measurements, even if the truly underlying composite system can not be described by a state in some bipartite Hilbert space[3]. On first sight, this already seems to solve the targeted problem of bounding the quantum set by an operational principle. However, the "locally-quantum"-assumption is incompatible with the aim for a completely device-independent characterisation of quantum distributions $\mathcal{Q}'$. In fact, the authors of [44] point out that the requirement for quantum subsystems is just a further witness for the need to supplement no-signaling with some other (locally constraining) principle. In that sense, [44] provided additional evidence for the conjecture in [1] that constraining non-locality (globally) is a necessary but insufficient part of justifying the quantum formalism.

After discovering that the no-signaling principle alone cannot fully characterize the set of physically realisable distributions $P(a,b|x,y)$, numerous other complementary principles have been proposed over the past few decades. This includes local orthogonality [45], macroscopic locality [46, 47], no-hypersignaling [48], macroscopic non-contextuality [14] and consistent exclusivity [49], to just name a few.

2.2.1 Almost-quantum distributions

The collapse of communication complexity is a particularly well-known, yet simple, example of a principle whereby many no-signaling boxes be-

[3]Within the framework of GPTs, this happens when choosing a tensor product other than the standard one.

yond quantum were found to give an implausible advantage in certain non-local games. [32, 50] Implausible is here understood in the sense that scaling of the non-local scenario to more inputs (i.e. information) does not impact the amount of communication that is needed to succeed in the game.
Although most of the principles mentioned above can be used to derive the maximum quantum violation of CHSH (i.e. the Tsirelson bound), it was shown that none of them is able to reconstruct the full boundary of the quantum set $\mathcal{Q}'$ within the space of distributions $P(a,b|x,y)$. [15, 32]

A proposal of a slight modification of the quantum set $\mathcal{Q}'$, called the almost-quantum set $\tilde{\mathcal{Q}}$ of distributions, played a particularly notable role in ruling out many of the aforementioned principles. Briefly stated, the requirement of commutation between measurement operators in definition 2.1.6 is relaxed[4] to apply only to specific quantum states, resulting in a larger set of distributions $P(a,b|x,y) \in \tilde{\mathcal{Q}}$. The idea is that there is no physical necessity for insisting on pairwise commutation between measurements in different subsystems with respect to *all* joint quantum states. In that sense, the almost-quantum formalism does not equal the widely used quantum formalism, but is very similar to it, even in terms of the (observable) predictions that it makes [15].

Checking membership of $P(a,b|x,y)$ in $\tilde{\mathcal{Q}}$ requires solving a hierarchy of feasibility SDPs, similar to the NPA hierachy for $\mathcal{Q}'$. In fact, the almost-quantum set is equivalent to the set obtained through the "$1+AB$" level of the NPA hierachy in the simplest Bell scenario. Hereby, "$1+AB$" is a special intermediate level between the first and second NPA level. [5]

By the physical similarity between $\tilde{\mathcal{Q}}$ and $\mathcal{Q}'$, this demonstrates that even quite low levels of NPA are quite reasonable approximations to $\mathcal{Q}'$. The effectiveness of the almost-quantum set in bounding quantum distributions was recently also confirmed in terms of relative volume with respect to the no-signaling set $\mathcal{NS}$. [43] Even for more complex bipartite Bell scenarios, the almost-quantum set stays a good approximation, though testing membership in almost-quantum is then not as efficient anymore. [15, 43]

[4] Thus $[\Pi_a^x, \Pi_b^y] = 0$ in definition 2.1.6 is replaced by $[\Pi_a^x, \Pi_b^y]|\psi\rangle = 0$, whereby without loss of generality ρ in def. 2.1.6 is redefined as $\rho = |\psi\rangle\langle\psi|$ with $|\psi\rangle \in \mathcal{H}$.

[5] "Levels" in the NPA hierachy can be understood as the "order of the outer approximation" to $\mathcal{Q}'$. A higher level allows the SDP optimisation to yield an approximation with higher precision. For details see [15, 32]

Meanwhile, nearly all operational principles have been shown to single out sets of distributions which are all strictly larger than the almost-quantum set $\tilde{\mathcal{Q}}$, and thus also the quantum set $\mathcal{Q}'$ itself. [14, 15, 32, 51] Only a few principles, like macroscopic non-contextuality, exactly correspond to the almost-quantum set [14]. This has prompted the conjecture that all bipartite, device-independent, and operational principles (including IC) may ultimately converge to a constraint that corresponds to $\tilde{\mathcal{Q}}$ rather than $\mathcal{Q}'$.[15, 32, 52]
In order to bound sets of distributions which are strictly contained in $\tilde{\mathcal{Q}}$, one probably needs to consider genuinely multipartite information-theoretic principles [17] or principles that also make explicit assumptions about the local state space structure [51, 52].

To our knowledge, information causality is the only operational principle for which it is still unknown whether it bounds a superset or a subset of $\tilde{\mathcal{Q}}$. Though there is some numerical evidence that it is not a subset [15], just like all the other principles. The lack of a universally tight bound for information causality has prevented a definite conclusion to date.
In the next subsection we describe how information causality was originally proposed and how the statement of this principle has developed since then.

2.2.2 The brief history of Information Causality

Information causality is an information-theoretic principle that follows from the fundamental properties of certain types of entropies in the context of a particular type of communication games, called information retrieval tasks. We start this subsection by first introducing this type of game.

Information retrieval & Random Access Codes

Consider a bipartite system with subsystem A associated with some party/agent A and subsystem B with some party B. Assume that a (noisy) communication channel with at most κ bits of (classical) capacity connects the two physical subsystems. Initially, party A receives n inputs $\overrightarrow{\alpha} = (\alpha_0, ..., \alpha_{n-1})$ (a.k.a. the "data"), whereby each input can take d possible values. Subsequently, the task for party B is to choose a single output g (a.k.a. the "retrieved information") from d possible values.
It is crucial that neither A or B have control about the data and the choice

of a query, such that they cannot conspire to make the task trivial by themselves. Therefore, it is easiest to imagine an additional external party, the verifier V, which distributes the inputs and checks whether B's retrieved information g was correct.

In such a scenario, an abstract information retrieval task is then described by some queries $Q \in \mathbb{Z}_d$ and "winning relations" $w_q : \mathbb{Z}_d^n \mapsto \mathbb{Z}_d$ which identify for each set of n inputs a correct[6] output (modulo d), given some query $q \in Q$. [54] Note that in this formalism, multiple "correct" output values g can exists for a fixed query q and a fixed input vector $\overrightarrow{\alpha}$. In this case, multiple winning conditions with the same label "w_q" are specified, each mapping to one of correct values.

A very popular example of an IR task in both Shannon and quantum information theory is the so-called Random Access Code (RAC). [54–56] In the bipartite version, with cooperating parties A and B, the goal is to transfer a piece of requested information about a randomized dataset from party A to party B. Concretely, given some random input dits $\overrightarrow{\alpha} \in [d]^n$ at party A, the aim for party B is to retrieve the β-th dit in the input string of party A. Hereby, $\beta \in [n]$ is a (randomized) integer that party B receives as input and outputs a guess g_β of α_β such that the success probability of the event $g_\beta = \alpha_\beta \mod d$ is maximized. The winning relations for any type of RAC then simply become $w_\beta = \alpha_\beta \mod d$ for all $\beta \in [n]$.

While party A and B might agree on a cooperative strategy before they have received their respective inputs, they are only allowed to communicate via their *limited* communication channel during each round of the game. Also, although two-way communication is technically allowed, we focus here on strategies involving only the most important direction of information transfer, namely from party A to party B. We denote the locally constructed message of A as μ, and the possibly noisy message μ' as received by party B.

A simple but natural figure-of-merit for RACs is the (average) success probability of party B retrieving the queried dit. [16, 57, 58] The idea is to repeat the game many times and calculate the frequency of successful rounds. This allows us to compare different protocols or strategies of party A and B within the given rules of the game.

[6] Thereby we assumed that the retrieval of information can either be a "success" or "failure". For information retrieval games with continuous scores, see [53].

The (classical) baseline for any guess that party B makes, is the uniform random probability of $\frac{1}{d}$. [57, 59] This success probability is realized in the trivial scenario when there is no communication allowed between the two parties, and B just guesses the requested dit at random.
In the other extreme, if the channel's capacity is unlimited, or at least n dits, the transfer of the full dit string $\overrightarrow{\alpha}$ from A to B is trivial and the success probability is 1.

For a fixed setting that lies in between the two extremes, say for $n = 2$, $d = 2$ and a limited channel capcacity of 1 bit, the challenge is to determine an optimal encoding and decoding strategy.
If party A always just sends the value of one of the two input bits (α_0, α_1) to B, while both $\overrightarrow{\alpha}$ and β are distributed uniformly, then the probability of success can become $\frac{3}{4}$ in the best case. Thus, indeed, the communication of 1 bit does help to improve over random guessing quite a bit, even though all the employed resources in the RAC were classical. Note that the given example only applies to the simplest and lowest values of $d = 2$ and $n = 2$. The more general upper bounds for higher n and d, and for non-uniform random inputs $\overrightarrow{\alpha}/\beta$, are much more involved.

The intuition of the communication advantage is that B does only care about getting the correct bit-value and not whether party A has send exactly the part of the bit-string $\overrightarrow{\alpha}$ that was queried with β. So even if party A sends the non-requested bit to B, as long as it has the same value, B just happened to be lucky. It is easy to see that this false-positive coincidence happens 25% of the time, which explains the deviation from random guessing.

The problem of finding an optimal trade-off between required communication resources and an acceptable success probability has been of great interest to research in information theory. [57, 60]

Non-Locality assisted RACs & the van Dam protocol

The exploitation of non-locality has turned out to yield a significant gain in the average success probability in RACs and made RACs very popular as a benchmark in quantum information theory. More concretely, there are two types of RACs that can make use quantum resources: Quantum RACs (QRACs) and Entanglement-Assisted RACs (EARACs).

In a Quantum RAC (QRAC), the two parties are equipped with a quantum channel, instead of a classical communication channel. [57–59, 61] The idea is to physically transfer qudits from A to B and make use of the non-classical properties of quantum information carriers. While QRACs are very powerful, they are practically much more resource intensive than regular RAC setups and protocols. Moreover, the intend of our discussion on operational principles was to avoid the explicit use of the quantum formalism, which is rather difficult for QRACs.

In contrast, an EARAC is much more like a standard RAC, with a classical communication channel, but the two parties share a no-signaling resource that corresponds to some bipartite distribution $P(a,b|x,y) \in \mathcal{NS}$. In a fully classical RAC, for example, they could have access to a Local-Determinstic box (eq. 2.2) or some classical source of shared randomness. In a RAC with quantum subsystems, however, the two parties could also share an entangled pair of qubits and use the outcomes of local measurements to adapt their game strategy based on the measurement outcomes (a,b).

Such a RAC, in which the two parties are equipped with a classical communication channel and a bipartite no-signaling box, should probably be called a non-locality assisted random access code (NARAC). However, for consistency, we will refer to them as EARACs (i.e. Entanglemenet assisted RACs) in this thesis. This quantum-restricted term is namely used more commonly throughout the literature. [55, 56, 58, 62]

For an EARAC, there is no Bell local $P(a,b|x,y) \in \mathcal{L}$ which can improve the average success probability compared to a standard RAC. For example, for $n = 2$, $d = 2$, and a single bit of classical communication, the average success probability of an EARAC stays between $\frac{1}{2}$ and $\frac{3}{4}$, which any (classical) strategy in a standard RAC can achieve as well. [55, 57]
Quantum and stronger non-local correlations, however, can increase the amount of potential information accessible to B. Curiously, it turns out that QRACs and EARACs complement each other in terms of achievable success probabilities. [58] This also means that there are special cases in which QRACs can achieve higher success probabilities than EARACs, but also many scenarios vice versa. [55, 58].

An important example of a EARAC with bipartite no-signaling distributions $P(a,b|x,y) \notin \mathcal{Q}'$ beyond the quantum set is the one involving PR-boxes $P_{PR}^{\mu\nu\sigma}$. The maximal non-locality of PR-boxes enables hereby gen-

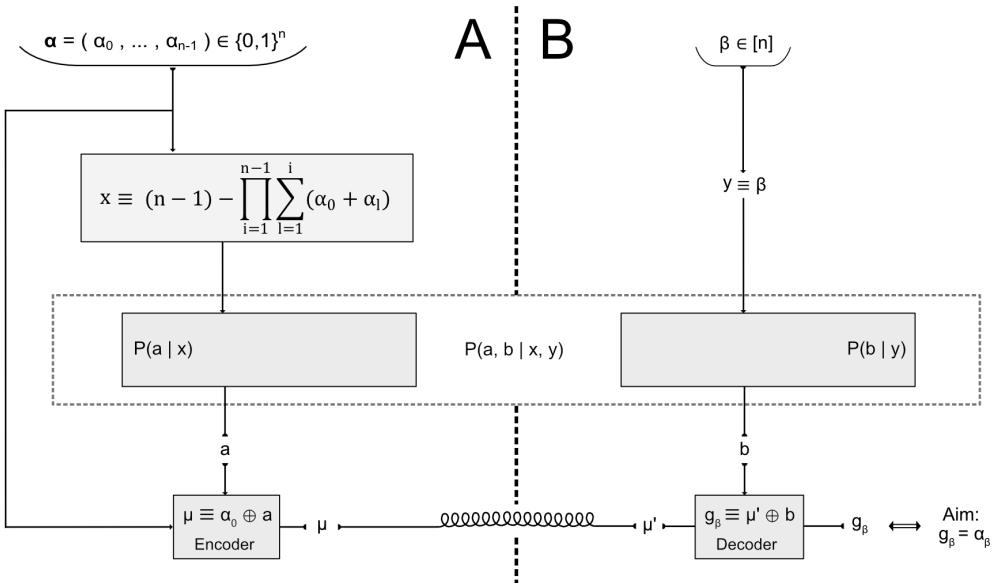


Figure 2.2: *The van Dam protocol applied to a bipartite Random Access Code (RAC) between space-like separated parties A and B in a $(n, n, 2, 2)$ Bell scenario. While A receives a dit-string $\overrightarrow{\alpha}$ of fixed length n, B is queried with only a single n-dimensional bit β. The latter indicates which dit in $\overrightarrow{\alpha}$ his final output dit g_β is supposed to recover. The two parties only share a noisy communication channel of $\kappa \equiv I(\mu : \mu')$ dits classical capacity.*

erally the highest success probabilities in RACs. Asymptotically, it even allows error-free guessing by party B if the number of shared PR-boxes grows towards infinity.

The most well-known protocol that can achieve error-free guessing via PR-boxes, is the so-called *van Dam protocol*. [32, 63] In the simplest scenario of two input bits $(\alpha_0, \alpha_1) \in 0, 1^2$ (i.e. $d = 2$ and $n = 2$) and a single bit of communication $\kappa = 1$, it goes as follows:
On receiving inputs $\overrightarrow{\alpha}$, party A starts by putting $x = \alpha_0 \oplus \alpha_1$ into part A of the shared no-signaling box $P(a, b|x, y)$ and retrieves the output a. Subsequently, A constructs the message $\mu = \alpha_0 \oplus a$ and sends it via the classical channel to B. Meanwhile, B got input β, evaluated $y = \beta$ on part B of the shared box and obtains b as output. Lastly, B receives the message μ' from A and makes the guess $g_\beta = \mu' \oplus b$. An illustration of the described protocol in a RAC setup is shown in figure 2.2.

If the classical communication channel is taken to be noiseless then $\mu' = \mu$, and so $g_\beta = \mu \oplus b = \alpha_0 \oplus a \oplus b$. Assuming that $P(a, b|x, y)$ is a

PR-box also allows us to substitute the relation $a \oplus b = xy$, which always holds by definition of the PR-box. Then $g_\beta = \alpha_0 \oplus xy = \alpha_0 \oplus (\alpha_0 \oplus \alpha_1)\beta$ gives indeed $g_0 = \alpha_0$ and $g_1 = \alpha_1$ with certainty, which corresponds to a success probability of 1 in the RAC game.

Motivation for information casuality as a physical principle

The above example, however, has an implausible consequence. Although only a single bit μ is communicated, B seems to have error-free access to whatever bit-index β is queried. Mathematically, the *potential information* shared between A and B is $I(\overrightarrow{\alpha} : \mu', B) \equiv H(\overrightarrow{\alpha}) + H(\mu', B) - H(\overrightarrow{\alpha}), \mu', B) = 2 - H(\overrightarrow{\alpha})|\mu', B)$, since for two uniformly distributed input bits $P(\overrightarrow{\alpha}) = \frac{1}{2}$ with Shannon entropy $H(\overrightarrow{\alpha}) = 2$. Party B was able to apply the van Dam protocol consistently to construct a correct guess by evaluating B and receiving μ' for either of the two inputs bits $\overrightarrow{\alpha}$. Thus clearly there is no uncertainty in $\overrightarrow{\alpha}$ left when B and μ' are known, so $H(\overrightarrow{\alpha})|\mu', B) = 2$ and $I(\overrightarrow{\alpha} : \mu', B) = 2$ bits. [7]

Intuitively, communicating one bit should at most give access to one of the two bits. This should be the case within classical systems but also in (multipartite) quantum systems. Even if the two parties share entangled quantum states, the well-known no-communication theorem in quantum mechanics forbids that quantum entanglement can transfer information and, therefore, quantum non-locality should not give an advantage in information retrieval tasks.
This limitation on the potentially accessible information by the communication capacity is the principle of information causality.

It is crucial to notice here the subtle difference between potential and retrievable information. Actually, it is similar to the Holevo bound in quantum information. Namely, while a multi-qubit state is specified by an exponential number of coefficients (i.e. potential information), one can obtain through measurements at most a linear number of bits (i.e. extractable information).
Indeed, even a PR-box does not allow party B to guess in a RAC more than one bit after a single bit of communication. It only enables a free, completely local choice of which bit to retrieve. As soon as the bit-index β is determined and B has evaluated his part of the no-signaling box on it, any information about the other input bits $\overrightarrow{\alpha}$ is lost.

[7] For a more rigorous proof, see [16]

2.2.3 Formalising IC in non-local games

Information causality is not merely the consequence of a physical axiom, but is actually a very fundamental information-theoretic property that follows from a few mathematical but intuitive assumptions about the notion of entropy.

The study of classical and quantum information theory has revealed several ways to quantify the information content of (complex) physical systems. In particular, mutual information $I(A : B) = H(A) + H(B) - H(A, B)$ based on Shannon entropy $H(\cdot)$ is a popular choice for quantifying the information that is shared between multiple (classical) random variables. While this type of mutual information has many interesting properties, most of them can actually be derived from just a few basic characteristics of the conditional mutual information, $I(A : B|C)$. Those include [27, 64]:

- *Non-negativity* : $I(A : B|C) \geq 0$ (2.9)
- *Markov independence* : If $A \to B \to C$ a Markov chain with respect to physically implementable transformations, then $I(A : C|B) = 0$. (2.10)
- *Data processing inequality* : If $B \to \tilde{B}$ is a physically implementable transformation (i.e. a permissible map), then $I(A : B) \geq I(A : \tilde{B})$. (2.11)
- *Chain rule* : $I(A, B : C) = I(B : C|A) + I(A : C)$ (2.12)
- *Consistency* : If A, B and C classical random variables, then $I(A : B|C)$ exactly matches the (conditional) Shannon mutual information of A and B conditioned on C. (2.13)

for any suitable random variables A,B, $\tilde{B}$ and C. These fundamental properties do hold for Shannon entropy with classical random variables but also for von Neumann entropy with A,B, $\tilde{B}$ and C the density matrices of quantum systems. More generally, the above properties are considered essential for an entropy measure to serve as a physical measure of (shared) information [64] and, as it happens, they are also sufficient for information causality to hold in any multipartite system with limited communication between parties. In other words, if there is a mutual information $I(\cdot, \cdot\cdot)$ that satisfies eq. 2.9 - 2.13 for some given random variables A,B, $\tilde{B}$ and C in a physical system, then the physical system satisfies information causality.

In this work, we focus on bipartite scenarios, where two subsystems share a no-signaling box $P(a,b|x,y) \in \mathcal{NS}$ and where a noisy one-way communication channel $\mu \mapsto \mu'$ exists.
The mathematical statement of information causality in such communication scenarios reads: [16]

$$I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu') \equiv \kappa \tag{2.14}$$

For completeness, we have included a derivation in the appendix A, which is a slightly modified variant of the derivation in [27].

Although the physical interpretation of IC in terms of mutual information is more intuitive, it is heavily depending on the definition of mutual information $I(X:Y) = H(X) + H(Y) - H(X,Y)$. Deriving and stating the IC principle directly in terms of entropy is thus a bit more fundamental. This was also remarked in [65] and they proposed to rewrite the above inequality as

$$H(\overrightarrow{\alpha}|\mu', B) \geq H(\overrightarrow{\alpha}) - H(\mu')$$

with $H(\mu') = \kappa + H(\mu'|\mu)$, which quantifies the minimal uncertainty that B should have after receiving the message μ'. [65] In fact, deriving this purely entropic inequality required only three properties for the entropy $H(\cdot)$, as opposed to the four properties (eq. 2.9 - 2.13) required for the mutual information $I(\cdot,\cdot)$.

