An Operational Environment for Quantum Self-Testing

Matthias Christandl, Nicholas Gauguin Houghton-Larsen, and Laura Mančinska

Department of Mathematical Sciences, University of Copenhagen,
Universitetsparken 5, 2200 Copenhagen, Denmark

Observed quantum correlations are known to determine in certain cases the underlying quantum state and measurements. This phenomenon is known as *(quantum) self-testing*.

Self-testing constitutes a significant research area with practical and theoretical ramifications for quantum information theory. But since its conception two decades ago by Mayers and Yao, the common way to rigorously formulate self-testing has been in terms of operator-algebraic identities, and this formulation lacks an operational interpretation. In particular, it is unclear how to formulate self-testing in other physical theories, in formulations of quantum theory not referring to operator-algebra, or in scenarios causally different from the standard one.

In this paper, we explain how to understand quantum self-testing operationally, in terms of causally structured dilations of the input-output channel encoding the correlations. These dilations model side-information which leaks to an environment according to a specific schedule, and we show how self-testing concerns the relative strength between such scheduled leaks of information. As such, the title of our paper has double meaning: we recast conventional quantum self-testing in terms of information-leaks to an environment — and this realises quantum self-testing as a special case within the surroundings of a general operational framework.

Our new approach to quantum self-testing not only supplies an operational understanding apt for various generalisations, but also resolves some unexplained aspects of the existing definition, naturally suggests a distance measure suitable for robust self-testing, and points towards self-testing as a modular concept in a larger, cryptographic perspective.

Matthias Christandl: christandl@math.ku.dk
Nicholas Gauguin Houghton-Larsen: nicholas.gauguin@gmail.com
Laura Mančinska: mancinska@math.ku.dk

Contents

1 Introduction

1.A History and Motivation

More than half a century ago, J. S. Bell understood that a maximally entangled pair of qubits may be subjected to local measurements such that the statistics of measurement outcomes predicted by quantum theory cannot be explained in terms of local or shared classical randomness [Bel64]. Almost three decades later, A. Ekert observed that this remarkable fact amounts to some of the randomness in the measurement outcomes being fundamentally unknowable in advance of the experiment in Bell's scenario, and therefore in principle useful for cryptographic purposes [Eke91].

These ideas were revisited when D. Mayers and A. Yao introduced the notion of *quantum self-testing* [MY98, MY04], according to which a collection of measurement statistics may in certain cases imply that a specific configuration of quantum state and measurements was used in the experiment.[1] Today, quantum self-testing has grown to a sizeable sub-field of quantum information theory, with applications in randomness generation [Col06, CK11, CY14, ŠASA16, MS16, BMP18], quantum key distribution [MS16, JMS20] and delegated quantum computation [RUV13, McK16, NV17, CGJV19, JMS20], and with theoretical significance for the understanding of quantum correlations [GKW+18, CS18, Col20] and quantum complexity theory [NW19, JNV+21].

Though the cardinal idea behind self-testing — that certain measurement statistics arise from a unique configuration of quantum state and measurements — is perhaps intuitively meaningful, it is not entirely trivial to formalise. This is because there are strictly speaking always several different state-measurement-configurations which realise the same measurement statistics. For example, starting from any initial configuration, a simultaneous unitary rotation of the measurements and state will leave the statistics invariant. So will adding an ancillary state which is discarded in measurements.

This paper is concerned with the proper mathematical formalisation of self-testing.

The conventional formal definition of self-testing was introduced in Ref. [MY04], standardised in Ref. [MYS12] and has recently been reviewed in Ref. [ŠB20]. This definition relates the different possible state-measurement-configurations with given measurement statistics to a fixed *canonical* configuration, by accounting for the possibility of unitary rotations and unmeasured ancillary states in the traditional operator-algebraic language for quantum theory.[2] As witnessed by the breadth of self-testing results, this definition has been used successfully for more than 15 years to progress our understanding of theoretical and practical aspects of self-testing.

However, because the definition is phrased in terms of operator-algebraic objects rather than information channels, it suffers from a lack of *operational meaning*.[3] This circumstance is not only theoretically unaesthetic, but also hinders the immediate generalisation of self-testing to scenarios causally different from Bell's [BGP10, BRGP12, Fri12], to theories beyond quantum theory [AC04, Sel04, Bae06, Bar07, BW11, BW16, CDP10, Har13], or

[1] Granted that the state in this configuration is pure and entangled, self-testing strengthens Ekert's observation (because no environment can hold information about a pure state), as well as Bell's (because no separable state can reproduce the measurement statistics).

[2] Specifically, this is the definition of 'full' self-testing, in which one makes a statement simultaneously about the state and the measurements; one can instead consider relaxed self-testing statements e.g. about the state alone [ŠB20].

[3] It is unfortunate that the syntactically similar 'operator' and 'operational' are here in semantic opposition to each other.

even to formalisations of quantum theory which transcend the operator-algebraic formalism [Har01, CDP11]. It is conceivable that quantum theory, or aspects thereof, will eventually be best understood and studied from an operational perspective free of operator-algebra, and thus it is highly relevant to have a definition of self-testing compatible with that mode of abstraction.

In this paper, we propose a new formalisation of self-testing which resolves these issues. Our formalisation is purely operational, and thus immediately interpretable in theories beyond quantum theory and independent of a specific formalism for quantum theory. It moreover naturally hints at an operational notion of 'closeness' between state-measurement-configurations apt for the definition of approximate (i.e. *robust)* self-testing [MYS12], a concept which for practical purposes of self-testing is very important. Finally, the new approach points to interesting relaxations of self-testing which eventually exhibit self-testing as the special case of a phenomenon definable in a general framework of causally structured networks of channels — as such, it automatically generalises self-testing to arbitrary causal scenarios. As a benefit, our new definition additionally clarifies certain vague narratives about the conventional operator-algebraic definition (e.g. a commonly employed restriction to projective measurements) and yields new, general and intuitive proofs of many known features related to self-testing (e.g. the facts that any configuration which can be self-tested uses a pure state and must give rise to extremal measurement statistics).

All of these results are discussed in Section 1.B in more detail, along with the main strokes of our formalisation.

1.B Main Idea and Results

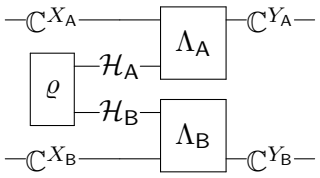
Consider two separated parties A and B in possession of Hilbert spaces $\mathcal{H}_\mathsf{A}$ and $\mathcal{H}_\mathsf{B}$, respectively, and a shared quantum state ϱ on $\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$. (Readers familiar with the CHSH-game can freely think of the case $\mathcal{H}_\mathsf{A} = \mathcal{H}_\mathsf{B} = \mathbb{C}^2$.) The two parties can locally apply measurements indexed by finite sets X_A and X_B, with outcomes in finite sets Y_A and Y_B, respectively. (Again, for concreteness, the reader may consider the case where $X_\mathsf{A} = X_\mathsf{A} = \{0, 1\}$ and $Y_\mathsf{A} = Y_\mathsf{B} = \{1, -1\}$, so that each party can choose between two measurements, each with two outcomes ± 1.)

For each index $x_\mathsf{A} \in X_\mathsf{A}$, the corresponding measurement for party A may be represented by a quantum channel $\Lambda_\mathsf{A}^{x_\mathsf{A}}$ from $\mathcal{H}_\mathsf{A}$ to $\mathbb{C}^{Y_\mathsf{A}}$, whose outputs are diagonal in the computational basis (for any finite set I, we denote by $\mathbb{C}^I$ the Hilbert space with computational basis $(|i\rangle)_{i\in I}$).[4] The measurement ensemble $(\Lambda_\mathsf{A}^{x_\mathsf{A}})_{x_\mathsf{A}\in X_\mathsf{A}}$ may in turn be described by a single quantum channel Λ_A from the system $\mathbb{C}^{X_\mathsf{A}} \otimes \mathcal{H}_\mathsf{A}$ to the system $\mathbb{C}^{Y_\mathsf{A}}$, which measures the system $\mathbb{C}^{X_\mathsf{A}}$ in the computational basis to read off the measurement choice $x_\mathsf{A} \in X_\mathsf{A}$ and then applies the corresponding measurement $\Lambda_\mathsf{A}^{x_\mathsf{A}}$. Likewise, the measurement ensemble for party B may be summarised by a single quantum channel Λ_B from $\mathbb{C}^{X_\mathsf{B}} \otimes \mathcal{H}_\mathsf{B}$ to $\mathbb{C}^{Y_\mathsf{B}}$. Altogether, the quantum channels Λ_A and Λ_B describe how classical information from the systems $\mathbb{C}^{X_\mathsf{A}}$ and $\mathbb{C}^{X_\mathsf{B}}$ along with quantum information from the systems $\mathcal{H}_\mathsf{A}$ and $\mathcal{H}_\mathsf{B}$ is transformed to classical information in the systems $\mathbb{C}^{Y_\mathsf{A}}$ and $\mathbb{C}^{Y_\mathsf{B}}$, respectively.

The quadruple $\mathfrak{B} = (X_\mathsf{A}, X_\mathsf{B}, Y_\mathsf{A}, Y_\mathsf{B})$ which defines the measurement indices and outcomes is called a *Bell-scenario*, and the triple $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ is called a *quantum strategy*

[4] As discussed in Section 2.B, the channel $\Lambda_\mathsf{A}^{x_\mathsf{A}}$ may be represented by a POVM on $\mathcal{H}_\mathsf{A}$, namely a collection $E_\mathsf{A}^{x_\mathsf{A}} = \big(E_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})\big)_{y_\mathsf{A}\in Y_\mathsf{A}}$ of positive operators on $\mathcal{H}_\mathsf{A}$ such that $\sum_{y_\mathsf{A}\in Y_\mathsf{A}} E_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A}) = \mathbb{1}_{\mathcal{H}_\mathsf{A}}$. Specifically, the measurement channel $\Lambda_\mathsf{A}^{x_\mathsf{A}}$ associated to the POVM $E_\mathsf{A}^{x_\mathsf{A}}$ is given by $\Lambda_\mathsf{A}^{x_\mathsf{A}}(T) = \sum_{y_\mathsf{A}\in Y_\mathsf{A}} \mathrm{tr}(E_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})T)\ |y_\mathsf{A}\rangle\langle y_\mathsf{A}|$ for $T \in \mathrm{End}(\mathcal{H}_\mathsf{A})$.

for $\mathfrak{B}$. To any quantum strategy S is an associated *(input-output) behaviour*, namely the channel from $\mathbb{C}^{X_\mathsf{A}} \otimes \mathbb{C}^{X_\mathsf{B}}$ to $\mathbb{C}^{Y_\mathsf{A}} \otimes \mathbb{C}^{Y_\mathsf{B}}$ given by $(\Lambda_\mathsf{A} \otimes \Lambda_\mathsf{B}) \circ (\mathrm{id}_{\mathbb{C}^{X_\mathsf{A}}} \otimes \varrho \otimes \mathrm{id}_{\mathbb{C}^{X_\mathsf{B}}})$, which takes as input any pair of measurement indices and yields as output the joint distribution of measurement outcomes. We may pictorially represent the behaviour channel as

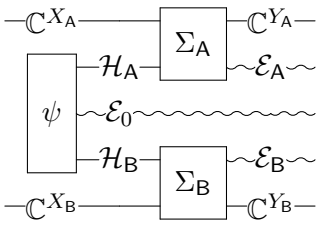


, (1)

with time flowing from left to right. This channel can be identified with a classical information channel from the set $X := X_\mathsf{A} \times X_\mathsf{B}$ to the set $Y := Y_\mathsf{A} \times Y_\mathsf{B}$ and as such the behaviour of a strategy is equivalent to a collection $P = (P^x)_{x \in X}$ of probability distributions on Y.[5] (In Section 1.A, such collections were referred to as 'measurement statistics' and the quantum strategies as 'configurations'.)

Now, the conventional formalisation of self-testing is roughly as follows. One says that the behaviour P *self-tests* the strategy $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_\mathsf{A}, \tilde{\Lambda}_\mathsf{B})$, if any other strategy $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ with behaviour P is 'reducible' to $\tilde{S}$, where *reducibility* is a certain relation defined among quantum strategies in terms of the constituents $\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B}, \tilde{\varrho}, \tilde{\Lambda}_\mathsf{A}$ and $\tilde{\Lambda}_\mathsf{B}$. The details of this definition are reviewed in Section 3, but for now it suffices to understand that the main issue addressed by our paper is that the definition is phrased using operator-algebraic objects which abstractly represent these constituents, rather than using operational features of the constituents themselves. For example, when the states ϱ and $\tilde{\varrho}$ are pure, represented by unit vectors $|\psi\rangle \in \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$ and $|\tilde{\psi}\rangle \in \tilde{\mathcal{H}}_\mathsf{A} \otimes \tilde{\mathcal{H}}_\mathsf{B}$, and when the measurements are all projective, represented by PVM-elements $\Pi_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})$, $\Pi_\mathsf{B}^{x_\mathsf{B}}(y_\mathsf{B})$, $\tilde{\Pi}_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})$ and $\tilde{\Pi}_\mathsf{B}^{x_\mathsf{B}}(y_\mathsf{B})$, then the reducibility condition requires the existence of isometries W_A, W_B and a unit vector $|\psi^{\mathrm{res}}\rangle$ such that $[W_\mathsf{A} \otimes W_\mathsf{B}][\Pi_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A}) \otimes \Pi_\mathsf{B}^{x_\mathsf{B}}(y_\mathsf{B})]\,|\psi\rangle = [\tilde{\Pi}_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A}) \otimes \tilde{\Pi}_\mathsf{B}^{x_\mathsf{B}}(y_\mathsf{B})]\,|\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle$ for all $x_\mathsf{A} \in X_\mathsf{A}$, $x_\mathsf{B} \in X_\mathsf{B}$, $y_\mathsf{A} \in Y_\mathsf{A}$ and $y_\mathsf{B} \in Y_\mathsf{B}$.

What we propose to do instead is the following. To each strategy $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$, we may associate the quantum channel



, (2)

algebraically given by $(\Sigma_\mathsf{A} \otimes \mathrm{id}_{\mathcal{E}_0} \otimes \Sigma_\mathsf{B}) \circ (\mathrm{id}_{\widehat{X}_\mathsf{A}} \otimes \psi \otimes \mathrm{id}_{\widehat{X}_\mathsf{B}})$, from $\mathbb{C}^{X_\mathsf{A}} \otimes \mathbb{C}^{X_\mathsf{B}}$ to $\mathbb{C}^{Y_\mathsf{A}} \otimes \mathbb{C}^{Y_\mathsf{B}} \otimes \mathcal{E}_\mathsf{A} \otimes \mathcal{E}_\mathsf{B} \otimes \mathcal{E}_0$, where Σ_A and Σ_B are Stinespring dilations of Λ_A and Λ_B, respectively, and where ψ is a Stinespring dilation (i.e. purification) of ϱ. (We review Stinespring dilations in Section 2.D; forming a Stinespring dilation is sometimes called 'going to the Church of the Larger Hilbert Space'.) The channel (2) will be called the *Stinespring implementation of* S.[6]

[5] This collection is sometimes referred to as the '*correlations*' produced by the strategy. The term 'behaviour', which we shall use throughout, is due to Cirelson [Cir93].

[6] Strictly speaking, the association of Stinespring implementations to strategies is not unique, since Stinespring dilations are not themselves unique; however, it is unique up to a triple of channels acting on the systems $\mathcal{E}_\mathsf{A}$, $\mathcal{E}_0$ and $\mathcal{E}_\mathsf{B}$, cf. Theorem 2.19. (If we were to take each Stinespring dilation minimal, the uniqueness would be even up to a triple of unitary channels acting on $\mathcal{E}_\mathsf{A}$, $\mathcal{E}_0$ and $\mathcal{E}_\mathsf{B}$.)

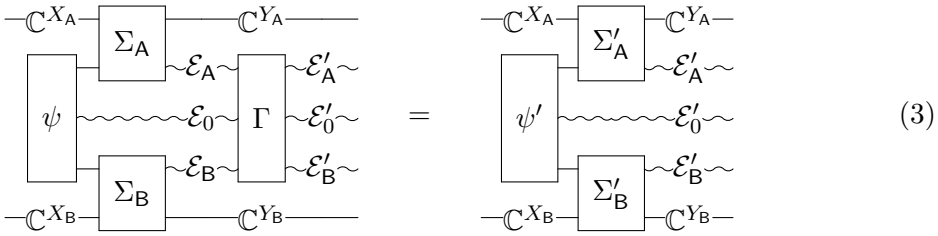
Now, the association of Stinespring implementations to quantum strategies does not look terribly operational, as Stinespring dilations are specific to quantum theory. However, for the purpose of redefining self-testing we shall only rely on the following operational feature of a Stinespring implementation: it represents *maximal side-information* leaking to an environment in the course of employing the strategy S, and as such is strongest among more general 'implementations' of the strategy (this will be made precise in Definition 4.2 and Lemma 4.11).

Specifically, the operational interpretation of a Stinespring implementation is that the system $\mathcal{E}_0$ holds the maximal amount of information extractable about the strategy S before knowing the inputs x_A and x_B, whereas for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the system $\mathcal{E}_\mathsf{P}$ holds the maximal information extractable after receiving the input x_P. For example, a copy of the input x_P will be derivable from the side-information in $\mathcal{E}_\mathsf{P}$ (though the information content is generally larger, as is it a Stinespring dilation and thus conveys also quantum information), and if the state ϱ is mixed then some classical non-trivial side-information will be available already in the system $\mathcal{E}_0$ (but again, its information content is generally larger).

The Stinespring implementation (2) of S is evidently a Stinespring dilation of the behaviour channel (1) when considering the system $\mathcal{E}_\mathsf{A} \otimes \mathcal{E}_0 \otimes \mathcal{E}_\mathsf{B}$ as one single environment; but the whole point will be that the concept of self-testing emerges by virtue of the difference between the three systems $\mathcal{E}_\mathsf{A}$, $\mathcal{E}_\mathsf{B}$ and $\mathcal{E}_0$ which is due to their causal relationships to the inputs. Indeed, the key insight in our paper is that the Stinespring implementations of strategies can be used to recast the conventional operator-algebraic notion of reducibility among strategies, and thus of self-testing, in a way which is purely operational.

In fact, we will introduce three different relations among strategies, all defined in terms of their Stinespring implementations (but definable in arbitrary theories), and all of which illuminate in new ways the conventional understanding of self-testing. We describe these three relations below, along with our results.

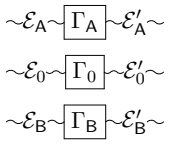
Let $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ and $S' = (\varrho', \Lambda'_\mathsf{A}, \Lambda'_\mathsf{B})$ be quantum strategies, and let ψ, Σ_A, Σ_B, ψ', Σ'_A and Σ'_B be Stinespring dilations of their respective constituents. Consider the condition



(3)

among the corresponding Stinespring implementations, for some channel Γ from $\mathcal{E}_\mathsf{A} \otimes \mathcal{E}_0 \otimes \mathcal{E}_\mathsf{B}$ to $\mathcal{E}'_\mathsf{A} \otimes \mathcal{E}'_0 \otimes \mathcal{E}'_\mathsf{B}$. (Here, we have dropped the labelling of systems $\mathcal{H}_\mathsf{A}$, $\mathcal{H}_\mathsf{B}$ and $\mathcal{H}'_\mathsf{A}$, $\mathcal{H}'_\mathsf{B}$ corresponding to the internal wires on the left and right hand side, respectively.) A channel Γ satisfying Eq. (3) exists if and only if S and S' have the same behaviour, indeed the identity expresses simply that two Stinespring dilations of the behaviour channel are equivalent (Theorem 2.19). If however we restrict the form that Γ may take in Eq. (3), then we obtain finer, non-trivial relations among strategies. In particular, we will introduce the following three operational relations and derive a number of results illuminating self-testing in terms of them.

- *S locally simulates S'* (we write $S \geq_{l.s.} S'$) if there exists Γ of the form

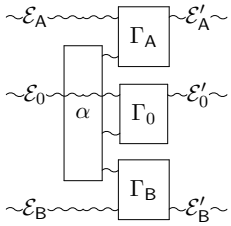


satisfying Eq. (3).

As we show, any quantum strategy S is equivalent under local simulation to a strategy $S_{\text{proj.}}$ with projective measurements (Proposition 4.13). In this formal sense, measurements may always without loss of generality be assumed projective. Furthermore, under a simple rank-assumption on the state of $\tilde{S}$, the operator-algebraic notion of S being reducible to $\tilde{S}$ is equivalent to the operational notion of S locally simulating $\tilde{S}$ (Theorem 4.8). This is our key result, and the only one with a technical proof.

Local simulation gives rise to a natural notion self-testing (Definition 4.7): We say that P *self-tests* $\tilde{S}$ *according to local simulation* (for short, *a.t.l.s.*) if any strategy S with behaviour P locally simulates $\tilde{S}$, i.e. $S \geq_{l.s.} \tilde{S}$. This notion allows to recast the conventional operator-algebraic formalisation of self-testing in operational terms (Theorem 4.9), giving a contribution to the understanding of self-testing.

- S *locally assisted simulates* S' (written $S \geq_{l.a.s.} S'$) if there exists Γ of the form

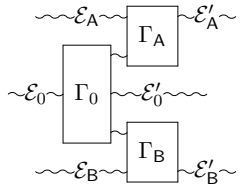


satisfying Eq. (3);

Local assisted simulation is coarser than local simulation and, somewhat surprisingly, symmetric and therefore an equivalence relation among quantum strategies (Theorem 4.23). In fact, local assisted simulation is the equivalence relation *generated by* local simulation (Remark 4.24). As a consequence of this, if a behaviour P self-tests a strategy $\tilde{S}$ in the conventional sense, then all strategies with behaviour P are equivalent in a formal sense, namely according to local assisted simulation. (For example, all optimal strategies for the CHSH-game are equivalent under local assisted simulation.)

One may define a new notion of *self-testing according to local assisted simulation* (for short, *self-testing a.t.l.a.s.*), which requires $S \leq_{l.a.s.} \tilde{S}$ for every strategy S with behaviour P. By the above results, the condition $S \leq_{l.a.s.} \tilde{S}$ weakens the relation $S \geq_{l.s.} \tilde{S}$ (since the latter implies $S \geq_{l.a.s.} \tilde{S}$ as well as $S \leq_{l.a.s.} \tilde{S}$). Therefore, this notion of self-testing relaxes the conventional notion of self-testing. It puts emphasis on upper bounding the abilities of unknown strategies in terms of the known strategy $\tilde{S}$ (in line with Ekert [Eke91]) rather than lower bounding the abilities of unknown strategies (in line with Bell [Bel64]).

- S *causally simulates* S' (written $S \geq_{c.s.} S'$) if there exists Γ of the form



satisfying Eq. (3).

This notion of simulation is again coarser than the previous one. And like local assisted simulation, causal simulation gives rise to a self-testing notion too, by requiring $S \leq_{c.s.} \tilde{S}$ for every strategy S with behaviour P. We call this *self-testing according*

to causal simulation (for short, *self-testing a.t.c.s.*).[7] The significance of self-testing a.t.c.s. is that it realises self-testing as part of a very general theory of *causal channels* and *causal dilations* introduced in the PhD thesis [HL21]. We discuss this connection in Section 4.C.

We emphasize that all three relations turn out be operationally definable; this follows from the operational 'maximal leak'-feature of Stinespring implementations, which we will make explicit in alternative characterisations of the relations (see Definition 4.2, Remark 4.20 and Remark 4.29). In particular, it will be clear how to interpret the three modes of simulation in any other theory of information channels which are serially and parallelly composable [HL21].

To the best of our knowledge, the operational account of general self-testing presented in this paper is the first of its kind to appear in the literature. It should be noted, however, that Ref. [SFK+20] has discussed operational formulations of self-testing of states. Specifically, the authors introduce a notion of extractability in terms of local operations and shared classical randomness, and they render a state ϱ self-tested by a behaviour B, if B can be extracted from ϱ and if ϱ is extractible from any other state σ from which B can be extracted. Though also based on an operational idea of extraction, this approach differs from ours in two important respects. Firstly, it differs in complexity by disregarding the measurements in a self-testing scenario; in our approach, the appearance of measurements is manifested in the paramount causal difference between outputs, which we discussed above (the inclusion of measurements is also what makes the equivalence to standard self-testing non-trivial to establish, cf. Lemma 4.17). Secondly, the two approaches differ in concept, since we consider dilations of the involved state (and measurements). As a consequence, certain mixed states are self-testable according to the definition in Ref. [SFK+20], whereas, according to ours, the inclusion of dilations ultimately turns out to demand purity of self-testable states (cf. Corollary 5.5).

After completion of the research in this work, we have been made aware of Ref. [RB22], where notions of self-testing are discussed in classical causal network settings. It is indeed also possible to consider our operational definition of self-testing in non-quantum theories, for example the theory of classical information (we call this theory **CIT**). We believe it would be interesting to conduct a study of self-testing in other theories, this was, however, out of scope for the work presented here. We refer the interested reader to Ref. [HL21].

1.C Outline of the Paper

- ▷ We shall start by making more precise what we mean by *operational concepts* in a physical theory such as quantum theory. This will be the topic of Section 2, which also serves as preliminary section to establish common notions, terminology and results from quantum information theory. We discuss in particular *compositional structure* of a theory (Section 2.A), the interplay with *classical notions* such as probabilities (Section 2.B), and we recall Naimark's theorem for measurements (Section 2.C) and Stinespring's dilation theorem (Section 2.D). Readers who are familiar with these results and already feel comfortable with the notion of operational structure may skip Section 2, referring to it if need be.

- ▷ In Section 3, we discuss the conventional perception of self-testing (Section 3.A)

[7] When the state of $\tilde{S}$ is pure, self-testing a.t.c.s. coincides with self-testing a.t.l.a.s. (since in that case $S \leq_{c.s.} \tilde{S}$ is equivalent to $S \leq_{l.a.s.} \tilde{S}$, by inspecting the structure of the channel Γ).

and recall in detail the standard operator-algebraic definition of reducibility and self-testing (Section 3.B). We then discuss the need for an alternative understanding of self-testing (Section 3.C), which is operational in the sense of Section 2.

- ▷ Our new operational approach to self-testing is presented in Section 4, which is the core section of the paper. We introduce here *local simulation* and the corresponding reformulation of self-testing, and we establish the equivalence to reducibility and the conventional notion of self-testing (Section 4.A). We also introduce the relaxation to *local assisted simulation* and discuss how it illuminates quantum self-testing in new ways (Section 4.B). Finally, we explain how the further relaxation to *causal simulation* allows us to realise quantum self-testing from an even larger perspective, namely within the framework of causal channels and dilations (Section 4.C).

- ▷ In Section 5, we end by showing how several features of self-testing can be given new, operational proofs based on our reformulations of self-testing.[8] In particular, we establish a relationship between local simulation and state extractability (Section 5.A), between local assisted simulation and measurement extractability (Section 5.B), and between causal simulation and the structure of convex decompositions of the behaviour (Section 5.C). We also observe an operational feature of self-testing which we believe has so far gone unnoticed, but which from our operational formulation is natural (Section 5.D). It suggests that self-testing is in certain cases characterised by the property of 'exhausting' the parties A and B of information.

Each of the above four sections begins with a more detailed section outline. In the concluding Section 6, we summarise the paper and discuss possible directions for future investigations.

2 The Operational Structure of Quantum Information Theory

This section has two purposes. Partly, it serves as a preliminary reminder of quantum information theory as cast in the mathematical language of completely positive trace-preserving maps (quantum channels). Partly, the style of presentation serves to exemplify what we mean by the *operational structure* of a physical theory, in line with a 'pragmatist' tradition of physics [CFS16]. We define a concept to be *operational* if it can be formulated by referring only to operational structure. Our definition of this (Definition 2.12) is not fully formalised, but we render it safe to adopt the viewpoint that it is evident when a concept has been cast as operational according to this definition (whereas it may <u>not</u> be evident when this is impossible to do). The section is divided into four subsections.

Roughly speaking, operational structure of a theory is comprised by (A) a purely *compositional structure* (as pioneered in categorical frameworks [AC04, Sel04, Bae06]), and additionally often (B) a connection to *classical notions*, e.g. probabilities (as usually formalised using Generalised Probabilistic Theories [Bar07, BW11, BW16]). These two pillars are discussed in Section 2.A and Section 2.B, respectively, with a special focus on quantum information theory. The compositional structure of a theory already sustains a notion of *compositional concepts* (informally captured in Definition 2.6), and by adding classical structure we then obtain the full notion of *operational concepts* (Definition 2.12). Readers who already feel intuitively comfortable with the notion of 'operational concepts' may skip Section 2.A and Section 2.B, and refer to these subsections later if necessary.

[8] Section 5 may, apart for the definition of causal simulation (Definition 4.28), be read independently of Section 4.C.

In Section 2.C we recall Naimark's theorem for quantum measurement, and, finally, in Section 2.D we make some observations about *isometric* quantum channels (Example 2.1), reviewing in particular Stinespring's dilation theorem [Sti55] from an operational perspective.