For independently distributed inputs $\overrightarrow{\alpha}$, we can express eq. 2.14 furthermore in terms of B's guesses g_β

$$\sum_{i=0}^{n-1} I(\alpha_i; g_\beta \,|\beta = i) \leq I(\overrightarrow{\alpha} : \mu', B) \leq \kappa \tag{2.15}$$

That is, the information shared between A's input data and all the information that is available to B for constructing the guess g_β is upper-bounded by the channel capacity κ.

In the $\kappa = 0$ case, we see that information causality reduces to the no-signaling principle. Even some properties of no-signaling, like the composition restrictions of quantum systems found in [44], are conjectured to hold for information causality as well. [24]
Moreover, in the special case $\kappa = 1$, violating the principle of non-trivial communication complexity implies violation of information causality.

For an extended discussion on interpreting the IC statement, see the last part of appendix A.

2.2.4 From the IC criterion to a bound on no-signaling distributions

Information causality was originally proposed in [16] as a candidate principle for recovering the quantum set of distributions $\mathcal{Q}'$ solely from information-theoretic constraints. In contrast to the above IC statement, it was specialised to the context of (EA)RAC games and also specifically assumed the van Dam protocol for party A and B. Instead of fixing the shared no-signaling box $P(a,b|x,y)$ to be a PR-box, they also considered a more general EARAC setup in which any box $P(a,b|x,y) \in \mathcal{NS}$ in a $(2,2,2,2)$ Bell scenario could be shared by the two parties.
So the setup considered in the original IC paper [16] was similar to figure 2.2 for $n = 2$. The only difference is that they restricted the setup to a noiseless channel (i.e. $\mu' = \mu$) with a capacity of $\kappa = 1$ bit.

From the definition of the van Dam protocol, it is easy to see that either α_0 or α_1 is guessed correctly by B whenever $a = b$, except if $y = \beta = 1$ and $x = \alpha_0 \oplus \alpha_1 = 1$. Thus the guessing biases, with respect to random guessing, for $\beta = 0$ and $\beta = 1$ are respectively: [16, 19]

$$E[\alpha_0 = g_0] = (P(a = b|0,0) - \frac{1}{2}) + (P(a = b|1,0) - \frac{1}{2}) = \frac{1}{2}(E_{00} + E_{10})$$
$$E[\alpha_1 = g_1] = (P(a = b|0,1) - \frac{1}{2}) + (P(a \neq b|1,1) - \frac{1}{2}) = \frac{1}{2}(E_{01} - E_{11})$$

with the probabilistic Bell correlators $E_{xy} = 2P(a = b|x,y) - 1$ and using $P(a = b|x,y) = 1 - P(a \neq b|x,y)$.
If α_0 and α_1 are independent and uniform random, then $H(\alpha_0) = H(\alpha_1) = 1$ and the IC statement (eq. 2.15) can be expressed conveniently in terms of those biases: [16, 32]

$$\begin{aligned}\kappa \geq \sum_{i=0}^{n-1} I(\alpha_i; g_\beta|\beta = i) &= \sum_{i=0}^{n-1} H(\alpha_\beta|\beta = i) + H(g_\beta|\beta = i) - H(\alpha_\beta, g_\beta|\beta = i) \\ &= 2 - \sum_{i=0}^{n-1} H(\alpha_\beta|g_\beta, \beta = i) = 2 - \sum_{i=0}^{n-1} H(\alpha_\beta = g_\beta|\beta = i) = 2 - \sum_{i=0}^{n-1} H\left(\frac{1 + E[\alpha_i = g_i]}{2}\right)\end{aligned} \tag{2.16}$$

where the second to last step uses a well-known result in information theory called Fano's inequality, which is an equality here.

To extend the applicability of their setup to n inputs, Pawlowski et al used in [16] a technique called concatenation that combines multiple copies of a $(2, 2, 2, 2)$-box. In particular, their approach involved an exponential number of no-signaling boxes, whereby A combines pairs of inputs as $\alpha_k \oplus \alpha_{k+1}$ and retrieves a single output a_k per pair. The outputs of two different pairs, a_k and $a_{k'}$, are then fed as input pair to another copy of the box. By repeating this over multiple steps, an exponential number of inputs $\overrightarrow{\alpha}$ can be reduced to a single message bit, which subsequently can be send over the classical channel to B. A similar step-wise evaluation of the boxes on B's side can then indeed be shown to recover any of A's n input bits when using PR-boxes. [16]

A combinatoric argument then showed that the evaluation of concatenated boxes actually increases the noise in the final message bit μ such that the information that B gains from the message decreases with increasing n. Consequently, the guessing biases $E[\alpha_i = g_i]$ asymptotically vanish for all $i \in [n]$. Expanding the entropy function $H(\cdot)$ in eq. 2.16 in the limit $n \to \infty$ then finally gives a simple quadratic constraint on the space of distributions $P(a, b|x, y)$ in terms of their corresponding correlators E_{xy}:

$$IC_{RAC}\left[P(a, b|x, y)\right] \equiv (E_{00} + E_{10})^2 + (E_{01} - E_{11})^2 \leq 4 \tag{2.17}$$

The above is actually a variant of the quadratic Uffink inequality [66] with permuted output labels, which is strictly stronger than the CHSH inequality. Therefore, we call 2.17 an Uffink-like inequality.

Note that distilling the bound 2.17 from 2.15 required the explicit specification of a protocol that A and B apply to determine a value for the guess g_β. The van Dam protocol was useful in this case since the PR-box, as an extremal no-signaling box, exactly saturates the perfect guessing probability of 1.0 and the maximum mutual information of $I(\alpha_0, \alpha_1 : \mu', B) = H(\alpha_0, \alpha_1) = 2$ bits. The maximally mixed box $P_{\mathbb{I}}(a, b|x, y) = \frac{1}{4}$, in contrast, only allows only random guessing (success probability of $\frac{1}{2}$) and thus $I(\alpha_0, \alpha_1 : \mu', B) = H(\alpha_0, \alpha_1) = 0$. So it seems plausible that, by using the van Dam protocol, some probabilistic mixture of P_{PR} and $P_{\mathbb{I}}$ will sit on the boundary between satisfying and violating the IC principle. In fact, the mixture $P(a, b|x, y) = \eta P_{PR} + (1 - \eta) P_{\mathbb{I}}$ with a maximal value of η $(0.0 \leq \eta \leq 1.0)$ such that $P(a, b|x, y) \in \mathcal{Q}'$ has a CHSH value of exactly

$2\sqrt{2}$, i.e. the Tsirelson bound. [16, 32] Remarkably, despite the minimal assumptions that IC makes on the mutual information, it was later shown that Tsirelson's bound can also be derived solely from assuming a generalised data-processing inequality 2.11. [67]

2.2.5 Generalising and simplifying the derivation of IC bounds on no-signaling distributions

The study of information causality has, since the proposal in [16], mostly hold on to the RAC-based formulation and rather focused on generalising and simplifying the protocol of the (EA)RAC game.

A key issue with the original protocol in [16] was the concatenation procedure. While it was effective for deriving the Tsirelson bound and a quadratic constraint, it required an infinite number of copies of the no-signaling boxes. This is unpractical for experimental tests of IC and makes the derivation itself quite laborious. Also it is unclear how to generalise the protocol with concatenation to boxes in scenarios other than $(2, 2, 2, 2)$. To resolve that issue, [26] suggested to replace the concatenation of boxes by making the communication channel noisy.

In concatenation, the bound was strongest for very large n because the sequential evaluation of box copies by party A resulted in diluting the information about input bits $\overrightarrow{\alpha}$. That way, the message bit μ becomes less informative and party B's guessing success becomes more dependent on the correlated side-information b that party B obtains by querying the no-signaling box with input β. Even boxes with weaker non-locality are then sufficient to decode the little amount of information that is left within the message bit μ, which is noisy by its construction.
By modelling the whole process of multiple box copies instead as a noisy channel with some noise parameter p_c, the exact same effect is achieved and the bound 2.17 is rederived. More concretely, a binary symmetric channel with error probability p_c is used, which implies a communication capacity of $\kappa = 1 - h(p_c)$. [27]

In addition to the already known results, they showed that the new protocol is applicable to an arbitrary (n, m, d, d) Bell scenario, though they explicitly only demonstrated it in the $(3, 3, 2, 2)$ case. The only relevant assumption is that all input bits $\overrightarrow{\alpha}$ are treated equally, in that sense any cooperative strategy of party A and B must yield the same guessing proba-

bility for all bit indices β.

Still the derivation in [26] relied somewhat on heuristic arguments and had to be tailored to a specific setting. Only recently, the noisy channel approach was made more systematic by Jain et al in [27].
In the first step, the conditional guessing probabilities $P(g_\beta|\overrightarrow{\alpha} = \overrightarrow{j}, \beta = i)$ for any $i \in [n]$ and $\overrightarrow{j} \in [d]^n$ are calculated in terms of the Bell correlators E_{xy} of the no-signaling box and the noise parameter p_c of the classical symmetric noise channel.

The mutual information terms $I(\overrightarrow{\alpha} : \mu', B)$ in 2.14 can then already be fully determined by those probabilities. However, their expressions still involve the highly non-linear Shannon entropy $H(\cdot)$. The method presented in [27] eliminates those by exploiting a key property of the entropy function, namely that the derivatives of $H(\cdot)$ are piecewise well approximated by polynomial functions. By taking the derivatives of both sides of 2.14 and taking the limit to a zero capacity channel, a polynomial inequality in terms of the box biases E_{xy} is derived. The inequality is still valid because the derivates are taken by applying L'Hoptials rule. In the limit of a fully noisy channel (e.g. $p_c \to \frac{1}{2}$ for a binary symmetric channel), both sides of 2.14 vanish and so L'Hopitals rule is applicable on their ratio.

Using this method, bounds for a whole family of $(n, n, 2, 2)$ scenarios was derived in a single calculation and the quadratic, Uffink-like inequality (eq. 2.17) was obtained as a special case. [27]
Furthermore the $(d, 2, d, d)$ and (n, n, d, d) families were studied to demonstrate the efficiency of their derivation method. Notably, the $(n, n, 2, 2)$ bounds are actually tighter than those found for the equivalent concatenation setup in [16] or via the noisy channel approach for $n = 3$ in [26]. The inequalities derived for the $(d, 2, d, d)$ case, in contrast, match those of an earlier work. [68]

Lastly, while all of the above works focused on the bipartite setting, information causality was also studied for tripartite distributions $P(a, b, c|x, y, z)$ in [69]. For this, however, they only considered the case of bipartite nonlocality between different bipartitions of the three parties.
Pollyceno et al, in contrast, recently proposed in [70] a genuine multipartite generalisation of the principle using the novel framework of quantum causal structures. Instead of one sender A and one receiver B, they consider multiple senders and a single receiver. The goal is compute a certain

function f_β which depends on the β-th bit in the data of each sender. [70]

2.2.6 IC bounds in the CHSH scenario

In [20, 71] it was observed that for a few symmetric[8] types of boxes, the quadratic IC bound (eq. 2.17) actually coincides perfectly with the quantum boundary within the space of bipartite no-signaling distributions $P(a, b|x, y)$. This is a remarkable observation since none of the other operational principles was yet found to be tightly bounding the quantum set $\mathcal{Q}'$ for any specific subset of distributions, except from the very special case of a noisy PR-box $P(a, b|x, y) = \eta P_{PR}^{\mu\nu\sigma} + (1 - \eta)/4$ for $0.0 \leq \eta \leq 1.0$. [11, 32] However, also the IC principle seems to fail on tightly bounding the quantum set $\mathcal{Q}'$ for many other subsets of (asymmetric) boxes when testing IC violations with the quadratic IC bound. [19, 20]

The main open question on the topic of IC thus remains to what extend the remaining gap between the IC bound 2.17 and the quantum boundary can be reduced by deriving stronger IC constraints within the space of bipartite distributions $P(a, b|x, y)$.

As previously discussed, progress has been made on tightening the bound for more complex Bell scenarios than the CHSH scenario, like $(n, n, 2, 2)$ and (n, n, d, d) [27], and for multipartite settings [70]. However, in the simplest $(2, 2, 2, 2)$ scenario, there is to date no IC bound that is equally strong or stronger than the quadratic IC inequality (eq. 2.17) across the whole no-signaling polytope. Even the novel technique presented by Jain et al ([27]) for finding non-locality bounds from IC criteria gave only improvements for higher dimensional distribution spaces, i.e. for $P(a, b|x, y)$ with $a, b \in [d]$, $x, y \in [n]$ and $d, n \geq 3$.

Importantly, the emphasis here lies on *universally* tighter bounds in the CHSH scenario. That is, bounds which are the strongest in any given no-signaling slice.
When disregarding this demand for universality, there are actually some examples of stronger bounds that were derived for specific no-signaling slices. However, those come at the cost of even worse performance than the quadratic IC inequality in other parts of the polytope. A specialised constraint from [27] for the case of strongly correlated inputs α_0 and α_1, for

[8] Symmetric boxes refers to so-called "isotropic boxes", which are mixtures of one or more PR-boxes and white noise ($P(a, b|x, y) = 1/4$)

example, can detect significantly more IC-violating non-local boxes than 2.17 when focusing on mixtures of two LD-boxes (eq. 2.2) and one PR-box (eq. 2.8). [27] However, as we will show in the next chapter, the criterion for strongly correlated inputs is nearly equivalent to the no-signaling boundary for the subset of boxes considered in [19]. [9]

Furthermore, a few *numerical* bounds have been suggested which are actually more restrictive than Uffink across the whole no-signaling set.
To construct those, however, the initial RAC-based formulation of IC had to be generalised. All of the above works namely rather held on to the original approach of baking the RAC game into the definition and derivation of IC.

One example of such a numerical bound was presented in the work by Yu et al [23]. They reformulated IC as an abstract RAC-independent retrieval task by using the notion of redundant information. This notion stems from the so-called partial information decomposition, which splits the information of multiple variables X_i about some target variable T up into three contributions: unique, synergistic and redundant information. If we consider that party B retrieves two pieces of information, G_1 and G_2, about the data $\overrightarrow{\alpha}$, then redundant information quantifies how much information G_2 contains about $\overrightarrow{\alpha}$ that was already available when retrieving G_1, and vice versa. The unique information gained by B from G_1 and G_2 about $\overrightarrow{\alpha}$ is then just the sum of mutual information terms $I(\overrightarrow{\alpha} : G_1)$ and $I(\overrightarrow{\alpha} : G_1, G_2)$, subtracted by the overlapping (i.e. redundant) information $I_{red}(G_1, G_2 \mapsto \overrightarrow{\alpha})$. Enforcing IC on this setup means that this potentially retrievable information is upper-bounded by the classical capacity of any communication channel between A and B, just like for the RAC.
The other generalisation in [23] is that B's query β no longer denotes a specific bit-index in A's input string $\overrightarrow{\alpha}$). Rather β can label any (partial) piece of information about the data $\overrightarrow{\alpha}$).

Unfortunately, also this modified IC statement by Yu et al does not prevent from choosing a specific protocol for A and B. The mutual and redundant information terms are fully determined by the probabilities $P(\overrightarrow{\alpha}, G_i | i = \beta)$, but expressing those in terms of the box correlations $P(a, b|x, y)$ requires knowing how the values of $a, b, x, y, \overrightarrow{\alpha}, G_i$ and β are related to each other.

[9] One of the authors of [27] pointed out that if an arbitrary correlation strength between the input bits is allowed, one can optimise over the correlation strength for any set of distributions $P(a, b|x, y)$ to obtain a bound that is universally stronger than Uffink.

In [23], Yu et al decided to keep using the van Dam protocol but tested the violation of IC with their modified and generalised statement. In doing so, they numerically demonstrated equal or tighter bounds in the same three slices that were already considered for the original formulation of IC [19]. [10] However, while their IC bound matched the quantum boundary in two of the three no-signaling slices, it did not in the third slice. In fact, the IC-quantum gap in the latter case exactly matched the one observed already for the Uffink-like IC boundary.

The quantum-tight IC bounds in [23] were hereby obtained for all kinds of isotropic[11] boxes, while the failing slice contains mixtures of a PR-box, a Local-Deterministic box and white noise ($P(a,b|x,y) = 1/4$). Surprisingly, an even earlier proposed generalisation of IC by Chaves et al [22] gave a stronger bound than the quadratic IC bound for those anisotropic [12] boxes. They formulated IC within the framework of so-called quantum causal structures, and leveraged SDPs to optimise the potential information of B that is still compatible with those information-theoretic structures. [22] A key consequence of this approach is that, like in [23], the potential information takes relative information about A's bits $\overrightarrow{\alpha}$ into account, which makes boxes with weaker non-locality sufficient to violate the IC bound.
However, although employing causal structures allows to detect more IC violating boxes than the original IC inequality, the improvement is only slight and the gap to the quantum boundary stays relatively large for anisotropic boxes. The principle of local orthogonality, in contrast, is in those parts of the no-signaling polytope significantly stronger and even nearly tight for the quantum set. [45]

[10] While preparing this work, it was remarked in private communication that the paper made a mistake in proving that the IC statement is satisfied by (all) quantum distributions $P(a,b|x,y) \in \mathcal{Q}'$. Nevertheless, we consider it relevant in discussing the different perspectives which have been taken on IC.

[11] I.e. Box mixtures with one or more PR-boxes and white noise ($P(a,b|x,y) = 1/4$), but *no* Local-Determinstic box.

[12] Box mixtures with one or more PR-boxes and white noise ($P(a,b|x,y) = 1/4$), but *no* Local-Determinstic box.

2.3 Non-Locality Distillation

All pure entangled quantum states are resources for establishing non-local correlations. In contrast, measurements on mixed states are not necessarily able to produce distributions $P(a,b|x,y)$ outside of the local set $\mathcal{L}$, even when the mixed state is inseparable [32] Some of the well-known Werner states, for example, are entangled but not all of them do exhibit non-locality by themselves.
However, it turns out that by combining multiple copies of those states, one can still simulate correlations beyond the local set. This is called super-activation of non-locality. More generally, one can take any no-signaling box $P(a,b|x,y)$ and increase the amount of observed non-locality by sharing multiple instances of it, which then goes by the name *non-locality distillation*.
Even if a box $P(a,b|x,y)$ already violated a Bell inequality, non-locality distillation can strengthen the violation. With respect to a CHSH functional $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$, this means that the value of $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ can be increased.

If now multiple copies $\{P(a_1,b_1|x_1,y_1),\ldots,P(a_N,b_N|x_N,y_N)\}$ of a box $P(a,b|x,y) \in \mathcal{NS}$ are shared between the two space-like separated parties, the inputs $\{(x_1,y_1),\ldots,(x_N,y_N)\}$ and outputs $\{(a_1,b_1),\ldots,(a_N,b_N)\}$ can be locally processed in such a way that effectively another single no-signaling box with distribution $Q(a,b|x,y) \in \mathcal{NS}$ is obtained. This is exactly what *wirings*, a specific form of non-locality distillation, are about. [71, 72]
The effective non-locality of the wired box $Q(a,b|x,y)$ can hereby be higher or lower than the non-locality in the original distribution $P(a.b|x,y)$. However, $Q(a,b|x,y)$ can never have more non-locality than the sum over the individual instances of $P(a,b|x,y)$. Thus, for any tuple of coefficients (s_1,s_2,s_3,s_4):

$$\sum_{c=1}^{N} S_{CHSH}^{(s_1,s_2,s_3,s_4)}\left(P(a_N,b_N|x_N,y_N)\right) \geq S_{CHSH}^{(s_1,s_2,s_3,s_4)}(Q(a,b|x,y)).$$

Although wirings are represented in various ways throughout the literature, we will refer to the following definition:

Definition 2.3.1. ***Wiring*** *— A wiring W between two boxes $P_1(a_1, b_1|x_1, y_1) \in \mathcal{NS}$ and $P_2(a_2, b_2|x_2, y_2) \in \mathcal{NS}$ in a bipartite $(2, 2, 2, 2)$ Bell scenario is a tuple of boolean functions $(f_{in}^{(1)}, f_{in}^{(2)}, f_{out}, g_{in}^{(1)}, g_{in}^{(2)}, g_{out})$ (i.e. the "wires") where*

$$f_{in}^{(1)} \mapsto f_{in}^{(1)}(x, a_2) \equiv x_1 \qquad f_{in}^{(2)} \mapsto f_{in}^{(2)}(x, a_1) \equiv x_2 \qquad f_{out} \mapsto f_{out}(x, a_1, a_2) \equiv a$$
$$g_{in}^{(1)} \mapsto g_{in}^{(1)}(y, b_2) \equiv y_1 \qquad g_{in}^{(2)} \mapsto g_{in}^{(2)}(y, b_1) \equiv y_2 \qquad g_{out} \mapsto g_{out}(y, b_1, b_2) \equiv b$$

for any $a, b, a_1, b_1, a_2, b_2, x, y \in \{0, 1\}$ such that the following non-cyclicity[13] conditions are satisfied: [72]

$$f_{in}^{(1)}(x, a_2) = f_{in}^{(1)}(x) \qquad or \qquad f_{in}^{(2)}(x, a_1) = f_{in}^{(2)}(x) \tag{2.18}$$

and equally for the $g_{in}^{(i)}$-functions.

The functions $f_{in}^{(k)}$ ($g_{in}^{(k)}$) hereby specify the input x_k (y_k) into the k-th box within subsystem A (B). On the other hand, f_{out} (g_{out}) gives the output of the (effective) wired box a (b) from subsystem A (B).

Applying a wiring W to two boxes results in another (effective) no-signaling box $W(P_1, P_2) = Q(a, b|x, y) \in \mathcal{NS}$. Consequently, N boxes can also be wired to a single box $W(P(a, b|x, y)^{\times N}) = R(a, b|x, y) \in \mathcal{NS}$ by an iterative process from two-box wirings. [72] Hereby, it is important to remark that wiring more than two boxes is an ambiguous process. Three instances of some box $P(a, b|x, y)$, for example, can be wired as $W(W(P, P), P)$ or $W(P, W(P, P))$, which generally do *not* result in the same box. The operation of a wiring W is namely non-associative and *not* symmetric in its arguments [72]. The former variant, however, can be shown to result in a box with the highest amount of non-locality. If we wire N copies of the same box $P(a, b|x, y)$, we will assume $W(...W(W(P, P), P)..., P)$ as the canonical way to perform the wiring. Furthermore, we call $N - 1$ the *wiring order*. whereby an increasing order implies a more complex wiring process.

In above definition, indexed values (a_1, a_2, b_1, b_2) denote the (hidden) outputs of boxes $P_1(a_1, b_1|x_1, y_1)$ and $P_2(a_2, b_2|x_2, y_2)$. The (hidden) inputs of

[13] Non-cyclicity prevents that the inputs of each box depends on the output of the other box. At least one box input (e.g. x_1 or x_2) must solely depend on the global input x. Otherwise there is a logical loop and none of the boxes could be queried in the first place.

P_1 and P_2, on the other hand, are implicitly determined by the wiring functions as $x_1 \equiv f_{in}^{(1)}(x, a_2)$ and $x_2 \equiv f_{in}^{(2)}(x, a_1)$ for party A's part of the wired box. Equivalently for B's side: $g_{in}^{(1)}(y, b_2) \equiv y_1$ and $g_{in}^{(2)}(y, b_1) \equiv y_2$.
We can understand (x, y) as the initial input values to the wired box $Q(a, b|x, y)$ and (a, b) as the final output values of the wired box. While the inputs (x, y) are given when querying the (global) wired box $Q(a, b|x, y)$, the outputs (a, b) result from post-processing the outputs (a_1, a_2) and (b_1, b_2) of the individual boxes for party A and B respectively. This happens according to the wiring functions, namely $a \equiv f_{out}(x, a_1, a_2)$ and $b \equiv g_{out}(y, b_1, b_2)$. [72, 73] The processing is separated between the parties, by f and g functions respectively, to preserve the no-signaling property of the wired boxes. For common examples of wirings and their insightful visualisation, we refer the reader to figure 3 in [72].

Note that wirings are somewhat different from the previously mentioned process of concatenating no-signaling boxes. Although the latter can also be used for non-locality distillation, there are n inputs $\overrightarrow{x} \in \{0, 1\}^n$ and $\overrightarrow{y} \in \{0, 1\}^n$ for each party that are distributed in parallel over the different boxes $\{P(a_1, b_1|x_1, y_1), \ldots, P(a_N, b_N|x_N, y_N)\}$. In contrast to wirings, the resulting setup as a whole can *not* be viewed as an (effective) no-signaling box. The number of inputs of a wired box, must namely match that of the individual boxes to preserve the structure of a bipartite output distribution $P(a, b|x, y)$. So wiring N for bipartite boxes, each party can only provide a single input to the wired box that is subsequently relayed to only a single box from $\{P(a_1, b_1|x_1, y_1), \ldots, P(a_N, b_N|x_N, y_N)\}$.

Studying wirings is extremely non-trivial since there are a huge number of valid ways to connect the inputs and outputs, even for just two boxes. This becomes only worse if one goes beyond deterministic wirings and takes also probabilistic mixtures of those wirings into consideration [72]. Fortunately, similar to the two compact sets of distributions, $\mathcal{L}$ and $\mathcal{NS}$, the set of wirings turns out to be convex and compact as well. [72, 74] Particularly for wirings from definition 2.3.1, every such wiring can be reproduced from a mixture of $82^4 = (6,724)^2 = 45,212,176$ extremal wirings. Hereby, 82 is the number of extremal wirings that one can choose from independently for each of 2 possible inputs x for party A and each of 2 inputs y for party B. All those 82 wirings can be composed from only 5 different parameterised types, labelled as "deterministic", "one-sided", "XOR-gated", "AND-gated" and "sequential" respectively. For details on the different types of extremal wirings, see Table 1 in [74].

The simplest non-trivial wiring types are the ones where the (independent) outputs of both boxes (a_1, a_2) $((b_1, b_2))$ are used in conjunction to produce the effective box output a (b). From the table of extremal wirings, we see that this involves either XOR-gating or AND-gating the outputs. Any other combination of the outputs can be produced as a mixture of these two atomic operations.
Lastly, the sequential wiring type is most notable type from the table as it is the only one that actually uses the output of one box as an input for the other box. The other four extremal types rather query the boxes P_1 and P_2 independently, and post-process their outputs to produce the final outputs (a, b) of the wired box $Q(a, b|x, y)$.

The importance of considering wirings lies in their potential to construct an effective box $R(a, b|x, y) \in \mathcal{NS}$ that might exceed the boundary of any restricted subset of distributions $S \subset \mathcal{NS}$ such that $Q(a, b|x, y) \notin \mathcal{S}$, even when the underlying individual boxes were contained in that same subset $\mathcal{S}$, i.e. $P_i(a_i, b_i|x_i, y_i) \in \mathcal{S}$ for all $i \in [1, n]$.
A set of distributions $\mathcal{S}$ is called *closed* under wirings if there does *not* exist any subset of boxes $P_1, P_2, \ldots \in \mathcal{S}$ and any wiring W for which the correlation lies outside the set $\mathcal{S}$, i.e. $W(P_1, P_2, \ldots) \notin \mathcal{S}$. Equally, we say that some constraint $\mathcal{C}(P(a, b|x, y)) \leq C$ on the correlation space is stable (under wirings) if and only if its enclosed subset is not closed.

On the one hand, wirings are plausible ("physical") local operations in the assumed LOCC framework, provided that a sufficient number of boxes N (or copies of a single box) are available. On the other hand, for any $\mathcal{S} \subset \mathcal{NS}$ which is not closed under wirings, it is impossible to define a physical self-consistent theory [14] that generates the distributions in $\mathcal{S}$. [75] This means that, starting from theory-consistent distributions, there would otherwise be the possibility to produce (by local operations that correspond to some suitable wiring) a distribution outside of the theory domain, where the theory is not well-defined anymore. Consequently, any constraint $\mathcal{C}(P(a, b|x, y)) \leq C$ that is not stable under wirings is an unphysical constraint.

Unfortunately, proving closure of a set $\mathcal{S}$ is very involved and tailored to each situation, making success highly dependent on the particular struc-

[14] "Theory" is meant here in the most abstract and general way possible, since it isn't well-defined within the device-independent framework of black-box distributions $P(a, b|x, y)$. The reader familiar with GPTs may, without loss of generality, assume that "theory" here refers to some constrained variant of the well-known "boxworld" GPT. [13]

ture of $\mathcal{S}$. Therefore, very few non-trivial closed subsets of $\mathcal{NS}$ are currently known for which the closure has actually been proven. [75] The no-signaling set $\mathcal{NS}$ itself is the most important example for which it has been shown. Also the quantum set $\mathcal{Q}$ (i.e. the one with the tensor product structure) is closed under wirings, but for the often considered alternative quantum set $\mathcal{Q}'$ it is still unknown. [75] Here we only consider the simplest bipartite Bell scenario where $\mathcal{Q}$ and $\mathcal{Q}'$ are nearly identical. In that sense, $\mathcal{Q}'$ can be considered closed throughout our discussion as well.

Although wirings have not yet been studied explicitly in the context of IC, the original variant of Uffinks inequality from [66] is very similar and was readily shown to be *not* closed under wirings in [71]. Specifically, they proposed a wiring that can distill multiple copies of some Uffink-satisfying mixtures between the canonical $P_{PR}^{(000)}$ PR-box (eq. 2.8), a local determinstic $P_{LD}^{(0101)}$-type box (eq. 2.2) and the maximally random box $P_{\mathbb{I}}(a,b|x,y) = \frac{1}{4}$ to distributions $Q(a,b|x,y) \in \mathcal{NS}$ which then violate the Uffink inequality. In terms of the extremal wiring types from table 1 in [74], the wiring proposed by Allcock et al is composed of AND-gating and XOR-gating the outputs of different box copies. The concrete wiring functions read:

$$\begin{aligned} f_{in}^{(1)}(x,a_2) &= x & g_{in}^{(1)}(y,b_2) &= y \\ f_{in}^{(2)}(x,a_1) &= x \oplus a_1 \oplus 1 & g_{in}^{(2)}(y,b_1) &= yb_1 \\ f_{out}(x,a_1,a_2) &= a_1 \oplus a_2 \oplus 1 & g_{out}(y,b_1,b_2) &= b_1 \oplus b_2 \oplus 1 \end{aligned} \tag{2.19}$$

(2.20)

In addition to this very specific combination of a wiring and a certain type of box mixture, several other counterexamples have been found for the non-closure of Uffink-like inequalities. For instance, when considering non-local boxes in various quantum voids. [73]
Since the quadratic IC inequality 2.17 is also just a variant of the original Uffink inequality with permutated output labels (i.e. swapped $\{0,1\}$ values for both outputs a and b), there must exist some similar wiring to 2.19 which distills IC-consistent boxes to ones that violate the IC principle.

We remark that despite the failure of Uffink, the stability under wirings of IC itself is still inconclusive. We lack a constraint that exactly describes which distributions $P(a,b|x,y)$ are consistent with the IC principle for all possible formulations of the principle and all possible strategies of the two parties. Only with such an IC-tight bound, we could start to assess the stability of IC itself.

Anyhow, if IC turns out to be stable under wirings, the set $\mathcal{IC}$ of IC consistent distributions will definitely be the maximal closed subset of the set enclosed by the Uffink inequality. [71, 75]
Beyond the IC principle, we remark that the set of distributions satisfying the macroscopic locality principle is already known to be closed under wirings. Also the local orthogonality principle is stable in simple bipartite non-locality scenarios, but actually becomes unstable in more general contextuality scenarios. [75]

Chapter 3

Methods, Experiments & Results

3.1 `ABoxWorld`: A modular numerical framework for no-signaling correlations

The formulation of the information causality principle in terms of mutual information (or, equivalently, entropy) is a real challenge. For deriving a direct constraint on the space of no-signaling distribution $P(a, b|x, y) \in \mathcal{NS}$ from IC statements (like 2.15) in a bipartite non-local scenario, one namely has to make the protocol explicit.
On the one hand, this is because the mutual information terms $I(\alpha_i : g_\beta|\beta = i)$ refer to random variables, α_i and g_β, which do generally not have a fixed relation to the box variables (a, b, x, y). On the other hand, it is not evident from the IC statement how exactly the communicated message μ is constructed by party A.

To compute the mutual information $I(\alpha_i : g_\beta|\beta = i)$, one thus first needs to specify the information retrieval task (e.g. RAC), the specific encoding of input data into a classical message μ of party A, the decoding routine for constructing the guess g_β of party B, and the type of (noisy) communication channel $\mu \mapsto \mu'$. This implies that one can always consider only the special case of a specific protocol that party A and B will follow in some non-local game. Unfortunately, this means that we have to test through all possible protocols to determine which protocol gives the strongest bound on IC. Thereby we use that maximising the strength of the bound corresponds to maximising the value of $I(\alpha_i : g_\beta|\beta = i)$ terms in 2.15 for a fixed channel capacity.

However, examining the information retrieval performance $\sum_i I(\alpha_i : g_\beta|\beta = i)$ for many different (family of) protocols would be a tedious task if we want to do it analytically, even when the most recent method by [27] is applied. This thus invites the use of numerical tools to explore the implications of information causality on constraining no-signaling distributions $P(a, b|x, y) \in \mathcal{NS}$.

Within the broader field of quantum information, some numerical libraries already exist for quantum circuits (e.g. Qiskit), simulation of open quantum systems (e.g. QuTiP, QuantumOptics.jl) and even (quantum) entanglement theory (e.g. Toqito). However, there has not yet been any general-purpose library or modular codebase for studying non-local correlations beyond quantum mechanics. Also, most code related to Bell non-locality is written in either proprietary languages, like Matlab, or in popular but inefficient languages, like Python. Therefore, we have implemented our own modular framework in Julia. All of the following numerical experiments have been performed in this new framework. [1]

At the core of the framework are the bipartite no-signaling boxes $P(a, b|x, y)$[2], stored in the compact Collins-Gisin (CG) representation [32, 76]. This representation reduces the $m_A m_B o_A o_B$ components of $P(a, b|x, y)$ in a bipartite (m_A, m_B, o_A, o_B) scenario to only $m_A(o_A - 1) + m_B(o_B - 1) + m_A(o_A - 1)m_B(o_B - 1)$. Because of the normalisation of $P(a, b|x, y)$ and the no-signaling conditions (eq. 2.3 & 2.4), we namely only need to store the marginals $P(a|x)$ and $P(b|y)$, as well as $n - 1$ components of $P(a, b|x, y)$ for each pair of inputs (x, y). So while the benefit in the simplest $(2, 2, 2, 2)$ Bell scenario is limited, this is already a lot more memory efficient for only slightly more complex scenarios.

Various operations can then be performed on the no-signaling boxes. This includes the testing of user-defined Bell inequalities in any bipartite Bell scenario and applying an arbitrary wiring of two boxes in the simplest CHSH-scenario. The implementation of wirings was hereby heavily based upon the python code [3] accompanying [72].

[1] Available on GitHub: `https://github.com/t-rothe/ABoxWorld`

[2] In addition to boxes $P(a, b|x, y)$ with a single input per party, there is also limited support for time-ordered (sequential) no-signaling boxes $P(a, b_1, \ldots, b_t|x, y_1, \ldots, y_t)$. For background information on time-ordered distributions, see appendix B

[3] `https://github.com/Pierre-Botteron/Algebra-of-Boxes-code/tree/main/non_local_boxes`

Since the exact quantum boundary is not representable by a compact constraint, we used the NPA hierachy of SDP programs [4] to test membership of boxes $P(a,b|x,y)\mathcal{NS}$ in a close approximation to the quantum set. [42] Throughout our numerical experiments, we thereby considered NPA at level 3 to be sufficient for detecting significant gaps between the quantum boundary, corresponding to the set $\mathcal{Q}'$, and any boundary implied by the IC constraints.
Two different NPA implementations, QuantumNPA.jl [5] and ncpol2sdpa [6], were used in combination with the proprietary solver Mosek [7]. The former implementation was often preferred since it was also written in Julia and up to 2 orders of magnitudes faster. The ncpol2sdpa python library, however, has a wider set of features and is thus more adaptable to problems whereby any variables, in addition to the NPA moment matrix components, need to be injected into the definition of the optimisation problem. This is the case, for example, when solving the membership problem as a more stable optimization problem, rather than as an possibly unstable feasibility problem.

3.2 Comparing IC bounds

So far, one common limitation of many works on IC has been the focus on demonstrating the strength of IC only in very specifically chosen no-signaling slices, i.e. very specific families of box mixtures. Curiously, no comprehensive comparison of all the known bounds has been made yet. To map out the strengths and weaknesses of each IC bound, we plotted the various IC bounds for the simplest bipartite scenario (i.e. $(2,2,2,2)$) side-by-side in figure 3.1.

Figures 3.1a - 3.1c correspond to the slices that were also chosen in [19] and [23]. The box mixtures represented in figure 3.1d, on the other hand, were first considered in the context of IC in [27]. For comparability and for demonstrating our numerical framework, we hereby held onto the style of the figures in the original works. Those were either line plots with respect to the values of two different types of CHSH-scores, or an area plot

[4] We refer any reader who is unfamiliar with SDPs to [77] for background information, and to [42] for more information specifically about the NPA hierachy of SDPs

[5] `https://github.com/ewoodhead/QuantumNPA.jl`

[6] `https://github.com/peterwittek/ncpol2sdpa`

[7] `https://github.com/MOSEK/Mosek.jl`

that directly represents the coefficient-space (η_1, η_2) of convex mixtures $P(a,b|x,y) \equiv \eta_1 P_1(a,b|x,y) + \eta_2 P_2(a,b|x,y) + (1-\eta_1-\eta_2)P_3(a,b|x,y)$. Due to the normalisation of mixture coefficients $\eta_1 + \eta_2 = 1$, all the 2D-plots in figure 3.1 correspond each to some family of mixtures of exactly three no-signaling distributions $(P_1(a,b|x,y), P_2(a,b|x,y), P_3(a,b|x,y))$.

Regarding the line plots in figures 3.1a - 3.1c, the lines separate distributions that satisfy the respective constraints, lying in the regions below, from those that violate them, lying in the regions above. All relevant distributions $P(a,b|x,y) \in \mathcal{NS}$ are situated in the lower-left triangle, below the black no-signaling boundary, while anything above the triangle can be ignored. Hereby, "relevant" thus refers to the no-signaling property from definition 2.1.2.

Equivalently, each colored area in figure 3.1d identifies all the distributions that are consistent with the corresponding IC or quantum constraint. The blank region in the lower-left violates all constraints (except no-signaling 2.1.2), while the blank area on the right of the figure only contains irrelevant distributions $P(a,b|x,y) \notin \mathcal{NS}$. Inevitably, colored areas overlap towards the top-right corner in this visualisation and thus only the quantum NPA region is fully visible.

The visualisation of the space of distributions $P(a,b|x,y)$ with respect to the values of two different ("orthogonal") CHSH functionals in figures 3.1a - 3.1c has the advantage that each distribution can be directly interpreted by the amount of non-locality that it involves.
Thereby, the two CHSH functionals are chosen such that each mixture of three no-signaling boxes $P(a,b|x,y) \equiv \eta_1 P_1(a,b|x,y) + \eta_2 P_2(a,b|x,y) + (1-\eta_1-\eta_2)P_3(a,b|x,y)$ can be uniquely identified by simply solving a system of linear equations.

Which of the eight CHSH functionals in eq. 2.7 to choose for each slice thus depends on the three extremal boxes in the mixture. Each CHSH functional $S_{CHSH}^{(s_1,s_2,s_3,s_4)}$ is namely associated to a facet of the local set of distributions $\mathcal{L}$ and to a PR-box (corresponding to distributions in 2.8) by which it is maximally violated.
Using this correspondence, the two CHSH variants for figure 3.1a and 3.1b were chosen such that the PR-boxes in the mixture saturate the optimum value for the two CHSH functionals respectively. For figure 3.1c, on the other hand, the CHSH functionals were selected based on whether the LD

box in the mixture (P_{LD}^{0000}) is contained in the associated facets of $\mathcal{L}$.

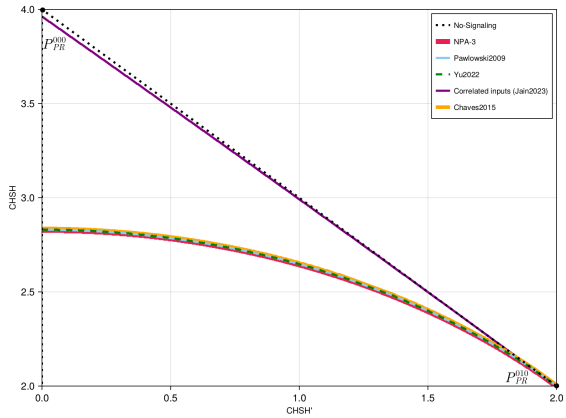
Concretely, the shown slices of $\mathcal{NS}$ are given with respect to the values of $S_{CHSH} \equiv S_{CHSH}^{(1,1,1,-1)}$ (i.e. the canonical CHSH functional), $S_{CHSH'} \equiv S_{CHSH}^{(1,-1,1,1)}$, and $S_{CHSH''} \equiv S_{CHSH}^{(-1,1,1,1)}$ as defined in eq. 2.7. For simplicity we labelled them as CHSH, CHSH', and CHSH'' respectively.