2.A The Compositional Structure of **QIT**

The compositional structure of a physical theory comprises a collection of *systems* and *transformations* and defines how these elements can be *composed*. Systems are interpreted as the basic information-carrying entities in the theory, and any two systems can be composed to form a single, joint system. Transformations from one system to another are interpreted as 'physical processes' or 'information channels' which modify the information carried by the systems, and transformations can be composed either *serially* (one performed after another) or *parallelly* (one performed alongside with another).

Mathematically, this structure is succinctly captured by a *symmetric monoidal category* [ML13], however it suffices for our purposes to intuitively understand this concept insofar as it concerns the theory of quantum information.

In quantum information theory, a system is identified with a non-zero finite-dimensional Hilbert space.[9] We generically typeset systems with capital calligraphic letters $\mathcal{H}, \mathcal{K}, \mathcal{L}, \ldots$. A transformation between systems is in quantum information theory identified with a so-called *quantum channel*. For any Hilbert space $\mathcal{L}$, we denote by $\mathrm{End}(\mathcal{L})$ the vector space of all linear operators on $\mathcal{L}$; a quantum channel from system $\mathcal{H}$ to system $\mathcal{K}$ is then by definition a linear map $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ which is *trace-preserving* (meaning that $\mathrm{tr}(\Lambda(A)) = \mathrm{tr}(A)$ for any $A \in \mathrm{End}(\mathcal{H})$) and *completely positive* (meaning that for any system $\mathcal{R}$ and any positive operator $A \in \mathrm{End}(\mathcal{H} \otimes \mathcal{R})$, the operator $(\Lambda \otimes \mathrm{id}_{\mathrm{End}(\mathcal{R})})(A) \in \mathrm{End}(\mathcal{K} \otimes \mathcal{R})$ is positive[10]). We generically typeset quantum channels with Greek letters $\Lambda, \Phi, \Sigma, \ldots$.

Transformations $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ and $\Phi : \mathrm{End}(\mathcal{K}) \to \mathrm{End}(\mathcal{L})$ can be composed serially by means of functional composition to yield the transformation $\Phi \circ \Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{L})$. The serial composition is associative (meaning that $\Xi \circ (\Phi \circ \Lambda) = (\Xi \circ \Phi) \circ \Lambda$), and there exists for every system $\mathcal{H}$ a distinguished identity transformation which acts as unit for the serial composition, namely the identity channel $\mathrm{id}_{\mathrm{End}(\mathcal{H})} : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H})$, $A \mapsto A$, which for brevity we denote $\mathrm{id}_{\mathcal{H}}$. These features of systems and transformations so far gives quantum information theory the mathematical structure of a *category* [ML13], namely a collection of morphisms which compose serially subject to natural requirements.

Example 2.1 (Isometric Quantum Channels)**.** If $V : \mathcal{H} \to \mathcal{K}$ is an *isometry*, i.e. a linear operator with $V^*V = \mathbb{1}_{\mathcal{H}}$, then the linear map $\Sigma_V : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ given by $\Sigma_V(A) = VAV^*$ is trace-preserving and completely positive, hence a quantum channel from $\mathcal{H}$ to $\mathcal{K}$. We call quantum channels of this form *isometric*. The identity transformation $\mathrm{id}_{\mathcal{L}}$ on a system $\mathcal{L}$ is isometric, and the serial composition of isometric channels is again isometric. ⧫

Quantum information theory has more structure than that of a category. Indeed, systems $\mathcal{H}_1$ and $\mathcal{H}_2$ can be composed by means of the tensor product to form the joint system

[9] One can define a version of quantum information theory in which infinite-dimensional systems are allowed, but the definition of quantum channels is then slightly more technical and we shall not need this.

[10] Note that $\Lambda \otimes \mathrm{id}_{\mathrm{End}(\mathcal{R})}$ can be thought of as a map $\mathrm{End}(\mathcal{H} \otimes \mathcal{R}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{R})$ by observing the canonical isomorphisms $\mathrm{End}(\mathcal{H} \otimes \mathcal{R}) \cong \mathrm{End}(\mathcal{H}) \otimes \mathrm{End}(\mathcal{R})$ and $\mathrm{End}(\mathcal{K} \otimes \mathcal{R}) \cong \mathrm{End}(\mathcal{K}) \otimes \mathrm{End}(\mathcal{R})$.

$\mathcal{H}_1 \otimes \mathcal{H}_2$, and transformations $\Lambda_1 : \mathrm{End}(\mathcal{H}_1) \to \mathrm{End}(\mathcal{K}_1)$ and $\Lambda_2 : \mathrm{End}(\mathcal{H}_2) \to \mathrm{End}(\mathcal{K}_2)$ can be composed parallelly by means of the tensor product to yield the transformation $\Lambda_1 \otimes \Lambda_2 : \mathrm{End}(\mathcal{H}_1 \otimes \mathcal{H}_2) \to \mathrm{End}(\mathcal{K}_1 \otimes \mathcal{K}_2)$. (Complete positivity ensures that the resulting map is a quantum channel.) The parallel composition of transformations is tied to the composition of system via the identities $\mathrm{id}_{\mathcal{H}_1 \otimes \mathcal{H}_2} = \mathrm{id}_{\mathcal{H}_1} \otimes \mathrm{id}_{\mathcal{H}_2}$. Parallel composition is commutative and associative among transformations as well and systems,[11] and there is a distinguished *trivial system*, namely the one-dimensional Hilbert space $\mathbf{1} := \mathbb{C}$, which physically represents 'nothing'. The trivial system acts as a unit for the composition of systems ($\mathcal{H} \otimes \mathbf{1} \cong \mathcal{H}$) and its identity transformation $\mathrm{id}_{\mathbf{1}}$ acts as a unit for the parallel composition. All these additional features of quantum theory promotes its mathematical structure to that of a *symmetric monoidal category* [ML13], namely a category augmented by a notion of parallel composition (the word 'symmetric' refers to the commutativity of parallel composition).

In our definition of compositional structure, there is one more feature of quantum information theory we wish to impose generally:[12] the trivial system $\mathbf{1}$ is such that from any system $\mathcal{H}$ there is a unique transformation to $\mathbf{1}$, namely the trace $\mathrm{tr}_{\mathcal{H}} : \mathrm{End}(\mathcal{H}) \to \mathbb{C}(\cong \mathrm{End}(\mathbf{1}))$. As the system $\mathbf{1}$ represents 'nothing', this physically corresponds to saying that there is precisely one way of discarding information in a system $\mathcal{H}$. This is important to us because it facilities a unique notion of *marginalisation* by way of discarding only part of a system (e.g. from $\mathcal{H}_1 \otimes \mathcal{H}_2$ to $\mathcal{H}_1$ we have the transformation $\mathrm{id}_{\mathcal{H}_1} \otimes \mathrm{tr}_{\mathcal{H}_2}$, called *partial trace* in quantum information theory). We shall rely heavily on marginalisation later on.

As such, we arrive at the following model for compositional structure:

Definition 2.2 (Compositional Theories)**.** A *compositional theory* $\mathbf{\Theta}$ is a symmetric monoidal category in which every system $\mathcal{H}$ admits precisely one transformation to the trivial system $\mathbf{1}$, denoted $\mathrm{tr}_{\mathcal{H}}$. We denote the serial composition in $\mathbf{\Theta}$ by $\circ$ and the parallel composition and composition of systems in $\mathbf{\Theta}$ by $\otimes$. ■

Remark 2.3. Symmetric monoidal categories form a well-established model for parallel and serial composition in physical theories [AC04, Sel04, Bae06], and the uniqueness requirement for the trivial system can be seen as formalising the impossibility of signalling from the future to the past [CDP10, CL13, Coe14]. Symmetric monoidal categories in which every system admits a unique transformation to $\mathbf{1}$ are sometimes called *semicartesian* symmetric monoidal categories. ✠

Example 2.4. We will denote by **QIT** the compositional theory of quantum information, which we defined above. ⧫

Example 2.5. We will denote by **CIT** the compositional theory of classical information, where systems are finite non-empty sets composing under the cartesian product $\times$, and where transformations from X to Y are classical channels (Markov kernels) from X to Y, i.e. families $T = (t^x)_{x \in X}$ of probability distributions t^x on Y. The serial and parallel composition of transformations, which we denote $\circ$ and $\times$, respectively, are defined in the obvious way (see e.g. [HL21] Example 1.1.9). This theory is also known as **FinStoch**, cf. Ref. [Fri20]. ⧫

[11] Or rather, commutative and associative 'up to natural isomorphisms'; for systems we have $\mathcal{H}_1 \otimes \mathcal{H}_2 \cong \mathcal{H}_2 \otimes \mathcal{H}_1$ and $(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3 \cong \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$, and for transformations similar conditions.

[12] Strictly speaking, this feature is not an additional piece of structure but rather an additional axiom about the existing structure.

Though we defined **QIT** using the language of Hilbert spaces and operators, it might be definable (up to a suitable notion of isomorphism) in ways which do not reference an operator-algebraic formalism. Indeed, it has been shown [Har01, CDP11] that there does exist such alternative definitions of **QIT**, and in this sense the operator-algebraic formalism for quantum theory should be considered a choice of 'coordinate system' for the theory, rather than identified with the theory itself.

This is true more generally for any compositional theory: Even when a priori presented using specific mathematical objects, the theory may admit alternative presentations which do not reference those objects. This circumstance prompts the idea that some concepts in a compositional theory might be definable *intrinsically*, using the compositional structure alone. Such concepts will necessarily be meaningful across different representations of the same theory, and also necessarily interpretable in every compositional theory:

Definition 2.6 (Compositional Concepts, Informally)**.** In a compositional theory, a *compositional concept* is one that can be defined by reference only to the composition of systems $\otimes$, serial and parallel composition of transformations $\circ$ and $\otimes$, the distinguished trivial system $\mathbf{1}$, and the distinguished identity and discarding transformations $\mathrm{id}_{\mathcal{H}}$ and $\mathrm{tr}_{\mathcal{H}}$. ■

For readers versed in mathematical logic, it is possible to make the definition of 'compositional concepts' more precise by introducing an appropriate first-order language apt for symmetric monoidal categories (in particular with binary symbols $\circ$, $\otimes$ and constant symbol $\mathbf{1}$). A compositional concept is then formally a predicate in this language.[13]

Rather than being explicit about such a formal language, we find it suitable to here provide instead a selection of examples which should indicate what it means to be a compositional concept.

Example 2.7 (Reversibles)**.** We call a transformation Λ from $\mathcal{H}$ to $\mathcal{K}$ *reversible* if it has a left-inverse, i.e. if there exists a transformation Λ^- from $\mathcal{K}$ to $\mathcal{H}$ such that $\Lambda^- \circ \Lambda = \mathrm{id}_{\mathcal{H}}$. Being a reversible transformation is a compositional concept.[14] ⧫

Example 2.8 (Isomorphisms and Automorphisms)**.** An *isomorphism*, i.e. a transformation Ξ from $\mathcal{H}$ to $\mathcal{K}$ for which there exists a two-sided inverse, is a compositional concept. So is an *automorphism*, i.e. an isomorphism from a system $\mathcal{H}$ to itself. ⧫

Example 2.9 (Dimensionality)**.** The systems in **QIT** are vector spaces and thus admit a notion of one system having larger dimension than another. The dimension of a vector space is however defined in terms of maximal sets of linearly independent vectors, and this concept is not evidently compositional as it references mathematical structure internal to the systems. Nevertheless, the concept that $\dim \mathcal{H} \leq \dim \mathcal{K}$ can be compositionally

[13] Let us however point out that such a formal language should have a built-in way of distinguishing e.g. in **QIT** between the systems $\mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathbb{C}^4$, which are isomorphic and thus a priori compositionally indistinguishable (see Footnote 16). This can be achieved for example by giving systems unique names and only allowing systems with distinct names to be composed. Moreover, we might wish to include elements of sets theory into the formal language, if for instance we wish to define certain concepts as compositional which involve infinite families of transformations.

[14] Formally, reversibility would be a predicate with one free variable, Λ, namely the predicate $Rev(\Lambda)$ given by

$$\exists\, \Lambda^- : \big(\mathrm{Dom}(\Lambda^-) = \mathrm{Cod}(\Lambda)\big) \wedge \big(\mathrm{Cod}(\Lambda^-) = \mathrm{Dom}(\Lambda)\big) \wedge \big(\Lambda^- \circ \Lambda = \mathrm{id}_{\mathrm{Dom}(\Lambda)}\big)$$

Here, 'Dom' and 'Cod' are unary function symbols supposed to assign domain and codomain, respectively, to a transformation.

defined: this condition holds precisely if there exists a reversible transformation from $\mathcal{H}$ to $\mathcal{K}$, and reversibility is compositional by Example 2.7.[15]

It follows that also e.g. the notion of a quantum system $\mathcal{H}$ being 2-dimensional can be compositionally defined, indeed it means that $\dim \mathbf{1} \leq \dim \mathcal{H}$, that $\dim \mathcal{H} \not\leq \dim \mathbf{1}$, and that $\dim \mathcal{H} \leq \dim \mathcal{H}'$ for any other system $\mathcal{H}'$ with those two properties. By induction, any concrete finite dimensionality can be defined as a compositional concept. (This is not to say that every compositional theory will actually exhibit systems of every concrete dimension.) ⧫

Example 2.10 (States)**.** The notion of a *state* is compositional, since states are definable as transformations from the trivial system $\mathbf{1}$. Specifically, in **QIT**, a state on the system $\mathcal{H}$ is usually defined as a *density operator* $\varrho \in \mathrm{End}(\mathcal{H})$ (i.e. a positive operator with unit trace), but density operators ϱ can be identified with quantum channels from $\mathbf{1}$ to $\mathcal{H}$ since these are precisely the linear maps $\mathbb{C}(\cong \mathrm{End}(\mathbf{1})) \to \mathrm{End}(\mathcal{H})$ given by $z \mapsto z\varrho$. (By abuse of notation, we will always identify states in **QIT** with density operators.) In the compositional theory **CIT**, this notion of state also translates to the intended notion, indeed a transformation in **CIT** from the one-element set $\mathbf{1}$ to a general set X corresponds to a probability distribution on X.

Provided we have a suitable naming of systems (see Footnote 13), the concept of *product states* $\varrho_1 \otimes \varrho_2$ on a system $\mathcal{H}_1 \otimes \mathcal{H}_2$ is also compositional. In the theory **CIT**, it translates to stochastic independence. ⧫

A quantum state is called *pure* if its density operator is a rank-one projection, i.e. of the form $|\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$. Pure states are in fact isometric channels (since an isometry $V : \mathbf{1} \to \mathcal{K}$ is essentially a unit vector $|\psi\rangle \in \mathcal{K}$). In Section 2.D we shall see that the concept of being an isometric channel (Example 2.1) — which is a priori defined using operator-algebra — is in fact also a compositional concept, albeit in disguise (Remark 2.25). In particular the concept of pure states is compositional.

Now, the intuitive understanding of what it means for a concept to be operational subsumes all concepts which are compositional. There are however concepts which are operational in the conventional, intuitive sense and which seem to fall outside the scope of compositional structure.

Two such examples are the concepts of *probability* and *measurement* (another example is that of *separable* states, i.e. convex combinations of product states). Formally, given a set Y, it is natural to define in **QIT** a 'probability distribution on Y' as a quantum state on the system $\mathbb{C}^Y$ which is diagonal in the computational basis, i.e. is of the form $\sum_{y \in Y} p(y)\, |y\rangle\langle y|$ for some function $p : Y \to [0, 1]$. Likewise, a 'measurement on $\mathcal{H}$ with outcomes in Y' may be defined as a quantum channel $M : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathbb{C}^Y)$ such that for every state ϱ on $\mathcal{H}$, the state $M(\varrho)$ is a probability distribution on Y. (As such, being a probability is actually a special case of being a measurement.) The concepts of 'probability distribution' and 'measurement' are instrumental to the operational structure of quantum theory, but they are not compositional in our sense.[16] Even disregarding this circumstance, it is unclear that the concept of being a specific probability distribution

[15] Formally, this concept is a predicate with two free variables, namely the predicate $Lar(\mathcal{H}, \mathcal{K})$ given by

$$\exists\, \Lambda : Rev(\Lambda) \wedge \big(\mathrm{Dom}(\Lambda) = \mathcal{H}\big) \wedge \big(\mathrm{Cod}(\Lambda) = \mathcal{K}\big)$$

where $Rev(\Lambda)$ is the predicate from before expressing reversibility of Λ.

[16] This can be shown rigorously by formalising the idea that every compositional concept is 'invariant under conjugation by automorphisms', i.e. if we pick an automorphism $\alpha_{\mathcal{H}}$ of each system $\mathcal{H}$, in a coherent

would be compositional, e.g. 'the distribution on $\{0,1\}$ which assigns probability $1/3$ to 0'. This concept is however also operational in the conventional sense, since we may test it (at least to arbitrary precision) by gathering statistics from a large number of independent experiments.

For these reasons, we need to add elements to a compositional theory to capture its full operational structure.

2.B Classical Notions in **QIT**

In order to capture operational concepts additional to the compositional ones, a 'classical structure' should be added to a compositional theory. This structure should describe how to interpret certain information as classical (e.g. measurement outcomes), and how to assign probabilities to such information. This has been formalised mathematically in the notion of Generalised Probabilistic Theories [Bar07, BW11, BW16] and more explicitly merged with category-based compositional structure in Refs. [CDP10, Har13]. Below, we give an alternative and for our purposes simpler treatment which works by virtue of a natural *embedding* of **CIT** into the target theory, in our case **QIT**.[17]

Recall from Example 2.5 that **CIT** is the compositional theory of classical information, whose systems are non-empty finite sets and whose transformations from X to Y are classical channels, i.e. families $T = (t^x)_{x\in X}$ of probability distributions t^x on Y. We now describe how the theory **CIT** embeds into the theory **QIT**. This embedding can then be used to define classical notions within **QIT**.

Any system X in **CIT** yields a system $\widehat{X} := \mathbb{C}^X$ in **QIT**, which is the quantum representation of the classical system X. We will call a quantum system of the form $\widehat{X}$ an *embedded classical system*. (It is not the intention that embedded classical systems can only hold states which are diagonal in the computational basis; being embedded classical simply means being equipped with a notion of which states are considered classical.) Note that this embedding preserves the composition of systems, since $\widehat{X_1 \times X_2} \cong \widehat{X}_1 \otimes \widehat{X}_2$, and that the trivial system in **CIT** embeds as the trivial system in **QIT**.

Any transformation $T = (t^x)_{x\in X}$ from X to Y in **CIT** yields a transformation $\widehat{T}$ from $\widehat{X}$ to $\widehat{Y}$ in **QIT**, namely the quantum channel given by $\widehat{T}(A) = \sum_{x\in X, y\in Y} \langle x|\, A\, |x\rangle\, t^x(y)\, |y\rangle\langle y|$. We will call a quantum channel of the form $\widehat{T}$ an *embedded classical channel*. In particular, a classical state is a state of the form $\sum_{y\in Y} p(y)\, |y\rangle\langle y|$ for a probability distribution $p : Y \to [0,1]$. Again, note that the embedding preserves the serial and parallel composition of transformations, since $\widehat{S \circ T} = \widehat{S} \circ \widehat{T}$ and $\widehat{T_1 \times T_2} = \widehat{T}_1 \otimes \widehat{T}_2$.

All in all, the above embedding of systems and transformations yields a faithful picture of the compositional theory **CIT** inside the compositional theory **QIT**.

The embedding of **CIT** into **QIT** does <u>not</u> preserve identities, i.e. the channel $\Delta_X := \widehat{\mathrm{id}_X}$ is distinct from the channel $\mathrm{id}_{\widehat{X}}$.[18] Explicitly, the channel is given by $\Delta_X(A) =$

way such that $\alpha_{\mathcal{H}_1\otimes\mathcal{H}_2} = \alpha_{\mathcal{H}_1} \otimes \alpha_{\mathcal{H}_2}$, then for any predicate $C(\Lambda_1, \ldots, \Lambda_n)$ with transformations $\Lambda_1, \ldots, \Lambda_n$ as free variables, the sentence $\forall \Lambda_1, \ldots, \Lambda_n : \big(C(\Lambda_1, \ldots, \Lambda_n) \Leftrightarrow C(\alpha_{\mathrm{Cod}(\Lambda_1)} \circ \Lambda_1 \circ \alpha^{-1}_{\mathrm{Dom}(\Lambda_1)}, \ldots, \alpha_{\mathrm{Cod}(\Lambda_n)} \circ \Lambda_n \circ \alpha^{-1}_{\mathrm{Dom}(\Lambda_n)})\big)$ is provable. On the other hand, the above defined concept of measurement is not invariant when an automorphism is applied to the codomain $\mathbb{C}^Y$.

[17] This idea was mentioned in Ref. [HL21], see Remarks 1.1.12 and 1.1.13 and Section 1.4 therein.

[18] As observed in [HL21] Remark 1.1.12, this is not a peculiarity of the particular embedding chosen — there simply does not exist any identity-preserving embedding (formally: injective strong monoidal functor)

$\sum_{x\in X} \langle x| A |x\rangle |x\rangle\langle x|$ and we will call it the *dephasing channel on* $\widehat{X}$. It is easy to verify that an arbitrary quantum channel Λ from $\widehat{X}$ to $\widehat{Y}$ is embedded classical (i.e. of the form $\widehat{T}$) if and only if $\Delta_Y \circ \Lambda \circ \Delta_X = \Lambda$. As such, the dephasing channels can be used to identify **CIT** as a sub-theory within **QIT**.

Altogether, the above ideas may be abstracted from the case of quantum information theory to yield a general notion of operational structure, just as we did for compositional structure.

Given compositional theories $\mathbf{\Theta}_0$ and $\mathbf{\Theta}$, an *embedding* of $\mathbf{\Theta}_0$ into $\mathbf{\Theta}$ means a map $\widehat{\ } : \mathbf{\Theta}_0 \to \mathbf{\Theta}$ which injectively takes systems to systems and transformations to transformations, and which preserves the trivial system, composition of systems, and serial and parallel composition of transformations. Let us moreover say that $\widehat{\ }$ *identifies* $\mathbf{\Theta}_0$ *within* $\mathbf{\Theta}$ if for any embedded systems $\widehat{X}$ and $\widehat{Y}$ in $\mathbf{\Theta}$, an arbitrary transformation Λ from $\widehat{X}$ to $\widehat{Y}$ is embedded (i.e. of the form $\widehat{T}$) if and only if it satisfies $\widehat{\mathrm{id}_Y} \circ \Lambda \circ \widehat{\mathrm{id}_X} = \Lambda$. Note that this requirement can be considered a natural relaxation of surjectivity: any embedded transformation $\Lambda = \widehat{T}$ satisfies $\widehat{\mathrm{id}_Y} \circ \widehat{T} \circ \widehat{\mathrm{id}_X} = \mathrm{id}_Y \circ T \circ \mathrm{id}_X = \widehat{T}$ by properties of the embedding, but the requirement guarantees that every transformation Λ satisfying $\widehat{\mathrm{id}_Y} \circ \Lambda \circ \widehat{\mathrm{id}_X} = \Lambda$ (i.e. every 'classical' transformation) is in the image of $\widehat{\ }$.

Definition 2.11 (Operational Theories)**.** An *operational theory* is a pair $(\mathbf{\Theta}, \widehat{\ })$, where $\mathbf{\Theta}$ is a compositional theory and where $\widehat{\ } : \mathbf{CIT} \to \mathbf{\Theta}$ is an embedding which identifies **CIT** within $\mathbf{\Theta}$. ■

Quantum information theory is an operational theory by virtue of the embedding $\widehat{\ } : \mathbf{CIT} \to \mathbf{QIT}$ described above. Classical information theory is an operational theory too, by virtue of the identity embedding $\widehat{\ } : \mathbf{CIT} \to \mathbf{CIT}$.

In an operational theory, we may successfully capture the previously considered operational concepts which were not compositional. For example, the dephasing[19] channels $\Delta_X = \widehat{\mathrm{id}_X}$ can be used to define all sorts of hybrid transformations:

Measurements. If $\widehat{Y}$ is an embedded classical system and M a transformation from $\mathcal{H}$ to $\widehat{Y}$, we say that M *has classical outcomes* if $\Delta_Y \circ M = M$. Alternatively, we call such a channel a *measurement on* $\mathcal{H}$ *with outcomes in* Y. In the operational theory **QIT**, it is easy to verify using the Kraus representation of the channel M, that if M is a measurement then there exists a Y-indexed *POVM (Positive Operator-Valued Measure)* on $\mathcal{H}$, i.e. a family $E = (E(y))_{y\in Y}$ of positive operators $E(y)$ on $\mathcal{H}$ with $\sum_{y\in Y} E(y) = \mathbb{1}_{\mathcal{H}}$, such that

$$M(A) = \sum_{y\in Y} \mathrm{tr}(E(y)A)\, |y\rangle\langle y| \quad \text{for all } A \in \mathrm{End}(\mathcal{H}). \tag{4}$$

The POVM is unique given the measurement M, and any POVM E defines a measurement as above. Thus, we can in **QIT** identify measurements with POVMs.

Ensembles of Channels. If $\widehat{X}$ is an embedded classical system and if Λ is a transformation from $\mathcal{H} \otimes \widehat{X}$ to $\mathcal{K}$, we say that Λ *has classical inputs on* $\widehat{X}$ if $\Lambda \circ (\mathrm{id}_{\mathcal{H}} \otimes \Delta_X) = \Lambda$. In **QIT**, it is easy to verify that this holds precisely if there exists a family $(\Lambda^x)_{x\in X}$ of

from **CIT** to **QIT**.

[19] In a general theory, there may be nothing 'dephasing' about a dephasing channel, however we anyway choose to import the term from quantum information theory.

channels $\Lambda^x : \mathcal{H} \to \mathcal{K}$ such that

$$\Lambda(A \otimes B) = \sum_{x \in X} \Lambda^x(A) \langle x| B |x\rangle \quad \text{for all } A \in \mathrm{End}(\mathcal{H}),\ B \in \mathrm{End}(\widehat{X}). \tag{5}$$

Thus, to specify such a channel is to specify an ensemble of channels $\Lambda^x : \mathcal{H} \to \mathcal{K}$, indexed by the set X. (In particular, a channel $\Lambda : \widehat{X} \to \mathcal{K}$ with classical inputs corresponds to an ensemble of states on $\mathcal{K}$.) We will often use the terminology that Λ 'measures' of 'reads off' the classical value x and applies the according channel Λ^x.

It should be obvious how to use the dephasing channels to define also quantum *instruments* (that is, channels with classical outcomes on some factor in the output system), ensembles of instruments, etc.

With a definition of operational structure in place, we can state more precisely what we mean by operational concepts:

Definition 2.12 (Operational Concepts, Informally)**.** In an operational theory, an *operational concept* is one that can be defined in terms of the compositional structure, the embedding $\widehat{\ }$ and any specific systems and transformations in **CIT**. ■

This definition subsumes compositional concepts and concepts defined in terms of dephasing channels (in particular, measurements and ensembles as defined above), but is also meant to subsume reference to specific embedded systems and probabilities (e.g. 'the probability distribution on $\{0,1\}$ which assigns probability $1/3$ to 0'). Again, readers versed in formal logic may entertain the exercise of constructing a more rigorous definition, but we shall here content ourselves with providing a few illustrative examples.

Example 2.13 (Biased Coins)**.** In **QIT**, the state on $\mathbb{C}^2$ with density operator $\varrho = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|1\rangle\langle 1|$ is an operational concept. Indeed, $\mathbb{C}^2$ can be defined as the embedded system $\widehat{\{0,1\}}$, and ϱ as the embedded state $\widehat{s}$, where s is the probability distribution in **CIT** on the system $\{0,1\}$ which assigns probability $\frac{1}{3}$ to outcome 0. Similarly, any other state $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ on $\mathbb{C}^2$ is operational when $p \in [0,1]$ is a definable real number. (The state is in fact compositional when $p = 1/2$, since it is then the unique state invariant under all automorphisms on $\mathbb{C}^2$.) ♦

Example 2.14 (Maximally Entangled)**.** The notion of being a *maximally entangled* state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ is operationally definable. Indeed, $\mathbb{C}^2 \otimes \mathbb{C}^2$ is the embedded system $\widehat{\{0,1\}} \otimes \widehat{\{0,1\}}$, and a state ψ on this system is maximally entangled precisely if it is pure and if for some local automorphism $\alpha_1 \otimes \alpha_2$ of $\mathbb{C}^2 \otimes \mathbb{C}^2$, the state $\psi' := [\alpha_1 \otimes \alpha_2](\psi)$ satisfies $(\Delta_{\{0,1\}} \otimes \Delta_{\{0,1\}})(\psi') = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$. ♦

As an exercise, we encourage the reader to consider how the *trace distance* $\delta(\varrho,\sigma) := \frac{1}{2}\|\varrho - \sigma\|_1$ between quantum states ϱ and σ may be realised as an operational concept, using that $\delta(\varrho,\sigma) = 2p^*(\varrho,\sigma) - 1$ where $p^*(\varrho,\sigma)$ is the optimal probability of distinguishing in a measurement between the states ϱ or σ if provided uniformly at random.