As mentioned previously, computing the mutual information for any of the known IC bounds requires fixing a specific protocol and (noisy) communication channel. For plotting the bounds in 3.1, we assumed the van Dam protocol and a binary symmetric channel. This setup was demonstrated to give optimal RAC success probabilities and the strongest IC bound for all extremal no-signaling distributions in the simplest bipartite scenario. [26, 27] While this does not necessarily imply optimality for mixtures of extremal distributions, there has not yet been found any protocol for the RAC setup with an IC bound that is stronger than the Uffink-like inequality obtained from the van Dam protocol.

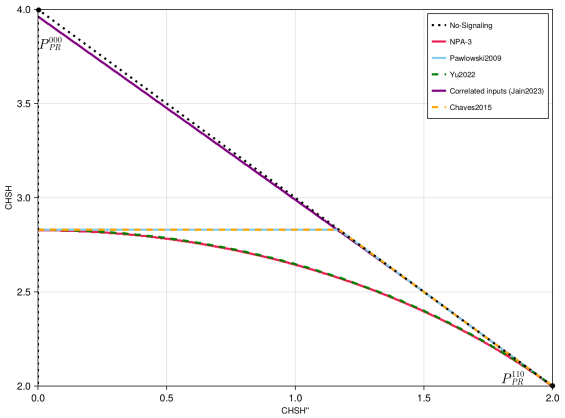
In figure 3.1a, we consider a family of so-called isotropic mixtures between the canonical PR-box P_{PR}^{000}, the maximally mixed box $P_{\mathbb{I}}(a,b|x,y) = \frac{1}{4}$ (i.e. white noise) and the PR-box P_{PR}^{010}. While the more general IC bounds from [16], [22] and [23] exactly reconstruct the quantum boundary, the bound for the special case of nearly perfectly correlated bits (α_0, α_1) from [27] is not much stronger than the no-signaling constraint. This observation is consistent across all figures, 3.1a - 3.1c, except from the slice in 3.1d that was also selected by the authors of [27] themselves to demonstrate their specialised bound. [8]

For a different family of isotropic distributions, $\eta_1 P_{PR}^{000} + \eta_2 P_{PR}^{110} + (1 - \eta_1 - \eta_2) P_{\mathbb{I}}$, we see in figure 3.1b that only the bound by Yu et al from [23] can recover the NPA-based quantum boundary. The original Uffink-like IC inequality and the bound by Chaves et al from [22], on the other hand, are both a horizontal line at the Tsirelson canonical CHSH value of $S_{CHSH} = 2\sqrt{2}$. In both cases the success probabilities, for the respective

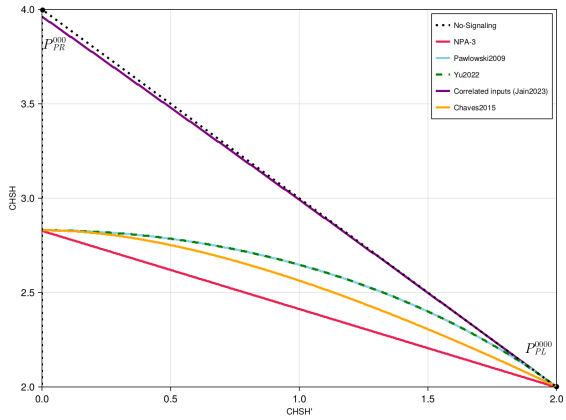
[8] The "specialised" bound for nearly perfectly correlated bits α_0 and α_1 can be generalised to arbitrary correlation strengths between A's data bits $\overrightarrow{\alpha}$. From that perspective, the Uffink-like bound from [16] is equally just a special case, namely the case of uncorrelated bits. The generalised bound matches Uffink for most $\mathcal{NS}$ distributions, but can be stronger in a few non-isotropic slices, like the one drawn in figure 3.1d. Most importantly, however, as a generalisation it can never be weaker than Uffink.



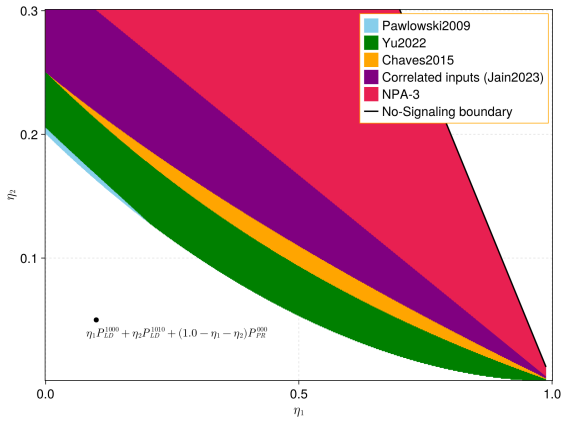
(a) *Projection space of isotropic mixtures* $\eta_1 P_{PR}^{000} + \eta_2 P_{PR}^{010} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$ *with* $0.0 \leq \eta_1, \eta_2 \leq 1.0$ *onto their* $S_{CHSH'}$ *(horizontal) and* S_{CHSH} *(vertical) values. All IC bounds, except the purple bound for nearly perfectly correlated inputs, exactly match the quantum boundary according to NPA level 3. Some lines have been slightly translated vertically to make all overlapping lines visible.*



(b) *Projection space of isotropic mixtures* $\eta_1 P_{PR}^{000} + \eta_2 P_{PR}^{110} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$ *with* $0.0 \leq \eta_1, \eta_2 \leq 1.0$ *onto their* $S_{CHSH'}$ *(horizontal) and* S_{CHSH} *(vertical) values. The IC bound by Yu et al from [23] reconstructs the NPA level 3 boundary, while Chaves2015 matches the original IC bound from [16].*



(c) *Projection space of mixtures* $\eta_1 P_{PR}^{000} + \eta_2 P_{LD}^{0000} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$ *with* $0.0 \leq \eta_1, \eta_2 \leq 1.0$ *onto their* $S_{CHSH'}$ *(horizontal) and* S_{CHSH} *(vertical) values. None of the IC bounds reconstructs the NPA level 3 boundary. While Chaves2015 is for these box mixtures stronger than all other IC constraints, it is not quantum-tight.*



(d) *Space of isotropic mixtures* $\eta_1 P_{PR}^{000} + \eta_2 P_{LD}^{0000} + (1-\eta_1-\eta_2)P_{LD}^{0101}$*, whereby the (partially overlapping) areas indicate the boxes which satisfy the respective constraints. No IC bound recovers the NPA boundary, but the specialised bound for nearly perfectly correlated inputs comes closest. Chaves2015 provides the strongest constraint of the bounds which enclose an Uffink-compatible set.*

Figure 3.1: *Various IC bounds and the quantum boundary across different no-signaling slices. Slices of 3.1a, 3.1b, and 3.1a reproduce & extend plots from [19, 23]. The area plot in 3.1d is based on figure 2 in [27].*

information retrieval tasks, are thus fully independent of the presence of a P_{PR}^{110}-type distribution within the mixture. In other words, the correlation resulting from a PR-box P_{PR}^{110} does not increase B's chances for making a correct guess beyond the success probability that a (classical) local deterministic box (eq. 2.2) would have given.

The remaining two figures, 3.1c and 3.1d, show the bounds for two non-isotropic types of mixtures with very similar observations. On the one hand, a mixture with a single LD-box P_{LD}^{0000} (eq. 2.2) in 3.1c and a mixture with two LD-boxes P_{LD}^{0000} and P_{LD}^{0101} in 3.1d. In both cases, the quantum boundary is simply linear (according to NPA at level 3), but clearly none of plotted bounds is able to reconstruct it. [9]
Nevertheless, in 3.1c the IC statement based on causal structures from [22] results in a boundary that is strictly closer to the quantum boundary than any of the other bounds. The generalisation of IC in terms of redundant information from [23] by Yu et al, in contrast, seems to have *no* advantage for the considered mixture and matches the Uffink-like IC bound in this slice. Thus, Chaves' [22] and Yu's [23] bounds are complementary, in the sense that each is the strictly strongest bound for some type of mixtures.

Although nearly the same ranking between the IC bounds, with respect to their strength, applies to the area plot in figure 3.1d, there are some subtle differences visible. Firstly, there is a growing gap between the original IC bound and the bound by Yu et al with an increasing proportion of the local deterministic box P_{LD}^{1010} in the mixture, i.e. higher η_2. This gap between the two bounds is hereby, in comparison to figure 3.1c, neither a consequence of numerical imprecision nor due to the alternative representation of the no-signaling slice in terms of coefficients (η_1, η_2). Secondly, the specialised bound from [27] for nearly perfectly correlated inputs (α_0, α_1) is now the strongest. While the gap of the correlated-inputs bound to Chaves' bound around the center of the plot ($\eta_1 = \eta_2 = 0.5$) is quite significant, the respective boundaries converge towards each other in the limits single LD-boxes in the mixtures, $\eta_1 \to 0$ or $\eta_2 \to 0$.

[9] It was remarked by one of the authors of [27] that the more general correlated-inputs bound, with arbitrary correlation strength between α_0 and α_1, does actually recover the quantum NPA boundary within this specific slice.

3.2.1 Generalising the observed differences between IC constraints

	Constraint on $I(\overrightarrow{\alpha} : \mu', B)$	Strongest for
RAC success *Pawlowski et al [16]*	$\sum_{i=0}^{n-1} I(\alpha_i; g_\beta \mid \beta = i) \leq I(\mu : \mu')$	$P_{PR}^{0\nu\sigma} + P_{PR}^{0\nu'\sigma'} + P_{\mathbb{I}}$
Causal Structures *Chaves et al [22]*	$\sum_{i=0}^{n-1} I(\alpha_i : g_\beta, \mu' \mid \beta = i)$ $+ \sum_{i=1}^{n-1} I(\alpha_0 : \alpha_i \mid g_\beta, \mu', \beta = i)$ $\leq I(\mu : \mu') + \sum_{i=1}^{n} H(\alpha_i) - H(\overrightarrow{\alpha})$	$P_{PR}^{0\nu\sigma} + P_{PR}^{0\nu'\sigma'} + P_{\mathbb{I}}$ $P_{PR}^{0\nu\sigma} + P_{LD}^{\alpha\gamma\beta\lambda} + P_{\mathbb{I}}$ $P_{PR}^{0\nu\sigma} + P_{LD}^{\alpha\gamma\beta\lambda} + P_{LD}^{\alpha'\gamma'\beta'\lambda'}$
Redundant information *Yu et al [23]*	$\sum_{i=0}^{M-1} I(\overrightarrow{\alpha} : g_\beta \mid \beta = i)$ $- I_{red}(g_0, \ldots, g_{M-1} \mapsto \overrightarrow{\alpha})$ $\leq I(\mu : \mu')$	$P_{PR}^{0\nu\sigma} + P_{PR}^{0\nu'\sigma'} + P_{\mathbb{I}}$ $P_{PR}^{\mu\nu\sigma} + P_{PR}^{\mu'\nu'\sigma'} + P_{\mathbb{I}}$

Table 3.1: *Comparison of the different IC statements proposed so far based on RACs ([16]/[27]), quantum causal structures ([22]) and redundant information ([23]) respectively. Blue and red colored parts in the inequalities indicate the differences in the formulation of IC with respect to the original IC statement (top row). Additionally, red identifies vanishing terms for all settings considered in this work. The $\mathcal{Q}'$-tight constraint on isotropic mixtures of the form $P_{PR}^{0\nu\sigma} + P_{PR}^{0\nu'\sigma'} + P_{\mathbb{I}}$ is obtained by all IC statements. Causal structures and redundant information have unique mixtures for which they achieve the strongest bound, though they are not necessarily IC-tight in those slices.*

By also considering various other mixtures between three extremal boxes, we can summarise and generalise the above observations for the three independent formulations of IC in table 3.1.
From the table, we can see that non-isotropic mixtures containing local deterministic distributions $P_{LD}^{\alpha\gamma\beta\lambda}$ (eq. 2.2) are constrained most strongly by IC through quantum causal structures as proposed by Chaves et al [22].

There are two key differences between Chaves' IC statement and the original IC statement from [16] that explain why they result in bounds of dif-

ferent strength. On the one hand, the mutual information terms in Chaves' IC statement have an additional and direct dependence on the (received) message μ', aside from the indirect dependence via g_β in the original IC statement 2.15. On the other hand, some new terms $I(\alpha_0 : \alpha_i \,|g_\beta, \mu', \beta = i)$ with $i \neq 0$ capture information about α_0 that B inherently obtains whenever he successfully guesses a different bit α_i.

Curiously, throughout all our experiments in $(2, 2, 2, 2)$ scenarios (i.e. $i \in \{0, 1\}$), we observed that $I(\alpha_0 : \alpha_1 \,|g_\beta, \mu', \beta = 1) = 0.0$ up to numerical precision. This suggests that the extra dependence on μ' in $I(\alpha_i : g_\beta, \mu'|\beta = i) > I(\alpha_i : g_\beta|\beta = i)$ is the main contribution of Chaves' improvement to the IC statement. On first sight, this is somewhat surprising since $g_\beta \equiv \mu' \oplus b$ for the van Dam protocol might imply that g_β should encode the same information about α_i as μ', such that

$$I(\alpha_i : g_\beta|\beta = i) = I(\alpha_i : g_\beta, \mu'|\beta = i)$$

for all $i \in \{0, 1\}$ and $P(a, b|x, y) \in \mathcal{NS}$.

However, using LD-boxes $P_{LD}^{\alpha\gamma\beta\lambda}$ introduces a certain asymmetry to the RAC protocol by non-uniform marginals $P(a|x)$ and $P(b|y)$. Consequently, $I(\alpha_i : g_\beta, \mu'|\beta = i) > I(\alpha_i : g_\beta|\beta = i)$ for some $i \in \{0, 1\}$ and $P(a, b|x, y) \in \mathcal{NS}$ since, roughly speaking, the bias in b partially erases information in μ' about α_0 or α_1 when computing $g_\beta \equiv \mu' \oplus b$. This would indeed imply that more boxes $P(a, b|x, y)$ violate IC and thus also result in a tighter IC bound when replacing $I(\alpha_i : g_\beta|\beta = i)$ with $I(\alpha_i : g_\beta, \mu'|\beta = i)$ in the IC statement.

The above reported observations provide numerical evidence that the IC statement by Chaves et al always results in a stronger IC bound than the original one for all mixtures which involve $P_{LD}^{\alpha\gamma\beta\lambda}$ boxes. These non-isotropic mixtures are highly non-intuitive and so the exact reason for the success of Chaves' bound in these no-signaling slices remains unclear to us.

In contrast to causal structures, the generalised IC statement from Yu et al [23] did not show any advantage over the original RAC-based IC formulation from [16] whenever a $P_{LD}^{\alpha\gamma\beta\lambda}$ box is part of a three-box mixture.
However, the optimality of the two bounds is reversed when considering mixtures which contain a $P_{PR}^{1\nu\sigma}$ box. So in that case, the result from Chaves et al [22] matches the Uffink-like IC bound, while the IC formulation by Yu et al from [23] is significantly stronger than Uffink (and is even quantum-tight). This is the complementary of these two IC formulations that we

already observed in the special cases of figures 3.1c and 3.1d.

Accepting partial information about the bits in $\overrightarrow{\alpha}$ is currently unique to the formulation of IC by Yu et al in [23]. In contrast, the mutual information terms in Pawlowski's [16] and Chaves' [22] IC formulations only capture the retrieval of complete and individual bits α_i in $\overrightarrow{\alpha}$.
Note that while the IC formulation by Chaves et al also has terms $I(\alpha_0 : \alpha_i \,|g_\beta, \mu', \beta = i)$ which account for correlated information between pairs of different bits in $\overrightarrow{\alpha}$, there is an important deficit. Rather than accepting parity as a piece of information on it's own, the $I(\alpha_0 : \alpha_i|\mu', g_\beta)$ term only asks whether B can improve his guessing probabilities on the value of α_0 *provided that* he already successfully retrieved α_i in isolation. In other words, the parity is only considered to be useful information if it is accompanied with a correct guess of the value of either α_0 or α_1.

To interpret the advantage of the bound by Yu et al [23], note that the key difference between $P_{PR}^{1\nu\sigma}$ [10] boxes and $P_{PR}^{0\nu\sigma}$ [11] boxes lies in which specific combinations of input and output values $(a, b, x.y)$ have a non-zero probability within the distribution $P(a, b|x, y)$. While the canonical PR distribution P_{PR}^{000}, for instance, is nonzero whenever $a \oplus b = xy$, the condition for P_{PR}^{110} reads $a \oplus b = xy \oplus x \oplus y$. The additional individual dependence on x and y biases the output b that party B uses to decode the received message μ'. Within a RAC with the van Dam protocol, this requires party B to correct his output b by undoing the addition of x and/or y. For his own input y, this is easy. However, he generally does not have access to the pure box input x of party A.
The IC statement by Yu et al [23] takes this into account by using a modified objective for the RAC game that is invariant under the addition of x to g. That is, if the RAC is successful for some guess $g = \hat{g}$, then it equally succeeds for the guess $g = \hat{g} \oplus x$.

[10] E.g. P_{PR}^{110}
[11] E.g. the canonical PR-box

3.3 Violating IC by wiring-based non-locality distillation

The preceding analysis and demonstration of IC bounds assumed that parties A and B share only a single instance of each no-signaling distribution $P(a,b|x,y) \in \mathcal{NS}$. However, in a more general scenario, if multiple copies of a box $P(a.b|x,y)$ are shared, a suitable wiring can result in an effective box $Q(a,b|x,y)$ with a larger CHSH value S_{CHSH} than an individual copy of $P(a,b|x,y)$. This suggests that even when a probability distribution $P(a,b|x,y)$ obeys the Uffink-like IC inequality $IC_{RAC}\left[P(a,b|x,y)\right] \leq 4$ (eq. 2.17), it might equivalently be possible to wire multiple instances of $P(a,b|x,y)$ to a distribution $Q(a,b|x,y) \equiv W(P(a,b|x,y)^{\times N})$ such that then $IC_{RAC}\left[Q(a,b|x,y)\right] > 4$.

It is generally a non-trivial task to find combinations $(W, P(a,b|x,y), N)$ for which this is the case. In particular, solving $IC_{RAC}\left[Q(a,b|x,y)\right] \leq 4$ analytically in terms of $Q(a,b|x,y)$ is already infeasible due to the non-linearity of IC_{RAC}. Consequently, finding expressions for wiring functions $(f_{in}^{(1)}, f_{in}^{(2)}, f_{out}, g_{in}^{(1)}, g_{in}^{(2)}, g_{out}) \equiv W$, a distribution $P(a,b|x,y)$, and N such that $IC_{RAC}\left[W(P(a,b|x,y)^{\times N})\right] \leq 4$ is only feasible by means of numerical methods.

The simplest approach would be to search among all possible wirings W by brute-force while focusing on a fixed subset of distributions $P(a,b|x,y)$. However, the number of valid wiring functions $W \equiv (f_{in}^{(1)}, f_{in}^{(2)}, f_{out}, g_{in}^{(1)}, g_{in}^{(2)}, g_{out})$ is huge. Even in the simplest $(2,2,2,2)$ Bell scenario and $N = 2$, a single iteration through all wirings would be inefficient, if not infeasible.
To reduce the search space, one could alternatively select a relevant subset of wirings that is iterable within reasonable time. Following this approach, the authors in [73] studied wirings in the context of the principle of non-trivial communication complexity, which limits the value of the CHSH functional S_{CHSH} to a certain constant. Thereby, they only considered wirings under which canonical PR-boxes stay invariant, i.e. $\{W : W(P_{PR}^{000}, P_{PR}^{000})\}$).
Overall, the wirings in this subset proved to be a good choice for increasing the S_{CHSH} value and, therefore, for violating the principal of non-trivial communication complexity at some point.
One instance from the mentioned subset is the wiring proposed by Allcock et al in [71], which we introduced earlier in eq. 2.19.

ht

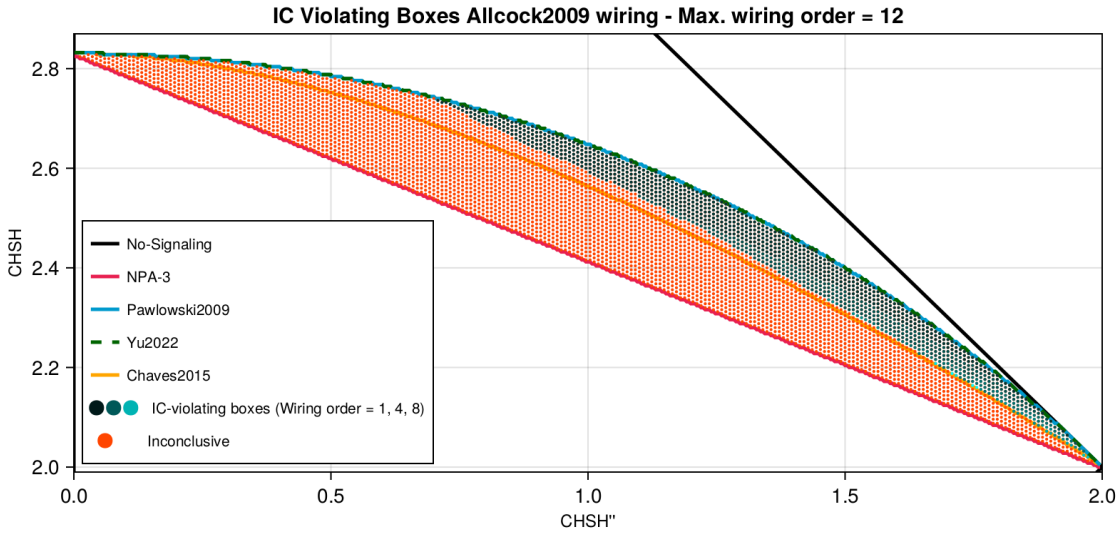


Figure 3.2: *Post-wiring violations of IC for the mixtures* $P(a,b|x,y) = \eta_1 P_{PR}^{000} + \eta_2 P_{LD}^{0101} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$ *with* $0 \leq \eta_1, \eta_2 \leq 1$ *and a wiring proposed in [71]. Each mixture with* (η_1, η_2) *is represented on the axes by their values of* $S_{CHSH} \equiv S_{CHSH'}^{(1,1,1,-1)}$ *(horizontal) and* $S_{CHSH} \equiv S_{CHSH}^{(1,1,1,-1)}$ *(vertical).*
Dots indicate assessed boxes $P(a,b|x,y)$*, whereby red dots signify boxes that could not be distilled to a IC-violating box by the wiring up to order* 12 *(*$\sim N = 13$ *box copies). Blue dots signify distillable boxes, with the luminance of the color indicating the minimally required wiring order* $N-1$ *(*$\sim$ *N boxes copies). While the wiring can help violating IC for the bounds of [16] and [23], no post-wiring violations are found for the bound in [22].*

3.3.1 Post-wiring violations of the quadratic IC inequality

For the original Uffink inequality [66], it was shown that this wiring alone can already distill many Uffink-consistent boxes $P(a,b|x,y)$ to boxes $Q(a,b|x,y)$ that violate Uffink's inequality. Therefore, it was anticipated that this would equally hold for the Uffink-like IC inequality 2.17.

To demonstrate this, and to illustrate the extension of our numerical framework to wirings, we plotted in figure 3.2 the post-wiring violations of the original IC inequality 2.17, specifically for the wiring of Allcock et al. The chosen mixture of boxes for this figure, $P(a,b|x,y) = \eta_1 P_{PR}^{000} + \eta_2 P_{LD}^{0101} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$, is hereby the same as in figure 6 of [71].

Instead of assessing all values $0 \leq \eta_1, \eta_2 \leq 1$, we focused on those boxes $P(a, b|x, y)$ within a grid with spacing $8 \cdot 10^{-4}$ which are within the so-called IC-Q-Gap, meaning that $P(a, b|x, y)$ satisfies the original IC inequality $IC_{RAC}[P(a, b|x, y)] \leq 4$ (eq. 2.17) and $P(a, b|x, y) \notin \mathcal{Q}'$. The latter restriction can be made since the quantum set $\mathcal{Q}'$ is known to be closed under wirings. Testing whether quantum boxes violate IC after applying a wiring is thus superfluous. Also if $IC_{RAC}[P(a, b|x, y)] > 4$, we do not need to be test the box anymore since it already violates IC without any wiring and for any formulation IC.
For comparison, figure 3.2 also shows the other two IC constraints from table 3.1.

Beside the IC bounds, the figure shows as blue and red dots all the boxes in the slice that we assessed. Those are thus the boxes $P(a, b|x, y)$ which have been wired together, not the final wired boxes $Q(a, b|x, y)$ which could also lie outside of the visualised slice.
Blue dots identify the boxes that did violate the IC inequality 2.17 after applying the wiring $W(P, P)$ at least once ($N \geq 2$). Hereby, the luminance of the blue color shows how many copies N of $P(a, b|x, y)$ were required to obtain the violation, i.e. the minimum N such that $IC_{RAC}[W(P(a, b|x, y)^{\times N})] > 4$. A lighter color means that more box copies were needed. The red dots, in contrast, indicate that IC was satisfied for this specific wiring up to the pre-determined maximum wiring order of 12. That is,

$$IC_{RAC}[W(P(a, b|x, y)^{\times N})] \leq 4$$

for any number of box copies $N \in [1, 13]$.
However, a box displayed in red does not necessarily rule out the possibility that applying the wiring to more than 12 copies of that box could result in an IC violation. Furthermore, a stronger constraint for IC than eq. 2.17 could lower the (non-locality) barrier for a violation, potentially allowing the wiring to distill a red box $P(a, b|x, y)$ to an IC-violating box $Q(a, b|x, y)$ after all.

Figure 3.2 shows that, for most boxes marked in blue, just 2 box copies are enough to achieve an IC violation with this single wiring. In the lower-right part of the figure, however, there are also boxes $P(a.b|x, y)$ for which the wiring needs at least 4 to 6 copies to distill boxes sufficiently far. In only very few exceptions, visible along the orange line, the wiring requires $N = 8$ to 10 box copies to violate IC.

We remark that, although the figure visually limits the wiring order to a maximum of 12, the underlying computations were actually carried out for up to $N = 25$ box copies. However, no additional IC violations were found when choosing N between 10 and 25.

When comparing the different IC bounds to the blue region of post-wiring violations, it is evident that the visually overlapping bounds proposed by the original IC paper [16] and by Yu et al. [23] are both not IC-tight constraints. After all, many boxes violate these bounds after applying the wiring to just two satisfying copies. In stark contrast, there is not a single box $P(a, b|x, y)$ below the IC bound by Chaves et al [22] which could be distilled with the considered wiring to a box above the other two bounds.

In fact, on the right side of the figure (i.e. higher $S_{CHSH''}$ values) Chaves' bound even turns out to be tangent to the lower edge of the blue-dotted region, at least up to numerical precision.
This is somewhat surprising because Chaves' IC bound, derived from the framework of causal structures, is entirely unrelated to the concept of wirings. Therefore, we might have evidence that Chaves' bound is actually IC-tight for the plotted family of box mixtures. IC-tightness of Chaves' bound could then explain why, despite the lack of connection between causal structures and wirings, the wiring's distillation capability ends exactly at Chaves' IC boundary.
Of course, the matching boundaries could also be a coincidence since the blue-dotted region and Chaves' IC bound do not coincide over the entire width of the figure. The differences become particularly significant for $S_{CHSH''} < 1.5$ in this slice. Moreover, when considering various other non-isotropic box mixtures, Chaves' bound is not always the strongest IC bound, nor does the region of post-wiring violations consistently reach this bound in all such slices.

3.3.2 Broadening the search for post-wiring violations

The single wiring as proposed by [71] can thus already reveal many post-violations of the quadratic IC inequality 2.17. There are, however, still many boxes $P(a, b|x, y)$ which do not violate IC after wiring them with this specific wiring. On the one hand, there is still a large gap between the quantum boundary and the lower edge of the blue-dotted region of figure 3.2. On the other hand, we found that the proportion of boxes within the

IC-Q-gap that show post-wiring violations of 2.17 is much smaller for other types of non-isotropic box mixtures than for the slice of figure 3.2.

To get a more complete picture of which boxes can definitely be wired to IC violating boxes, we thus have to consider far more wirings. In searching for more post-wiring violations, it would be particularly interesting to find a counter-example for the instability of the causal structures IC bound from [22] under wirings, specifically for the case of non-isotropic box mixtures. After all, the single wiring of figure 3.2 was already sufficient to rule out IC-tightness of the Uffink-like IC inequality 2.17 and the bound by Yu et al from [23]. Therefore, finding a suitable combination of a wiring and a box for disproving IC-tightness of Chaves' IC bound formed the primary motivation to extend the search for post-wiring violations.

Based on the preceding observations and additional experiments with other wirings from the previously mentioned set of PR-box preserving wirings, we anticipated that a completely different class of wirings would be necessary to identify post-wiring violations capable of surpassing Chaves' IC boundary.
Unfortunately, there does not seem to be a clear heuristic for selecting a class of wirings specifically aimed at violating IC bounds.
Already for the quadratic and relatively compact IC inequality 2.17, it is not clear what kind of wiring might increase the value of the non-linear functional $IC_{RAC}[Q(a,b|x,y)]$, for example. This becomes only worse when considering any of the other IC inequality statements (see column 2 in table 3.1) with highly non-linear mutual information terms.

In addition to the difficulty of identifying a useful class of wirings, focusing exclusively on a specific class might discard too many wirings that could prove unexpectedly useful in a particular information retrieval task, such as a RAC. Therefore, a more appropriate strategy might be to first broadly explore the entire set of wirings before narrowing it down to certain optimal classes.
To achieve this, we can exploit that the set of wirings, as defined in 2.3.1, is a convex and compact set such that it forms a polytope, just like the set of distributions $\mathcal{NS}$ and $\mathcal{L}$. This means that there are a finite number of extremal points which characterise and generate the full set of wirings. Even better, each of those extremal wirings can be generated from a composition of only 5 types of partial wirings, as listed in table 1 in [74]. The total amount of extremal wirings, $45,212,176$, is then still very large but iterable within a manageable runtime, especially on a sufficiently powerful

compute cluster.
After several trial runs, we nevertheless further reduced the search space by discarding 2 of the 5 types of partial extremal wirings from [74], specifically the deterministic and one-sided types. By studying the changes in the values of S_{CHSH} and $IC_{RAC}[Q(a,b|x,y)]$, we observed that all wirings constructed with the two discarded types were namely not capable of increasing the amount of non-locality for 10 or fewer copies of any box $P(a,b|x,y)$ within the slice represented in figure 3.2.

While iterating through all of the remaining $72^4 = (5,184)^2 = 26,873,856$ extremal wirings, we recovered the previously considered wiring from [71] for the exact same box mixtures as in figure 3.2. Additionally, we discovered a few other extremal wirings that appeared to be very similar, as they exhibited the same or fewer post-wiring violations for boxes $P(a,b|x,y)$ within the same slice.

Unfortunately, despite conducting extensive searches in various non-isotropic slices, specifically those with one or two LD-boxes in the three-box mixtures, we did not find any wiring that could distill boxes above Chaves' IC bound. Thus, it remains inconclusive whether the set of boxes compatible with Chaves' IC bound is closed under wirings or not. In the end, there are still various possibilities for disproving the stability of IC within causal structures. For example, by wiring different boxes that satisfy the bound, instead of multiple copies of the same box.

One novel method that allows to efficiently optimise also over all convex combinations of extremal wirings has recently been presented in [72]. Concretely, they employ projected gradient descent with an adaptive learning rate to search for wirings within the continuous space of *mixed* wirings. Mixed wirings are just convex mixtures of the deterministic wirings that we introduced in def. 2.3.1, whereby the wiring functions $(f_{in}^{(1)}, f_{in}^{(2)}, f_{out}, g_{in}^{(1)}, g_{in}^{(2)}, g_{out})$ take the same binary arguments as before but now each output nuous values in the range $[0,1]$ instead of a binary value. In particular, these continuous values denote each the *probability* for the box inputs (x_1, x_2, y_1, y_2) and the effective box outputs (a,b) having a value of 0 respectively. So all box inputs or outputs still take binary values, but now with a certain probability.
Because of this, $Q(a,b|x,y) \equiv W(P(a,b|x,y)^{\times N})$ and $S_{CHSH}(Q(a,b|x,y)$ become also continuous functions with respect to the wiring function values $f_{in}^{(1)}(x,a_2), f_{in}^{(2)}(x,a_1), f_{out}(x,a_1,a_2), g_{in}^{(1)}(y,b_2), g_{in}^{(2)}(y,b_1)$, and $g_{out}(y,b_1,b_2)$.

Consequently, the gradient of an objective function, such as $S_{CHSH}(Q(a,b|x,y))$ or $IC_{RAC}[Q(a,b|x,y)]$, can be defined. Well-established machine learning techniques, like gradient descent, can then be used to find an optimal wiring W with respect to the objective.

We successfully re-implemented this method in our numerical framework with the quadratic expression $IC_{RAC}[Q(a,b|x,y)]$ from eq. 2.17 as the objective to be maximised. We performed a few trial runs, focusing on non-isotropic box mixtures with a single LD-box, such as the one from figure 3.2.
Despite extra efforts to avoid local maxima[12], the optimisation often got stuck on wirings that yield no violations of the original IC bound (eq. 2.17) for any of the assessed box mixtures. This could either be due to the non-linearity of the objective or suboptimal hyperparameters for gradient descent and line search, with the latter being more probable since $IC_{RAC}[Q(a,b|x,y)]$ is not significantly more complicated than the expression for the CHSH functional S_{CHSH}.
Nevertheless, in a few runs, we actually found some wirings with post-wiring violations for the box mixtures $P(a,b|x,y) = \eta_1 P_{PR}^{000} + \eta_2 P_{LD}^{0101} + (1-\eta_1-\eta_2)P_{\mathbb{I}}$, though we had already discovered these wirings previously while iterating over the set of extremal wirings.

Regarding our primary aim of finding post-wiring violations of Chaves' IC bound, we must resort to using the left-hand side of the IC inequality in the center cell of Table 3.1 as the objective function. Unfortunately, the above method breaks down when considering this objective based on mutual information. This is because the mutual information terms in Chaves' IC statement, as functions of the distributions $P(a,b|x,y)$, only vary in large discrete steps of 1.0 (i.e. whole bits). This raises the general difficulty of optimizing discrete objectives with gradient-based methods

An additional problem is that for all boxes $P(a,b|x,y)$ relevant to our search, namely those in the IC-Q-gap[13], the mutual information terms on the left-hand side of Chaves' IC inequality take the same value. While boxes on the boundary between IC-satisfying and IC-violating boxes saturate Chaves' IC inequality by definition, any quantum box $P(a,b|x,y) \in \mathcal{Q}'$ with $P(a,b|x,y) \notin \mathcal{L}$ does saturate it as well. In that sense, boxes within the IC-Q-gap are indistinguishable with respect to the LHS of Chaves' IC

[12] By using multi-start optimisation and an adaptive learning rate
[13] That is, any $P(a,b|x,y)$ such that $IC_{RAC}[P(a,b|x,y)] > 0$ and $P(a,b|x,y) \notin \mathcal{Q}'$

inequality and consequently, the idea of being close to (or far from) violating this inequality is not well-defined.

Gradient-based optimisation of wirings for the purpose of violating Chaves' IC bound is thus not possible with the given formulation in terms of mutual information. However, it might become possible if one derives a direct polynomial constraint on the space of distributions $P(a, b|x, y)$ from Chaves' IC inequality. For now, only the Uffink-like IC inequality seems to be compatible with this method, using $IC_{RAC}[Q(a, b|x, y)]$ or the average success probability of an EARAC as the objective function.

Chapter 4

Discussion & Conclusion

This work offers a comprehensive review of research on bounding non-locality with the information causality (IC) principle in a bipartite setting. Rather than focusing on one specific formulation of IC, our analysis intends to elucidate IC from various perspectives. On the one hand, the strengths and weaknesses of different generalised formulations of IC are examined. On the other hand, starting from any IC formulation, various ways to generalise the derivation of explicit constraints on distributions $P(a,b|x,y) \in \mathcal{NS}$ are explored. Overall, our study emphasises the remarkable power of the (bipartite) IC principle in approximating the quantum boundary for many isotropic boxes. However, it also showcases the heavy deficiencies in our understanding to what extent IC bounds can be strengthened further and to what extent the gap between the IC and quantum boundary can be closed.
Although we also provide new insights about the stability of certain IC formulations under non-locality distillation, this work mainly serves as a survey of various attempts to derive tight IC bounds for setups beyond the standard EARAC.

4.1 Implications

• **Unifying the view on generalised formulations of IC**
First, the IC bounds that follow from the generalised formulations of IC in [22, 23, 27] are studied in a unified notation and for box mixtures beyond the ones that were specifically selected in the original works. For the first time, this allows an overview of the strengths and weaknesses of each approach that has been proposed so far.

Particularly, we chose to compare all the IC bounds which are universally stronger than the quadratic Uffink-like constraint, which follows from the original RAC-based formulation of IC in [16, 26].

A more detailed analysis in [22] of the causal relationships between variables in an EARAC, for example, has significantly improved our ability to accurately quantify the amount of information $I(\overrightarrow{\alpha} : g_\beta|\beta)$ that party B can obtain from party A's data $\overrightarrow{\alpha}$ for a given box $P(a,b|x,y) \in \mathcal{NS}$. Each time a bit α_β is guessed correctly, there might namely be additional information left in message μ' that allows to slightly increase the guessing probability of another bit $\alpha_{\beta'}$ ($\beta' \neq \beta$) as well. While the upper-bounding channel capacity is held constant, the retrievable information increases and violates the IC inequality for many more boxes, which implies a stronger bound.

• IC is powerful for symmetric distributions, but shows significant weaknesses for boxes with biased outputs
The comparison of IC bounds demonstrates their astonishing capability in reconstructing quantum boundaries for most families of symmetric and unbiased boxes, while it also highlights the significant gaps to the quantum boundary in non-isotropic slices. The size of the gap and the difficulty to strengthen IC bounds hereby seems to increase with the number of Local-Deterministic boxes in the mixture.

Indeed, this is not surprising as we assumed uniform (i.e. unbiased) and i.i.d. data bits $\overrightarrow{\alpha}$ while Local-Deterministic boxes do imply a certain bias in the value of party B's guess g_β.
Only for the IC bound from [27], the usual assumption of uncorrelated data bits (α_0, α_1) was dropped and EARACs with nearly perfect correlation between α_0 and α_1 were shown to give an IC bound that is much stronger than Uffink's inequality for mixtures of two Local-Deterministic boxes and a single PR-box. The choice of nearly perfect correlation is, however, not optimal for other type of boxes and even results in IC bounds which are hardly stronger than the no-signaling conditions.
This shows that IC can also work well in less symmetric scenarios of non-isotropic boxes but it requires more specialised fine-tuning to the setup of the information retrieval task in which IC is formulated.

• **The two generalised (bipartite) formulations of IC from [22, 23] are complementary**
Furthermore, we demonstrate that there is currently no bound known that is IC-tight for isotropic boxes while also being the strongest bound for non-isotropic boxes.
In particular, IC within the framework of causal structures and the IC formulation proposed by [23] complement each other. While the former bound from [22] is the strongest IC bound for non-isotropic boxes, it is not stronger than the original Uffink-like IC inequality for any family of isotropic boxes. In contrast, the latter bound from [23] closes the gap between the IC and quantum boundary for certain families of isotropic boxes but performs equivalently to the Uffink-like IC inequality for all non-isotropic boxes.

• **Focus of IC research on improving RAC protocol and simplifying derivations of IC constraints**
The biggest contributions to IC so far, however, have not been made on re-formulating IC, but rather on simplifying the derivation of explicit IC constraints in terms of distributions $P(a, b|x, y)$ and improving the protocol that parties A and B apply within their bipartite information retrieval task.

For instance, the proposal in [26] to replace the box concatenation procedure by a noisy communication channel and the recently presented method from [27] to make the derivation of explicit IC bounds more systematic, have been most disruptive in simplifying the computation of the quadratic Uffink-like IC bound. Although the resulting IC bound is the same[1], this is a first step to enable the study of IC for more general and more complex variations on the regular EARAC scenario, which may reveal many more IC-violating boxes compared to the original IC formulation from [16].
The above mentioned specialised IC bound for nearly perfectly correlated data bits from [27] is thereby early evidence for the potential strength that can still be gained for IC bounds when generalising the setup of the EARAC[2]. This is the case even when using one of the existing (but possibly incomplete) formulations of IC in table 3.1.

[1] At least in the default setup that we assume throughout this thesis: The standard EARAC as the bipartite information retrieval task + shared no-signaling boxes in the simplest $(2, 2, 2, 2)$ Bell scenario

[2] or any other bipartite information retrieval task

• **Wired box copies can violate IC bounds (from [16, 23]) that do not exploit the full information in messages**
One other possible generalisation of a standard EARAC with a single no-signaling box is giving the two parties access to multiple boxes or multiple copies of the same box. In combination with arbitrary local processing, this generalised scenario allows to distill non-locality through wirings. In our numerical experiments, we specifically studied the impact of wirings on the ability of boxes $P(a, b|x, y)$ to violate any of the IC bounds. Although this has been done before on a superficial level in works like [73] for the quadratic Uffink-like IC inequality (eq. 2.17, our experiments were focused on searching for wirings that are optimal for catalysing the violation of any given bound, specifically in the context of IC. As no universally IC-tight bound has yet been identified, it is of particular interest to find boxes that, while satisfying all proposed IC bounds for individual box copies, can violate at least one of the bounds when wiring multiple copies. Discovering any such post-wiring violation would then indicate the minimum extent to which IC bounds can be tightened further towards the quantum boundary.
Unfortunately, we found no such post-wiring violation for up to 12 box copies, despite an extensive search for wirings within the set of extremal (deterministic) wirings and boxes within a few specific families of non-isotropic boxes. However, some extremal wirings did still give many post-wiring violations when focusing on either the quadratic Uffink-like IC bound or the bound proposed by Yu et al in [23]. [3] We demonstrated this for a certain wiring that was first considered in [71] and that belongs to a class[4] of wirings which was already known from [73] to violate the Uffink-like IC inequality. For the bound by Yu et al, however, this was not explicitly demonstrated before.