This concludes our presentation of what we mean by operational concepts in quantum information theory (or indeed any operational theory). As mentioned in the beginning, we circumvent an excruciating formalisation of Definition 2.12 by adopting the viewpoint that it is intuitively clear when a concept has been successfully given an operational definition according to this definition.

2.C Naimark's Theorem

In this subsection, we recall a useful structure-theorem for quantum measurements by M. Naimark. In general, a measurement which correspond to the POVM $E = (E(y))_{y\in Y}$ is called *projective* if all of the positive operators $E(y)$ are projections, i.e. satisfy $E(y)^2 = E(y)$. (We do not know whether the notion of being a projective measurement is operational or not, but this is not relevant for our purposes.) Naimark's theorem [Nai40] states that any measurement can be realised using a projective measurement on a larger system:

Theorem 2.15 (Naimark)**.** *For any quantum measurement $M : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\widehat{Y})$, there exists a Hilbert space $\mathcal{K}^{\mathrm{Nai}}$, a projective quantum measurement $M^{\mathrm{Nai}} : \mathrm{End}(\mathcal{H} \otimes \mathcal{K}^{\mathrm{Nai}}) \to \mathrm{End}(\widehat{Y})$ and a pure state ϕ^{Nai} on $\mathcal{K}^{\mathrm{Nai}}$ such that*

$$M = M^{\mathrm{Nai}} \circ (\mathrm{id}_{\mathcal{H}} \otimes \phi^{\mathrm{Nai}}). \tag{6}$$

Remark 2.16. It is not uncommon in statements of Naimark's theorem to take $\mathcal{K}^{\mathrm{Nai}} = \mathbb{C}^m$ for some $m \in \mathbb{N}$ and fix ϕ^{Nai} to be the state $|0\rangle\langle 0|$ (this can always be achieved by a unitary rotation). However, this level of specification is irrelevant for our purposes. ✠

Remark 2.17 (Generalisation to Ensembles)**.** Naimark's theorem can straightforwardly be generalised to measurement ensembles. Specifically, any measurement ensemble $\Lambda : \mathrm{End}(\widehat{X} \otimes \mathcal{H}) \to \mathrm{End}(\widehat{Y})$ is of the form $\Lambda^{\mathrm{Nai}} \circ (\mathrm{id}_{\widehat{X}} \otimes \mathrm{id}_{\mathcal{H}} \otimes \phi^{\mathrm{Nai}})$ for some pure state ϕ^{Nai} on a system $\mathcal{K}^{\mathrm{Nai}}$ and some ensemble Λ^{Nai} of projective measurements on $\mathcal{H} \otimes \mathcal{K}^{\mathrm{Nai}}$. The proof is by induction on the cardinality $|X|$, with the case $|X| = 1$ corresponding to Naimark's original theorem. ✠

2.D Isometric Channels and Stinespring's Dilation Theorem

Recall from Example 2.1 that an *isometric* quantum channel Σ from $\mathcal{H}$ to $\mathcal{K}$ is one of the form $A \mapsto VAV^*$ for some isometry $V : \mathcal{H} \to \mathcal{K}$. Note that an isometric channel determines its underlying isometry up to a phase, that is, the isometries V and $e^{i\theta}V$ give rise to the same isometric channel for any $\theta \in [0, 2\pi)$, but this is the only ambiguity. In particular, pure states on $\mathcal{K}$ (which are simply isometric channels from $\mathbf{1}$ to $\mathcal{K}$) determine their underlying unit vectors $|\psi\rangle \in \mathcal{K}$ up to a phase. We will denote pure states by letters $\psi, \phi, \chi, \ldots$, and corresponding vector representatives by $|\psi\rangle, |\phi\rangle, |\chi\rangle, \ldots$ whenever we write out equations whose content is insensitive to the choice of phase.

We start by observing the following useful fact about isometric channels (recall from Example 2.7 the notion of reversibility):

Lemma 2.18. *Every isometric channel $\Sigma : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ is reversible.*

Proof. Let V be an isometry representing Σ. It is tempting to define a left-inverse $\Sigma^- : \mathrm{End}(\mathcal{K}) \to \mathrm{End}(\mathcal{H})$ as $B \mapsto V^*BV$. However, even though Σ^- reverses the action of Σ, it does not define a quantum channel as it is not trace-preserving (except when V is unitary). Thus, we instead pick an arbitrary state τ on $\mathcal{H}$ and define Σ^- by $\Sigma^-(B) = V^*BV + \tau\,\mathrm{tr}(\sqrt{\mathbb{1}_{\mathcal{H}} - VV^*}B\sqrt{\mathbb{1}_{\mathcal{H}} - VV^*})$. This map is completely positive and trace-preserving, thus a quantum channel, and it satisfies $\Sigma^-(\Sigma(A)) = A$ for every $A \in \mathrm{End}(\mathcal{H})$. □

Isometric channels are famous for their appearance in W. Stinespring's ubiquitous dilation theorem [Sti55], which we now recall in the context of quantum information theory.

Given any quantum channel $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$, a *(one-sided) dilation* [CDP10, HL21] of Λ is a quantum channel $\Phi : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{E})$ such that $(\mathrm{id}_{\mathcal{K}} \otimes \mathrm{tr}_{\mathcal{E}}) \circ \Phi = \Lambda$. We call $\mathcal{E}$ the *environment* of the dilation Φ. The concept of dilation is compositional, as it is definable in terms of serial and parallel composition and the distinguished discarding transformations. In general, a dilation can be interpreted as describing side-information which escapes to a hidden system (namely the environment).

A dilation Σ of Λ is called a *Stinespring dilation* if is it isometric.

Theorem 2.19 (Stinespring)**.** *Every quantum channel $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ has a Stinespring dilation. It is unique up to transformations on the environment, in the sense that if $\Sigma : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{E})$ and $\Sigma' : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{E}')$ are both Stinespring dilations of Λ, then there exists a quantum channel $\Gamma : \mathrm{End}(\mathcal{E}) \to \mathrm{End}(\mathcal{E}')$ such that $\Sigma' = (\Gamma \otimes \mathrm{id}_{\mathcal{K}}) \circ \Sigma$.*

Remark 2.20. If $\dim \mathcal{E} \leq \dim \mathcal{E}'$, then the channel Γ can be taken isometric — in fact, this is how the uniqueness clause is usually stated. The statement in Theorem 2.19 follows from the usual uniqueness statement by Lemma 2.18. ✠

Remark 2.21. In the case where $\mathcal{H} = \mathbf{1}$, the channel Λ corresponds to a density operator $\varrho \in \mathrm{End}(\mathcal{K})$ and its Stinespring dilations are pure states whose corresponding density operators ψ satisfy $(\mathrm{id}_{\mathcal{H}} \otimes \mathrm{tr}_{\mathcal{E}})(\psi) = \varrho$. These Stinespring dilations are more commonly called *purifications* of ϱ. ✠

Stinespring's dilation theorem implies the following property, which in particular lends itself to the interpretation that Stinespring dilations leak maximal information to the environment:

Theorem 2.22 (Completeness of Stinespring Dilations)**.** *Let $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ be a quantum channel. A Stinespring dilation $\Sigma : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{E})$ is a* complete dilation *[HL21] of Λ, meaning that if $\Phi : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{F})$ is any dilation of Λ, then there exists a quantum channel $\Gamma : \mathrm{End}(\mathcal{E}) \to \mathrm{End}(\mathcal{F})$ such that $\Phi = (\Gamma \otimes \mathrm{id}_{\mathcal{K}}) \circ \Sigma$.*

Proof. The channel $\Phi : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K}) \otimes \mathrm{End}(\mathcal{F})$ has a Stinespring dilation $\Sigma' : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K} \otimes \mathcal{F} \otimes \mathcal{G})$. Now, Σ' is also a Stinespring dilation of Λ (with environment $\mathcal{F} \otimes \mathcal{G}$). Therefore, by Theorem 2.19, there exists a channel $\tilde{\Gamma} : \mathrm{End}(\mathcal{E}) \to \mathrm{End}(\mathcal{F} \otimes \mathcal{G})$ such that $\Sigma' = (\tilde{\Gamma} \otimes \mathrm{id}_{\mathcal{K}}) \circ \Sigma$. Tracing out $\mathcal{G}$, we see that $\Phi = (\Gamma \otimes \mathrm{id}_{\mathcal{K}}) \circ \Sigma$, with $\Gamma = (\mathrm{id}_{\mathcal{F}} \otimes \mathrm{tr}_{\mathcal{G}}) \circ \tilde{\Gamma}$. □

Remark 2.23 (General Completeness)**.** In the theory **CIT**, it is also true that every channel has a complete dilation; it is given by copying the input and output to the environment. See Ref. [HL21] for further details and a systematic study of dilational completeness. ✠

Corollary 2.24 (Isometric implies Pure.)**.** *A isometric quantum channel $\Sigma : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{K})$ is* dilationally pure *[Chi14, HL21], meaning that every dilation of Σ is of the form $\Sigma \otimes \varrho$ for some state ϱ.*

Proof. To say that a channel is dilationally pure is to say that it is a complete dilation of itself. However if Σ is isometric then it is evidently a Stinespring dilation of itself, and therefore by Theorem 2.22 complete. □

Remark 2.25 (Isometric means Pure)**.** It can be shown quite easily that the dilationally pure quantum channels are precisely those which are isometric ([HL21] Cor. 2.3.32) — the

idea behind the proof is that if Σ requires more than one Kraus operator in its Kraus representation, then we may construct from this representation a dilation which is not of the form $\Sigma \otimes \varrho$. Dilational purity is evidently a compositional concept, and isometric channels may thus, perhaps surprisingly, be given a compositional definition. ✠

3 The Standard Perception of Quantum Self-Testing

In this section, we recall the standard operator-algebraic perception of quantum self-testing and discuss the need for an alternative, operational perspective. The section is divided into three subsections.

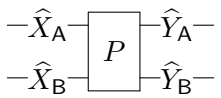
In Section 3.A we describe in broad terms what quantum self-testing aims to say. We shall here discuss Bell-scenarios, behaviours, and (quantum) strategies of a general kind. This discussion can be cast in operational terms cf. Section 2, and the broad aim of self-testing may as such be interpreted in general operational theories beyond quantum information theory. In Section 3.B we give the standard definition of quantum self-testing [MY04, MYS12, ŠB20], which is to say we lay out how the aim of self-testing is conventionally formalised for the theory **QIT**. The fundamental problem addressed by our paper is that this standard formalisation leaves the realm of operational structure; it is cast in operator-algebraic language that does not allow for natural generalisations to other theories, formalisms or settings beyond the Bell-scenario.

Throughout, we restrict attention to the case of self-testing in bipartite Bell-scenarios (i.e. with two parties, A and B), but the entire discussion easily generalises.

3.A Bell-Scenarios and the Aim of Self-Testing

By a **(bipartite) Bell-scenario** we mean a quadruple of finite non-empty sets, $\mathfrak{B} = (X_\mathsf{A}, X_\mathsf{B}, Y_\mathsf{A}, Y_\mathsf{B})$. The interpretation is that two parties, A and B, are separated and that party $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$ receives an input x_P from the set X_P and is expected to return, based on that input and without communicating with the other party, an output y_P from the set Y_P. We use throughout the notation $X := X_\mathsf{A} \times X_\mathsf{B}$ and $Y := Y_\mathsf{A} \times Y_\mathsf{B}$, and denote pairs $(x_\mathsf{A}, x_\mathsf{B})$ and $(y_\mathsf{A}, y_\mathsf{B})$ by x and y, respectively.

However the parties A and B proceed, it is natural to summarise their input-output behaviour by specifying for each input-pair $x = (x_\mathsf{A}, x_\mathsf{B})$ a probability distribution P^x which assigns to each output-pair $y = (y_\mathsf{A}, y_\mathsf{B})$ the probability $P^x(y)$ of output y on input x. Any such collection $P = (P^x)_{x \in X}$ of probability distributions on $Y = Y_\mathsf{A} \times Y_\mathsf{B}$ indexed by $X = X_\mathsf{A} \times X_\mathsf{B}$ will be called a **behaviour** for the Bell-scenario $\mathfrak{B}$.[20] It is worth observing that, as such, a behaviour P is nothing but a transformation in the theory **CIT** from the set X to the set Y. In a general operational theory, we may draw it as



, (7)

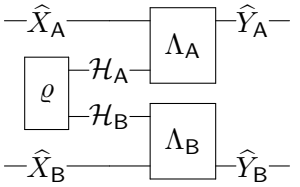
recalling from Section 2.B that $\widehat{X}_\mathsf{P}$ and $\widehat{Y}_\mathsf{P}$ are the embeddings of the classical systems X_P and Y_P into the theory (in **QIT**, they are given by $\widehat{X}_\mathsf{P} = \mathbb{C}^{X_\mathsf{P}}$ and $\widehat{Y}_\mathsf{P} = \mathbb{C}^{Y_\mathsf{P}}$). Intuitively, any theory of physics will specify which behaviours can be realised by non-communicating parties in the Bell-scenario $\mathfrak{B}$. Indeed, a given theory will determine a viable class of 'strategies' (or 'protocols') for obtaining outputs based on inputs, and to each such strategy will

[20] The term 'behaviour' is due to Cirelson [Cir93]. The term 'correlations' is used synonymously in the literature.

be associated a behaviour. Roughly speaking, self-testing is concerned with deducing characteristics of a strategy from its behaviour alone.

The notions of 'strategy' and 'realisable behaviour' can be made precise by referring to the operational structure of the theory. Namely, a behaviour can be realised in a given theory precisely if, as a channel, it can be built using serial and parallel compositions of channels allowed by the theory, in such a way that no information flows from A's input to B's output or from B's input to A's output.

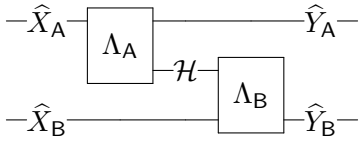
For example, a channel of the form



(8)

where ϱ is a state on a bipartite system $\mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$ and where Λ_A and Λ_B are arbitrary channels, constitutes a realisable behaviour. The concrete circuit of channels, i.e. the triple $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$, constitutes a strategy which realises the behaviour. The circuit physically corresponds to the situation in which the parties A and B share a state which they process locally in an input-dependent way to obtain their outputs.

On the other hand, a circuit of the form



(9)

does not a priori lead to a realisable behaviour, because information flows from A's input to B's output.

A general strategy for the Bell-scenario $\mathfrak{B} = (X_\mathsf{A}, X_\mathsf{B}, Y_\mathsf{A}, Y_\mathsf{B})$ could be defined as any circuit of transformations which has open inputs corresponding to the embeddings $\widehat{X}_\mathsf{A}$ and $\widehat{X}_\mathsf{B}$, open outputs corresponding to the embeddings $\widehat{Y}_\mathsf{A}$ and $\widehat{Y}_\mathsf{B}$, and such that in the directed acyclic graph describing the circuit structure there are no paths from $\widehat{X}_\mathsf{A}$ to $\widehat{Y}_\mathsf{B}$ or from $\widehat{X}_\mathsf{B}$ to $\widehat{Y}_\mathsf{A}$. It can however be shown ([HL21] Example 4.1.27) that every behaviour realisable by a general strategy is realisable by a triple-strategy $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ as depicted in Eq. (8). Since the total input-output behaviour is a classical embedded channel, it may additionally be assumed that Λ_P has classical inputs on $\widehat{X}_\mathsf{P}$ and classical outputs on $\widehat{Y}_\mathsf{P}$ (cf. Section 2.B).

Later, in Section 4.C, we shall make a point out of reconsidering general strategies described by arbitrary circuits, but in the standard description of self-testing one restricts attention to triple-strategies and the term 'strategy' is used exclusively for those:

Definition 3.1 (Strategies)**.** Consider (cf. Definition 2.11) an operational theory $(\boldsymbol{\Theta}, \hat{}\,)$, and let $\mathfrak{B} = (X_\mathsf{A}, X_\mathsf{B}, Y_\mathsf{A}, Y_\mathsf{B})$ be a Bell-scenario. A *strategy for* $\mathfrak{B}$ is a triple $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$, where ϱ is a state on a bipartite system $\mathcal{H} := \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$ in $\boldsymbol{\Theta}$ and where, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, Λ_P is a transformation in $\boldsymbol{\Theta}$ from $\widehat{X}_\mathsf{P} \otimes \mathcal{H}_\mathsf{P}$ to $\widehat{Y}_\mathsf{P}$ with classical inputs on $\widehat{X}_\mathsf{P}$ and with classical outputs. The *behaviour associated to* S is the classical channel $P = (P^x)_{x \in X}$ defined by Eq. (8). ■

Example 3.2 (Strategies in **CIT**: Local Hidden Variables and Bell's Theorem)**.** In the theory of classical information, **CIT**, a strategy for $\mathfrak{B}$ is a triple $(p, L_\mathsf{A}, L_\mathsf{B})$ where p is a

probability distribution on some product set $H = H_{\mathsf{A}} \times H_{\mathsf{B}}$, and where $L_{\mathsf{A}} : X_{\mathsf{A}} \times H_{\mathsf{A}} \to Y_{\mathsf{A}}$ and $L_{\mathsf{B}} : X_{\mathsf{B}} \times H_{\mathsf{B}} \to Y_{\mathsf{B}}$ are classical channels. Bell's theorem [Bel64] is the statement that certain behaviours realisable by quantum strategies are not reproducible by such classical strategies. This result remains a landmark in our understanding of quantum theory, as it rules out an underlying description of certain experiments in terms of 'hidden' random variables (encoded by the state p) and locally applied random functions (encoded by L_{A} and L_{B}). ♦

Example 3.3 (Strategies in **QIT**: Operator-Algebraic Representations)**.** In quantum information theory, **QIT**, a strategy for $\mathfrak{B}$ is a triple $(\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ where ϱ is a quantum state (density operator) on a tensor-product Hilbert space $\mathcal{H} = \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$, and where, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the channel $\Lambda_{\mathsf{P}} : \mathrm{End}(\widehat{X_{\mathsf{P}}} \otimes \mathcal{H}_{\mathsf{P}}) \to \mathrm{End}(\widehat{Y_{\mathsf{P}}})$ is an X_{P}-indexed ensemble of measurements on $\mathcal{H}_{\mathsf{P}}$ with outcomes in Y_{P}. Each measurement in the ensemble can be identified with a POVM, cf. Section 2.B; as such, the channel Λ_{P} may be identified with a collection $E_{\mathsf{P}} = (E_{\mathsf{P}}^{x_{\mathsf{P}}})_{x_{\mathsf{P}} \in X_{\mathsf{P}}}$ of POVMs $E_{\mathsf{P}}^{x_{\mathsf{P}}} = (E_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}}))_{y_{\mathsf{P}} \in Y_{\mathsf{P}}}$ on $\mathcal{H}_{\mathsf{P}}$. An equivalent definition of quantum strategies is thus as triples $(\varrho, E_{\mathsf{A}}, E_{\mathsf{B}})$ with ϱ a density operator and E_{A} and E_{B} ensembles of POVMs; this definition of quantum strategies is indeed more common in the literature on self-testing [ŠB20].

Given a quantum strategy $S = (\varrho, E_{\mathsf{A}}, E_{\mathsf{B}})$ and an input-pair $x = (x_{\mathsf{A}}, x_{\mathsf{B}})$, let us denote by E^x the POVM on $\mathcal{H} = \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ defined by $E^x(y) = E_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}}) \otimes E_{\mathsf{B}}^{x_{\mathsf{B}}}(y_{\mathsf{B}})$ for output-pairs $y = (y_{\mathsf{A}}, y_{\mathsf{B}})$. The POVM ensemble $(E^x)_{x \in X}$ then represents the channel $\Lambda_{\mathsf{A}} \otimes \Lambda_{\mathsf{B}}$, and the behaviour associated to the strategy S, i.e. the channel $P = (P^x)_{x \in X}$ defined by Eq. (8), is given by

$$P^x(y) = \mathrm{tr}(E^x(y)\varrho) \quad \text{for } x \in X,\ y \in Y. \tag{10}$$

(If $\varrho =: \psi$ is a pure state, this expression reduces to $\langle\psi| E^x(y) |\psi\rangle$.) ♦

We can now describe what quantum self-testing aims to do. Every quantum strategy for the Bell-scenario $\mathfrak{B}$ gives rise to a behaviour, as expressed by Eq. (8), or equivalently Eq. (10). In its broadest possible definition, quantum self-testing is the enterprise of deriving statements about a strategy from knowing just its behaviour. Mathematically, if

$$B : \mathrm{Strat}(\mathfrak{B}) \to \mathrm{Bhv}(\mathfrak{B}) \tag{11}$$

denotes the map which associates behaviours to strategies, the question is what can be said about strategies in the pre-image $B^{-1}(\{P\}) \subseteq \mathrm{Strat}(\mathfrak{B})$ given a quantum behaviour $P \in \mathrm{Bhv}(\mathfrak{B})$. In a narrower sense, self-testing refers to the situation in which one can derive from the behaviour essentially uniquely what the strategy is. This is indeed the sense in which self-testing was introduced [MY04] and standardised [MYS12], and it is what we mean throughout when referring to the standard, or conventional, notion of self-testing. Specifically, it is concerned with the situation in which there exists a fixed *canonical* strategy $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ in $B^{-1}(\{P\})$ such that every strategy $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ in $B^{-1}(\{P\})$ is 'reducible' to the canonical strategy $\tilde{S}$. One says in this case that the behaviour P *self-tests* the strategy $\tilde{S}$. As such, to define self-testing properly, one needs only define what is meant by one strategy being 'reducible' to another.

Now, there is nothing whatsoever about the aim of self-testing which seems to rely on quantum theory: the broad question of determining the form of strategies from their behaviour is meaningful in every operational theory. Nevertheless, any currently known definition of what it means for a strategy S to be *reducible* to another strategy $\tilde{S}$ is

cast in terms of their operator-algebraic constituents described in Example 3.3. We shall present the standard definition in Section 3.B below. It is not clear how to translate it to operational language, and this is the fundamental issue which we tackle in our paper and resolve by providing a new and purely operational definition.

3.B The Operator-Algebraic Formulation of Quantum Self-testing

We now present the standard definition of reducibility among quantum strategies and quantum self-testing [MYS12, ŠB20]. Some authors [Kan17] consider a slightly alternative reducibility definition (in fact in line with Ref. [MY98]), but it too relies on operator-algebra. We discuss in Remark 3.10 why we choose to make operational the standard definition rather than such alternative definitions.

Recall from Example 3.3 that a quantum strategy for the Bell-scenario $\mathfrak{B}$ is formally a triple of state and channels, $(\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$, or equivalently a triple of state and POVM-ensembles, $(\varrho, E_{\mathsf{A}}, E_{\mathsf{B}})$. Let us employ the following terminology:

Definition 3.4 (Types of Quantum Strategies)**.** A quantum strategy $S = (\varrho, E_{\mathsf{A}}, E_{\mathsf{B}})$ is called

- *pure-state* if the state ϱ is pure;
- *projective* if all the POVM-elements $E_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})$ are projections;
- *full-rank* if the marginal states ϱ_{A} and ϱ_{B} have full ranks on $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$, respectively.

■

It is worth observing that every behaviour realisable by a quantum strategy can be realised by some projective pure-state strategy, by picking any initial strategy, Naimark-extending all of its measurements (cf. Remark 2.17) and purifying its state.[21]

One would expect the reducibility relation to be a pre-order, $\geq_{red.}$, on the class of all strategies, $\mathrm{Strat}(\mathfrak{B})$, with $S \geq_{red.} \tilde{S}$ signifying that S is reducible to $\tilde{S}$. In reality, however, it is quite standard in expositions on self-testing to restrict attention to projective strategies [ŠB20], and sometimes even projective pure-state strategies [MYS12]. The sufficiency of this restricted domain of attention has been explained by various arguments which we now briefly discuss. Note that restricting the form of the fixed canonical strategy $\tilde{S}$ is conceptually different from restricting the form of the variable strategy S.

As for the canonical strategy $\tilde{S}$, it has been suggested [GKW+18] that it is impossible to self-test a strategy which is not pure-state and projective. A formal proof of such a statement would however require the reducibility relation to be already defined for arbitrary strategies, and to the best of our knowledge such a definition has not been given explicitly before. Instead, we adopt the viewpoint that it just so happens that self-testing has only been studied in the case of a projective pure-state canonical strategy, and that consequently a reducibility definition for a general $\tilde{S}$ has not been explored.[22]

[21] However, it is fully possible that not every quantum realisable behaviour can be realised by a projective full-rank strategy.

[22] After presenting our new operational version of reducibility, which applies to all strategies, we will prove formally that the corresponding version of self-testing is indeed only possible for pure-state strategies (Corollary 5.5). One may take it furthermore projective without loss of generality (Proposition 4.13).

Restricting the form of the variable strategy S, on the other hand, cannot be justified in a similar way; indeed, the entire spirit of self-testing is to characterise arbitrary, unknown strategies with a given behaviour. Nevertheless, a restriction to projective strategies is usually justified by a reference to the possibility of Naimark-extending the measurements [ŠB20]. Referring additionally to the possibility of purifying the state, S is sometimes restricted to be also pure-state, in line with Ref. [MYS12]. Now, if one has a particular application of the self-testing phenomenon in mind, it may or may not be valid to consider only projective pure-state strategies, depending on what that application is. We believe, however, that restrictions on the unknown strategy S are in general ill-motivated if one wants to characterise all strategies with a given behaviour.[23] In fact, assuming strategies to be always pure-state can lead to downright misleading conclusions. For example, one then readily arrives at the ridiculous statement that entanglement is necessary to produce any non-product behaviour (since every non-entangled pure state is product). Also, one concludes incorrectly that no adversarial environment can hold pre-existing information correlated with the measurement outcomes of the strategy. We do not know that the restriction to projective measurements should cause similar issues.[24]

For now, we shall refrain from further discussing the quality of the motivations behind these restrictions, and rather state the reducibility and self-testing definitions which we will henceforth unambiguously refer to as the standard definitions (specifically, they appear as Def. 2 in the survey [ŠB20]). Throughout, we will denote projective POVMs by the letter Π rather than E.

Definition 3.5 (Operator-Algebraic Reducibility among Quantum Strategies)**.** Let $S = (\varrho, \Pi_{\mathsf{A}}, \Pi_{\mathsf{B}})$ be a projective strategy and let $\tilde{S} = (\tilde{\psi}, \tilde{\Pi}_{\mathsf{A}}, \tilde{\Pi}_{\mathsf{B}})$ be a projective pure-state strategy. Let ψ be a purification of ϱ, with purifying system $\mathcal{P}$. We say that S *is reducible to* $\tilde{S}$, written $S \geq_{red.} \tilde{S}$, if there exist Hilbert spaces $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}}$, $\mathcal{H}_{\mathsf{B}}^{\mathrm{res}}$, isometries $W_{\mathsf{A}} : \mathcal{H}_{\mathsf{A}} \to \tilde{\mathcal{H}}_{\mathsf{A}} \otimes \tilde{\mathcal{H}}_{\mathsf{A}}^{\mathrm{res}}$ and $W_{\mathsf{B}} : \mathcal{H}_{\mathsf{B}} \to \tilde{\mathcal{H}}_{\mathsf{B}} \otimes \tilde{\mathcal{H}}_{\mathsf{B}}^{\mathrm{res}}$, and a pure state ψ^{res} on $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}} \otimes \mathcal{P}$, such that

$$\forall x \in X, y \in Y : \ [W \Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle = \tilde{\Pi}^x(y) |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle \, , \tag{12}$$

where $W = W_{\mathsf{A}} \otimes W_{\mathsf{B}}$, and where $\Pi^x(y) = \Pi_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}}) \otimes \Pi_{\mathsf{B}}^{x_{\mathsf{B}}}(y_{\mathsf{B}})$ and $\tilde{\Pi}^x(y) = \tilde{\Pi}_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}}) \otimes \tilde{\Pi}_{\mathsf{B}}^{x_{\mathsf{B}}}(y_{\mathsf{B}})$ for $x = (x_{\mathsf{A}}, x_{\mathsf{B}})$ and $y = (y_{\mathsf{A}}, y_{\mathsf{B}})$. ■

Remark 3.6. The reducibility condition is independent of the choice of purification ψ of ϱ, since the state ψ^{res} can be changed accordingly. ✠

Definition 3.7 (Self-Testing according to Reducibility)**.** Suppose that $\tilde{S} = (\tilde{\psi}, \tilde{\Pi}_{\mathsf{A}}, \tilde{\Pi}_{\mathsf{B}})$ is a projective pure-state strategy with behaviour P. We say that P *self-tests* $\tilde{S}$ if every projective strategy S with behaviour P is reducible to $\tilde{S}$. ■

The inner logic to this self-testing definition is that any strategy S which does abide to Eq. (12) has the same behaviour as $\tilde{S}$; therefore we cannot hope to prove anything stronger about the form of an unknown strategy S with the same behaviour as $\tilde{S}$. The reducibility condition is often put into words by saying that 'up to local isometric embeddings, the measurements act on the state in S as if $\tilde{S}$ is appended with an unmeasured residual state ψ^{res}'.