In the end, only for the IC bound based on causal structures from [22] our numerical experiments leave open whether any valid wiring can distill a set of compatible boxes[5] to some effective box that can violate the bound. Because of our extensive search for wirings, using multiple numerical methods, we consider our observations to be even weak (numerical) evidence for the IC-tightness of the causal structures bound for box mixtures with a single LD-box. Though, if this turns out to be correct, explaining the

[3] Note that the two bounds are completely equivalent in this case of non-isotropic box mixtures

[4] Concretely, this is the class of wirings W which preserved PR-boxes, i.e. $W(P_{PR}(a, b|x, y)^{\times N})$ for any number of copies $N \geq 2$. (See chapter 3

[5] or, even better, multiple copies of a single box

preference of the bound for this specific type of non-isotropic boxes seems a bit difficult since the bound is not the strongest one for non-isotropic boxes with two LD-boxes, for example.

• **Conjecturing stability under wirings of the IC bound based on causal structures (from [22])**
Although we found no examples of post-wiring violations of the Uffink-like IC bound for boxes satisfying the causal structures IC bound, we did actually find many examples of post-wiring violations for boxes very close to the causal structures bound[6]. Somehow, the existence of post-wiring violations thus seems to be exactly restricted to those boxes that violate the causal structures bound, at least for the two non-isotropic slices that we considered for our wiring experiments.[7] In other words, the causal structures bound seems to be setting a limit on the extent to which wirings can be applied to violate IC bounds. Overall, our observations then suggest that the set of boxes[8] which satisfy the causal structures bound is closed under wirings and, therefore, that the IC bound based on causal structures must be stable under wirings.

At this point, one might wonder whether using causal structures to formulate IC is, in contrast to the other two bounds, inherently already incorporating scenarios which make of use wirings. This would then justify the potential stability of the bound under wirings. However, we could not identify any clear relation between wirings and the IC bound based on causal structures. Although we can not completely rule out such a relation, the absence would mean that we need some alternative explanation for why the causal structure bound and the region of post-wiring violation of IC bounds happen to coincide[9] so well in our experiments. Since any information theoretical principle, including IC, is expected to be stable under wirings [75], IC-tightness of the causal structures bound within the no-signaling slices of our experiments could explain the seemingly perfect match as well.

[6]That is, boxes which violate the bound but are hardly distinguishable from boxes satisfying it.

[7]As mentioned in the previous chapter, those observations were not restricted to the wiring from [71] that we studied for figure 3.2, but also hold for other (similar) wirings.

[8]Strictly speaking, our experiments give only evidence for boxes which correspond to a mixture that contains a single LD-box, but we believe that the stability of the bound also holds more generally.

[9]like in figure 3.2

We emphasise that both conjectures about the causal structures bound, IC-tightness and stability under wirings for non-isotropic boxes with a single LD-box, are only supported by extensive numerical evidence but not proven. Future work thus might aim for a rigorous analytical proof and a justification for the specific benefit of the causal structures bound in the specific case of non-isotropic box mixtures with a single LD-box.

• **Providing a modular numerical framework for device-independent non-locality**
Lastly, this work also yielded a methodological and IC-independent contribution in form of a modular and efficient numerical framework for studying bipartite non-locality within the device-independent paradigm. [10] It features the convenience of a natural syntax for working with bipartite no-signaling boxes, the NPA hierachy, Bell inequalities and wirings. Although common Bell inequalities, IC bounds and wirings are readily implemented, it can be easily extended with custom Bell inequalities or wirings.

To our knowledge, a flexible codebase for working with bipartite no-signaling distributions $P(a, b|x, y)$ has not existed yet. While some specialised code on the broader topic of non-locality has been published for MATLAB and Python in the past, we aim to encourage efforts towards establishing a more efficient, open-source, and high-level codebase for (device-independent) non-locality research. Considering that this field of research involves a lot of optimisation tasks, we believe this would be best achieved using more suitable languages such as Julia.

4.2 Limitations & Future directions

In the last few years, research on IC has gathered new momentum and has thereby discovered increasingly larger sets of no-signaling boxes that violate the principle. However, both in proposing more general formulations of IC and in generalising the derivation of explicit constraints on distributions $P(a, b|x, y)$, there are still several aspects and scenarios that one could account for when applying IC as a bounding principle.

[10] Accessible via GitHub: `https://github.com/t-rothe/ABoxWorld`

IC formulations

First, regarding the generalisation of the RAC-based IC formulation, only the works by Chaves et al [22] and Yu et al [23] readily contributed a few ideas[11] with new perspectives on IC. To follow up on those, we would suggest to find IC formulations which also address the following two fundamental aspects:

- **Restriction to EARACs as information retrieval task —** EARACs are an important type of bipartite information retrieval task for which the advantage of non-locality can be easily demonstrated. Therefore, EARACs are also a natural framework to formalise IC, as has been done throughout the literature on this topic. However, it would be of interest to study IC also within the context of other information retrieval tasks for which the success of the game is primarily determined by the *amount* of information that is communicated between the parties. This could yield universal tighter IC bounds, yield specialised IC bounds that are tighter in specific no-signaling slices, or make the derivation of IC bounds less dependent on the chosen protocol in the game.
  The nonlocal torpedo game, presented in [54], might be an interesting variation on a standard EARAC, but other bipartite games might be even more exciting.
  Especially the choice of a suitable objective could make a crucial difference. A limitation of the EARAC objective $g = \alpha_\beta$ is namely the boolean criterion for defining success. That is, instead of maximising some continuous "score", party B can either completely win or completely fail in each round of the EARAC game. This causes also the retrievable information $\sum_i I(g : \alpha_\beta|\beta = i)$ on the left hand side of 2.14 to be a (piecewise) discontinous function with respect to distributions $P(a, b|x, y)$. Using instead an information retrieval game with a smooth, differentiable objective would simplify the optimisation of IC bounds over any complex search space (such as wirings) by employing gradient-based methods.

  Of course, decoupling the formulation IC completely from any specific information retrieval task would be an the best solution.
  In their proposal for a re-formulation of IC, Yu et al made in [23] a

[11] I.e. respectively maximising the information gain from received messages by exploiting correlation between data bits, and accounting for the possibility of the retrieval of partial or relative information

first attempt on this. However, to derive a constraint solely in terms of $P(a, b|x, y)$ distributions, they ultimately still need to reintroduce the EARAC. Moreover, the resulting constraint is not an explicit analytical expression; rather, it requires numerically solving a convex optimisation problem for each $P(a, b|x, y)$ box to assess the violation of IC.

- **Fixed entropy measure & Independence of rounds —** All proposed formulations of IC (see 3.1) are stated as inequalities that involve some form of mutual information, satisfying the properties in eq. 2.9 - 2.13. In [65], an equivalent entropic variant of the original IC formulation in 2.14 was introduced which requires slightly fewer assumptions about the considered entropy measure H.

  All works on IC to date, including the experiments in this thesis, have implicitly chosen Shannon entropy as the relevant measure of information for computing the explicit IC bounds with respect to distributions $P(a, b|x, y)$. To some extent this choice makes sense as it quantifies the "average" information content of random variable g about α_β, which is coherent with the implicit assumption that, in an experimental realisation of an EARAC, the success probability $P(g = \alpha_\beta|\beta)$ is estimated by repeating the non-local game throughout multiple *independent rounds*.
  Alternatively, however, one could also study IC for EARACs in a one-shot scenario. Concretely, party B could make each guess g for query β with knowledge of all previous guesses and queries that he made. In this case, using either a more conservative or a more optimistic entropy measure can make a difference and might even be more appropriate to maximise the strength of IC bounds. In the context of cryptographic applications, for example, the so-called min-entropy is preferred.

  Renyi entropies form a broad family of entropies that are commonly used in (quantum) information theory. The Shannon entropy is a special case within this family, specifically it is equivalent to the RÃ©nyi entropy of order $\alpha = 1$. Unfortunately, there is no commonly agreed definition for Renyi mutual information that could be used in the original IC formulation (eq. 2.14) [78]. This is because the generalisation of Shannon entropy to Renyi entropies weakens some of the strong properties of Shannon entropy, like the chain rule $H(X|YZ) = H(XY|Z) - H(Y|Z)$. Nevertheless, one might carelessly

consider to just substitute Renyi entropies into the purely entropic IC inequality from [65]. However, also in that case the proof of the validity of IC fails on the lack of the (strong) chain rule. Therefore, the general validity of IC with respect to an arbitrary type of Renyi entropy is inconclusive.

In the end, one might rightfully question the need for considering other entropy measures by referring to the success of current IC formulations with Shannon entropy. However, we object to this since there has not yet been given any rigorous justification for why it is the correct entropy measure in the context of IC. To some degree the choice of Shannon entropy seems arbitrary and the impact of alternative entropies should be studied in future work.

(Derivation of) IC bounds

Subsequently, regarding the derivation of explicit constraints on no-signaling distributions $P(a,b|x,y)$ from any given IC formulation, numerous assumptions about the applied protocol and wirings are necessary. Such specific choices then naturally lead to some substantial limitations. Notably:

• **Fixed protocol —** Throughout this work, we assumed that the two parties in an EARAC always use the van Dam protocol. While this protocol is optimal for EARACs with a single PR-box, it is possible that other no-signaling boxes could perform better with different deterministic, possibly even probabilistic, protocols. Equally when applying wirings, the relation between the effective box outputs (a,b) and box inputs (x,y) can heavily change such that different protocols perform much better than van Dam. However, to fulfill the necessity to specify a protocol, whilst ensuring an optimal success probability in an EARAC, one would have to optimise the protocol for each box $P(a,b|x,y)$ individually. This seems intractable when computing IC bounds analytically and has a dramatic impact on the runtime when computing them numerically.
Among many possible future directions, one could start by surveying the IC bounds obtained for various universally[12] applied protocols. Beyond this basic approach, one could also try to develop an efficient adaptive protocol that depends on the box $P(a.b|x,y)$, but without requiring optimization for each individual box

[12] I.e. independent of the box $P(a.b|x,y)$

Eventually, the exploration of various protocols might also be key in constructing IC bounds that are independent of any specific protocol.[13] However, it is unclear whether such general bounds can exist at all. The protocol dependence is namely due to the generally undefined random variables $(\mathbf{ff}, g, \mu, \mu')$ that appear in all of the IC formulations. To reduce the IC inequality $I(\mathbf{ff} : g) \leq I(\mu : \mu')$ to a constraint solely in terms of $P(a, b|x, y)$, one thus needs to exactly specify how the variables $(\mathbf{ff}, g, \mu, \mu')$ are related to the box inputs and outputs (x, y, a, b).
The big question is then whether any IC formulation can be defined that purely relies on (x, y, a, b). While this is possible for most other operational principles, we expect that this is inherently impossible for IC.

• **Fixed type of communication channel —** Related to the choice of a protocol is the selection of the type of (noisy) channel between party A and B. Throughout this thesis, we only used the binary symmetry channel, as originally proposed by Miklin and Pawlowski in [26]. In that same work [26], it was noticed that generalised dit-input dit-output communication channels yield the strongest IC bound for a non-zero channel capacity $\kappa \neq 0$ if $d \geq 3$. In contrast, the optimal channel capacity for the binary symmetric channel ($d = 2$) was found in the limit of a vanishing capacity $\kappa \to 0$. While the reason for this discrepancy between bit- and dit-channels remains open, this curious finding shows the high potential for learning something new about IC by simply considering other (noisy) channel types, particularly those with higher-dimensional inputs and outputs.

While preparing this thesis, we started[14] experimenting with more complex channels as well. See appendix C for a preliminary description and schematic of our (somewhat deviating) approach.

Alternatively, future work could also study EARACs for the case of *asymmetric* bit-channels. We tried to use the method in [27] to derive IC bounds for the binary erasure and (asymmetric) Z-channel, for example, but found that it does not apply to those cases. Seeking a generalisation of the method in [27] to a broader set of channel types is thus also desirable.

In either case, by exploring various types of communication channels, we

[13] This would also better fit the device-independent paradigm in which we would rather like to treat the physical system completely as a black box.

[14] I.e. Yet to be completed

might deepen our understanding about how exactly no-signaling correlations are used in the decoding process of an EARAC. That is, what "type" of information [15] should be transmitted via the channel for maximising the information retrieval, and what exact role the correlated information between box-output a and b play in the decoding process.

- **Focus on simple wirings of up to $N = 24$ box copies —** In our unsuccessful search for post-wiring violations of the causal structures bound from [22], we only considered a certain type of wirings that is described by combinations of a single wiring W, a single box P, and the number N of identical copies of P.[16] Despite this subset of wirings being powerful for identifying violations of the other two IC bounds, this type of wirings is rather simple. There are many more complex wirings for $N \geq 3$ for which post-wiring violations of the causal structures bound could exist.
On the one hand, one could use different wirings between different pairs of boxes, rather than applying the same wiring W repeatedly. When $N = 3$, this gives the wired box $W_2(W_1(P, P), P))$ for two wirings W_1 and W_2, for example. On the other hand, instead of wiring N copies of the same box $P(a, b|x, y)$, one could also wire N different no-signaling boxes $\{P_1, \ldots, P_N\}$.[17]

Unfortunately, in both of these cases the search space grows exponentially with N, compared to the constant number of wiring-box combinations in our experiments. Even with our restriction to extremal wirings, our search space was vast. Thus, for the more complex types of wirings, a brute-force approach through all possible (extremal) wirings and boxes becomes intractable.
Hence, more efficient methods for the wiring optimisation are required to make progress on testing the stability of the causal structures bound under wirings. In [72], for example, the authors proposed an efficient gradient-based method for finding post-wiring violations of the principle of non-trivial communication complexity by maximising the CHSH functional S_{CHSH} (eq. 2.6) over the full set of wirings. To apply this to any IC bound, however, one would first have to find a continuous and differentiable objective function that is maximised whenever the strength of the IC bounds is maximised.

[15] Or, equivalently, what part of the information in the data $(\alpha_0, ..., \alpha_n)$

[16] Recall that while W can only wire two boxes at a time, N box copies can be wired by iterative application of the wiring. I.e. for $N = 3$, either $W(W(P, P), P)$ or $W(P, W(P, P))$.

[17] In that case, all those boxes must satisfy the bound before the wiring is applied.

• **Wirings can (probably) not distill isotropic boxes —** While our discussion has primarily concentrated on the tightness of IC bounds for non-isotropic slices of the no-signaling polytope, there are also a few isotropic slices for which no tight IC bound is known yet. Furthermore, one could argue that having tight bounds for isotropic boxes is more relevant than for non-isotropic boxes since many of the potential applications of IC in quantum information processing involve symmetric setups and objectives. In some quantum key distribution (QKD) protocols, for example, symmetries in the bipartite distribution $P(a, b|x, y)$ are crucial and biases in the box outputs (a, b) could potentially even form a security risk.
Future research should thus focus specifically on generalising IC such that the bounds are strengthened specifically for isotropic boxes, even if the bounds become somewhat weaker for non-isotropic boxes. As before, this might either be done by proposing a new formulation of IC or by optimising the EARAC protocol.

As before, wirings promise here to be a simple method to identify new IC-violating boxes and to explore the minimum extent to which IC bounds can be strengthened further. However, strong evidence suggests that wirings cannot distill the non-locality of isotropic boxes [75]. Therefore, in contrast to non-isotropic boxes, we cannot use wirings to strengthen IC bounds in isotropic slices.
An interesting alternative direction in this context could be the use of non-locality recycling, rather than non-locality distillation. Instead of providing the parties in an EARAC with more than one copy of a no-signaling box, the parties are given a single box that they are allowed to query multiple times in a sequential manner[18]. In the appendix, chapter B, we briefly introduce the concept of time-ordered distributions as a generalisation of no-signaling boxes. Furthermore, we illustrate the use of time-ordered boxes in EARACs and derive a trivial generalisation of the quadratic Uffink-like IC bound (eq. 2.17) on the set of time-ordered distributions. Finally, we propose a formulation of IC that could be used for a numerical study of IC in the context of time-ordered correlations.

To our knowledge, not a single study has yet examined whether the temporal correlations in time-ordered boxes offer any advantages over stan-

[18]Within the quantum framework this corresponds to multiple sequential measurements on one and the same quantum state. I.e. without preparing a fresh copy of the state in between the measurements.

dard no-signaling boxes in non-local games like EARACs. Consequently, whether IC bounds on isotropic boxes can actually be strengthened via non-locality recycling remains an open but relevant question.

4.3 Conclusion

In conclusion, many open questions about IC's ability to constrain the set of no-signaling distributions $P(a, b|x, y)$ remain, even a decade after its proposal.

While for other operational principles, such as Local Orthogonality or Macroscopic Locality, general and tight bounds were found, an IC-tight bound probably remains out of reach for a while. Currently, the most accurate approach to certifying the validity of IC for a given box is the use of two complementary IC bounds by Yu et al [23] and Chaves et al [22]. Yu's generalised bound provides the strongest constraint for symmetric bipartite distributions $P(a, b|x, y)$ (i.e. isotropic boxes), while Chaves' generalisation excels for more asymmetric distributions (i.e. non-isotropic boxes). Although the original quadratic IC bound from [16] is the weakest of all IC bounds, it remains useful because of its explicit and simple polynomial form.
In the end, neither of the two generalised bounds from [22, 23] turns out to be IC-tight, and it remains unclear how closely any IC bounds can ultimately approach the quantum boundary.

Furthermore, we have seen that the extend to which non-locality distillation can be used to violate IC bounds with a few copies of a given non-local box is limited. Only for the original Uffink-like IC bound from [16] and Yu's generalised bound from [23], we did find boxes which satisfy these IC bounds but violate them after applying certain wirings. For Chaves' bound, in contrast, no such post-wiring violations were found, forming evidence for its stability under wirings. However, validating this evidence for stability by extending the search to wirings beyond the extremal (deterministic) ones we focused on in this work, as well as to more complex combinations of multiple wirings, will be very challenging. Despite our focus on simple wirings, which led to a radical reduction of the wiring search space, the search for post-wiring violations within the massive number of possible wirings remained computationally very demanding.

Moreover, the broader question of whether[19] the IC principle itself is stable under wirings cannot be assessed until an IC-tight and explicit constraint is identified.

To make any substantial progress on IC and to address the mentioned limitations, it will be necessary to (paritally) overcome some of the unique challenges associated to IC in comparison to other operational principles. Especially the protocol-dependence and the formulation in terms of very non-linear quantities, such as entropies or mutual information, have led to a rather slow progress in generalizing and strengthening IC bounds. Concretely, these challenges prevent us from obtaining general, yet simple and explicit polynomial IC bounds, thereby heavily complicating the analytical study of this topic.

It would be particularly interesting to see whether one can describe and formulate IC within bipartite and non-local information retrieval games that are very different from RACs. [20]. But even when restricted to RACs, many variations on the standard EARAC scenario remain to be explored in more depth. This includes, but is not restricted to, EARACs with other communication channels or more complex LOCC-compatible[21] operations (e.g. by allowing parties to sequentially query their part of the no-signaling box multiple times).
Although IC research began to shift its focus towards multipartite scenarios [25, 69, 70], there remains so much to be learned about the bipartite case. Specifically, what are the most resource efficient strategies to optimise performance in information retrieval tasks? and what aspects of such non-local games have the biggest impact on the strength of IC bounds?

[19] Probably, yes, since it's a physically motivated information-theoretic principle [75]

[20] RACs are restricted by the rather arbitrary objective to require exact retrieval of data bits with a certain queried index, i.e. requiring $g_\beta = a_\beta$ as a boolean indicator of success.

[21] i.e. Local Operations and Classical Communication

Acknowledgements

The successful completion of this project is causally linked to the wisdom and support of several individuals, with whom I fortunately shared an open and effective communication channel.

First, I wish to express my deepest gratitude to my primary advisor, Jordi Tura, for putting forward this project, trusting in my abilities, and repeatedly challenging me intellectually. I admire your commitment to supervision and the freedom you gave me to pursue my own path throughout this project, giving me a déjà vu of the rewarding excitement I felt during my BSc. project with you. Your mentorship has been vital in shaping both this work and my growth as a researcher more generally.

Secondly, I owe sincere gratitude to my daily advisor and esteemed mentor, Jan Li. Your unwavering support has been fantastic, not only for matters related to this project but also in catalysing my overall professional growth. Thank you for the time and confidence you invested in me, particularly during moments of uncertainty. Your mentorship has been invaluable in dealing with the complexities in this research.