[23] Some works do obtain such a characterisation, see e.g. Ref. [Kan17].

[24] After presenting our operational version of reducibility, which applies to all strategies, we will prove that it suffices to consider projective strategies (Proposition 4.13), thus giving substance to this perception.

Remark 3.8 (Pure-State Strategies)**.** In the case where the state ϱ is pure, with vector-representative $|\psi\rangle$, the reducibility condition $S \geq_{red.} \tilde{S}$ reads

$$\forall x \in X, y \in Y : \; W\Pi^x(y)\,|\psi\rangle = \tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle \otimes |\psi^{\text{res}}\rangle\,. \tag{13}$$

This is a common way of stating the reducibility condition in treatments which assume pure-state strategies throughout. ✠

Remark 3.9 (Terminology)**.** The term 'reducible' is not standard in the literature, indeed the reducibility condition is rarely separated from the self-testing definition and explicitly named. A more common terminology is that 'S is *equivalent* to $\tilde{S}$', however the relation $\geq_{red.}$ is a not symmetric and therefore no equivalence relation. We shall in Section 4.B discuss how the relation $\geq_{red.}$ (or rather, our operational version of it) generates an equivalence relation. This generated equivalence relation is in a sense very close to the original relation, and thus provides some substance to the term 'equivalent'. See Theorem 4.23 and Remark 4.24. ✠

Remark 3.10 (Other Self-Testing Definitions)**.** Some authors consider an alternative self-testing definition in which statements are made separately about the state and measurements (see e.g. [Kan17] Proposition A.3 and [ŠB20] Definition 3). In our opinion, it is better to consider statements which, like Definition 3.5, mixes state and measurements, since statements separately about measurements will not generalise adequately to an approximate setting (essentially because the marginal states ϱ_{A} and ϱ_{B} may be supported on subspaces corresponding to small eigenvalues). To our knowledge, Definition 3.5 above is the only definition with this property which currently exists. Therefore, this is the definition we choose to modify into an operational definition. ✠

3.C Summary and Critique of the Standard Approach

In summary, the conventional narrative of quantum self-testing in a Bell-scenario $\mathfrak{B}$ is as follows:

- An input-output behaviour for the Bell-scenario is identified with a collection of probability distributions $P = (P^x)_{x \in X}$.
- Any imaginable strategy for producing an input-output behaviour without communication can be reduced to a standardised strategy (Definition 3.1), in which each party makes an input-dependent measurement on a shared state and returns as output the measurement outcome.
- A standardised strategy can, by the operator-algebraic formalism of quantum theory, be identified with a triple $S = (\varrho, E_{\mathsf{A}}, E_{\mathsf{B}})$ of operator-algebraic objects as prescribed in Example 3.3.
- Quantum self-testing refers to the situation in which every strategy S with behaviour P can be *reduced* to a single 'canonical' strategy $\tilde{S}$. Formally, reducibility is a relation between strategies stated in operator-algebraic language (Definition 3.5). The condition which defines the relation is motivated by the observation that any strategy $\tilde{S}$ can be subjected to a number of behaviour-preserving transformations, e.g. tensoring with unmeasured states, local unitary rotations and local embeddings into larger spaces.

As discussed in Section 3.B, the reducibility relation $S \geq_{red.} \tilde{S}$ is usually only stated and enforced when S is projective and $\tilde{S}$ is projective and pure-state. The significance of this is, at best, unexplained, but the message of our paper is of course that the operator-algebraic narrative of self-testing is deceiving in its entirety; it fails to give operational meaning to self-testing in line with Section 2.

In the remainder of the paper, we set out to recast quantum self-testing in a purely operational language. The definition we propose (Definition 4.7) has several benefits over the conventional definition.

It will provide a meaning to self-testing interpretable in any operational theory. It will also clarify the significance of restricting to projective strategies (Proposition 4.13), yield a sense in which different strategies for a self-testing behaviour are genuinely equivalent (cf. Remark 3.9), and innately suggest an approximate relaxation suitable for robust self-testing (Remark 4.18). Moreover, by relaxing the definition and ultimately relating it to the framework of causal channels and dilations [HL21], it will additionally clarify the significance of standardised strategies (cf. the second item above) and point to a sensible definition of self-testing in experimental scenarios with other causal structures than Bell-scenarios. Most importantly, perhaps, this connection to a larger framework uncovers quantum self-testing as a special instance of a general and natural phenomenon (Section 4.C).

Our operational definition affects one more issue, which we now invite the reader to consider.

The logic which drives the operator-algebraic approach to self-testing is that the behaviour-preserving transformations on $\mathrm{Strat}(\mathfrak{B})$ – e.g. unitary conjugations and tensorings with unmeasured states – should determine and justify the shape of the formal definition of reducibility (Definition 3.5). The mindset which motivates this definition is that each such behaviour-preserving transformation witnesses an ambiguity in the operator-algebraic description of the physical system, and that *therefore* the definition of self-testing should be no stricter than to include this transformation as an allowed modification to the canonical strategy [ŠB20].

Now, however, suppose that – in addition to tensorings and unitary conjugations – a third kind of behaviour-preserving transformation $\mathscr{T} : \mathrm{Strat}(\mathfrak{B}) \to \mathrm{Strat}(\mathfrak{B})$ were to be discovered, such that for some behaviour P and for any strategy $S \in \mathrm{Strat}(\mathfrak{B})$ with behaviour P, the strategy $\mathscr{T}(S)$ is not reducible to S in the sense of Definition 3.5. This situation is not at all hypothetical: The operation of *transposing* (in some basis) the elements of a strategy preserves its behaviour, while it does not necessarily result in a strategy which is reducible in the sense of Definition 3.5 to the initial one (see e.g. Section 3.7.1 in Ref. [ŠB20]; in this example, the state is invariant under transposition). Likewise, Refs. [HM15, CFH21] present examples of strategies related by partial transposition (on one of the systems $\mathcal{H}_{\mathsf{P}}$); these strategies have identical behaviours too, but their states have radically different amounts of entanglement[25] and they are therefore likely not reducible to one another.

One possible reaction to the discovery of such a new behaviour-preserving transformation, is to conclude that P then constitutes a counterexample to self-testing. The alternative reaction is to conclude that the definition of self-testing itself should be weakened, so as to additionally incorporate the transformation $\mathscr{T}$ which was so far overlooked.

[25] Since the states have a positive partial transpose (PPT), both the original state and its partial transpose are valid states.

Now, the conventional operator-algebraic approach to self-testing is really unable to decide which of the two reactions to choose. Both are equally valid from an operator-algebraic perspective. It seems that, generally, the reaction of altering the definition is favoured (see e.g. Section 3.7.1 of Ref. [ŠB20]); this might be said to be also the natural choice from the operator-algebraic viewpoint, since it harmonises with the logic and mindset which motivated the reducibility definition in the first place.

However, our operational approach to self-testing shows that a completely different mindset is possible. The operational definition of self-testing presented in this work will turn out to be equivalent (cf. Theorem 4.8) to the definition induced by Definition 3.5 — without transpositions, indeed without *any* additional behaviour-preserving transformations. While we leave it as an interesting open question whether there exists another operational definition which corresponds to the inclusion of e.g. transpositions, our work demonstrates that a criterion external to the operator-algebraic formalism *can* be used to take a stand on the question of whether additional transformations should be absorbed into the self-testing definition or not. By our approach, the original reducibility definition, Definition 3.5, is justified and enforced not only *algebraically* but indeed *operationally*.

4 Reworking Quantum Self-Testing

In this section, we present a purely operational definition of quantum self-testing and prove its equivalence to the standard notion of self-testing. We then elaborate on this operational formalisation to eventually explain how quantum self-testing can be realised from an even more general perspective. This section is the core section of our paper, and it is divided into three subsections.

Our new definition and the link to conventional self-testing is established in Section 4.A. The new definition works by replacing the operator-algebraic reducibility relation (Definition 3.5) with an operationally defined pre-order between strategies, which we will call *local simulation* (Definition 4.2). The corresponding notion of self-testing (Definition 4.7) will be referred to as *self-testing according to local simulation* (for short, *self-testing a.t.l.s.*), whereas the conventional self-testing notion (Definition 3.7) is for emphasis referred to as *self-testing according to reducibility* (for short, *self-testing a.t.r.*). We then prove the main theorems relating local simulation to reducibility (Theorem 4.8) and self-testing a.t.l.s. to self-testing a.t.r. (Theorem 4.9). Strictly speaking, we can only prove the two notions equivalent under the technical assumption that the pure state of the canonical strategy $\tilde{S}$ has locally full rank.[26] We know of no examples of self-testing where this condition is not met.

In Section 4.B, in order to improve our understanding of local simulation, we introduce a slightly coarser relation, *local assisted simulation* (Definition 4.19). Like local simulation, local assisted simulation is a pre-order on the class of all strategies, and it is interpretable in any operational theory. It turns out that local assisted simulation is in quantum theory an equivalence relation (Theorem 4.23), in fact the equivalence relation *generated by* the relation of local simulation (Remark 4.24). This circumstance reveals in particular that quantum self-testing can be thought of as a two-fold phenomenon and it leads us to define an alternative notion of self-testing based on local assisted simulation (see Definition 4.25 and the preceding discussion).

Section 4.C takes our operational approach to self-testing one step further. By coarsening local assisted simulation further to *causal simulation* (Definition 4.28), we ultimately

[26] This is a non-trivial condition, since $\tilde{S}$ is assumed projective.

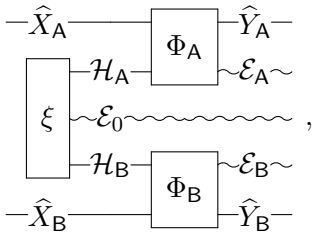
realise quantum self-testing as an instance of the general theory of causal channels and dilations [HL21]. This framework in particular explains how to perceive of quantum strategies and self-testing in a modular environment of general circuits, and how to generalise self-testing to scenarios with arbitrary causal structure. The presentation in Section 4.C is kept somewhat brief, and we refer the reader to the PhD thesis [HL21] for further details (the general version of self-testing is called 'rigidity' therein), and for comparisons to other frameworks involving general circuits of information channels [CDP09, Per17, KU17].

4.A A Purely Operational Formulation of Self-Testing

Let $\mathfrak{B} = (X_{\mathsf{A}}, X_{\mathsf{B}}, Y_{\mathsf{A}}, Y_{\mathsf{B}})$ be a Bell-scenario. Recall from Definition 3.1 the notion of a strategy for $\mathfrak{B}$, namely a triple $(\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ consisting of a state and channels representing measurement ensembles. Recall also that this notion of strategy is purely operational, as it refers only to the channels of the theory and to the operational notion of being a measurement ensemble.

Finally, recall from Section 2.D that a *dilation* of a channel Λ from $\mathcal{H}$ to $\mathcal{K}$ is a channel Φ from $\mathcal{H}$ to $\mathcal{K} \otimes \mathcal{E}$ whose marginal, when discarding the *environment* $\mathcal{E}$, is Λ.

Definition 4.1 (Component-wise Dilations and Implementations)**.** Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ be a strategy. A *component-wise dilation of* S is a triple $(\xi, \Phi_{\mathsf{A}}, \Phi_{\mathsf{B}})$ where ξ is a dilation of ϱ, Φ_{A} a dilation of Λ_{A} and Φ_{B} a dilation of Λ_{B}. An *implementation of* S is a channel of the form



, (14)

where $(\xi, \Phi_{\mathsf{A}}, \Phi_{\mathsf{B}})$ is a component-wise dilation of S. ■

The *trivial implementation* of $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ is the implementation corresponding to the component-wise dilation $(\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ of S. Note that this is simply the behaviour of the strategy. In the theory **QIT**, a *Stinespring implementation* of S is an implementation corresponding to a *component-wise Stinespring dilation*, i.e. a component-wise dilation $(\psi, \Sigma_{\mathsf{A}}, \Sigma_{\mathsf{B}})$ where Σ_{A} and Σ_{B} are Stinespring dilations of Λ_{A} and Λ_{B}, respectively, and where ψ is a Stinespring dilation (purification) of ϱ.

The concepts of component-wise dilation and implementation of a strategy are operational, in fact compositional, cf. Section 2.

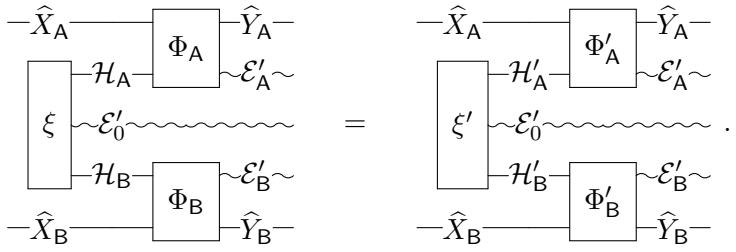
Dilations of a channel represent side-information escaping to an environment during the execution of the channel. Thus, given a strategy $(\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$, a viable interpretation of the component-wise dilation $(\xi, \Phi_{\mathsf{A}}, \Phi_{\mathsf{B}})$ is that ξ represents pre-existing side-information about the state ϱ, and that, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, Φ_{P} represents side-information formed locally about the input $x_{\mathsf{P}} \in X_{\mathsf{P}}$ and the subsystem $\mathcal{H}_{\mathsf{P}}$ upon applying the channel Λ_{P}. We think of the implementation given by Eq. (14) as representing the total information processing which may occur when taking the environment into account: the side-information residing in system $\mathcal{E}_0$ is released to the environment in advance of seeing the inputs x_{A} and x_{B}, whereas the side-information residing in system $\mathcal{E}_{\mathsf{P}}$ is leaked upon seeing input x_{P}. Note that when ϱ is a pure state, then (by Corollary 2.24) any dilation ξ factors as $\varrho \otimes \sigma$ for

some state σ, so any implementation factors too — this signifies that pre-existing side-information is independent from information generated upon the inputs x_{A} and x_{B}.

The implementation (14) is always a dilation of the behaviour channel P, with environment $\mathcal{E}_{\mathsf{A}} \otimes \mathcal{E}_0 \otimes \mathcal{E}_{\mathsf{B}}$. It is a Stinespring dilation of P when the implementation is a Stinespring implementation. As we shall see, however, the distinction between these three parts of the environment $\mathcal{E}_{\mathsf{A}}$, $\mathcal{E}_{\mathsf{B}}$ and $\mathcal{E}_0$ is instrumental to the significance of implementations.

Definition 4.2 (Local Simulation)**.** Let S and S' be any two strategies for $\mathfrak{B}$. We say that S *locally simulates* S', written $S \geq_{l.s.} S'$, if every implementation of S' is also an implementation of S.

Explicitly, $S \geq_{l.s.} S'$ if for any component-wise dilation $(\xi', \Phi'_{\mathsf{A}}, \Phi'_{\mathsf{B}})$ of S' there exists a component-wise dilation $(\xi, \Phi_{\mathsf{A}}, \Phi_{\mathsf{B}})$ of S such that the respective environments match up and their corresponding implementations coincide, i.e.



(15)

■

The relation $S \geq_{l.s.} S'$ seeks to express that any side-information which can be generated during the execution of S' may also be generated during the execution of S. The reader may wonder why in Definition 4.2 we have chosen the name 'local simulation' (and what it has to do with the description in Section 1), but this will become clear in Lemma 4.11. The formulation in Definition 4.2 has the advantage that it is obviously interpretable in every compositional theory.

Remark 4.3. If $S \geq_{l.s.} S'$, then the trivial implementation of S' is also the trivial implementation of S. In other words, S and S' must have the same behaviour. ✠

Proposition 4.4. *Local simulation is a pre-order on the class of strategies for* $\mathfrak{B}$*,* $\mathrm{Strat}(\mathfrak{B})$.

Proof. Reflexivity and transitivity are both obvious. □

Example 4.5 (Full-Rank Strategies)**.** Any strategy $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ is equivalent under local simulation to the strategy $\bar{S} = (\bar{\varrho}, \bar{\Lambda}_{\mathsf{A}}, \bar{\Lambda}_{\mathsf{B}})$ which arises from S by cutting down the systems $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$ to the local supports of the state ϱ (i.e. $S \geq_{l.s.} \bar{S}$ and $\bar{S} \geq_{l.s.} S$). In this sense, a strategy may without loss of generality be assumed to be full-rank. ⧫

Example 4.6 (Augmentation by Ancillary States)**.** Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ be a strategy and let γ be a state on a bipartite system $\mathcal{K}_{\mathsf{A}} \otimes \mathcal{K}_{\mathsf{B}}$. Define the *$\gamma$-augmented strategy* $S[\gamma]$ as the one which arises from S by appending γ to the state ϱ and locally discarding the system $\mathcal{K}_{\mathsf{P}}$ when measuring. Explicitly, the new state is $\varrho \otimes \gamma$ on the bipartite system $(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{K}_{\mathsf{A}}) \otimes (\mathcal{H}_{\mathsf{B}} \otimes \mathcal{K}_{\mathsf{B}})$, and the new measurement channels are $\Lambda_{\mathsf{A}} \otimes \mathrm{tr}_{\mathcal{K}_{\mathsf{A}}}$ and $\Lambda_{\mathsf{B}} \otimes \mathrm{tr}_{\mathcal{K}_{\mathsf{B}}}$. It is easy to check that $S[\gamma] \geq_{l.s.} S$. (The converse is generally false, as will be clear from Proposition 5.1.) ⧫

We now state our new, operational self-testing definition:

Definition 4.7 (Self-Testing according to Local Simulation)**.** Let $\tilde{S}$ be any strategy with behaviour P. We say that P *self-tests* $\tilde{S}$ *according to local simulation* (for short, *a.t.l.s.*) if any strategy S with behaviour P locally simulates $\tilde{S}$, i.e. $S \geq_{l.s.} \tilde{S}$. ■

Contrary to the usual notion of self-testing according to reducibility (a.t.r.), our notion of self-testing a.t.l.s. unproblematically quantifies universally over <u>all</u> strategies S, projective or not, and moreover does not a priori assume the canonical strategy $\tilde{S}$ to be pure-state and projective.

The main results of this subsection are the following:

Theorem 4.8 (Reducibility versus Local Simulation)**.** *Let $\tilde{S}$ be a projective pure-state strategy. Assume moreover that $\tilde{S}$ is full-rank. Then, for any projective strategy S, it holds that $S \geq_{red.} \tilde{S}$ if and only if $S \geq_{l.s.} \tilde{S}$.*

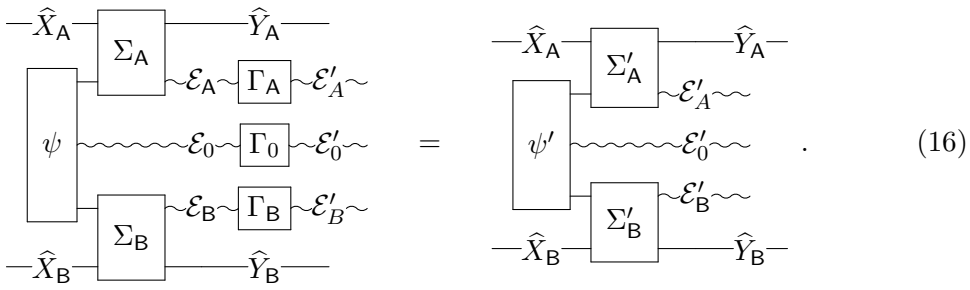
Theorem 4.9 (Self-Testing a.t.r. versus Self-Testing a.t.l.s.)**.** *Let $\tilde{S}$ be a projective pure-state and full-rank strategy, and let P denote its behaviour. Then, P self-tests $\tilde{S}$ a.t.r. if and only if P self-tests $\tilde{S}$ a.t.l.s.*

Remark 4.10. In connection with Theorem 4.8, it is worth remarking that $S \geq_{red.} \tilde{S}$ implies $S \geq_{l.s.} \tilde{S}$ regardless of the rank-assumption on the state. ✠

The remainder of this section is devoted to proving Theorem 4.8 and Theorem 4.9. Note that the 'if'-direction in Theorem 4.9 follows immediately from Theorem 4.8. The 'only if'-direction is not immediate, since self-testing a.t.r. quantifies only over projective strategies; it will however follow if we can prove that every strategy is $\geq_{l.s.}$-equivalent to a projective strategy. Thus, we prove this (Proposition 4.13 below) and we prove Theorem 4.8.

First, we need to better understand the notion of local simulation among strategies. A key step is the observation that local simulation may in **QIT** be characterised in terms of Stinespring implementations:[27]

Lemma 4.11 (Stinespring Characterisation of Local Simulation)**.** *Let S and S' be strategies, and let $(\psi, \Sigma_A, \Sigma_B)$ and $(\psi', \Sigma'_A, \Sigma'_B)$ be component-wise Stinespring dilations of S and S', respectively. Then, $S \geq_{l.s.} S'$ if and only if there exist channels Γ_0, Γ_A and Γ_B such that*



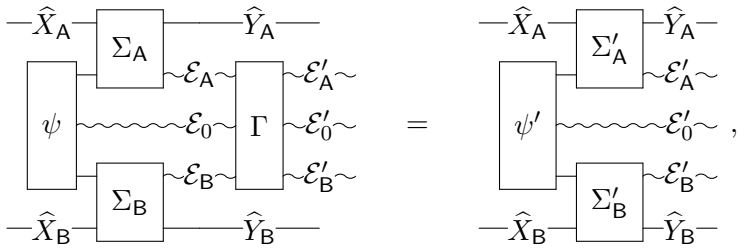
(16)

Proof. By Theorem 2.22, any component-wise dilation (ξ, Φ_A, Φ_B) of S can be obtained by applying channels Γ_A, Γ_B and Γ_0 to the elements of the component-wise Stinespring dilation

[27] As the proof of Lemma 4.11 shows, local simulation may more generally be characterised in terms of implementations corresponding to *complete* dilations, cf. the discussion in Section 2.D.

$(\psi, \Sigma_{\mathsf{A}}, \Sigma_{\mathsf{B}})$. Hence, any implementation of the strategy S is of the form displayed on the left hand side in Eq. (16), for some choice of Γ_{A}, Γ_{B} and Γ_0. Similarly, any implementation of S' is obtained by applying some channels Γ'_{A}, Γ'_{B} and Γ'_0 to the Stinespring implementation on the right hand side. As such, local simulation amounts to finding for each choice of Γ'_{A}, Γ'_{B} and Γ'_0 a matching choice of Γ_{A}, Γ_{B} and Γ_0. However, it suffices to do this when Γ'_{A}, Γ'_{B} and Γ'_0 are all identity channels, since a general choice of channels Γ_i is the matched by the compositions $\Gamma'_i \circ \Gamma_i$, for $i \in \{\mathsf{A}, \mathsf{B}, 0\}$. □

Remark 4.12. When strategies S and S' have the same behaviour, there is always by Theorem 2.22 some channel Γ such that



(17)

since a Stinespring implementation of a strategy is in particular a Stinespring dilation of its behaviour. Local simulation is a non-trivial relation between Stinespring implementations because it requires a channel Γ which factors as $\Gamma_{\mathsf{A}} \otimes \Gamma_0 \otimes \Gamma_{\mathsf{B}}$. ✠

With Lemma 4.11 at our disposal, local simulation has become a more tractable relation. (The reader may find it interesting to consult Example 4.5 and Example 4.6 in this light.) In particular, we will use it now to prove that any strategy is equivalent under local simulation to a projective strategy — once that result is established, Theorem 4.9 will follow automatically from Theorem 4.8. The result is however also interesting by itself, since it gives, for the first time, mathematical substance to the common perception that one may 'without loss of generality' assume quantum strategies to be projective:

Proposition 4.13. *Any strategy S is equivalent under local simulation to a projective strategy S'.*

Proof. Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$. For each $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, it follows from Naimark's theorem (see Remark 2.17) that the measurement ensemble Λ_{P} can be written as $\Lambda^{\mathrm{Nai}}_{\mathsf{P}} \circ (\mathrm{id}_{\widehat{X}_{\mathsf{P}}} \otimes \mathrm{id}_{\mathcal{H}_{\mathsf{P}}} \otimes \phi^{\mathrm{Nai}}_{\mathsf{P}})$ with $\phi^{\mathrm{Nai}}_{\mathsf{P}}$ a pure state on some system $\mathcal{K}^{\mathrm{Nai}}_{\mathsf{P}}$, and $\Lambda^{\mathrm{Nai}}_{\mathsf{P}}$ an ensemble of projective measurements on $\mathcal{H}_{\mathsf{P}} \otimes \mathcal{K}^{\mathrm{Nai}}_{\mathsf{P}}$. We show that S is $\geq_{l.s.}$-equivalent to the projective strategy $S' := (\varrho \otimes \phi^{\mathrm{Nai}}_{\mathsf{A}} \otimes \phi^{\mathrm{Nai}}_{\mathsf{B}}, \Lambda^{\mathrm{Nai}}_{\mathsf{A}}, \Lambda^{\mathrm{Nai}}_{\mathsf{B}})$. To this end, observe that if ψ is a purification of ϱ and $\Sigma^{\mathrm{Nai}}_{\mathsf{P}}$ a Stinespring dilation of $\Lambda^{\mathrm{Nai}}_{\mathsf{P}}$, then $(\psi \otimes \phi^{\mathrm{Nai}}_{\mathsf{A}} \otimes \phi^{\mathrm{Nai}}_{\mathsf{B}}, \Sigma^{\mathrm{Nai}}_{\mathsf{A}}, \Sigma^{\mathrm{Nai}}_{\mathsf{B}})$ is a component-wise Stinespring dilation of S' and $(\psi, \Sigma^{\mathrm{Nai}}_{\mathsf{A}} \circ (\mathrm{id}_{\widehat{X_{\mathsf{A}}}} \otimes \mathrm{id}_{\mathcal{H}_{\mathsf{A}}} \otimes \phi^{\mathrm{Nai}}_{\mathsf{A}}), \Sigma^{\mathrm{Nai}}_{\mathsf{B}} \circ (\mathrm{id}_{\widehat{X_{\mathsf{B}}}} \otimes \mathrm{id}_{\mathcal{H}_{\mathsf{B}}} \otimes \phi^{\mathrm{Nai}}_{\mathsf{B}}))$ is a component-wise Stinespring dilation of S. These corresponding Stinespring implementations are moreover identical. Thus, it follows immediately from Lemma 4.11 that $S \geq_{l.s.} S'$ and $S' \geq_{l.s.} S$. □

It remains now only to prove Theorem 4.8. To this end, we first translate local simulation to operator-algebraic language in the case of projective strategies:[28]

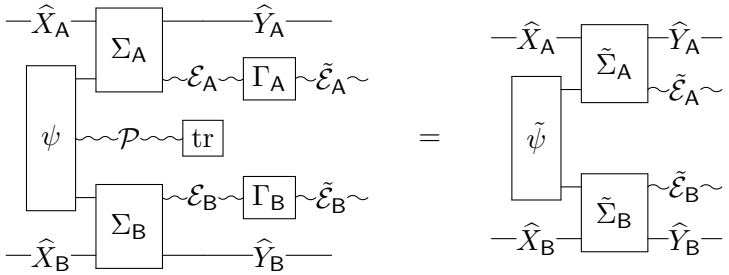
[28] We additionally assume that the state of the simulated strategy is pure, because the result in this case takes on a particularly simple form, and because we shall not need the result in the case of a general mixed state.

Proposition 4.14 (Operator-Algebraic Characterisation of Local Simulation)**.** *Let* $S = (\varrho, \Pi_{\mathsf{A}}, \Pi_{\mathsf{B}})$ *and* $\tilde{S} = (\tilde{\psi}, \tilde{\Pi}_{\mathsf{A}}, \tilde{\Pi}_{\mathsf{B}})$ *be projective strategies, with* $\tilde{\psi}$ *a pure state. Let* ψ *be a purification of* ϱ *with purifying space* $\mathcal{P}$*. Then,* $S \geq_{l.s.} \tilde{S}$ *if and only if there exist Hilbert spaces* $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}}$, $\mathcal{H}_{\mathsf{B}}^{\mathrm{res}}$*, a pure state* ψ^{res} *on* $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}} \otimes \mathcal{P}$*, and isometries* $V_{\mathsf{A}} : \widehat{X}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{A}} \to \widehat{X}_{\mathsf{A}} \otimes \tilde{\mathcal{H}}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{A}}^{\mathrm{res}}$ *and* $V_{\mathsf{B}} : \widehat{X}_{\mathsf{B}} \otimes \mathcal{H}_{\mathsf{B}} \to \widehat{X}_{\mathsf{B}} \otimes \tilde{\mathcal{H}}_{\mathsf{B}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}}$ *such that, with* $V = V_{\mathsf{A}} \otimes V_{\mathsf{B}}$*, we have*

$$[V(|x\rangle \otimes \Pi^x(y)) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = |x\rangle \otimes \tilde{\Pi}^x(y)|\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle\,. \tag{18}$$

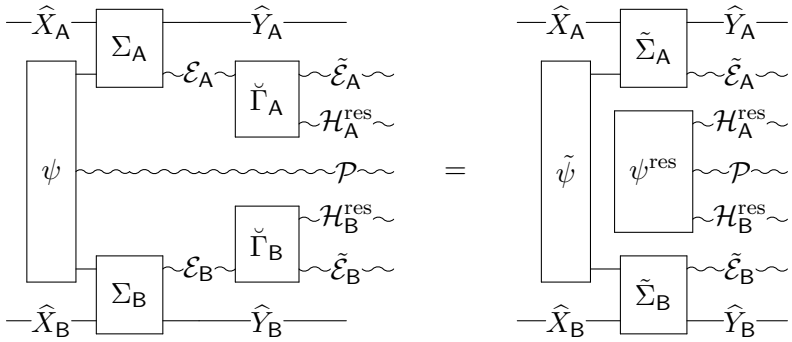
Remark 4.15. Note that, contrary to the conventional notion of reducibility (in Definition 3.5), the local isometries V_{P} in Eq. (18) are allowed to 'look at x_{P}' on the left hand side, and must as well 'return x_{P}' on the right hand side. This distinction is conceptually interesting, and we suspect it to be important for approximate versions of local simulation (Remark 4.18); in the exact case, however, it turns out to vanish (Lemma 4.17). ✠

Proof (of Proposition 4.14). There are two key ideas in this proof. By Lemma 4.11, the condition $S \geq_{l.s.} \tilde{S}$ is equivalent, for any choice of Stinespring dilations Σ_{A}, Σ_{B}, $\tilde{\Sigma}_{\mathsf{A}}, \tilde{\Sigma}_{\mathsf{B}}$ of $\Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}}$, to the existence of Γ_{A} and Γ_{B} such that



(19)

(the trace $\mathrm{tr}_{\mathcal{P}}$ is the only possible choice for Γ_0). The first idea is to realise that this condition is equivalent to the existence of isometric channels $\breve{\Gamma}_{\mathsf{A}}$, $\breve{\Gamma}_{\mathsf{B}}$ and a pure state ψ^{res} such that



. (20)

Indeed, Eq. (19) follows trivially from Eq. (20) by tracing out $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}}$ and $\mathcal{H}_{\mathsf{B}}^{\mathrm{res}}$. More interestingly, Eq. (20) follows from Eq. (19) by observing that if on the left hand of Eq. (19) we replace the trace by an identity and, for each $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the channels Γ_{P} by Stinespring dilations $\breve{\Gamma}_{\mathsf{P}}$, then the resulting channel is a dilation of the right hand side, which is isometric and therefore by Corollary 2.24 dilationally pure; consequently, that channel must take the form of tensoring with some state ψ^{res}, and this state must be pure since the total channel is now isometric. All in all Eq. (20) thus follows. It is worth highlighting that the condition (20) may be thought of as a 'purified' recharacterisation of local simulation, which might have been worthy of a separate lemma, given unbounded space.

Given the purified characterisation (20), the second idea in the proof is to choose the Stinespring dilations $\Sigma_{\mathsf{A}}, \Sigma_{\mathsf{B}}$, $\tilde{\Sigma}_{\mathsf{A}}$ and $\tilde{\Sigma}_{\mathsf{B}}$ conveniently. For $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the projective measurement ensemble $\Lambda_{\mathsf{P}} : \widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \to \widehat{Y}_{\mathsf{P}}$ corresponding to the PVM Π_{P} is given by

$$\Lambda_{\mathsf{P}}(F \otimes G) = \sum_{x_{\mathsf{P}} \in X_{\mathsf{P}}, y_{\mathsf{P}} \in Y_{\mathsf{P}}} \langle x_{\mathsf{P}}| F |x_{\mathsf{P}}\rangle \operatorname{tr}(\Pi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})G) \, |y_{\mathsf{P}}\rangle\langle y_{\mathsf{P}}| \text{ for } F \in \operatorname{End}(\widehat{X}_{\mathsf{P}}),\, G \in \operatorname{End}(\mathcal{H}_{\mathsf{P}}). \tag{21}$$

It can be easily checked that the operator $R_{\mathsf{P}} := \sum_{x_{\mathsf{P}} \in X_{\mathsf{P}}, y_{\mathsf{P}} \in Y_{\mathsf{P}}} |x_{\mathsf{P}}\rangle \otimes \Pi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}}) \otimes |y_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}|$ is an isometry from $\widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}$ to $\widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \otimes \widehat{Y_{\mathsf{P}}}$ (exploiting that each $\Pi_{\mathsf{P}}^{x_{\mathsf{P}}}$ is a PVM). Furthermore, the isometric channel $\Sigma_{\mathsf{P}}(\cdot) := R_{\mathsf{P}}(\cdot)R_{\mathsf{P}}^*$ is a dilation of Λ_{P} with environment $\mathcal{E}_{\mathsf{P}} := \widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}$. Hence, Σ_{P} is a Stinespring dilation of Λ_{P}. By the same argument, $\tilde{R}_{\mathsf{P}} := \sum_{x_{\mathsf{P}} \in X_{\mathsf{P}}, y_{\mathsf{P}} \in Y_{\mathsf{P}}} |x_{\mathsf{P}}\rangle \otimes \tilde{\Pi}_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}}) \otimes |y_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}|$ is an isometry giving rise to a Stinespring dilation $\tilde{\Sigma}_{\mathsf{P}}$ of $\tilde{\Lambda}_{\mathsf{P}}$, the measurement ensemble corresponding to the PVM $\tilde{\Pi}_{\mathsf{P}}$.

With the above choices for Σ_{P} and $\tilde{\Sigma}_{\mathsf{P}}$, the two Stinespring implementations of S and $\tilde{S}$ are the isometric channels represented by the isometries

$$\sum_{x \in X, y \in Y} |x\rangle \otimes [\Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle\langle x| \,, \quad \text{respectively} \quad \sum_{x \in X, y \in Y} |x\rangle \otimes \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle \otimes |y\rangle\langle x| \,. \tag{22}$$

As such, the condition (20) expresses the existence of isometries $V_{\mathsf{P}} : \widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \to \widehat{X}_{\mathsf{P}} \otimes \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ (representing $\breve{\Gamma}_{\mathsf{P}}$) and a unit vector $|\psi^{\mathrm{res}}\rangle \in \mathcal{H}_{\mathsf{A}}^{\mathrm{res}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}} \otimes \mathcal{P}$ (representing ψ^{res}) such that

$$\sum_{x \in X, y \in Y} [V(|x\rangle \otimes \Pi^x(y)) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle\langle x| = \sum_{x \in X, y \in Y} |x\rangle \otimes \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle \otimes \, |\psi^{\mathrm{res}}\rangle \otimes |y\rangle\langle x| \,, \tag{23}$$

with $V = V_{\mathsf{A}} \otimes V_{\mathsf{B}}$. Since $\big(\, |y\rangle\langle x| \,\big)_{x \in X, y \in Y}$ forms a linearly independent system of operators, Eq. (23) is equivalent to the condition (18), as desired. □

Remark 4.16 (Simulation for Non-Projective Strategies)**.** It is possible to state a version of Proposition 4.14 for non-projective strategies. The only step in the proof where we relied on the strategies being projective was when we chose the Stinespring dilations Σ_{P} and $\tilde{\Sigma}_{\mathsf{P}}$ of the measurement ensembles Λ_{P} and $\tilde{\Lambda}_{\mathsf{P}}$. Suppose now that the measurement ensembles $\tilde{\Lambda}_{\mathsf{A}}$ and $\tilde{\Lambda}_{\mathsf{B}}$ are still projective, but that, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the ensemble Λ_{P} corresponds to general POVMs $(E_{\mathsf{P}}^{x_{\mathsf{P}}})_{x_{\mathsf{P}} \in X_{\mathsf{P}}}$. For the Stinespring dilation Σ_{P}, we are then forced to include in its environment a copy of $\widehat{Y}_{\mathsf{P}}$, choosing as Stinespring isometry from $\widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}$ to $\widehat{X}_{\mathsf{P}} \otimes \widehat{Y}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \otimes \widehat{Y}_{\mathsf{P}}$ the operator

$$R_{\mathsf{P}} := \sum_{x_{\mathsf{P}} \in X_{\mathsf{P}}, y_{\mathsf{P}} \in Y_{\mathsf{P}}} |x_{\mathsf{P}}\rangle \otimes |y_{\mathsf{P}}\rangle \otimes \sqrt{E_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})} \otimes |y_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}| \,. \tag{24}$$

As such, the Stinespring implementation corresponding to $(\psi, \Sigma_{\mathsf{A}}, \Sigma_{\mathsf{B}})$ is represented by the isometry $\sum_{x \in X, y \in Y} |x\rangle \otimes |y\rangle \otimes [\sqrt{E^x(y)} \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle\langle x|$, and the simulation condition translates as the existence of a unit vector $|\psi^{\mathrm{res}}\rangle \in \mathcal{H}_{\mathsf{A}}^{\mathrm{res}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}} \otimes \mathcal{P}$ and isometries $V_{\mathsf{P}} : \widehat{X}_{\mathsf{P}} \otimes \widehat{Y}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \to \widehat{X}_{\mathsf{P}} \otimes \widehat{Y}_{\mathsf{P}} \otimes \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ such that, with $V = V_{\mathsf{A}} \otimes V_{\mathsf{B}}$, we have

$$\forall x \in X, y \in Y : \; [V(|x\rangle \otimes |y\rangle \otimes \sqrt{E^x(y)}) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle = |x\rangle \otimes |y\rangle \otimes \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle \,. \tag{25}$$

✠

Proposition 4.14 is the link between our operational notion of local simulation and the conventional operator-algebraic notion of reducibility — indeed, the condition (18) looks very similar to the reducibility condition (12). Ignoring for a moment the difference that (18) allows the local isometry to observe x_{P}, the proof of Proposition 4.14 explains in an entirely new way the significance of the residual state, the embedding isometries and the projected states $\Pi^x(y)\,|\psi\rangle$ and $\tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle$: the projected states signify side-information generated in the course of implementing the strategies, and the embedding connects side-information in one strategy to side-information in the other strategy.

We shall now prove that the difference between reducibility and condition (18) is (at least under a rank-assumption) only one of appearance. This result marks the technical highlight of the subsection, by finally establishing Theorem 4.8. We state in the result three equivalent conditions, of which the first is Eq. (18) and the third is the conventional reducibility condition (12). The second is an intermediate condition which mostly serves to smoothen the proof.

Lemma 4.17. *(Technical Result for Theorem 4.8).*
Let $S = (\varrho, \Pi_{\mathsf{A}}, \Pi_{\mathsf{B}})$ and $\tilde{S} = (\tilde{\psi}, \tilde{\Pi}_{\mathsf{A}}, \tilde{\Pi}_{\mathsf{B}})$ be projective strategies, for which $\tilde{\psi}$ is a pure state with locally full rank. Let ψ be a purification of ϱ with purifying space $\mathcal{P}$. Finally, let ψ^{res} be a pure state on a system of the form $\mathcal{H}_{\mathsf{A}}^{\mathrm{res}} \otimes \mathcal{H}_{\mathsf{B}}^{\mathrm{res}} \otimes \mathcal{P}$. The following are equivalent:

(1) There exist isometries $V_{\mathsf{P}} : \widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}} \to \widehat{X}_{\mathsf{P}} \otimes \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ such that, with $V = V_{\mathsf{A}} \otimes V_{\mathsf{B}}$, we have

$$\forall x \in X, y \in Y : \; [V(|x\rangle \otimes \Pi^x(y)) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = |x\rangle \otimes \tilde{\Pi}^x(y)|\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle\,. \tag{26}$$

(2) There exist isometries $W_{\mathsf{P}}^{x_{\mathsf{P}}} : \mathcal{H}_{\mathsf{P}} \to \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ for $x_{\mathsf{P}} \in X_{\mathsf{P}}$, such that, with $W^x = W_{\mathsf{A}}^{x_{\mathsf{A}}} \otimes W_{\mathsf{B}}^{x_{\mathsf{B}}}$ for $x = (x_{\mathsf{A}}, x_{\mathsf{B}})$, we have

$$\forall x \in X, y \in Y : \; [W^x \Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = \tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle \otimes \; |\psi^{\mathrm{res}}\rangle\,. \tag{27}$$

(3) There exist isometries $W_{\mathsf{P}} : \mathcal{H}_{\mathsf{P}} \to \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ such that, with $W = W_{\mathsf{A}} \otimes W_{\mathsf{B}}$, we have

$$\forall x \in X, y \in Y : \; [W \Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = \tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle \otimes \; |\psi^{\mathrm{res}}\rangle\,. \tag{28}$$

Proof. It is clear that (3) implies (2), by letting $W_{\mathsf{P}}^{x_{\mathsf{P}}} = W_{\mathsf{P}}$ independently of x_{P}. It is also clear that (2) implies (1), by letting $V_{\mathsf{P}} = \sum_{x_{\mathsf{P}} \in X_{\mathsf{P}}} |x_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}| \otimes W_{\mathsf{P}}^{x_{\mathsf{P}}}$. We show that (1) implies (2), and that (2) implies (3). We begin, however, by showing that if (1) holds then for each party $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, the local support $\mathrm{supp}(\varrho_{\mathsf{P}})$ of the P-marginal of ϱ is invariant under all of the operators $\Pi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})$, i.e. $\Pi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})[\mathrm{supp}(\varrho_{\mathsf{P}})] \subseteq \mathrm{supp}(\varrho_{\mathsf{P}})$ for $x_{\mathsf{P}} \in X_{\mathsf{P}}$ and $y_{\mathsf{P}} \in Y_{\mathsf{P}}$.

Let us assume without loss of generality that $\mathsf{P} = \mathsf{A}$. If in Eq. (26) we sum over $y_{\mathsf{B}} \in Y_{\mathsf{B}}$, then we obtain

$$\forall x \in X, y_{\mathsf{A}} \in Y_{\mathsf{A}} : \; [V(|x\rangle \otimes \Pi_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}}) \otimes \mathbb{1}_{\mathcal{H}_{\mathsf{B}}}) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = |x\rangle \otimes [\tilde{\Pi}_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}}) \otimes \mathbb{1}_{\tilde{\mathcal{H}}_{\mathsf{B}}}]|\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle\,. \tag{29}$$

By marginalising to the A-systems, Eq. (29) implies that

$$V_{\mathsf{A}}(\,|x_{\mathsf{A}}\rangle\langle x'_{\mathsf{A}}| \otimes \Pi_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}})\varrho_{\mathsf{A}}\Pi_{\mathsf{A}}^{x'_{\mathsf{A}}}(y'_{\mathsf{A}}))V_{\mathsf{A}}^* = |x_{\mathsf{A}}\rangle\langle x'_{\mathsf{A}}| \otimes \tilde{\Pi}_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}})\tilde{\varrho}_{\mathsf{A}}\tilde{\Pi}_{\mathsf{A}}^{x'_{\mathsf{A}}}(y'_{\mathsf{A}}) \otimes \varrho_{\mathsf{A}}^{\mathrm{res}} \tag{30}$$

for all $x_\mathsf{A}, x'_\mathsf{A} \in X_\mathsf{A}$ and all $y_\mathsf{A}, y'_\mathsf{A} \in Y_\mathsf{A}$. Here, $\tilde{\varrho}_\mathsf{A}$ denotes the A-marginal of $\tilde{\psi}$ and $\varrho_\mathsf{A}^{\mathrm{res}}$ the A-marginal of ψ^{res}. If in Eq. (30) we let $x'_\mathsf{A} = x_\mathsf{A}$ and sum over $y'_\mathsf{A} \in Y_\mathsf{A}$, then we obtain

$$V_\mathsf{A}(\,|x_\mathsf{A}\rangle\langle x_\mathsf{A}| \otimes \Pi_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})\varrho_\mathsf{A})V_\mathsf{A}^* = |x_\mathsf{A}\rangle\langle x_\mathsf{A}| \otimes \tilde{\Pi}_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})\tilde{\varrho}_\mathsf{A} \otimes \varrho_\mathsf{A}^{\mathrm{res}} \tag{31}$$

for all $x_\mathsf{A} \in X, y_\mathsf{A} \in Y_\mathsf{A}$. If additionally we sum over $y_\mathsf{A} \in Y_\mathsf{A}$, we obtain

$$V_\mathsf{A}(\,|x_\mathsf{A}\rangle\langle x_\mathsf{A}| \otimes \varrho_\mathsf{A})V_\mathsf{A}^* = |x_\mathsf{A}\rangle\langle x_\mathsf{A}| \otimes \tilde{\varrho}_\mathsf{A} \otimes \varrho_\mathsf{A}^{\mathrm{res}} \tag{32}$$

for all $x_\mathsf{A} \in X_\mathsf{A}$. Now fix, $x_\mathsf{A} \in X_\mathsf{A}$ and $y_\mathsf{A} \in Y_\mathsf{A}$. What we wish to show is that $\Pi^{x_\mathsf{A}}(y_\mathsf{A})_\mathsf{P}[\mathrm{supp}(\varrho_\mathsf{A})] \subseteq \mathrm{supp}(\varrho_\mathsf{A})$, and this amounts to showing that $\mathrm{Im}\,\Pi_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})\varrho_\mathsf{A} \subseteq \mathrm{Im}\,\varrho_\mathsf{A}$ where 'Im' denotes the image (range) of an operator. By virtue of Eqs. (31) and (32), however, this follows if we can show that $\mathrm{Im}\,\tilde{\Pi}_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})\tilde{\varrho}_\mathsf{A} \subseteq \mathrm{Im}\,\tilde{\varrho}_\mathsf{A}$. But that follows directly from the assumption that $\tilde{\varrho}_\mathsf{A}$ has full rank. (This is the only step in the proof where we need the rank-assumption on $\tilde{\varrho}$.)[29]

With this conclusion in place, we are ready to prove that (1) implies (2), and that (2) implies (3).

First, assume that (1) holds, with isometries V_A and V_B. For $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$ and $x_\mathsf{P} \in X_\mathsf{P}$, let

$$W_\mathsf{P}^{x_\mathsf{P}} := (\mathbb{1}_{\tilde{\mathcal{H}}_\mathsf{P} \otimes \mathcal{H}_\mathsf{P}^{\mathrm{res}}} \otimes \langle x_\mathsf{P}|)V_\mathsf{P}(\mathbb{1}_{\mathcal{H}_\mathsf{P}} \otimes |x_\mathsf{P}\rangle). \tag{33}$$

It follows directly from Eq. (26) that Eq. (27) holds, however it is not clear that the operators $W_\mathsf{P}^{x_\mathsf{P}}$ are isometries. Fix $x_\mathsf{A} \in X_\mathsf{A}$ and $x_\mathsf{B} \in X_\mathsf{B}$. It is enough to show that $W_\mathsf{P}^{x_\mathsf{P}}$ acts isometrically on the subspace $\mathcal{H}_\mathsf{P}^0 \subseteq \mathcal{H}_\mathsf{P}$ generated[30] from $\mathrm{supp}(\varrho_\mathsf{P})$ by the operators $\left(\Pi_\mathsf{P}^{x_\mathsf{P}}(y_\mathsf{P})\right)_{y_\mathsf{P} \in Y_\mathsf{P}}$, since we may always modify the action of $W_\mathsf{P}^{x_\mathsf{P}}$ outside this subspace. However, by our introductory observation the subspace $\mathcal{H}_\mathsf{P}^0$ is simply $\mathrm{supp}(\varrho_\mathsf{P})$ itself. To show that $W_\mathsf{P}^{x_\mathsf{P}}$ acts isometrically on $\mathrm{supp}(\varrho_\mathsf{P})$, observe that

$$[W_\mathsf{A}^{x_\mathsf{A}} \otimes W_\mathsf{B}^{x_\mathsf{B}} \otimes \mathbb{1}_\mathcal{P}]\,|\psi\rangle = |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle\,, \tag{34}$$

by summing over $y \in Y$ in Eq. (27). Consequently,

$$\langle\psi|\,[|W_\mathsf{A}^{x_\mathsf{A}}|^2 \otimes |W_\mathsf{B}^{x_\mathsf{B}}|^2 \otimes \mathbb{1}_\mathcal{P}]\,|\psi\rangle = 1, \tag{35}$$

where $|W_\mathsf{P}^{x_\mathsf{P}}|^2 = (W_\mathsf{P}^{x_\mathsf{P}})^* W_\mathsf{P}^{x_\mathsf{P}}$. The operators $|W_\mathsf{P}^{x_\mathsf{P}}|^2$ are positive, and by definition of $W_\mathsf{P}^{x_\mathsf{P}}$ satisfy $|W_\mathsf{P}^{x_\mathsf{P}}|^2 \leq \mathbb{1}_{\mathcal{H}_\mathsf{P}}$. Therefore, Eq. (35) entails that $|W_\mathsf{P}^{x_\mathsf{P}}|^2$ acts as $\mathbb{1}_{\mathcal{H}_\mathsf{P}}$ on the subspace $\mathrm{supp}(\varrho_\mathsf{P})$, which is precisely to say that $W_\mathsf{P}^{x_\mathsf{P}}$ acts isometrically on this subspace. Altogether, (2) holds as desired.

Next, assume that (2) holds; we show (3). Since (2) implies (1), our introductory observation still applies: the subspace $\mathcal{H}_\mathsf{P}^0$ generated from $\mathrm{supp}(\varrho_\mathsf{P})$ by the operators $\Pi_\mathsf{P}^{x_\mathsf{P}}(y_\mathsf{P})$ coincides with $\mathrm{supp}(\varrho_\mathsf{P})$ itself. Therefore, in order to show the existence of x_P-independent isometries W_P satisfying Eq. (28), it suffices to show that for any two $x_\mathsf{P}, x'_\mathsf{P} \in X_\mathsf{P}$, the isometries $W_\mathsf{P}^{x_\mathsf{P}}$ and $W_\mathsf{P}^{x'_\mathsf{P}}$ from Eq. (27) act identically on the subspace $\mathrm{supp}(\varrho_\mathsf{P})$. Let us assume without loss of generality that $\mathsf{P} = \mathsf{A}$. Let $\sum_{j=1}^r \sqrt{p(j)}\,|\psi_\mathsf{A}(j)\rangle \otimes |\psi_{\neg\mathsf{A}}(j)\rangle$

[29] Moreover, the conclusion relies only on the weaker circumstance that $\mathrm{supp}(\tilde{\varrho}_\mathsf{A})$ is invariant under the operators $\tilde{\Pi}_\mathsf{A}^{x_\mathsf{A}}(y_\mathsf{A})$ for $x_\mathsf{A} \in X_\mathsf{A}, y_\mathsf{A} \in Y_\mathsf{A}$, and similarly for B. Thus, the lemma could alternatively have been stated under this weaker assumption.

[30] The subspace *generated* from a subspace $\mathcal{K}$ by a family of operators $(F_i)_{i\in I}$ is by definition the smallest subspace containing $\mathcal{K}$ and invariant under F_i for every $i \in I$.

be a Schmidt decomposition of $|\psi\rangle$ with $p(1), \dots, p(r) > 0$, relative to the factorisation $\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\neg\mathsf{A}}$, where $\mathcal{H}_{\neg\mathsf{A}} = \mathcal{H}_{\mathsf{B}} \otimes \mathcal{P}$. The condition (2) entails Eq. (34), which now reads

$$\sum_{j=1}^{r} \sqrt{p(j)} W^{x_{\mathsf{A}}} |\psi_{\mathsf{A}}(j)\rangle \otimes [W_{\mathsf{B}}^{x_{\mathsf{B}}} \otimes \mathbb{1}_{\mathcal{P}}] |\psi_{\neg\mathsf{A}}(j)\rangle = |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle \tag{36}$$

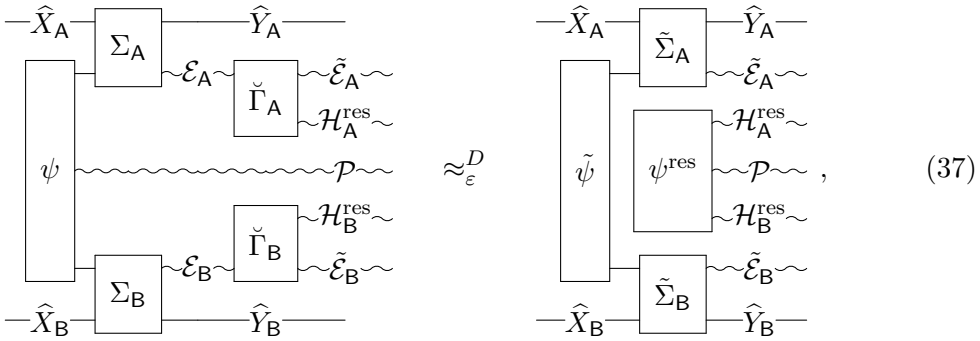
for any $x_{\mathsf{A}} \in X_{\mathsf{A}}$ and $x_{\mathsf{B}} \in X_{\mathsf{B}}$. Fix some $x_{\mathsf{B}} \in X_{\mathsf{B}}$. By Eq. (36), the vector $W^{x_{\mathsf{A}}} |\psi_{\mathsf{A}}(j)\rangle$ must be independent of x_{A}, for each $j = 1, \dots, r$; this follows by applying the operator $\mathbb{1}_{\tilde{\mathcal{H}}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{A}}^{\mathrm{res}}} \otimes \langle \psi_{\neg\mathsf{A}}(j)| [(W_{\mathsf{B}}^{x_{\mathsf{B}}})^* \otimes \mathbb{1}_{\mathcal{P}}]$ to both sides. Observing that $\mathrm{span}\{|\psi_{\mathsf{A}}(j)\rangle \mid j = 1, \dots, r\} = \mathrm{supp}(\varrho_{\mathsf{A}})$, we thus conclude that the operator $W_{\mathsf{A}}^{x_{\mathsf{A}}}$ restricted to $\mathrm{supp}(\varrho_{\mathsf{A}})$ is independent of x_{A}. This shows the desired, and altogether finishes the proof. □

Lemma 4.17 concludes the presentation of our new operational self-testing definition and its relation to conventional self-testing. Let us end by briefly considering how our operational definition might lend itself to the setting of robust self-testing.

Remark 4.18 (Approximate Simulation and Robust Self-Testing)**.** The aim of *robust self-testing* [MYS12, ŠB20] is to establish that if a strategy S has a behaviour which is 'close' to the behaviour P, then S is 'close' to being reducible to some canonical strategy $\tilde{S}$. This notion calls for a definition of approximate reducibility, and this is conventionally achieved by considering the Hilbert space norm of the difference between the left and right hand sides in the reducibility condition (12). Of course, the operational interpretation of this may be even more questionable than that of the exact case, and since (12) comprises several identities (one for each $x \in X$ and $y \in Y$) it is not obvious how to form a single numerical quantity which indicates closeness.

Our operational definition of local simulation, on the other hand, quite naturally lends itself to approximate generalisations. If D is a metric on quantum channels, then Definition 4.2 suggests the following notion of *approximate local simulation (w.r.t. D)*: For $\varepsilon \geq 0$, we say that S *locally ε-simulates $\tilde{S}$ (w.r.t. D)*, if for any implementation of $\tilde{S}$ there exists an implementation of S which is ε-close to it, as measured by D.

If the metric D is sufficiently well-behaved, then the chain of arguments which proved the equivalence between Definition 4.2 and the 'purified' condition (20) will carry over to the approximate setting. Specifically (assuming for simplicity that $\tilde{S}$ is pure-state), it will hold that S locally ε-simulates $\tilde{S}$ if and only if there exist isometric channels $\breve{\Gamma}_{\mathsf{A}}$, $\breve{\Gamma}_{\mathsf{B}}$ and a pure state ψ^{res} such that



, (37)

where '$\approx_{\varepsilon}^{D}$' signifies that the two channels are ε-close as measured by D. Examples of a metric which is well-behaved in this sense are the *Bures distance* [KSW08a, KSW08b] and the *purified diamond-distance* [HL21] on quantum channels. (For a systematic discussion

of metrics and their properties, see Chapter 3 in Ref. [HL21].[31])

Now, in a usual self-testing scenario, the inputs provided to the classical embedded systems $\widehat{X}_{\mathsf{A}}$ and $\widehat{X}_{\mathsf{B}}$ are all classical, so it is natural to consider instead of Eq. (37) the weaker condition $\max_{x\in X} D(\phi_x, \tilde{\phi}_x) \leq \varepsilon$, where ϕ_x and $\tilde{\phi}_x$ are the pure states which result from inserting input $|x\rangle\langle x| = |x_{\mathsf{A}}\rangle\langle x_{\mathsf{A}}| \otimes |x_{\mathsf{B}}\rangle\langle x_{\mathsf{B}}|$ into the channels on the left and right hand sides of Eq. (37), respectively. In other words, if $V_{\mathsf{P}} : \mathcal{E}_{\mathsf{A}} \to \tilde{\mathcal{E}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ are isometries representing $\breve{\Gamma}_{\mathsf{P}}$, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, and if $V = V_{\mathsf{A}} \otimes V_{\mathsf{B}}$, then the condition

$$\inf_{V_{\mathsf{A}}, V_{\mathsf{B}}, |\psi^{\mathrm{res}}\rangle} \max_{x\in X} D(\phi_x, \tilde{\phi}_x) \leq \varepsilon \tag{38}$$

is indicative of local ε-simulation of $\tilde{S}$ by S, where

$$|\phi_x\rangle = \sum_{y\in Y} [V(|x\rangle \otimes \Pi^x(y)) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle \quad \text{and} \quad |\tilde{\phi}_x\rangle = \sum_{y\in Y} |x\rangle \otimes \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle \otimes |y\rangle. \tag{39}$$

(We here assume the convenient choice of Stinespring dilations from the proof of Proposition 4.14, cf. Eqs. (22) and (23). In particular, $\mathcal{E}_{\mathsf{P}} = \widehat{X}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}$ and $\tilde{\mathcal{E}}_{\mathsf{P}} = \widehat{X}_{\mathsf{P}} \otimes \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$.)