Furthermore, I am equally thankful to Serge Fehr for co-supervising this project and for the inspiring discussions we've had. Having had the privilege of attending your Quantum Information Theory course—which was a highlight of my educational journey—I knew you were the ideal co-advisor for my project. I truly appreciate the time and effort you dedicated to guiding me through a thesis that was perhaps less mathematical than expected. Thank you for your immense patience and for helping me clarify and restructure my thoughts on the content of this thesis.

Lastly, I appreciate the insightful conversations with Nicholas Brunner, Marcin Pawlowski and Nikolas Miklin, which greatly enriched the final stages of this project.

As I reflect on the journey that led to this thesis, I recognise how each interaction with you and the advice you offered laid the foundation for this work and all that is yet to come.

Thank you.

Bibliography

[1] S. Popescu and D. Rohrlich, *Quantum nonlocality as an axiom*, Foundations of Physics **24**, 379 (1994).

[2] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, Physics Physique Fizika **1**, 195 (1964).

[3] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Physical Review **47**, 777 (1935).

[4] I. Supic and J. Bowles, *Self-testing of quantum systems: a review*, Quantum **4**, 337 (2020).

[5] R. Arnon-Friedman, R. Renner, and T. Vidick, *Simple and Tight Device-Independent Security Proofs*, SIAM Journal on Computing **48**, 181 (2019).

[6] J. Barrett, L. Hardy, and A. Kent, *No Signaling and Quantum Key Distribution*, Physical Review Letters **95**, 010503 (2005).

[7] A. AcÃn, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, *Device-Independent Security of Quantum Cryptography against Collective Attacks*, Physical Review Letters **98**, 230501 (2007).

[8] S. Pironio, A. Acin, N. Brunner, N. Gisin, S. Massar, and V. Scarani, *Device-independent quantum key distribution secure against collective attacks*, arXiv (2009).

[9] S. Pironio, A. Acin, S. Massar, A. B. d. l. 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 (2010).

[10] R. Colbeck and R. Renner, *Free randomness can be amplified*, Nature Physics **8**, 450 (2012).

[11] S. Popescu, *Nonlocality beyond quantum mechanics*, Nature Physics **10**, 264 (2014).

[12] S. Popescu and D. Rohrlich, *Causality and Nonlocality as Axioms for Quantum Mechanics*, arXiv (1997).

[13] M. MÃŒller, *Probabilistic theories and reconstructions of quantum theory*, SciPost Physics Lecture Notes , 028 (2021).

[14] J. Henson and A. B. Sainz, *Macroscopic noncontextuality as a principle for almost-quantum correlations*, Physical Review A **91**, 042114 (2015).

[15] Z.-X. Luo, Y.-Z. Xing, Y.-C. Ling, A. Kleinhammes, and Y. Wu, *Almost quantum correlations*, Nature Communications **6**, 6288 (2015).

[16] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, *Information Causality as a Physical Principle*, arXiv **461**, 1101 (2009).

[17] R. Gallego, L. E. WÃŒrflinger, A. AcÃn, and M. NavascuÃ©s, *Quantum Correlations Require Multipartite Information Principles*, Physical Review Letters **107**, 210403 (2011).

[18] M. E. Cuffaro, *Information causality, the Tsirelson bound, and the 'being-thus' of things*, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **72**, 266 (2020).

[19] J. Allcock, N. Brunner, M. Pawlowski, and V. Scarani, *Recovering part of the boundary between quantum and nonquantum correlations from information causality*, Physical Review A **80**, 040103 (2009).

[20] M. Ringbauer, A. Fedrizzi, D. W. Berry, and A. G. White, *Information Causality in the Quantum and Post-Quantum Regime*, Scientific Reports **4**, 6955 (2014).

[21] D. Cavalcanti, A. Salles, and V. Scarani, *Macroscopically local correlations can violate information causality*, Nature Communications **1**, 136 (2010).

[22] R. Chaves, C. Majenz, and D. Gross, *Information-theoretic implications of quantum causal structures*, Nature Communications **6**, 5766 (2015).

[23] B. Yu and V. Scarani, *Information causality beyond the random access code model*, arXiv (2022).

[24] R. K. Patra, S. G. Naik, E. P. Lobo, S. Sen, G. L. Sidhardh, M. Alimuddin, and M. Banik, *Principle of Information Causality Rationalizes Quantum Composition*, Physical Review Letters **130**, 110202 (2023).

[25] L. Pollyceno, A. Chaturvedi, C. Raj, P. R. Dieguez, and M. Pawlowski, *Monogamy of nonlocality from multipartite information causality*, arXiv (2024).

[26] N. Miklin and M. Pawlowski, *Information Causality without Concatenation*, Physical Review Letters **126**, 220403 (2021).

[27] P. Jain, M. Gachechiladze, and N. Miklin, *Information causality as a tool for bounding the set of quantum correlations*, arXiv (2023).

[28] M. PlÃ¡vala, *General probabilistic theories: An introduction*, arXiv (2021).

[29] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. AcÃn, *Unbounded randomness certification using sequences of measurements*, Physical Review A **95**, 020102 (2017).

[30] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Bell nonlocality*, Reviews of Modern Physics **86**, 419 (2014).

[31] F. J. Curchod, *Nonlocal resources for quantum information tasks*, PhD thesis, Universitat PolitÃšcnica de Catalunya. Institut de Ciencies Fotoniques, 2018, URI: http://hdl.handle.net/10803/663448.

[32] V. Scarani, *Bell nonlocality*, Oxford University Press, Oxford, first edition edition, 2019.

[33] A. Fine, *Hidden Variables, Joint Probability, and the Bell Inequalities*, Physical Review Letters **48**, 291 (1982).

[34] S. Popescu and D. Rohrlich, *Generic quantum nonlocality*, Physics Letters A **166**, 293 (1992).

[35] 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 (1969).

[36] R. Ramanathan, J. Tuziemski, M. Horodecki, and P. Horodecki, *No Quantum Realization of Extremal No-Signaling Boxes*, Physical Review Letters **117**, 050401 (2016).

[37] K. T. Goh, J. Kaniewski, E. Wolfe, T. VÃ©rtesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, *Geometry of the set of quantum correlations*, Physical Review A **97**, 022104 (2018).

[38] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, *MIP*=RE*, arXiv (2020).

[39] B. S. Cirel'son, *Quantum generalizations of Bell's inequality*, Letters in Mathematical Physics **4**, 93 (1980).

[40] A. Rai, C. Duarte, S. Brito, and R. Chaves, *Geometry of the quantum set on no-signaling faces*, Physical Review A **99**, 032106 (2019).

[41] L. Masanes, *Necessary and sufficient condition for quantum-generated correlations*, arXiv (2003).

[42] M. Navascues, S. Pironio, and A. Acin, *Bounding the set of quantum correlations*, arXiv (2006).

[43] P.-S. Lin, T. VÃ©rtesi, and Y.-C. Liang, *Naturally restricted subsets of nonsignaling correlations: typicality and convergence*, Quantum **6**, 765 (2022).

[44] H. Barnum, S. Beigi, S. Boixo, M. B. Elliott, and S. Wehner, *Local Quantum Measurement and No-Signaling Imply Quantum Correlations*, Physical Review Letters **104**, 140401 (2010).

[45] T. Fritz, A. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. AcÃn, *Local orthogonality as a multipartite principle for quantum correlations*, Nature Communications **4**, 2263 (2013).

[46] M. Navascus and H. Wunderlich, *A glance beyond the quantum model*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **466**, 881 (2010).

[47] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, *Limits on nonlocal correlations from the structure of the local state space*, New Journal of Physics **13**, 063024 (2011).

[48] M. DallArno, S. Brandsen, A. Tosini, F. Buscemi, and V. Vedral, *No-Hypersignaling Principle*, Physical Review Letters **119**, 020401 (2017).

[49] J. Henson, *Quantum contextuality from a simple principle?*, arXiv (2012).

[50] G. Brassard, H. Buhrman, N. Linden, A. A. MÃ©thot, A. Tapp, and F. Unger, *Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial*, Physical Review Letters **96**, 250401 (2006).

[51] A. B. Sainz, Y. Guryanova, A. AcÃn, and M. NavascuÃ©s, *Almost-Quantum Correlations Violate the No-Restriction Hypothesis*, Physical Review Letters **120**, 200402 (2018).

[52] 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).

[53] H. Guo, J. Zhang, and G. J. Koehler, *A survey of quantum games*, Decision Support Systems **46**, 318 (2008).

[54] P.-E. Emeriau, M. Howard, and S. Mansfield, *Quantum Advantage in Information Retrieval*, PRX Quantum **3**, 020307 (2022).

[55] M. Pawlowski and M. Zukowski, *Entanglement-assisted random access codes*, Physical Review A **81**, 042326 (2010).

[56] A. Chaturvedi, M. Pawlowski, and K. Horodecki, *Random access codes and nonlocal resources*, Physical Review A **96**, 022125 (2017).

[57] M. Farkas, N. Miklin, and A. Tavakoli, *Simple and general bounds on quantum random access codes*, arXiv (2023).

[58] A. Hameedi, D. Saha, P. Mironowicz, M. Pawlowski, and M. Bourennane, *Complementarity between entanglement-assisted and quantum distributed random access code*, Physical Review A **95**, 052345 (2017).

[59] A. Tavakoli, A. Hameedi, B. Marques, and M. Bourennane, *Quantum Random Access Codes Using Single d-Level Systems*, Physical Review Letters **114**, 170502 (2015).

[60] D. Saha, D. Das, A. K. Das, B. Bhattacharya, and A. S. Majumdar, *Measurement incompatibility and quantum advantage in communication*, Physical Review A **107**, 062210 (2023).

[61] A. Tavakoli, B. Marques, M. Pawlowski, and M. Bourennane, *Spatial versus sequential correlations for random access coding*, Physical Review A **93**, 032336 (2016).

[62] Y. Xiao, X.-H. Han, X. Fan, H.-C. Qu, and Y.-J. Gu, *Widening the sharpness modulation region of an entanglement-assisted sequential quantum random access code: Theory, experiment, and application*, Physical Review Research **3**, 023081 (2021).

[63] W. v. Dam, *Implausible consequences of superstrong nonlocality*, Natural Computing **12**, 9 (2013).

[64] M. M. Wilde, *From Classical to Quantum Shannon Theory*, arXiv (2011).

[65] S. W. Al-Safi and A. J. Short, *Information causality from an entropic and a probabilistic perspective*, Physical Review A **84**, 042323 (2011).

[66] J. Uffink, *Quadratic Bell Inequalities as Tests for Multipartite Entanglement*, Physical Review Letters **88**, 230406 (2002).

[67] O. C. O. Dahlsten, D. Lercher, and R. Renner, *Tsirelson's bound from a generalized data processing inequality*, New Journal of Physics **14**, 063024 (2012).

[68] M. Gachechiladze, B. Bak, M. Pawlowski, and N. Miklin, *Quantum Bell inequalities from Information Causality - tight for Macroscopic Locality*, Quantum **6**, 717 (2022).

[69] T. H. Yang, D. Cavalcanti, M. L. Almeida, C. Teo, and V. Scarani, *Information-causality and extremal tripartite correlations*, New Journal of Physics **14**, 013061 (2012).

[70] L. Pollyceno, R. Chaves, and R. Rabelo, *Information causality in multipartite scenarios*, Physical Review A **107**, 042203 (2023).

[71] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. VÃ©rtesi, *Closed sets of nonlocal correlations*, Physical Review A **80**, 062107 (2009).

[72] P. Botteron, A. Broadbent, R. Chhaibi, I. Nechita, and C. Pellegrini, *Algebra of Nonlocal Boxes and the Collapse of Communication Complexity*, Quantum **8**, 1402 (2024).

[73] S. G. A. Brito, M. G. M. Moreno, A. Rai, and R. Chaves, *Nonlocality distillation and quantum voids*, Physical Review A **100**, 012102 (2019).

[74] T. Short, S. Popescu, and N. Gisin, *Entanglement swapping for generalized non-local correlations*, arXiv (2005).

[75] B. Lang, T. VÃ©rtesi, and M. NavascuÃ©s, *Closed sets of correlations: answers from the zoo*, Journal of Physics A: Mathematical and Theoretical **47**, 424029 (2014).

[76] D. Collins and N. Gisin, *A relevant two qubit Bell inequality inequivalent to the CHSH inequality*, Journal of Physics A: Mathematical and General **37**, 1775 (2004).

[77] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*, IOP Publishing, 2023.

[78] G. Aishwarya and M. Madiman, *Remarks on RÃ©nyi versions of conditional entropy and mutual information*, 2019 IEEE International Symposium on Information Theory (ISIT) **00**, 1117 (2019).

[79] R. Gallego, L. E. WÃŒrflinger, R. Chaves, A. AcÃn, and M. NavascuÃ©s, *Nonlocality in sequential correlation scenarios*, New Journal of Physics **16**, 033037 (2014).

Appendix A

Deriving & rationalising Information Causality

Given the rules of an (EA)RAC, information causality is implied from natural properties of mutual information, which we stated in eq. 2.9 - 2.13:

$$\begin{aligned} & I(\overrightarrow{\alpha} : \mu', B) - I(\overrightarrow{\alpha}, B : \mu'|\mu) = I(\overrightarrow{\alpha} : \mu', B) - I(\mu.\overrightarrow{\alpha}, B : \mu') + I(\mu, \mu') \\ & = I(\overrightarrow{\alpha} : \mu'|B) + I(\overrightarrow{\alpha} : B) - I(\mu : \mu'|\overrightarrow{\alpha}, B) - I(\overrightarrow{\alpha}, B : \mu') + I(\mu : \mu') \\ & \leq I(\overrightarrow{\alpha} : \mu'|B) - I(\overrightarrow{\alpha}, B : \mu') + I(\mu : \mu') \\ & = I(\overrightarrow{\alpha} : \mu'|B) - I(\overrightarrow{\alpha} : \mu'|B) - I(B : \mu') + I(\mu : \mu') \\ & = I(\mu : \mu') - I(B : \mu') \leq I(\mu : \mu') \equiv \kappa \end{aligned} \tag{A.1}$$

Note that the above is just a slight modification of the derivation presented in Appendix A of [27]. At the moment of writing this thesis, we were namely not able to follow and reproduce the derivation in the pre-print of that paper.

The first two steps are simply the chain rule eq. 2.12 applied $I(\overrightarrow{\alpha}, B : \mu'|\mu)$ and $I(\mu.\overrightarrow{\alpha}, B : \mu')$ respectively. From the second to third line we drop $I(\overrightarrow{\alpha} : B)$ since B's subsystem B, before B's measurement, is completely uncorrelated with input data $\overrightarrow{\alpha}$ at A's side. Also we drop $I(\mu : \mu'|\overrightarrow{\alpha}, B)$ in the same step in exchange for an "$\leq$" by using non-negativity of the mutual information. From the third to fourth line, we apply the chain rule once again to $I(\overrightarrow{\alpha}, B : \mu')$. In the final step we use non-negativity to exchange $I(\overrightarrow{\alpha}, B : \mu')$ for a looser inequality, and we notice that the amount of information $I(\mu : \mu')$ shared between the original and received message is ideally equal to the constant channel capacity κ.
The second term in the first line, $I(\overrightarrow{\alpha}, B : \mu'|\mu)$, is difficult to interpret.

However, it is merely a mathematical catalyst in the derivation and can be dropped in the final expression. This is possible because it vanishes by itself. The received message μ' neither depends upon B's subsystem B nor contains it any extra information about $\overrightarrow{\alpha}$ relative to the original message μ, so indeed $I(\overrightarrow{\alpha}, B : \mu'|\mu)$.

The statement of information causality in a EARAC setup then reads: [16]

$$I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu') \equiv \kappa \tag{A.2}$$

Which is exactly eq. 2.14 in chapter 2 of the main text.

It is worth highlighting the non-constant, but more restrictive, inequality arising from not making the final step in the above derivation (eq. A.1):

$$I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu') - I(B : \mu')$$

In fact, this inequality makes the key point of information causality explicit, namely that all information of B about A's input data must flow via the communication channel. Even if the shared no-signaling correlations would encode some structural information about the data, they can't be decoded without an equivalent *amount* of information in the signaled message. The reverse idea is the basis for quantum key distribution whereby the no-signaling correlations are used as part of the key to access the encrypted message on the possibly exposed communication channel. In fact, the inequality $I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu') - I(B : \mu')$ leads to an equivalent to the classical Wyner-Ziv compression constraint in the rate-distortion framework. A discussion of this equivalence is beyond the scope of this thesis, but it can help to think of information causality as a trade-off between the limiting quantity $I(\mu : \mu') - I(B : \mu')$ as the "rate" and the "distortion" $H(\overrightarrow{\alpha}|\mu', B)$.

We mentioned the tighter statement $I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu') - I(B : \mu')$ of IC. Note that it involves the information arriving at B via the communication channel and the information flow from the correlated side-information resulting from B's measurements on his part the no-signaling box. The statement of information causality in all usual works [16, 27] simplify this inherent trade-off by implicitly assuming that no information about $\overrightarrow{\alpha}$ is allowed to come from B's local measurement on B, such that $I(B : \mu')$ is small and the inequality $I(\overrightarrow{\alpha} : \mu', B) \leq \kappa$ nearly tight.
However, while this is a fully reasonable assumption within the framework of quantum mechanics and when converging towards the quantum set of

correlations, it is in conflict with our aim for a theory agnostic characterisation of the correlations bounds implied by information causality. Moreover, it is a well-known approach in related fields of research, like quantum field theory, to introduce virtual sources of some physical quantities that we can let vanish later. Although it here rather complicates calculations, it might help to investigate constraints more holistically and interpret extreme demands on the main communication resource in the considered information retrieval task.

Appendix B

Information causality in sequential measurement scenarios

B.1 Non-Locality Recycling

It was found that by performing sequentially multiple measurements on the same copy of a quantum state, one can extract non-local statistics from states which do not show non-locality in distributions resulting from any single measurement. This type of genuinely spatio-temporal correlations has since been called "hidden" non-locality.

The conditional distributions for bipartite experiments with s and t sequential measurements in subsystem A and B respectively have the form $P(a_1, \ldots, a_s, b_1, \ldots, b_t | x_1, \ldots, x_s, y_1, \ldots, y_t)$. There are multiple types of sequential no-signaling correlations, depending on which operations can be performed in between different measurements.
Post-selected sequential correlations are the most general type since they allow arbitrary local transformations of the underlying physical resource between measurements, which can also depend on previous outputs a_i (b_i) and inputs x_i (y_i) of the sequential no-signaling box.

The simplest type of sequential correlations, on which we will focus, are the so-called *time-ordered* correlations. For those, no operations on the physical resource can be performed in between the measurements. Nevertheless, the post-measurement state of the physical resource is assumed to encode information about previous outputs a_i (b_i) and inputs x_i (y_i), it just can not be deliberately processed further before the next measurement happens.
For (s, t) time-ordered scenarios, the measurements are assumed to be per-

formed in a definite order and *outcomes* of any measurement may depend on previous inputs / outcomes, e.g. $b_i \neq b_i(y_i, y_j, b_j)$ for any $j < i$ (but *not* vice-versa).

Definition B.1.1. ***Time-Ordered distributions*** *—*
A distribution $P(a_1, \ldots, a_s, b_1, \ldots, b_t | x_1, \ldots, x_s, y_1, \ldots, y_t) \equiv P(a\ \overrightarrow{b}\ |x\ \overrightarrow{y})$ *is called Time-Ordered if it satisfies the so-called Arrow-Of-Time (AoT) condiitions*

$$P(b_1|y_1) = \sum_a \sum_{b_2} \cdots \sum_{b_t} P(a\ \overrightarrow{b}\ |x\ \overrightarrow{y})$$

and

$$P(b_j|\overleftarrow{y}_j \overleftarrow{b}_{j-1}) = \sum_a \sum_{b_{j+1}} \cdots \sum_{b_t} P(a\ \overrightarrow{b}\ |x\ \overrightarrow{y}) \tag{B.1}$$

for all $j \in \{2, \ldots, t-1\}$ *whereby* $\overleftarrow{v}_k \equiv v_1, \ldots, v_k$.

The above is for a single-sided B, a $(1, t)$-sequential scenario, but can easily be generalised to a mono-sequential A or two-sided sequential A and B.
The AoT conditions ensure that the statistics at any point in the sequence is independent of A's side (full no-signaling) and independent of future/-subsequent measurement outcomes/incomes of B.