The details of the condition (38) now of course depend on the metric D, but a general phenomenon is that the local isometry V_{P} is allowed to observe the local input x_{P}. More formally, suppose we restrict the infimum in Eq. (38) to isometries of the form $V_{\mathsf{P}} = \sum_{x_{\mathsf{P}}\in X_{\mathsf{P}}} |x_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}| \otimes W_{\mathsf{P}}^{x_{\mathsf{P}}}$ with $W_{\mathsf{P}}^{x_{\mathsf{P}}} : \mathcal{H}_{\mathsf{P}} \to \tilde{\mathcal{H}}_{\mathsf{P}} \otimes \mathcal{H}_{\mathsf{P}}^{\mathrm{res}}$ an isometry for each $x_{\mathsf{P}} \in X_{\mathsf{P}}$. We then have $|\phi_x\rangle = \sum_{y\in Y} |x\rangle \otimes [W^x \Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle$, so $|x\rangle$ factors out from both $|\phi_x\rangle$ and $|\tilde{\phi}_x\rangle$. Natural metrics D will be invariant under a fixed tensoring with $|x\rangle$, so the condition (38) holds if merely

$$\inf_{\substack{(W_{\mathsf{A}}^{x_{\mathsf{A}}})_{x_{\mathsf{A}}\in X_{\mathsf{A}}}, \\ (W_{\mathsf{B}}^{x_{\mathsf{B}}})_{x_{\mathsf{B}}\in X_{\mathsf{B}}}, |\psi^{\mathrm{res}}\rangle}} \max_{x\in X} D\left(\sum_{y\in Y} [W^x \Pi^x(y)) \otimes \mathbb{1}_{\mathcal{P}}] \, |\psi\rangle \otimes |y\rangle \, , \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle \otimes |\psi^{\mathrm{res}}\rangle \otimes |y\rangle \right) \leq \varepsilon. \tag{40}$$

Now, the infimum in Eq. (40) might be notably smaller than the infimum over x_{P}-independent isometries W_{P}, indeed the argument used in the proof of Lemma 4.17 to eliminate x_{P}-dependence in the exact case ($\varepsilon = 0$) does not immediately carry over to the general approximate case ($\varepsilon > 0$). Thus, our notion of approximate simulation might ultimately yield robust self-testing results genuinely different from ones based on the conventional reducibility definition with input-independent isometries.

To finish off with an idea of what the quantity (40) might be like, let us take D as the purified diamond distance $P_\diamond$ [HL21], so that the D-distances appearing in Eq. (40) are purified distances between states [TCR10, Tom12]. As the involved states are already pure, the purified distances coincide with trace distances, and the trace distance between pure states is always upper bounded by the Hilbert space distance between vector representatives.[32]

[31] In fact, it follows from the monotonicity and dilational properties considered therein that ε-approximate version of Definition 4.2 is equivalent to Eq. (37) for some isometric channels $\breve{\Gamma}_{\mathsf{A}}$, $\breve{\Gamma}_{\mathsf{B}}$ and some (not necessarily pure) state ψ^{res}; this is a straightforward modification of the exact argument. One can then argue that the state ψ^{res} may be assumed pure without increasing the distance.

[32] In fact, we have for unit vectors $|\chi\rangle$ and $|\chi'\rangle$ the identity $\delta_1(|\chi\rangle, |\chi'\rangle) = \frac{1+|\langle\chi|\chi'\rangle|}{2} \inf_{\theta\in[0,2\pi)} \left\| |\chi\rangle - e^{i\theta} |\chi'\rangle \right\|_2$, so we have $\delta_1(|\chi\rangle, |\chi'\rangle) \approx \inf_{\theta\in[0,2\pi)} \left\| |\chi\rangle - e^{i\theta} |\chi'\rangle \right\|_2$ when $|\chi\rangle$ and $|\chi'\rangle$ have large overlap.

Hence,

$$D\left(\sum_{y\in Y}\left[W^x\Pi^x(y))\otimes\mathbb{1}_{\mathcal{P}}\right]|\psi\rangle\otimes|y\rangle\,,\tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle\otimes\,|\psi^{\mathrm{res}}\rangle\otimes|y\rangle\right) \tag{41}$$

$$\leq\left\|\sum_{y\in Y}\left[W^x\Pi^x(y))\otimes\mathbb{1}_{\mathcal{P}}\right]|\psi\rangle\otimes|y\rangle-\tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle\otimes\,|\psi^{\mathrm{res}}\rangle\otimes|y\rangle\right\|_2 \tag{42}$$

$$=\sqrt{\sum_{y\in Y}\left\|\left[W^x\Pi^x(y))\otimes\mathbb{1}_{\mathcal{P}}\right]|\psi\rangle-\tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle\otimes\,|\psi^{\mathrm{res}}\rangle\right\|_2^2}, \tag{43}$$

and (40) therefore follows if

$$\inf_{\substack{(W_{\mathsf{A}}^{x_{\mathsf{A}}})_{x_{\mathsf{A}}\in X_{\mathsf{A}}},\\(W_{\mathsf{B}}^{x_{\mathsf{B}}})_{x_{\mathsf{B}}\in X_{\mathsf{B}}},|\psi^{\mathrm{res}}\rangle}}\ \max_{x\in X}\sqrt{\sum_{y\in Y}\left\|\left[W^x\Pi^x(y))\otimes\mathbb{1}_{\mathcal{P}}\right]|\psi\rangle-\tilde{\Pi}^x(y)\,|\tilde{\psi}\rangle\otimes\,|\psi^{\mathrm{res}}\rangle\right\|_2^2}\leq\varepsilon. \tag{44}$$

✠

For other considerations about the metric aspects of simulation, in particular as regards the development to come in Section 4.C, the reader is referred to the introductory remarks in Chapters 4 and 5 of Ref. [HL21]. In particular, it is discussed there that an additional potentially desirable property of a metric is monotonicity under *contraction* of certain input wires with output wires, corresponding to a non-increasing distance in Eq. (37) if e.g. we perform the procedure of processing and feeding the output from $\widehat{Y}_{\mathsf{A}}$ as input to $\widehat{X}_{\mathsf{B}}$.

4.B Symmetrising Simulation and Turning the Tables...

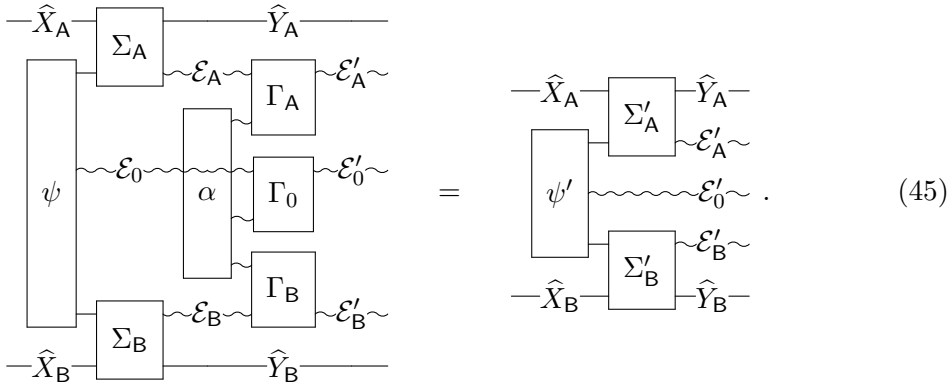
In Section 4.A, we introduced the operational notions of local simulation and of self-testing a.t.l.s., and we proved a formal relationship to the conventional operator-algebraic notions of reducibility and self-testing. From now on, we shall no longer look back to the operator-algebraic notions but rather work entirely with the operational counterparts.

This subsection serves to introduce a relaxation of local simulation, which we call *local assisted simulation* (Definition 4.19). Like local simulation, it is an operationally defined pre-order on strategies. By features of Stinespring dilations, however, local assisted simulation is in quantum theory in fact an equivalence relation (Theorem 4.23).

As such, the relation of local assisted simulation, which as far as we know has not been studied from an operator-algebraic viewpoint, uncovers quantum self-testing as an explicitly two-fold phenomenon — see in particular Remark 4.24, Definition 4.25 and the surrounding discussion.

Definition 4.19 (Local Assisted Simulation)**.** Let S and S' be strategies for $\mathfrak{B}$, with respective component-wise Stinespring dilations $(\psi,\Sigma_{\mathsf{A}},\Sigma_{\mathsf{B}})$ and $(\psi',\Sigma'_{\mathsf{A}},\Sigma'_{\mathsf{B}})$. We say that S *locally assisted simulates* S', written $S\geq_{l.a.s.}S'$, if there exist channels Γ_0, Γ_{A} and Γ_{B}

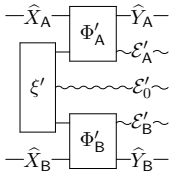
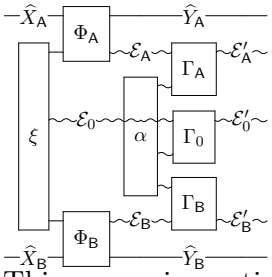
and some state α such that



(45)

(The reader is to imagine that the middle wiggly wire on the right hand side passes above the state α, so that the pictorial representation is really three-dimensional.) ■

The interpretation of local assisted simulation is that, using a pre-distributed state α as 'assistant', side-information of the strategy S' can be generated from side-information of the strategy S.

Remark 4.20 (On the Definition of Local Assisted Simulation)**.** Local assisted simulation can be given an alternative definition in the style of Definition 4.2, quantifying over general implementations: In fact, $S \geq_{l.a.s.} S'$ precisely if there exists a state α such that any general implementation  of S' is matched by  for some implementation of S and some choice of Γ_A, Γ_B and Γ_0. This means in particular that local assisted simulation can be defined in general operational theories (it also means that Definition 4.19 does not depend on the choice of Stinespring dilations). We chose the formulation in Definition 4.19 so as to avoid having to prove an equivalent of Lemma 4.11 afterwards. ✠

Proposition 4.21. *Local assisted simulation is a pre-order on the class of strategies for* $\mathfrak{B}$*,* $\mathrm{Strat}(\mathfrak{B})$*. It coarsens the pre-order of local simulation, i.e. the relation* $S \geq_{l.s.} S'$ *implies the relation* $S \geq_{l.a.s.} S'$*.*

Proof. Reflexivity and transitivity are obvious. The second statement follows by taking α product. □

Remark 4.22 (Local Assisted Simulation in Terms of Local Simulation)**.** Recall from Example 4.6 the concept of *augmenting* a strategy S by a state γ to form the new strategy $S[\gamma]$. It is worth observing that S locally assisted simulates S' with assistant α if and only if $S[\alpha_{\mathsf{AB}}]$ locally simulates S' (without assistance), where α_{AB} is the marginal state arising from α by discarding the middle system. When α is pure, this statement is visually evident from Eq. (45), by grouping ψ and α together as a single state which is then a purification of $\varrho \otimes \alpha_{\mathsf{AB}}$. It follows from this observation that $S \geq_{l.a.s.} S'$ if and only if there exists some state γ such that $S[\gamma] \geq_{l.s.} S'$. ✠

The following result relies on the features of Stinespring dilations exposed in Section 2.D:

Theorem 4.23. *In quantum theory,* $\geq_{l.a.s.}$ *is symmetric and thus an equivalence relation on* $\mathrm{Strat}(\mathfrak{B})$.

Proof. Suppose that $S \geq_{l.a.s.} S'$. As in the proof of Proposition 4.14, we may on the left hand side of Eq. (45) replace Γ_0, Γ_{A} and Γ_{B}, by respective Stinespring dilations $\breve{\Gamma}_0$, $\breve{\Gamma}_{\mathsf{A}}$, and $\breve{\Gamma}_{\mathsf{B}}$; the resulting channel is then a dilation of the right hand side, which is isometric and hence dilationally pure, and it must therefore equal the channel obtained by tensoring the right hand side with a (pure) state, say β. Now, since $\breve{\Gamma}_0$, $\breve{\Gamma}_{\mathsf{A}}$, and $\breve{\Gamma}_{\mathsf{B}}$ are isometric, they are reversible by Lemma 2.18, and can thus be cancelled by left-inverses $\breve{\Gamma}_0^-$, $\breve{\Gamma}_{\mathsf{A}}^-$, and $\breve{\Gamma}_{\mathsf{B}}^-$. If we do this and then trace out the state α, we exhibit the relation $S' \geq_{l.a.s.} S$ with β as assistant. This proves symmetry, as desired. □

By virtue of Theorem 4.23, we will sometimes use the notation $\sim_{l.a.s.}$ rather than $\geq_{l.a.s.}$. Specifically, we use the notation $\sim_{l.a.s.}$ when discussing the relation in quantum theory, whereas we resort to the original notation $\geq_{l.a.s.}$ when discussing the relation generally, where the order might matter.

Remark 4.24. Theorem 4.23 can be strengthened to assert that $\sim_{l.a.s.}$ is the equivalence relation *generated by* the pre-order $\geq_{l.s.}$, that is, it is the finest equivalence relation which coarsens $\geq_{l.s.}$. Indeed, let $\cong$ be any equivalence relation which coarsens $\geq_{l.s.}$. We must show that $S \sim_{l.a.s.} S'$ implies $S \cong S'$. By Remark 4.22, $S[\gamma] \geq_{l.s.} S'$ for some state γ. By Example 4.6, $S[\gamma] \geq_{l.s.} S$. As $\cong$ coarsens $\geq_{l.s.}$, these two conditions imply $S' \cong S[\gamma]$ and $S[\gamma] \cong S$, respectively. Consequently, $S \cong S'$. ✠

An important consequence of Theorem 4.23 is that local simulation implies equivalence under local assisted simulation. This fact substantiates the intuition of authors who use the term 'equivalence of strategies' when referring to reducibility (cf. Remark 3.9): though reducibility is not itself a symmetric relation, quantum strategies related by reducibility really are mathematically equivalent, because they are related by local simulation and thus equivalent under local assisted simulation. For example, consider the CHSH-game. Since the canonical strategy $\tilde{S}$ can be self-tested according to local simulation, our result implies that all strategies with the maximal violation are equivalent under local assisted simulation. According to Remark 4.22, every optimal strategy can therefore be locally simulated by the canonical one after tensoring an appropriate assisting state to the canonical (singlet) state.

From a different perspective, this observation reveals that quantum self-testing of a canonical strategy $\tilde{S}$ may be considered a two-fold phenomenon: it is not just the phenomenon that any strategy S with behaviour P is 'at least as strong as $\tilde{S}$' (namely, $S \geq_{l.s.} \tilde{S}$), it is also the phenomenon that any S with behaviour P is 'no stronger than $\tilde{S}$' (namely, $\tilde{S} \geq_{l.a.s.} S$). As such, self-testing on the one hand lower bounds the capabilities of unknown strategies and on the other hand upper bounds these same capabilities. In quantum theory, it just so happens that the lower bound implies an upper bound automatically.

To isolate the effect of the upper bound, it is beneficial to introduce the following notion of self-testing:

Definition 4.25 (Self-Testing according to Local Assisted Simulation.)**.** Let $\tilde{S}$ be any strategy with behaviour P. We say that P *self-tests* $\tilde{S}$ *according to local assisted simulation* (for short, *a.t.l.a.s.*) if any strategy S with behaviour P is locally assisted simulated by $\tilde{S}$, i.e. $S \leq_{l.a.s.} \tilde{S}$. ■

Remark 4.26. In quantum theory, self-testing a.t.l.a.s. is equally well defined by the requirements $S \leq_{l.a.s.} \tilde{S}$ or $S \sim_{l.a.s.} \tilde{S}$ for strategies S with behaviour P. Indeed, these two requirements are identical, by Theorem 4.23. In other theories, however, this may not be so and we intend the definition of self-testing a.t.l.a.s. to mean '$S \leq_{l.a.s.} \tilde{S}$' rather than equivalence under local assisted simulation. ✠

If P self-tests a strategy $\tilde{S}$ a.t.l.s., then P self-tests $\tilde{S}$ a.t.l.a.s. In fact, P in this case self-tests every strategy $\tilde{S}'$ with behaviour P a.t.l.a.s., since all such strategies are equivalent. We shall see in Section 5 that certain features of self-testing which have traditionally been attributed to the conventional reducibility relation can actually be derived from our weaker notion of self-testing a.t.l.a.s. (see in particular Proposition 5.6 and Proposition 5.14).

The following we leave as an open problem:

Open Problem 4.27. *Suppose that P self-tests a strategy $\tilde{S}$ a.t.l.a.s. Does there necessarily exist a strategy $\tilde{S}_0$ such that P self-tests $\tilde{S}_0$ a.t.l.s.?*

If Open Problem 4.27 has a positive answer, then quantum self-testing a.t.l.a.s. may be considered as fundamental as quantum self-testing a.t.l.s.(and a.t.r., by Theorem 4.9). The only difference between self-testing a.t.l.a.s. and a.t.l.s. is then that the latter highlights a certain canonical strategy as especially simple. We emphasize that the canonical CHSH-strategy (and indeed any other conventionally self-tested strategy) do not constitute a counterexample to the open problem, since these strategies already satisfy the strongest notion of self-testing according to local simulation.

4.C ... all the Way: Quantum Self-Testing in a Larger Perspective

We have now seen how to recast the conventional definition of quantum self-testing in an operational language (Section 4.A), and how this language uncovers quantum self-testing as a two-fold phenomenon (Section 4.B). The operational reformulation in Section 4.A suggests an interpretation in terms of side-information which escapes to an environment according to a particular causal structure. Our considerations in Section 4.B, meanwhile, led to the idea of a self-testing notion which aims exclusively to upper bound the power of an unknown strategy and is unconcerned with lower bounding its capabilities, cf. Definition 4.25. In this final subsection, we explain how to take these ideas even further by realising quantum self-testing in a general framework of causally structured circuits of channels and dilations.

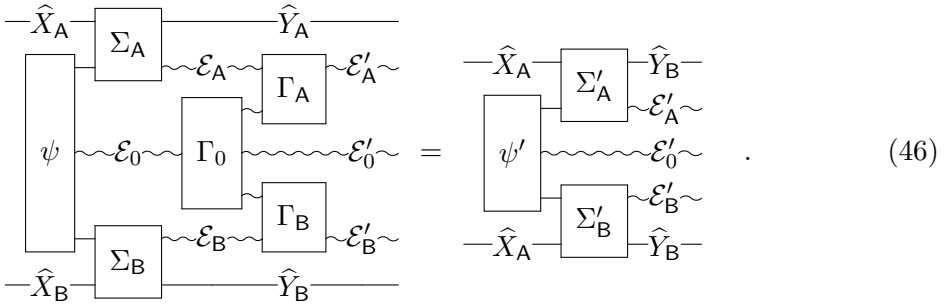
We start in Section 4.C.1 by introducing a relaxation of local assisted simulation, which we call *causal simulation* (Definition 4.28), and consider to be the ultimate mode of simulation among strategies. It gives rise to a notion of *self-testing according to causal simulation* (for short, *a.t.c.s.*), cf. Definition 4.30. When the canonical strategy $\tilde{S}$ is pure-state causal simulation coincides with local assisted simulation (Remark 4.31), however in general it is less restrictive as illustrated by Example 4.32. The point is now that self-testing a.t.c.s. is an instance of a phenomenon which belongs to a larger theory of *causal channels* and *derivability* among *causal dilations*, which we explain in Section 4.C.2 and Section 4.C.3. We show in Section 4.C.4 how to recover quantum self-testing according to causal simulation as a special case within this framework.

The presentation of the framework of causal channels and causal dilations and of its connection to quantum self-testing is brief and at points somewhat informal. For further details, and comparisons to other frameworks for networks of channels [CDP09, Per17, KU17], the reader is referred to Chapters 4 and 5 in Ref. [HL21].

4.C.1 Causal Simulation

Let us introduce the following relaxation of local assisted simulation:

Definition 4.28 (Causal Simulation)**.** Let S and S' be strategies for the Bell-scenario $\mathfrak{B}$, with respective component-wise Stinespring dilations $(\psi, \Sigma_\mathsf{A}, \Sigma_\mathsf{B})$ and $(\psi', \Sigma'_\mathsf{A}, \Sigma'_\mathsf{B})$. We say that S *causally simulates* S', written $S \geq_{c.s.} S'$, if there exist channels Γ_0, Γ_A and Γ_B such that



. (46)

■

Remark 4.29. As usual, the Stinespring implementations can be considered stand-ins for general implementations, and the definition of causal simulation is therefore meaningful in all operational theories (and independent of the choice of Stinespring dilations). ✠

The interpretation of causal simulation is that the side-information generated from S' can be derived from the side-information generated by S in a way which preserves causality. On the left hand side, the structure of the channel comprised by Γ_0, Γ_A and Γ_B is such that every output requires the exact same inputs as on the right hand side — therefore, the information can be leaked to the environment according to the same schedule on both sides.

Definition 4.30 (Self-Testing according to Causal Simulation.)**.** Let $\tilde{S}$ be any strategy with behaviour P. We say that P *self-tests* $\tilde{S}$ *according to causal simulation* (for short, *a.t.c.s.*) if any strategy S with behaviour P is causally simulated by $\tilde{S}$, i.e. $S \leq_{c.s.} \tilde{S}$. ■

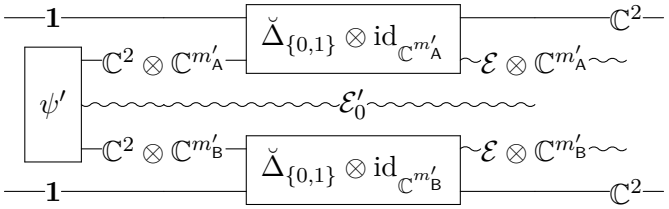
Remark 4.31 (Causal Simulation by Pure-State Strategies)**.** Note that when a $\tilde{S}$ is a pure-state strategy, the condition $\tilde{S} \geq_{c.s.} S$ is equivalent to the condition $\tilde{S} \geq_{l.a.s.} S$, for any strategy S. This is because the purifying system $\tilde{\mathcal{E}}_0$ of $\tilde{\varrho}$ may be taken trivial in a Stinespring implementation of $\tilde{S}$, and as such the channel Γ_0 witnessing causal simulation has trivial input and must in fact be a state. For this reason, it also holds that P self-tests $\tilde{S}$ a.t.l.c. if and only if P self-tests $\tilde{S}$ a.t.l.a.s. ✠

The following example serves to show that causal simulation is generally distinct from local assisted simulation. In fact, it constitutes an example of a strategy $\tilde{S}$ which is self-tested a.t.c.s., but which cannot possible be self-tested a.t.l.a.s. (in particular cannot be self-tested in the conventional operator-algebraic sense), since by Corollary 5.15 to be established later its behaviour would then be extremal.

Example 4.32 (Unconventional Self-Testing)**.** Consider the Bell-scenario $\mathfrak{B}$ for which $X_\mathsf{A} = X_\mathsf{B} = \{0\}$ and $Y_\mathsf{A} = Y_\mathsf{B} = \{0, 1\}$. It corresponds to an experimental set-up in which two separated parties are required to respond with bits $y_\mathsf{A} \in \{0, 1\}$ and $y_\mathsf{B} \in \{0, 1\}$. There is only one possible input for each party, and we interpret the local reception of this input as a query for the bit. A behaviour for the Bell-scenario $\mathfrak{B}$ is simply a probability distribution

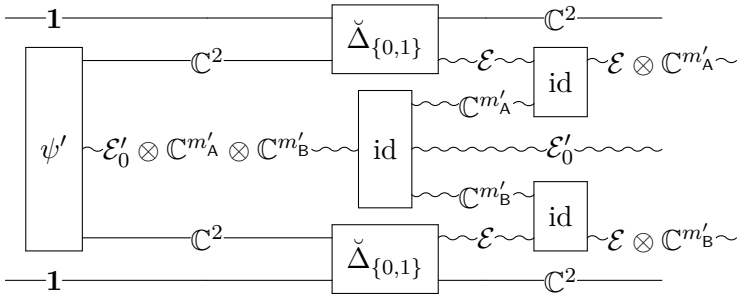
P on the set $Y_\mathsf{A} \times Y_\mathsf{B} = \{0,1\} \times \{0,1\}$. Let us consider the behaviour corresponding to perfectly correlated uniformly random outputs, i.e. the behaviour which as a classical embedded state on $\widehat{Y}_\mathsf{A} \otimes \widehat{Y}_\mathsf{B}$ is given by $P = \kappa := \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$. Consider also the strategy $\tilde{S}$ whose state is $\tilde{\varrho} = \kappa$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$ and whose measurement channels are $\Lambda_\mathsf{A} = \Lambda_\mathsf{B} = \Delta_{\{0,1\}}$, which measure in the computational basis of $\mathbb{C}^2$. We claim that P self-tests $\tilde{S}$ according to causal simulation.

Let S be a strategy with behaviour P. We must show that $S \leq_{c.s.} \tilde{S}$. Without loss of generality, we may assume that the local quantum systems $\mathcal{H}_\mathsf{A}$ and $\mathcal{H}_\mathsf{B}$ in the strategy S are of the form $\mathbb{C}^2 \otimes \mathbb{C}^{m_\mathsf{A}}$ and $\mathbb{C}^2 \otimes \mathbb{C}^{m_\mathsf{B}}$ for some $m_\mathsf{A}, m_\mathsf{B} \in \mathbb{N}$, and that the measurement channels correspond to the projective measurements given by $\Pi_\mathsf{A}(y_\mathsf{A}) = |y_\mathsf{A}\rangle\langle y_\mathsf{A}| \otimes \mathbb{1}_{\mathbb{C}^{m_\mathsf{A}}}$ and $\Pi_\mathsf{B}(y_\mathsf{B}) = |y_\mathsf{B}\rangle\langle y_\mathsf{B}| \otimes \mathbb{1}_{\mathbb{C}^{m_\mathsf{B}}}$. This is by Proposition 4.13 and the fact that any projective strategy is equivalent under local simulation to one of the said form (by local isometric embedding and a local unitary change of basis). Now, observe that we may even without loss of generality assume that $m_\mathsf{A} = m_\mathsf{B} = 1$, i.e. that S in fact consists in measuring a pair of qubits in the computational bases. This is because a strategy S' with general values of $m'_\mathsf{A}, m'_\mathsf{B} \in \mathbb{N}$ is causally simulated by one with $m_\mathsf{A} = m_\mathsf{B} = 1$. Indeed, the measurement channels Λ'_P are given by $\Delta_{\{0,1\}} \otimes \mathrm{tr}_{\mathbb{C}^{m'_\mathsf{P}}}$, so if $\breve{\Delta}_{\{0,1\}} : \mathrm{End}(\mathbb{C}^2) \to \mathrm{End}(\mathbb{C}^2 \otimes \mathcal{E})$ is a Stinespring dilation of $\Delta_{\{0,1\}}$ then $\breve{\Delta}_{\{0,1\}} \otimes \mathrm{id}_{\mathbb{C}^{m'_\mathsf{P}}}$ is a Stinespring dilation of Λ'_P. Hence, S' has Stinespring implementation



(47)

which can trivially be rewritten as



(48)

thus showing the asserted causal simulation by a qubit strategy S. (Operationally, information which in S' was not leaked before the measurements has now been moved to the register of information which exists in the environment before receiving the bit queries.)

To show the desired, it therefore altogether suffices to show that $\tilde{S}$ causally simulates any qubit strategy $S = (\varrho, \Delta_{\{0,1\}}, \Delta_{\{0,1\}})$ with behaviour P whose measurements are in the computational basis. Let us characterise the purifications ψ of states ϱ in such strategies. First, write $\varrho = \sum_{k=1}^n p_k \psi_k$ as a convex combination of pure states. Then, observe that for the strategy S to have behaviour P means that $\kappa = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| = (\Delta_{\{0,1\}} \otimes \Delta_{\{0,1\}})(\varrho) = \sum_{k=1}^n p_k (\Delta_{\{0,1\}} \otimes \Delta_{\{0,1\}})(\psi_k)$. This can only happen if for $k = 1, \ldots, n$ we have

$$(\Delta_{\{0,1\}} \otimes \Delta_{\{0,1\}})(\psi_k) = q_k |00\rangle\langle 00| + (1 - q_k)|11\rangle\langle 11| \tag{49}$$

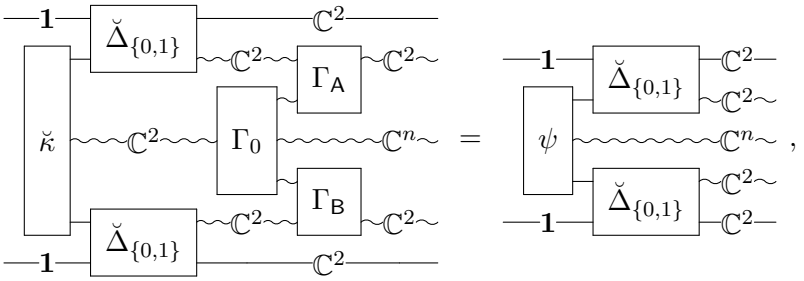
for some $q_1, \dots, q_n \in [0,1]$ with $\sum_{k=1}^n p_k q_k = \frac{1}{2}$. From this we conclude that each ψ_k must have a vector representative of the form $\sqrt{q_k}\,|00\rangle + \sqrt{1-q_k}e^{i\theta_k}\,|11\rangle$ with $\theta_k \in [0, 2\pi)$. As such, a purification of ϱ, with purifying space $\mathbb{C}^n$, is given by

$$|\psi\rangle := \sum_{k=1}^m \sqrt{p_k}\,|\psi_k\rangle \otimes |k\rangle = \sum_{k=1}^m \sqrt{p_k}(\sqrt{q_k}\,|00\rangle + \sqrt{1-q_k}e^{i\theta_k}\,|11\rangle\,) \otimes |k\rangle \tag{50}$$

$$= \frac{1}{\sqrt{2}}\,|00\rangle \otimes |\chi_0\rangle + \frac{1}{\sqrt{2}}\,|11\rangle \otimes |\chi_1\rangle\,, \tag{51}$$

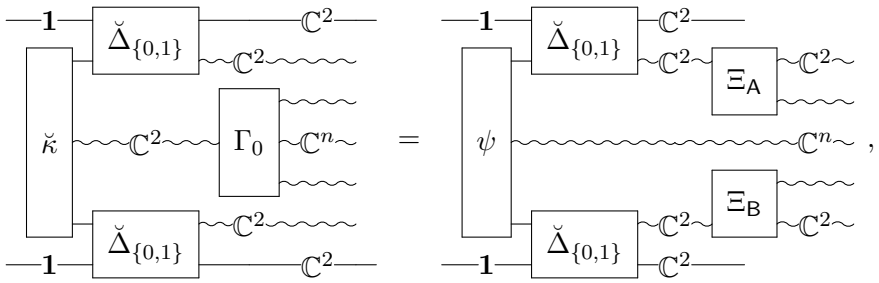
with $|\chi_0\rangle := \sum_{k=1}^n \sqrt{2p_k q_k}\,|k\rangle$ and $|\chi_1\rangle := \sum_{k=1}^n \sqrt{2p_k(1-q_k)}e^{i\theta_k}\,|k\rangle$. Since $\sum_{k=1}^n p_k q_k = 1/2$, both $|\chi_0\rangle$ and $|\chi_1\rangle$ are unit vectors. In conclusion, for any such strategy with the correct behaviour, a purification of the state ϱ is given by $|\psi\rangle = \frac{1}{\sqrt{2}}\,|00\rangle \otimes |\chi_0\rangle + \frac{1}{\sqrt{2}}\,|11\rangle \otimes |\chi_1\rangle$ with $|\chi_0\rangle\,, |\chi_1\rangle \in \mathbb{C}^n$ unit vectors; conversely, whenever ϱ has a purification ψ of this form it actually has the correct behaviour. (Our canonical strategy $\tilde{S}$ incidentally belongs to this class, with $|\chi_j\rangle = |j\rangle \in \mathbb{C}^2$.)

To finish the proof, we must therefore show that for any such pure state ψ there exist channels $\Gamma_0, \Gamma_\mathsf{A}, \Gamma_\mathsf{B}$ such that



(52)

where $\breve{\kappa}$ denotes the purification of κ with vector representative $|\breve{\kappa}\rangle = \frac{1}{\sqrt{2}}\,|00\rangle \otimes |0\rangle + \frac{1}{\sqrt{2}}\,|11\rangle \otimes |1\rangle \in (\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes \mathbb{C}^2$. Indeed, this condition precisely expressed that $\tilde{S}$ causally simulates S. It suffices for this to find isometric channels Γ_0, Ξ_A, Ξ_B such that



(53)

since by reversibility of isometric channels (Lemma 2.18), Ξ_A and Ξ_B may then be cancelled to yield Eq. (52). Satisfying Eq. (53) is easy, however, as we may simply choose Γ_0 corresponding to the isometry $|j\rangle \mapsto |j\rangle \otimes |\chi_j\rangle \otimes |j\rangle$ and Ξ_A and Ξ_B both corresponding to the isometry $|j\rangle \mapsto |j\rangle \otimes |j\rangle$ (choosing the systems of the unlabelled wiggly wires as $\mathbb{C}^2$). This concludes the argument and proves that P self-tests $\tilde{S}$ according to causal simulation. ♦

It is worth observing that the strategy $\tilde{S}$ considered in Example 4.32 is essentially classical and can thus be considered as a strategy in the operational theory of classical information, **CIT** (cf. Section 2). Since our formulation of self-testing is operational, it is meaningful to ask whether it is also the case that the behaviour P self-tests $\tilde{S}$ a.t.c.s. *in the theory* **CIT**. This is indeed case; for a general discussion, see Sec. 4.4.C in Ref. [HL21], in

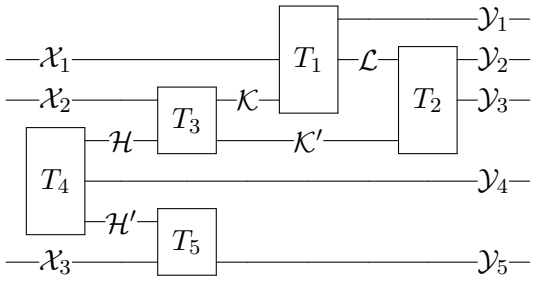
particular the statement is implied by Cor. 4.4.24 (where the term 'rigidity' is synonymous with 'self-testing a.t.c.s.').

In the following, we explain how causal simulation among quantum strategies may be considered a special case of a much more general phenomenon.

4.C.2 Causal Channels

Let $\mathbf{\Theta}$ be any compositional theory (e.g. $\mathbf{\Theta} = \mathbf{QIT}$). In short, causal channels provide a compact way of dealing with circuits of channels in $\mathbf{\Theta}$. We have already used pictures of circuits extensively when describing implementations and the various modes of simulation between quantum strategies, e.g. as in Eq. (46). In such situations, we always understood that the picture represented a single channel from the input systems on the left to the output systems on the right, rather than representing the specific circuit structure and the individual channels appearing in it.

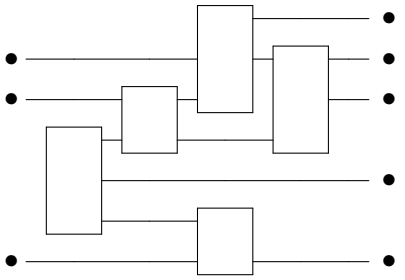
For example, a depiction such as



(54)

would represent a channel T from the system $\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_3$ to the system $\mathcal{Y}_1 \otimes \ldots \otimes \mathcal{Y}_5$, rather than representing the structure of boxes and wires and the individual channels $T_1, \ldots, T_5$ filling these boxes. (This is completely analogous to the way in which an expression such as '2×3' is usually taken to represent the number 6 rather than the sequence of symbols '2', '$\times$' and '3'.)

If we wish to address instead the circuit of channels depicted in Eq. (54), we must refer to two pieces of information. One of these is purely graph-theoretic, namely the *stencil*



, (55)

which is formally a directed acyclic graph (the direction is from left to right) whose edges we think of as 'wires' and whose nodes come in two types, which we think of respectively as 'open ports' and 'boxes'. The second piece of information is the *filling*, which assigns to each wire in the stencil a system in $\mathbf{\Theta}$ and to each box in the stencil a channel in $\mathbf{\Theta}$ between the adjacent systems. In general, a *circuit* is thus a pair $(\mathfrak{F}, G)$, where G is a stencil described by a directed acyclic graph and where $\mathfrak{F}$ is a filling of the stencil G.

To any circuit $(\mathfrak{F}, G)$ is an associated *(input-output) behaviour*, denoted $\mathfrak{F}[G]$, namely the total channel from the system corresponding to the open input ports to the to the

system corresponding to the open output ports. There is also an associated *causal specification*, namely the map $\mathscr{C}_G$ which to every open output port in G associates the set of input ports in G which are ancestral to that output port. We think of these input ports as the *causes* of the output port. (In the stencil (55), for example, the top three output ports each has the top two input ports as causes, the fourth output port has no causes, and the bottom output port has the bottom input port as cause.)

The main idea behind causal channels is to consider the following equivalence relation among circuits:

Definition 4.33 (Congruence of Circuits)**.** Two circuits $(\mathfrak{F}, G)$ and $(\mathfrak{F}', G')$ are called *congruent* if they have the same behaviour and the same causal specification, i.e. if $\mathfrak{F}[G] = \mathfrak{F}'[G']$ and $\mathscr{C}_G = \mathscr{C}_{G'}$. ■

Clearly, there is a very natural way of representing the congruence class of a given circuit $(\mathfrak{F}, G)$, namely in terms of the pair $(T, \mathscr{C}) = (\mathfrak{F}[G], \mathscr{C}_G)$ consisting of the behaviour T and the causal specification $\mathscr{C}$.

Definition 4.34 (Constructible Causal Channels)**.** A *constructible causal channel* in $\mathbf{\Theta}$ is a pair of the form $(T, \mathscr{C})$ where $T = \mathfrak{F}[G]$ and $\mathscr{C} = \mathscr{C}_G$ for some circuit $(\mathfrak{F}, G)$ in $\mathbf{\Theta}$. ■

Remark 4.35 (Constructibility)**.** The treatment in Ref. [HL21] involves causal channels $(T, \mathscr{C})$ of an abstract kind, which need not admit a representation in terms of circuits, but where T merely complies to certain non-signalling conditions suggested by $\mathscr{C}$. The term 'constructible' then precisely refers to those causal channels which arise from circuits. Here, we shall only consider constructible causal channels, but we often include the term for emphasis. ✠

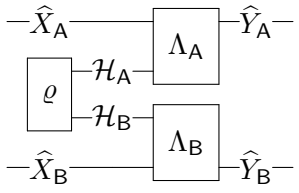
A constructible causal channel is a mathematically much simpler object than a circuit. The reason why it is viable to make this simplification is that congruence of circuits respects the various operations one might wish to perform on the level of circuits:

It is easy to see that if we compose two circuits in series or parallel, then the congruence class of the resulting composition can be determined from the congruence classes of the initial circuits. In other words, serial and parallel composition can be considered as operations on the level of constructible causal channels themselves.

More remarkably, it is also true in most[33] theories that the congruence class of a circuit which arises from contraction of wires in an initial circuit can be derived from the congruence class of this initial circuit. (An output wire can be contracted with an input wire whenever the resulting graph remains acyclic.) In other words, contraction of wires can also be performed a the level of constructible causal channels.

Though we shall not directly need the latter result here, the above facts are ultimately what justifies in manipulations the replacement of circuits by causal channels.

Example 4.36 (Bell-Channels)**.** Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ be a strategy for a Bell-scenario $\mathfrak{B}$. The strategy S is really the description of a circuit, and the diagram



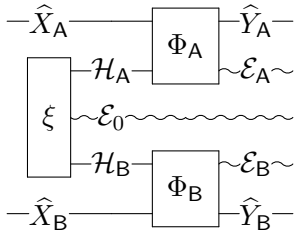
(56)

[33] Specifically, this is true under the so-called *universality* axiom of Ref. [HL21], see Theorem 4.2.6 therein. The universality axiom holds in the theories **QIT** and **CIT**.

gives rise to a constructible causal channel $(P, \mathscr{C})$ which in addition to the behaviour P of the strategy also holds a causal specification $\mathscr{C}$ which to the output port y_P associates the set of input ports $\mathscr{C}(\mathsf{y}_\mathsf{P}) = \{\mathsf{x}_\mathsf{P}\}$, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$. Any other strategy S' with behaviour P gives rise to the same causal channel, but the circuit-representation is different.

A constructible causal channel $(P, \mathscr{C})$ will be called a *Bell-channel* if it admits a circuit-representation of the form (56). It can be shown [HL21] that if $(P, \mathscr{C})$ is any constructible causal channel for which P is embedded classical and $\mathscr{C}$ is given by $\mathscr{C}(\mathsf{y}_\mathsf{P}) = \{\mathsf{x}_\mathsf{P}\}$ for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, then $(P, \mathscr{C})$ is in fact a Bell-channel. ♦

Example 4.37 (Implementations)**.** Let $(\xi, \Phi_\mathsf{A}, \Phi_\mathsf{B})$ be a components-wise dilation of a strategy $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ for a Bell-scenario $\mathfrak{B}$. Then, the appropriate way of thinking about the implementation



(57)

is as a constructible causal channel $(L, \mathscr{E})$, where L is the input-output behaviour, and where the causal specification $\mathscr{E}$ is given by $\mathscr{E}(\mathsf{e}_0) = \emptyset$ and $\mathscr{E}(\mathsf{e}_\mathsf{P}) = \mathscr{E}(\mathsf{y}_\mathsf{P}) = \{\mathsf{x}_\mathsf{P}\}$ for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$. ♦

4.C.3 Causal Dilations and the Derivability Pre-Order

The causal channel in Example 4.37 which corresponds to an implementation of a strategy S has the following property: if we trace out the three ports e_A, e_B and e_0, then we obtain precisely the causal channel $(P, \mathscr{C})$ from Example 4.36 describing the behaviour of the strategy. Specifically, the channel L is a dilation of the channel P, and the causal specification $\mathscr{E}$ assigns to any output port y, which is not traced out, the same set of causes as does $\mathscr{C}$.

This is an instance of the following concept:

Definition 4.38 (Constructible Causal Dilations)**.** Let $(T, \mathscr{C})$ be a causal channel. A *construcible causal dilation of* $(T, \mathscr{C})$ is a constructible causal channel $(L, \mathscr{E})$ whose set of input ports coincides with that of $(T, \mathscr{C})$, whose set of output ports contains that of $(T, \mathscr{C})$, and such that L is a dilation of T and $\mathscr{E}(\mathsf{y}) = \mathscr{C}(\mathsf{y})$ for any output port y in $(T, \mathscr{C})$.

The set of output ports in $(L, \mathscr{E})$ which are not output ports of $(T, \mathscr{C})$ are called the *environment of* $(L, \mathscr{E})$. ■

On the level of circuits, a constructible causal dilation of $(T, \mathscr{C}) = (\mathfrak{F}[G], \mathscr{C}_G)$ corresponds to a circuit which after having the output ports in its environment traced out becomes congruent to $(\mathfrak{F}, G)$. As such, the constructible causal dilations of $(T, \mathscr{C})$ represent all the ways in which the circuit $(\mathfrak{F}, G)$ may possibly be realised in a larger environment. Side-information leaks to this environment according to a schedule prescribed by the causal dependence of outputs on inputs.

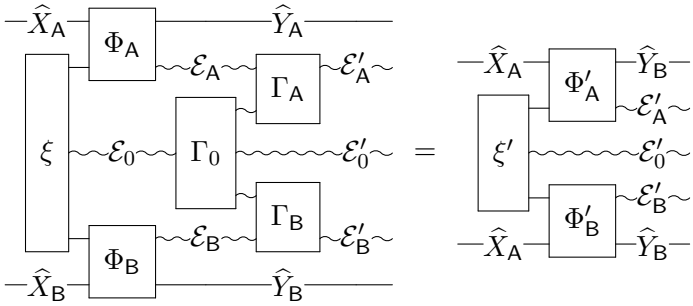
There is a natural notion of relative strength between causal dilations:

Definition 4.39 (Derivability)**.** Let $(L, \mathscr{E})$ and $(L', \mathscr{E}')$ be constructible causal dilations of $(T, \mathscr{C})$. We say that $(L', \mathscr{E}')$ *is derivable from* $(L, \mathscr{E})$, or that $(L, \mathscr{E})$ *derives* $(L', \mathscr{E}')$, if we can obtain $(L', \mathscr{E}')$ by composing $(L, \mathscr{E})$ with some constructible causal channel $(H, \mathscr{D})$ which acts only in the environment. We write in this case $(L, \mathscr{E}) \trianglerighteq (L', \mathscr{E}')$. ■

Again, derivability can be described on the level of circuits: $(L, \mathscr{E})$ derives $(L', \mathscr{E}')$ if a circuit corresponding to $(L, \mathscr{E})$ can be composed with some circuit, acting only in the environment, to yield a circuit corresponding to $(L', \mathscr{E}')$. As such, derivability of $(L', \mathscr{E}')$ from $(L, \mathscr{E})$ means that all side-information which escapes to the environment according to $(L', \mathscr{E}')$ can be extracted — even conforming to the correct schedule — if we possess $(L, \mathscr{E})$.

Example 4.40. If $(T, \mathscr{C})$ is a constructible causal channel, then $(T, \mathscr{C})$ is a constructible causal dilation of itself. We call it the *trivial* dilation of $(T, \mathscr{C})$, and it is derivable from any other constructible causal dilation, by simply tracing out each output in the environment. ⧫

Example 4.41. Let S and S' be strategies for a Bell-scenario with behaviour P, and let (ξ, Φ_A, Φ_B) and (ξ', Φ'_A, Φ'_B) be respective component-wise dilations. We have seen in Example 4.37 that the corresponding implementations are causal dilations of the causal channel $(P, \mathscr{C})$. The causal dilation corresponding to (ξ', Φ'_A, Φ'_B) is derivable from the causal dilation corresponding to (ξ, Φ_A, Φ_B) if there exist channels Γ_0, Γ_A and Γ_B such that



. (58)

Indeed, the channels Γ_0, Γ_A and Γ_B make up a circuit in the environment which derives one implementation-circuit from the other, up to congruence of circuits (in particular preserving the causal dependence of outputs on inputs). ⧫

Remark 4.42 (Modularity)**.** The notion of derivability among causal dilations enjoys several 'composability', or 'modularity', features. For instance, if causal dilations $(L_1, \mathscr{E}_1)$ of $(T_1, \mathscr{C}_1)$ and $(L_2, \mathscr{E}_2)$ of $(T_2, \mathscr{C}_2)$ derive the causal dilations $(L'_1, \mathscr{E}'_1)$ and $(L'_2, \mathscr{E}'_2)$, respectively, then the parallel [serial] composition of $(L_1, \mathscr{E}_1)$ and $(L_2, \mathscr{E}_2)$ derives the parallel [serial] composition of $(L'_1, \mathscr{E}'_1)$ and $(L'_2, \mathscr{E}'_2)$. For precise and general versions of these statements, see [HL21] Theorems 4.3.24 and 4.3.25. ✠

Now, for any constructible causal channel $(T, \mathscr{C})$, the derivability relation is a pre-order on the class of constructible causal dilations of $(T, \mathscr{C})$. It is of interest to know this class and pre-order, since this provides an understanding of the various ways in which side-information may leak to the environment and how these relate to each other. In fact, it is already interesting to identify a *dense* subclass of causal dilations, i.e. a subclass such that any causal dilation may be derived from a member of this subclass. A very special case of this is when every dilation is derivable from a single one:

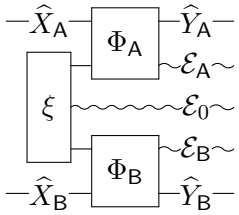
Definition 4.43 (Complete Causal Dilations)**.** A constructible causal dilation $(K, \mathscr{F})$ of $(T, \mathscr{C})$ is called *(causally) complete* if it is a $\trianglerighteq$-largest element, i.e. if for any constructible causal dilation $(L, \mathscr{E})$ of $(T, \mathscr{C})$, we have $(K, \mathscr{F}) \trianglerighteq (L, \mathscr{E})$. ■

A complete causal dilation of $(T, \mathscr{C})$ formalises the notion of a strongest possible information-leak to the environment. What we now wish to illustrate is that in the theory **QIT**, in the special causal case of Bell-scenarios, causal completeness is essentially the occurrence of quantum self-testing according to causal simulation.

4.C.4 The Special Case of Bell-Scenarios

Let $\mathfrak{B} = (X_\mathsf{A}, X_\mathsf{B}, Y_\mathsf{A}, Y_\mathsf{B})$ be a Bell-scenario. Recall from Example 4.36 that a Bell-channel is a constructible causal channel $(P, \mathscr{C})$ which admits a circuit-representation of the form (56), or, equivalently, one whose behaviour P is classical and whose causal specification is given by $\mathscr{C}(\mathsf{y}_\mathsf{A}) = \{\mathsf{x}_\mathsf{A}\}$ and $\mathscr{C}(\mathsf{y}_\mathsf{B}) = \{\mathsf{x}_\mathsf{B}\}$. We have the following result, which holds in any operational theory:

Theorem 4.44 (Density Theorem for Bell-Channels)**.** *Let $(P, \mathscr{C})$ be a Bell-channel. Then, any constructible causal dilation of $(P, \mathscr{C})$ is derivable from a constructible causal dilation corresponding to a circuit of the form*



(59)

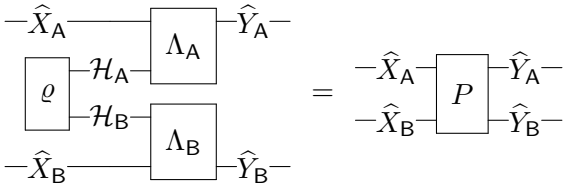
for some state ξ and some channels Φ_A and Φ_B.

Remark 4.45. In the theory **QIT**, it follows from Theorem 2.22 that we may strengthen the statement in Theorem 4.44 by assuming the state ξ to be pure and the channels Φ_A and Φ_B to be isometric. ✠

For a proof of Theorem 4.44, the reader is referred to Theorem 4.4.8 in Ref. [HL21]. The proof is essentially graph-theoretic, and the significance of Theorem 4.44 is that any circuit which implements $(P, \mathscr{C})$ in a larger environment, however complicated, can be derived from a circuit congruent to one with the simple triple-structure displayed in Eq. (59).

It would seem that Theorem 4.44 explains the significance of triple-strategies relative to more general strategies, cf. our earlier discussion in Section 3.A. This is not quite true, however, due to the following curious state of affairs:

It may happen that the circuit (59) *is not the implementation of a strategy $S = (\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ with behaviour P.* The subtle reason for this circumstance is that that we defined a strategy S to consist of a state ϱ together with measurement ensembles Λ_A and Λ_B. Whereas it is obviously true that $(\xi, \Phi_\mathsf{A}, \Phi_\mathsf{B})$ in Eq. (59) is a component-wise dilations of some triple $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ for which



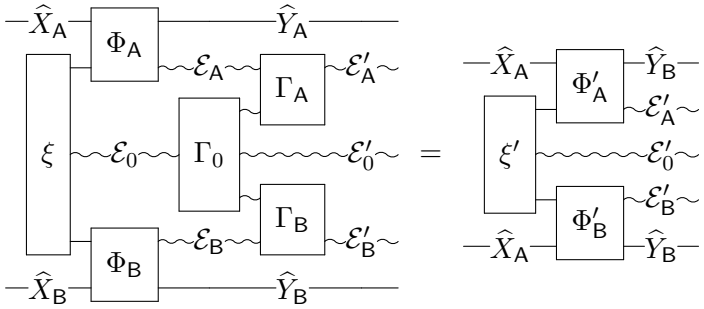
, (60)

the fact that P is classical does not guarantee that the channels Λ_A and Λ_B are measurement ensembles, i.e. have classical outputs and classical inputs on $\widehat{X}_\mathsf{A}$ and $\widehat{X}_\mathsf{B}$.

Let us call any triple $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ a *quasi-strategy with behaviour P* if it satisfies Eq. (60). *Implementations* of a quasi-strategy $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$ are defined in the obvious way, namely as causal dilations of the form (59), where $(\xi, \Phi_\mathsf{A}, \Phi_\mathsf{B})$ is a component-wise dilation of $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$. *Causal simulation* can be also be defined in the obvious way between quasi-strategies.

Now, Theorem 4.44 expresses that every constructible causal dilation of a Bell-channel $(P, \mathscr{C})$ is derivable from the implementation of some quasi-strategy with behaviour P. The appearance of quasi-strategies constitutes an annoyance in relating quantum self-testing to general causal dilations, and we shall return to this problem very shortly. First, however, we state the following two results, which again hold in any operational theory:

Theorem 4.46 (Derivability within the Dense Class for Bell-Channels)**.** *Let $(P, \mathscr{C})$ be a Bell-channel, and let $(L, \mathscr{E})$ and $(L', \mathscr{E}')$ be constructible causal dilations which are implementations of quasi-strategies, corresponding to the component-wise dilations $(\xi, \Phi_{\mathsf{A}}, \Phi_{\mathsf{B}})$ and $(\xi', \Phi'_{\mathsf{A}}, \Phi'_{\mathsf{B}})$, respectively. Then, $(L, \mathscr{E}) \trianglerighteq (L', \mathscr{E}')$ if and only if there exist channels Γ_0, Γ_{A} and Γ_{B} such that*



(61)

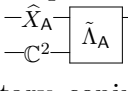
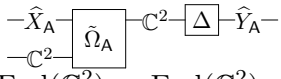
Corollary 4.47 (Characterisation of Completeness for Bell-Channels)**.** *The Bell-channel $(P, \mathscr{C})$ has a complete causal dilation if and only if there exists a quasi-strategy $\tilde{S}$ with behaviour P which causally simulates every quasi-strategy S with behaviour P.*

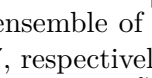
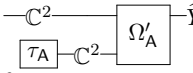
For a proof of Theorem 4.46, see Theorem 4.8.10 in Ref. [HL21]. It consists in demonstrating that the most general causal channel which may derive $(L', \mathscr{E}')$ from $(L, \mathscr{E})$ is of the form represented in Eq. (61) by the channels Γ_0, Γ_{A} and Γ_{B}, and the argument is again graph-theoretic. Corollary 4.47 follows immediately from Theorem 4.46 and Theorem 4.44.

Now, were it not for the appearance of quasi-strategies rather than ordinary strategies, Corollary 4.47 would provide the final link between quantum self-testing and the general theory of derivability among causal dilations: Quantum self-testing (according to causal simulation) simply signifies the existence of a complete causal dilation of the Bell-channel.

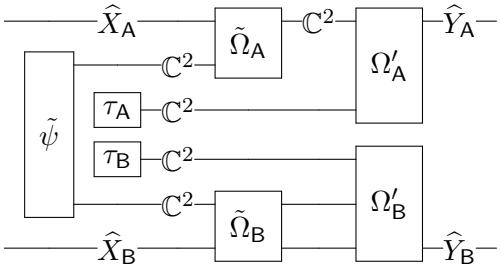
As the following example shows, however, some quasi-strategies cannot be treated as ordinary strategies:

Example 4.48. (A Genuine Quasi-Strategy.) Consider the Bell-scenario $\mathfrak{B}$ with $X_{\mathsf{A}} = X_{\mathsf{B}} = \{0, 1\}$ and $Y_{\mathsf{A}} = Y_{\mathsf{B}} = \{+1, -1\}$. Let $\tilde{S} = (\tilde{\psi}, \tilde{\Lambda}_{\mathsf{B}}, \tilde{\Lambda}_{\mathsf{B}})$ be an ordinary strategy for $\mathfrak{B}$ whose state is a pure state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ and whose measurement ensembles are projective, with each projection $\Pi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})$ having rank one. (For example, $\tilde{S}$ could be the ideal strategy for the CHSH-game [CHSH69, ŠB20].) From the ordinary strategy $\tilde{S}$, we may construct a quasi-strategy $\tilde{S}'$ which is not an ordinary strategy:

First, observe that every projective measurement with rank-one projections may be realised as a unitary conjugation followed by a measurement in the computational basis. As such, the measurement ensemble $\tilde{\Lambda}_{\mathsf{A}}$ may be rewritten as  =  where $\tilde{\Omega}_{\mathsf{A}}$ is an X_{A}-indexed ensemble of unitary conjugations $\mathrm{End}(\mathbb{C}^2) \to \mathrm{End}(\mathbb{C}^2)$, and where Δ is the measurement in the computational basis of $\mathbb{C}^2$ (equivalently, the dephasing channel on $\widehat{Y}_{\mathsf{A}}$ under some natural identification of $\widehat{Y}_{\mathsf{A}}$ with $\mathbb{C}^2$).

Next, the channel Δ equals the uniform mixture $\frac{1}{2}\mathrm{id}_{\mathbb{C}^2} + \frac{1}{2}Z$, where $Z : \mathrm{End}(\mathbb{C}^2) \to \mathrm{End}(\mathbb{C}^2)$ denotes the unitary conjugation by the Pauli matrix $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. As such, we may substitute  by  , where τ_{A} is the state $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$ and where Ω'_{A} is the ensemble of unitary conjugations which when reading off $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ applies $\mathrm{id}_{\mathbb{C}^2}$ or Z, respectively.