Bell locality for (one-sided) time-ordered distributions then means the following [Def. 2 in [79]]:

Definition B.1.2. ***Time-Ordered Bell locality*** *— A conditional distribution* $P(a, b_1, b_2|x, y_1, \ldots, y_t)$*, which is time-ordered in the sense of def. B.1.1, is called Time-Ordered Bell local (TO-local) if and only if*

$$P(a, b_1, \ldots, b_t|x, y_1, \ldots, y_t) = \int_\Lambda d\lambda Q(\lambda) P_\lambda(a|x) P_\lambda(b_1, \ldots, b_t|y_1, \ldots, y_t) \tag{B.2}$$

given some ensemble of marginal distributions ("processes")

$$\{(Q(\lambda), P_\lambda(a|x), P_\lambda(a|x)P_\lambda(b_1, \ldots, b_t|y_1, \ldots, y_t))\,|\lambda \in \Lambda\}$$

, such that $Q(\lambda) \geq 0$ *and* $\sum_\lambda Q(\lambda) = 1$*. Furthermore* $P_\lambda(b_1, \ldots, b_t|y_1, \ldots, y_t)$ *must satisfy the AoT conditions, i.e.*

$$P_\lambda(b_1|y_1) = \sum_{b_2} P_\lambda(b_1, b_2|y_1, y_2)$$

While the CHSH family of Bell inequalities was sufficient to detect Bell non-local distributions in the single measurement case, three different types of Bell inequalities are needed for time-ordered case. [79]
For illustrative purposes, we consider only two one-sided sequential measurement for B (i.e. distributions $P(a.b_1, b_2|x, y_1, y_2)$ and only list here a single instance of each type of Bell inequality. For the following, it helps to interpret the sequential subsystem B as two (time-)separated subsystems B1 and B2.

The first type of Bell inequality just accounts for non-locality in the first-measurement of subsystem B and completely ignores the second measurement (outcome) b_2. Alternatively, instead of ignoring any measurements in subsystem B2, we can also consider all possible input-output combinations (y_2, b_2). Both views are fully equivalent.
However, sticking to the view of ignoring subsystem B2, we can simply write down the canonical CHSH inequality between subsystems A and B1 as an example for first type of inequality: [79]

$$E_{00} + E_{01} + E_{10} - E_{11} \leq 2$$

Where we have used the usual Bell correlators between subsystems A and B1

$$E_{kl} \equiv 2P(a = b_1|x = k, y_1 = l) - 1 = 2\left(\sum_i \sum_j P(a = i, b_1 = i, b_2 = j|x = k, y_1 = l, y_2)\right) - 1$$
$$= \sum_{a=b_1} P(a, b_1|x = k, y_1 = l) - \sum_{a\neq b_1} P(a, b_1|x = k, y_1 = l)$$

which we call marginal Bell correlators within the sequential scenario. Note that the value of y_2 in the last part can be chosen arbitrarily because of the no-signaling property of TO distributions $P(a, b_1, b_2|x, y_1, y_2)$.

For the second inequality, we only consider non-locality with respect to the second measurement in subsystem B2. However, since the second measurement depends upon what happens in subsystem B1, we now have to fix a specific input-output pair (y_1, b_1). Thus for every (y_1, b_1) there is a separate inequality: [79]

$$E^{b_1}_{0\,y_1 0} + E^{b_1}_{0\,y_1 1} + E^{b_1}_{1\,y_1 0} - E^{b_1}_{1\,y_1 1} \leq 2$$

With the post-selected Bell correlators

$$E^{b_1}_{k\,y_1 l} \equiv 2P(a = b_2|x = k, y_2 = l, (y_1, b_1)) - 1$$
$$= \left(\sum_{a,b_2} P(a, b_1, b_2|x = k, y_1, y_2 = l)\right)^{-1} \left(\sum_{a=b_2} P(a, b_1, b_2|x = k, y_1, y_2 = l) - \sum_{a\neq b_2} P(a, b_1, b_2|x = k, y_1, y_2 = l)\right)$$

for each fixed (y_1, b_1). These correlators thus only consider correlations within the bipartition of subsystem A and the (post-measurement) subsystem $B^{(2)}$ after the first measurement.
Experimentally, fixing (y_1, b_1) would correspond to post-selecting the measurement outcomes in subsystem B2 on those cases where B1 had input y_1 and got output b_1. In the first inequality, in contrast, the results in subsystem B2 simply did not matter and could have been mixed, as can be seen from the definitions of E_{xy_1} and $E^{b_1}_{x\,y_1y_2}$ respectively.

The third and last type of Bell inequality for TO distributions assesses a special type of non-locality by taking also (joint) Bell correlators

$$E_{k\,lv} \equiv P(a = 0, b_1 = b_2|x = k, y_1 = l, y_2 = v) - P(a = 1, b_1 = b_2|x = k, y_1 = l, y_2 = v)$$
$$+ P(a = 1, b_1 \neq b_2|x = k, y_1 = l, y_2 = v) - P(a = 0, b_1 \neq b_2|x = k, y_1 = l, y_2 = v)$$

between all three subsystems A, B1 and B2 into account. Although this definition looks confusing on first sight, An CHSH-like instance of the third inequality type is then: [79]

$$E_{0\,00} + E_{0\,01} - E_{0\,10} + E_{0\,11} - E_{1\,00} - E_{1\,01} - E_{1\,10} + E_{1\,11}$$
$$- \left(\sum_{b_1} P(b_1|y_1 = 0)\left[E^{b_1}_{0\,00} - E^{b_1}_{0\,01} - E^{b_1}_{1\,00} + E^{b_1}_{1\,01}\right]\right.$$
$$\left. + P(b_1|y_1 = 1)\left[E^{b_1}_{0\,10} + E^{b_1}_{0\,11} + E^{b_1}_{1\,10} + E^{b_1}_{1\,11}\right]\right) \leq 2$$

with $P(b_1|y_1) \equiv \sum_{a,b_2} P(a, b_1, b_2|a, y_1, y_2)$ for arbitrary $(y_2, b_2) \in \{0, 1\}^2$ because of the AoT and no-signaling conditions.

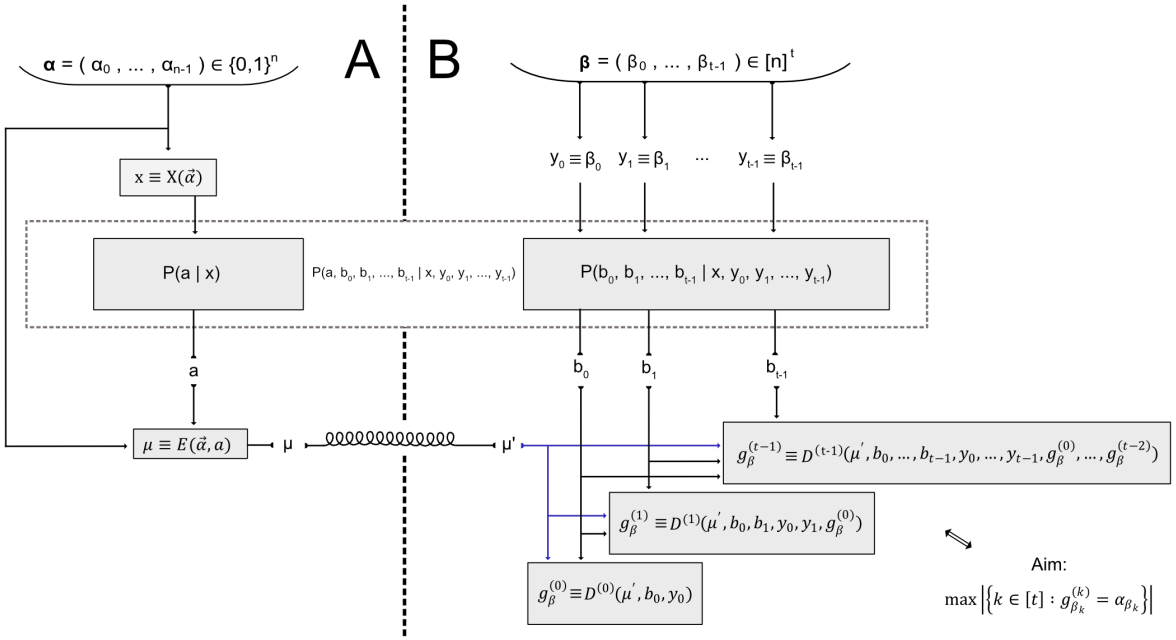


Figure B.1: *The van Dam protocol generalised to a Time-Ordered $(1,t)$-sequential scenario within a bipartite EARAC setup between A and B in a $(n,n,2,2)$ Bell scenario. While A receives a dit-string $\overrightarrow{\alpha}$ of fixed length n, B is queried with t n-dimensional bits $\overrightarrow{\beta}$. Each of the latter β_i indicates which dit in $\overrightarrow{\alpha}$ is supposed to recovered by final output dits $g_\beta^{(i)}$. The two parties only share a noisy communication channel of $\kappa \equiv I(\mu : \mu')$ dits classical capacity.*
The aim of maximising the number of correct guesses, $g_\beta^{(i)} = \alpha_\beta$, is only a simple example.

B.2 Bounding information retrieval in RACs with sequential measurements

The statement of the IC principle stays the following:

$$I(\overrightarrow{\alpha} : \mu', B) \leq I(\mu : \mu')$$

Where κ is the capacity of the noisy communication channel. The proof that this IC statement holds for Shannon and quantum mutual information makes no reference to the properties of time-ordered boxes, Therefore, it is identical to the proof of IC in [27] for the case of noisy communication channels.
Similarly to non-sequential case, the mutual information $I(\overrightarrow{\alpha} : \mu', B)$ can be expressed more concretely in terms of the mutual information with

respect to the guesses $(g^{(1)}, g^{(2)}$, instead of the abstract subsystem B and received message μ'. In a $(1,t)$ sequential scenario with binary inputs (x, y_1, y_2) and binary outputs (a, b_1, b_2):

$$\begin{aligned} I(\overrightarrow{\alpha} : \mu', B) &\geq I(\overrightarrow{\alpha} : \mu', B\,|y_0) \\ &\geq I(\overrightarrow{\alpha} : \mu', B, g^{(0)}, B_{(b_0|y_0)}\,|y_0) \\ &\geq \dots \\ &\geq I(\overrightarrow{\alpha} : \mu', B_{(\overleftarrow{b}_{t-1}|\overleftarrow{y}_{t-1})}\,|\overleftarrow{g}^{(t-1)}, \overleftarrow{y}_{t-1}) + \sum_{k=0}^{t-1} I(vec\alpha : g^{(k)}\,|\overleftarrow{g}^{(k-1)}, \overleftarrow{y}_k) \\ &\geq \sum_{k=0}^{t-1} I(vec\alpha : g^{(k)}\,|\overleftarrow{g}^{(k-1)}, \overleftarrow{y}_k) \\ &\geq \dots \\ &\geq \sum_{k=0}^{t-1}\sum_{i=0}^{n-1} I(\alpha_i : g^{(k)}\,|\overleftarrow{g}^{(k-1)}, \overleftarrow{y}_k) \\ &= \sum_{k=0}^{t-1}\sum_{i=0}^{n-1} I(\alpha_i : g^{(k)}\,|\overleftarrow{g}^{(k-1)}, y_k = i, \overleftarrow{y}_{k-1}) \end{aligned} \tag{B.3}$$

For simplicity we assumed that all guesses are made by different instances of party B such that each guess g^k is independent of all previous guesses $\overleftarrow{g}^{k-1}$. Otherwise, the information from earlier guesses for the same input value y_k might be used to improve the current one. Note that this is only assumed for the guessed values $\overleftarrow{g}^{k-1}$, the post-measurement state $B_{(\overleftarrow{b}_{k-1}|\overleftarrow{y}_{k-1})}$ might thus still transfer some information about previous outputs $\overleftarrow{b}_{k-1}$ to the next output b_k.

We then arrive at the time-orderd sequential generalisation of the Uffink-like IC-bound from [27]:

$$\sum_{k=1}^{t-1} 2^{-(k-1)} \sum_{\overleftarrow{j}_{(k-1)}} (e^{(k)}_{0,(0,\overleftarrow{j}_{(k-1)})} + e^{(k)}_{1,(0,\overleftarrow{j}_{(k-1)})})^2 + (e^{(k)}_{0,(1,\overleftarrow{j}_{(k-1)})} + e^{(k)}_{1,(1,\overleftarrow{j}_{(k-1)})})^2 \leq 4 \tag{B.4}$$

whereby we have defined the time-ordered box biases

$$e^{(z)}_{i,\overleftarrow{l}_z} \equiv 2P(a = b_z | x = i, \overleftarrow{y}_z = \overleftarrow{l}_z) - 1 \tag{B.5}$$

$$\stackrel{\text{A.O.T}}{=} 2P(a = b_z | x = i, \overleftarrow{y}_t = \overleftarrow{l}_t) - 1 \tag{B.6}$$

More general protocol, including inter-dependencies between guesses and partial box outputs b. We can calculate the mutual information $I(\overrightarrow{\alpha},)$ as

$$
\begin{aligned}
I(\overrightarrow{\alpha} : \mu', B) &\geq \sum_{z=0}^{t-1} I(\overrightarrow{\alpha} : g^{(z)} \,|\overleftarrow{y}_z, \overleftarrow{g}^{(z-1)}) \qquad \text{(B.7)}\\
&= \sum_{z=0}^{t-1} n - \Big(H(g^{(z)}, \overrightarrow{\alpha} \,|\overleftarrow{y}_z, \overleftarrow{g}^{(z-1)}) - H(g^{(z)} \,|\overleftarrow{y}_z, \overleftarrow{g}^{(z-1)})\Big)\\
&= \sum_{z=0}^{t-1} n - \Bigg[\sum_i \sum_{\overleftarrow{j}_{z-1}} \frac{1}{n^z}\Big(H(g^{(z)}, \overrightarrow{\alpha} \,|y_z = i, \overleftarrow{y}_{z-1} = \overleftarrow{j}_{z-1}, \overleftarrow{g}^{(z-1)})\\
&\quad - H(g^{(z)} \,|y_z = i, \overleftarrow{y}_{z-1} = \overleftarrow{j}_{z-1}, \overleftarrow{g}^{(z-1)})\Big)\Bigg]
\end{aligned}
$$

with $t = 2$ and $n = 2$, we can calculate the mutual information by specifying the following conditional probabilities:

$$
\begin{aligned}
&P(g^{(0)} = c|\,\overrightarrow{\alpha} = \overrightarrow{v}, y_0 = i) = \frac{1}{2} + \frac{e_c}{2}\Bigg(\sum_{a,b_0} \delta\Big[D^{(0)}(E(\overrightarrow{v}, a), b_0, i) = c\Big]\\
&\Bigg(\sum_{k=0}^{n-1} \delta[X(\overrightarrow{v}) = k]\; P(a, b_0|x = k, y_0 = i)\Bigg)\Bigg)\\
&P(g^{(1)} = c, g^{(1)} = r_0|\,\overrightarrow{\alpha} = \overrightarrow{v}, y_1 = i, y_0 = j_0) = \frac{1}{2}\\
&+ \frac{e_c}{2}\Bigg(\sum_{a,b_0,b_1} \delta\Big[D^{(1)}(E(\overrightarrow{v}, a), b_1, (j_0, i), r_0) = c\Big]\delta\Big[D^{(0)}(E(\overrightarrow{v}, a), b_0, (j_0, i), c) = r_0\Big]\\
&\cdot \Bigg(\sum_{k=0}^{n-1} \delta[X(\overrightarrow{v}) = k]\; P(a, b_0, b_1|x = k, y_1 = i, y_0 = j_0)\Bigg)\Bigg)
\end{aligned}
$$

with the kronecker delta $\delta[* = **] \equiv \delta_{*,**}$. The above is fully determined by the time-ordered no-signaling box $P(a, b_0, b_1|x, y_0, y_1)$. Note that through the many interdependencies and the resulting kronecker delta factors, we can study the situation only numerically in this generality.

Appendix C

Multi-bit channels in EARACs

Recall from the main text that in [26], it was noticed that generalised dit-input dit-output communication channels yield the strongest IC bound for a non-zero channel capacity $\kappa \neq 0$ if $d \geq 3$. In contrast, the optimal channel capacity for the binary symmetric channel ($d = 2$) was found in the limit of a vanishing capacity $\kappa \to 0$. While the reason for this discrepancy between bit- and dit-channels remains open, this curious finding shows the high potential for learning something new about IC by simply considering other (noisy) channel types, particularly those with higher-dimensional inputs and outputs.

While preparing this thesis, we started experimenting with more complex channels as well. However, we considered a different type of higher-dimensional channel since the introduction of dit-channels generally only makes sense when simultaneously no-signaling boxes $P(a, b|x, y)$ with non-bit outputs are used. That is, when $a, b \in [d]$ with d the dimensionality of the input and output of the communication channel. The equivalent of a PR-box in a (n, n, d, d) Bell scenario could otherwise not reach optimal performance in an EARAC anymore.
The focus of this thesis, however, are boxes in the simplest $((2, 2, 2, 2))$ Bell scenario. To adhere to this focus, we can alternatively use multiple bit-channels in such a way that they simulate the channels with higher-dimensional inputs and outputs. Therefore, we started to investigate the case of channels with d bit-valued inputs and d bit-valued outputs, which can be considered equivalent to channels with a single 2^d-dimensional input and output for some $d \in \mathbb{N}$. Notice that we must ensure that the number n of data bits $(\alpha_0, ..., \alpha_n)$ is larger than the maximal channel ca-

pacity. Otherwise, if $n = d$, the RAC is a trivial task for party A and B.[1] For a full schematic of the modified EARAC setup, see figure C.1 in the appendix.

The open question is then whether also in this setting the strongest IC-bound is obtained for a non-zero channel capacity κ whenever $d \geq 2$. If yes, why? If no, we may conclude that the optimality of a non-zero channel capacity for EARACs with dit-channels is not simply due to the higher transmission rate of the channel, but rather a consequence of using boxes in more complex (n, n, d, d) Bell scenarios.

[1] Consequently, to avoid more complex no-signaling boxes with n-dimensional input-bits (i.e. $(n, n, 2, 2)$-boxes) for $n > 2$, we must also allow the parties to share multiple copies of any box in the $(2, 2, 2, 2)$ scenario.

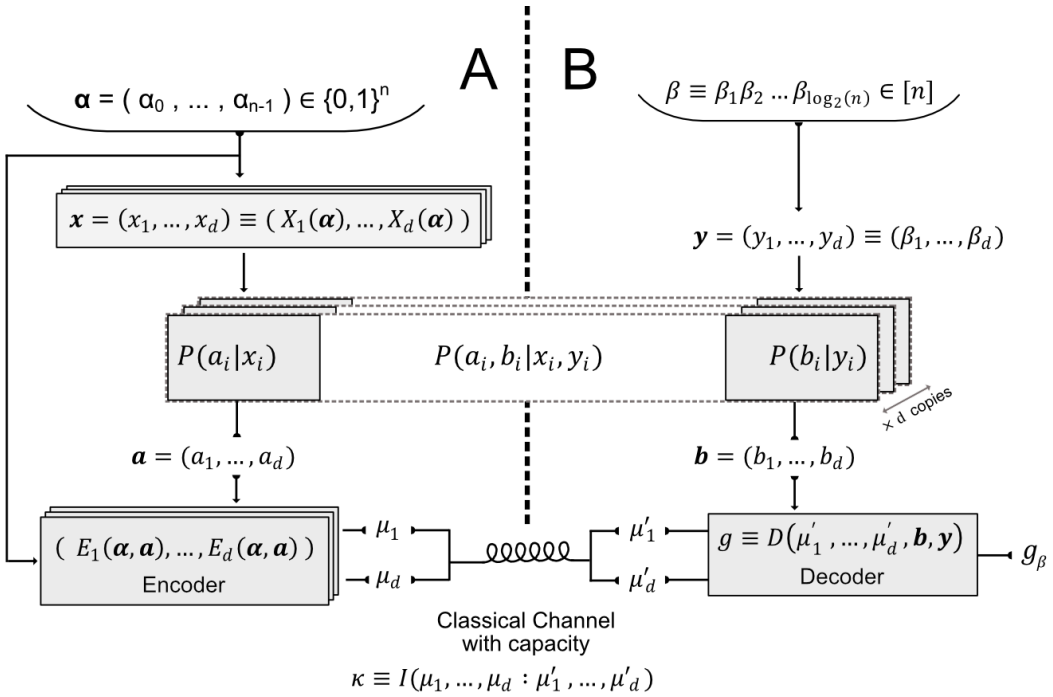


Figure C.1: *Schematic of a bipartite Random Access Code (RAC) between space-like separated parties A and B communicating with a noisy communication channel with d inputs and d outputs. Additionally, d copies of a no-signaling box $P(a,b|x,y) \in \mathcal{NS}$ in a $(2,2,2,2)$ scenario are shared between the parties. Each box copy has a unique identifier $i \in \{1,\dots,d\}$ and takes one independently produced pair of inputs (x_i, y_i) with the corresponding index, and subsequently outputs the respective pair of outputs (a_i, b_i) into the tuples of independent outputs $(a_1,\dots,a_d)$ and $(b_1,\dots,b_d)$. I.e. box copy i is fully represented by distribution $P(a_i, b_i|x_i, y_i)$.*
While A receives a bit-string $\overrightarrow{\alpha}$ of fixed length n, B is queried with only a single n-dimensional bit β. The latter indicates the index of the bit in $\overrightarrow{\alpha}$ which the final "guess" g_β is supposed to recover. In particular, it is assumed that $n = 2^d$ such that $\beta = \beta_1\beta_2\dots\beta_d$ is the complete binary expansion of $\beta \in [d]$