The measurement ensemble $\tilde{\Lambda}_{\mathsf{B}}$ may be rewritten in a similar way, with ensembles $\tilde{\Omega}_{\mathsf{B}}$ and Ω'_{B} of unitary conjugations, and the state τ_{B} again given by $\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$. Altogether, the behaviour P of $\tilde{S}$ can therefore be written as



(62)

Now, we may regroup components so as to form a quasi-strategy $S' = (\tilde{\varrho}', \tilde{\Lambda}'_{\mathsf{A}}, \tilde{\Lambda}'_{\mathsf{B}})$, where $\tilde{\varrho}'$ is the state $\tau_{\mathsf{A}} \otimes \tilde{\psi} \otimes \tau_{\mathsf{B}}$ on $(\mathbb{C}^2 \otimes \mathbb{C}^2) \otimes (\mathbb{C}^2 \otimes \mathbb{C}^2)$, and where $\tilde{\Lambda}'_{\mathsf{P}} : \mathrm{End}(\widehat{X}_{\mathsf{P}} \otimes \mathbb{C}^2 \otimes \mathbb{C}^2) \to \mathrm{End}(\widehat{Y}_{\mathsf{P}})$ is the channel given by the composition of $\tilde{\Omega}_{\mathsf{P}}$ and Ω'_{P}. The quasi-strategy $\tilde{S}'$ is not an ordinary strategy, since the channels $\tilde{\Lambda}'_{\mathsf{P}}$ are not measurement ensembles. In fact, for $x_{\mathsf{P}} \in \{0,1\}$ and $k \in \{0,1\}$, the channel $F \mapsto \tilde{\Lambda}'_{\mathsf{P}}(|x_{\mathsf{P}}\rangle\langle x_{\mathsf{P}}| \otimes F \otimes |k\rangle\langle k|)$ is one of four possible unitary conjugations from $\mathbb{C}^2$ to $\widehat{Y}_{\mathsf{P}}$, so a suitable choice of input results in a state on $\widehat{Y}_{\mathsf{P}}$ which is not diagonal in the computational basis. ♦

While mathematically sensible, the quasi-strategy in Example 4.48 is physically somewhat obscure. Rather than performing a measurement to obtain an output, each party $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$ performs a random unitary rotation on the P-part of the state $\tilde{\psi}$ (the randomness residing in the state τ_{P}) and returns the resulting state as output — the total input-output behaviour then happens to be classical 'on average'. The quasi-strategy is obscure for the same reason that a measurement in the computational basis of $\mathbb{C}^2$ (i.e. the channel Δ) cannot seriously be employed by flipping a coin and applying one of the unitary conjugations $\mathrm{id}_{\mathbb{C}^2}$ and Z, even though it is mathematically true that $\Delta = \frac{1}{2}\mathrm{id}_{\mathbb{C}^2} + \frac{1}{2}Z$.

It could seem that the problematic aspect of considering this procedure as a measurement ultimately stems from the fact that the outcome of the coin flip may stay in memory (this is addressed more formally in [HL21] Example 4.3.9). We do not wish to discuss philosophical aspects of quasi-strategies here, but we need to somehow exclude them from consideration in order to establish a formal link to ordinary quantum self-testing. This is not only because quasi-strategies do not occur as valid strategies in ordinary self-testing, but also because including them would in some cases actually constitute an obstruction to self-testing in terms of complete causal dilations: whereas for example the canonical optimal strategy $\tilde{S}$ for the CHSH-game [CHSH69, ŠB20] is self-tested a.t.r., and therefore causally simulates any ordinary strategy with the correct behaviour, it does not causally simulate the implementation corresponding to the quasi-strategy $\tilde{S}'$ in Example 4.48. Indeed, $\tilde{S}'$ has an implementation in which the states τ_{A} and τ_{B} are dilated by means of copying (this precisely corresponding to 'keeping the coin flip in memory', as mentioned above). Conditioned on the value of these copies, the output on $\widehat{Y}_{\mathsf{A}} \otimes \widehat{Y}_{\mathsf{B}}$ is one of four pure entangled states, and since their mixture is a classical state the copies must be correlated non-trivially with these states (i.e. the states must be different). On the other hand,

the state $\tilde{\psi}$ in the ordinary canonical strategy $\tilde{\psi}$ is pure, and therefore $\tilde{S}$ cannot causally simulate any strategy with this feature (see also Section 5.C).

In Ref. [HL21], the exclusion of non-ordinary quasi-strategies is achieved by introducing a notion of *classically bound* causal dilations, in an attempt to formalise the occurrence of a measurement which cannot be causally dilated e.g. by keeping a record of a unitary conjugation (see Hou21 Definition 5.2.10 and the surrounding). We shall not explain this notion further here, but merely remark that the classically bound causal dilations are precisely those which can be derived from implementations of ordinary strategies ([HL21] Proposition 5.2.12). It thus achieves the desired restriction of Theorem 4.44, and allows for the following result which concludes our presentation of how to recover quantum self-testing from the larger perspective of causal channels and causal dilations:

Theorem 4.49 (Self-Testing a.t.c.s. as an Instance of Causal Completeness)**.** *Let $(P, \mathscr{C})$ be a Bell-channel. Then the following are equivalent:*

1. *There is a classically bound constructible causal dilation of $(P, \mathscr{C})$ from which every classically bound constructible dilation can be derived.*

2. *P self-tests some strategy according to causal simulation.*

In fact, if $\tilde{S}$ is a strategy self-tested by P a.t.c.s., then its associated Stinespring implementation defines a classically bound constructible causal dilation from which all others can be derived.

Remark 4.50. When $\tilde{S}$ is pure-state, causal simulation may in Theorem 4.49 be replaced by local assisted simulation (see Remark 4.31). In the language of causal dilations, the pure-state assumption on $\tilde{S}$ corresponds to the requirement that side-information in the environment which has no causes — i.e. which may be leaked in advance of the inputs — must factor from the remaining information. (See Section 5.C and Chapter 5 in Ref. [HL21] for more formal discussions of this.) ✠

5 Utility and Applications of the Operational Formulation

In this final section of the paper, we explore how our operational notions of simulation from Section 4.A, Section 4.B and Section 4.C.1 can be used to prove features related to quantum self-testing. Some of these features are already known, but their proofs have previously relied on operator-algebra. The proofs presented in this section are all operational. Moreover, they clarify that some of the results hold even when replacing local simulation (reducibility) by the weaker notions of local assisted simulation or causal simulation.

In Section 5.A we show how local simulation of $\tilde{S}$ by S implies that the state $\tilde{\varrho}$ is locally extractable from the state ϱ (Proposition 5.1). This is well-understood in the operator-algebraic framework [ŠB20] (see Remark 5.2 below), albeit under the assumption that $\tilde{\varrho}$ is pure. Proposition 5.1 is phrased for general states and implies rigorously that only pure-state strategies can be self-tested according to local simulation (Corollary 5.5).

In Section 5.B, we show how local assisted simulation of S by $\tilde{S}$ implies that the measurement ensembles of S can be extracted from those of $\tilde{S}$ (Proposition 5.6). This result is related to certain formulations of self-testing of measurements alone [Kan17, ŠB20], but we believe that it has not been stated in this form before.

In Section 5.C, we demonstrate that causal simulation has implications for the structure of convex decompositions of behaviours (Lemma 5.12 and Proposition 5.14). This

in particular recovers from a general perspective the result of Ref. [GKW+18] that a behaviour which self-tests some strategy a.t.r. must be extremal in the convex set of realisable behaviours. In fact, our version of this result (Corollary 5.15) holds under the weaker assumption of self-testing a.t.l.a.s.

Finally, in Section 5.D, we report a feature of self-testing which to our best knowledge has not been observed before. We will focus our view on the *environment* of strategies implementing a certain behaviour. In the case of pure projective rank one strategies, we find that the respective environments of the players will hold exactly a copy of their input - no more and no less. We do not currently know of any concrete applications of this feature, but rather consider it as illustrative in its own right.

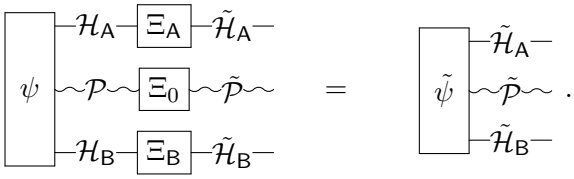
5.A Local Simulation and State Extractability

It is well-known that the operator-algebraic reducibility relation $S \geq_{red.} \tilde{S}$ (Definition 3.5) implies that the state $\tilde{\psi}$ of $\tilde{S}$ is locally extractible from the state ϱ of S. Indeed, if in the reducibility conditions $[W\Pi^x(y) \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = \tilde{\Pi}^x(y)|\tilde{\psi}\rangle \otimes |\psi^{\text{res}}\rangle$ we sum over y, then we obtain the relation $[W \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = |\tilde{\psi}\rangle \otimes |\psi^{\text{res}}\rangle$, or, unfolding W as $W_{\mathsf{A}} \otimes W_{\mathsf{B}}$,

$$[W_{\mathsf{A}} \otimes W_{\mathsf{B}} \otimes \mathbb{1}_{\mathcal{P}}]\,|\psi\rangle = |\tilde{\psi}\rangle \otimes |\psi^{\text{res}}\rangle\,. \tag{63}$$

Below, we prove that a similar state extractibility result follows from the relation $S \geq_{l.s.} \tilde{S}$. Our proof is entirely compositional (it is in terms of diagrams), and it works even when the involved strategies are not pure-state or projective:

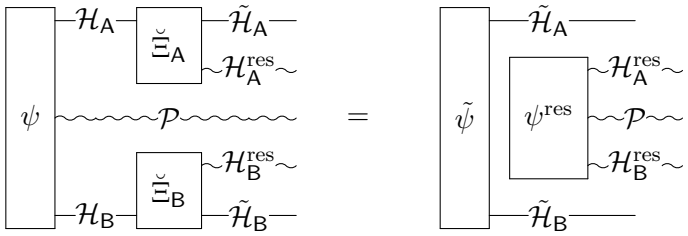
Proposition 5.1 (State Extractibility)**.** *Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ and $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ be strategies. Let ψ and $\tilde{\psi}$ be purifications of ϱ and $\tilde{\varrho}$, respectively. If $S \geq_{l.s.} \tilde{S}$, then there exist channels Ξ_0, Ξ_{A} and Ξ_{B} such that*



(64)

We make a few remarks before presenting the proof.

Remark 5.2 (Relation to Conventional State Extractibility)**.** If the state $\tilde{\varrho}$ is already pure we may choose $\tilde{\psi} = \tilde{\varrho}$, and the channel Ξ_0 is then the trace. Employing our usual trick of replacing Ξ_{A} and Ξ_{B} by Stinespring dilations $\breve{\Xi}_{\mathsf{A}}$ and $\breve{\Xi}_{\mathsf{A}}$ and replacing the trace by an identity, we thus conclude from Eq. (64) that

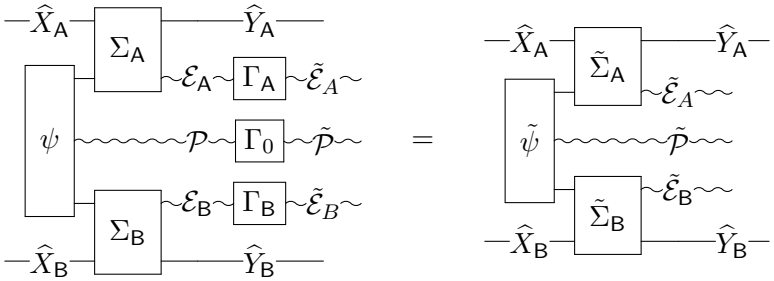


(65)

for some pure state ψ^{res}, and this is precisely Eq. (63), with W_{A} and W_{B} representing $\breve{\Xi}_{\mathsf{A}}$ and $\breve{\Xi}_{\mathsf{B}}$. ✠

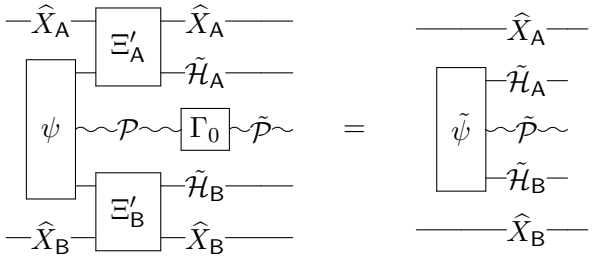
Remark 5.3 (Generalisation to the Approximate Case)**.** It will be clear from the proof of Proposition 5.1 that the result translates without loss to the case of approximate simulation (cf. Remark 4.18) provided the metric is monotone, i.e. distance does not increase under post-processing. ✠

Proof (of Proposition 5.1). Choose Stinespring dilations Σ_{P} of Λ_{P} and $\tilde{\Sigma}_{\mathsf{P}}$ of $\tilde{\Lambda}_{\mathsf{P}}$, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$. By Lemma 4.11, there exist channels Γ_0, Γ_{A} and Γ_{B} such that



(66)

The channels $\tilde{\Sigma}_{\mathsf{A}}$ and $\tilde{\Sigma}_{\mathsf{B}}$ are isometric and thus reversible by Lemma 2.18. Applying left-inverses on both sides of Eq. (66) yields



(67)

for some channels Ξ'_{A} and Ξ'_{B}. If we now, for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, trace out the output systems $\widehat{X}_{\mathsf{P}}$ and insert arbitrary states σ_{P} into the input systems $\widehat{X}_{\mathsf{P}}$, then the desired identity (64) follows by defining $\Xi_{\mathsf{P}} := (\mathrm{tr}_{\widehat{X}_{\mathsf{P}}} \otimes \mathrm{id}_{\tilde{\mathcal{H}}_{\mathsf{P}}}) \circ \Xi'_{\mathsf{P}} \circ (\sigma_{\mathsf{P}} \otimes \mathrm{id}_{\mathcal{H}_{\mathsf{P}}})$ for $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$ and $\Xi_0 := \Gamma_0$. □

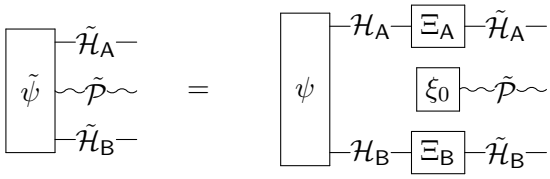
We immediately conclude the following:

Corollary 5.4. *Suppose that the strategy $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ is self-tested according to local simulation by its behaviour P. Let $\tilde{\psi}$ be a purification of $\tilde{\varrho}$. For any strategy $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ with behaviour P, and for any purification ψ of ϱ, there exist channels Ξ_0, Ξ_{A} and Ξ_{B} such that Eq. (64) holds.*

Since Corollary 5.4 is a statement about canonical strategies with arbitrary state, it can be used to rigorously show that certain states simply cannot appear in self-testable strategies:

Corollary 5.5 (All a.t.l.s. Self-Testable Strategies are Pure-State)**.** *Suppose that $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ is a strategy which is self-tested a.t.l.s. by its behaviour. Then, the state $\tilde{\varrho}$ must be pure.*

Proof. There is some strategy S with behaviour P whose state, say ψ, is pure. Indeed, such a strategy can be obtained if in any given strategy S_0 we purify the state and hand the purification to one of the parties, say A. By Corollary 5.4, a purification $\tilde{\psi}$ of $\tilde{\varrho}$ now satisfies



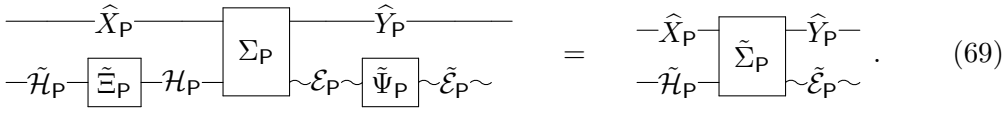
(68)

for some channels Ξ_{A}, Ξ_{B} and some state ξ_0 (since the purifying system for ψ may be taken trivial). In other words, $\tilde{\psi}$ factors across its purifying system and it follows that $\tilde{\varrho}$ must already be pure. □

5.B Local Assisted Simulation and Measurement Extractibility

Next, we prove that local assisted simulation implies an extractibility result for the measurement ensembles under a suitable rank-assumption. As mentioned, this result connects to alternative formulations of self-testing [Kan17, ŠB20], but as far as we know the general statement below (Proposition 5.6) has not been presented before. Like for state extractibility, our proof is purely compositional in terms of diagrams.

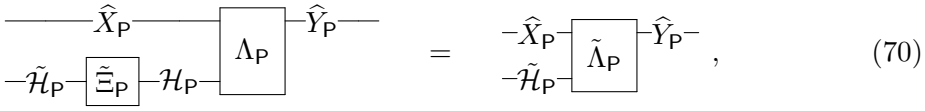
Proposition 5.6 (Measurement Extractibility)**.** *Let* $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ *and* $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ *be strategies such that* $S \leq_{l.a.s.} \tilde{S}$*. Let* $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$*, and let* Σ_{P} *and* $\tilde{\Sigma}_{\mathsf{P}}$ *be Stinespring dilations of* Λ_{P} *and* $\tilde{\Lambda}_{\mathsf{P}}$*, respectively. If the* P*-marginal* $\tilde{\varrho}_{\mathsf{P}}$ *of* $\tilde{\varrho}$ *has full rank, then there exist channels* $\tilde{\Xi}_{\mathsf{P}}$ *and* $\tilde{\Psi}_{\mathsf{P}}$ *such that*



(69)

Before presenting the proof, we again make a few remarks.

Remark 5.7. Tracing out the environment on each side of Eq. (69) leaves the identity



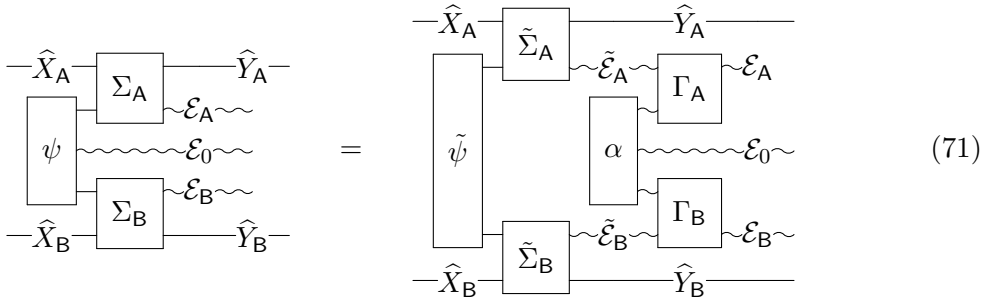
(70)

which expresses that the ensemble $\tilde{\Lambda}_{\mathsf{P}}$ can be extracted from the ensemble Λ_{P}. Eq. (69) is thus a strong version of extractability, according to which even dilations of $\tilde{\Lambda}_{\mathsf{P}}$ can be extracted from dilations of Λ_{P}. (The same, incidentally, can be said about Proposition 5.1 regarding dilations of states.) ✠

Remark 5.8. The rank-assumption in Proposition 5.6 is necessary since there is no way of determining the action of $\tilde{\Lambda}_{\mathsf{P}}$ outside the support of $\tilde{\varrho}_{\mathsf{P}}$ given only a Stinespring implementation of $\tilde{S}$. ✠

Remark 5.9 (Generalisation to the Approximate Case)**.** As follows from the very last argument in the proof of Proposition 5.6, approximate versions of this result might lose factors, with the loss depending on the spectrum of the density operator $\tilde{\varrho}_{\mathsf{P}}$. ✠

Proof (of Proposition 5.6). Let us assume without loss of generality that $\mathsf{P} = \mathsf{A}$. Let us also assume for simplicity that the state $\tilde{\varrho} =: \tilde{\psi}$ is pure (the idea of the proof is not conceptually different in the general case, but less clearly exposed). Since $S \leq_{l.a.s.} \tilde{S}$, we know that

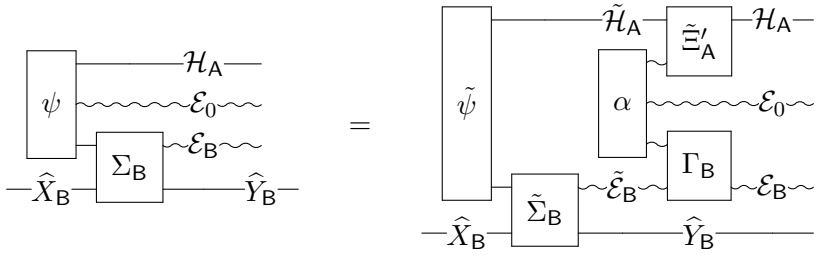


(71)

for some state α and some channels Γ_{A} and Γ_{B} (the channel Γ_0 appearing in the definition of local assisted simulation can be absorbed into the assistant state α because it has no

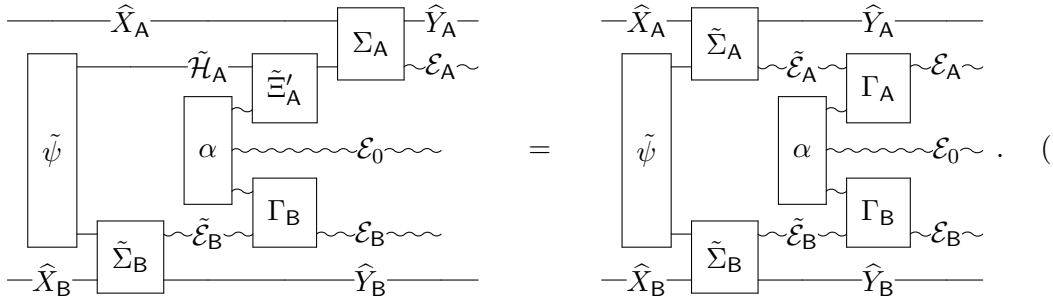
input from a purifying system for $\tilde{\psi}$). We may assume without loss of generality that Γ_{B} is reversible. (If not, we replace Γ_{B} by a Stinespring dilation and absorb the pure state arising on the left hand side into ψ, thus effectively replacing the strategy S by a strategy S' for which the measurement ensemble corresponding to A is still Λ_{A}.)[34]

Now, by Lemma 2.18 the isometric channel Σ_{A} is reversible. If we apply a left-inverse on both sides of Eq. (71), the resulting top output system is $\widehat{X}_{\mathsf{A}}$. When we trash this output and insert an arbitrary state on input $\widehat{X}_{\mathsf{A}}$, we obtain



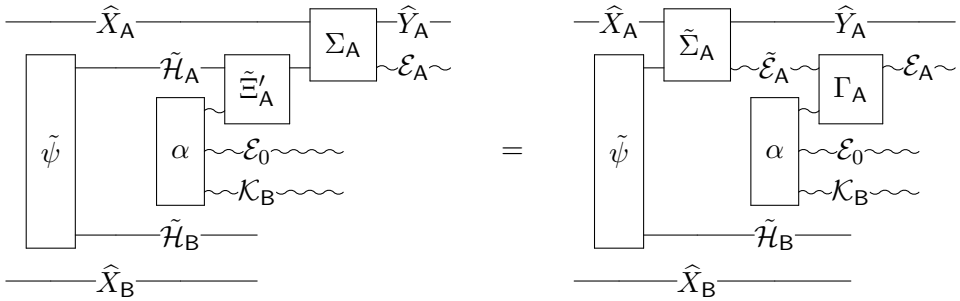
(72)

for some channel $\tilde{\Xi}'_{\mathsf{A}}$. By inserting this fragment back into the left hand side of Eq. (71), we find



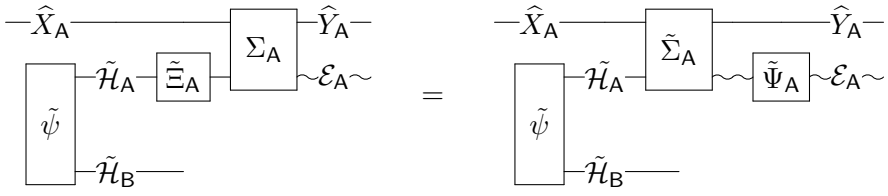
. (73)

Now, Γ_{B} was assumed reversible, so we may cancel it by a left-inverse.[35] Afterwards we may then cancel the reversible channel $\tilde{\Sigma}_{\mathsf{B}}$, which yields the identity



. (74)

If we now insert an arbitrary state into $\widehat{X}_{\mathsf{B}}$ and trace out the output $\widehat{X}_{\mathsf{B}}$ along with the systems $\mathcal{E}_0$ and $\mathcal{K}_{\mathsf{B}}$, we finally obtain



(75)

[34] In the case of a general state $\tilde{\varrho}$, there would be as well a channel Γ_0 in Eq. (71) which may by a similar argument also be assumed reversible.

[35] In the case of general $\tilde{\varrho}$, we would cancel Γ_0 as well.

for some channels $\tilde{\Xi}_{\mathsf{A}}$ and $\tilde{\Psi}_{\mathsf{A}}$.[36] Lastly, since the A-marginal of $\tilde{\psi}$ has full rank, the state $\tilde{\psi}$ gives rise to a Choi-Jamiołkowski-like isomorphism, and for any input to $\widehat{X}_{\mathsf{A}}$ Eq. (75) expresses the equality of two such Choi-Jamiołkowski states. We thus conclude by injectivity that the two channels in Eq. (69) acts identically on any input to $\widehat{X}_{\mathsf{A}}$, and therefore must be identical, as desired. □

We immediately conclude the following:

Corollary 5.10. *Suppose that $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$ is a full-rank strategy which is self-tested according to local assisted simulation by its behaviour P. Let $\tilde{\Sigma}_{\mathsf{A}}$, $\tilde{\Sigma}_{\mathsf{B}}$ be Stinespring dilations of $\tilde{\Lambda}_{\mathsf{A}}$, respectively $\tilde{\Lambda}_{\mathsf{B}}$. For any strategy $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ with behaviour P, and for any Stinespring dilation Σ_{P} of Λ_{P}, $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$, there exist channels $\tilde{\Xi}_{\mathsf{P}}$ and $\tilde{\Psi}_{\mathsf{P}}$ such that Eq. (69) holds.*

Remark 5.11 (Equivalence of Measurement Ensembles)**.** Since assisted simulation is in quantum theory an equivalence relation (Theorem 4.23), the extractibility converse to Proposition 5.6 can be deduced as well if the state ϱ has full local rank. In other words, if $\tilde{S}$ is full-rank and self-tested a.t.l.a.s. by the behaviour P, then for any full-rank strategy S with behaviour P the measurement ensembles of S are *equivalent* to the canonical measurement ensembles, in the sense that either can be extracted from the other. ✠

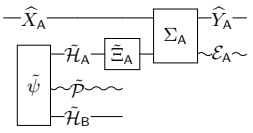
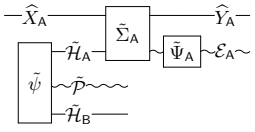
5.C Causal Simulation and Convex Decompositions — Extremality

It is known that a necessary condition for a behaviour P to self-test a strategy a.t.r. is extremality of P in the convex set of all behaviours realisable by quantum strategies. A formal proof of this has been given in Ref. [GKW+18], based on operator-algebraic considerations.

Below, we present a new proof of the statement in operational language. Our proof branches into pieces which collectively reveal that a more general result holds for arbitrary operational theories, and in fact under the weaker notion of self-testing a.t.c.s. (Proposition 5.14). The extremality result for quantum self-testing is obtained as an immediate consequence of this (Corollary 5.15).

Let P be a realisable behaviour. By a *convex decomposition of P* we mean formally a sequence $\big((p_k, P_k)\big)_{k=1,\dots,n}$ such that $p_1, \dots, p_n$ are strictly positive real numbers and such that $P_1, \dots, P_n$ are realisable behaviours with $\sum_{k=1}^n p_k P_k = P$ (it follows that necessarily $\sum_{k=1}^n p_k = 1$).

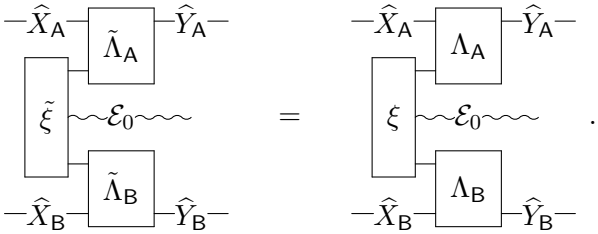
Given a strategy $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ with behaviour P, every convex decomposition of the state ϱ gives rise to a convex decomposition of P. Specifically, if $\varrho = \sum_{k=1}^n p_k \varrho_k$ and if we define P_k as the behaviour of the strategy $(\varrho_k, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$, then clearly $P = \sum_{k=1}^n p_k P_k$. Let us say that a convex decomposition $\big((p_k, P_k)\big)_{k=1,\dots,n}$ is *visible from S* if it arises in this way. (If for example S is pure-state, then all convex decompositions visible from S are trivial, meaning with $P_1 = \ldots = P_n$; indeed, any convex decomposition of the pure state must have $\varrho_1 = \ldots = \varrho_n$.)

[36] For general $\tilde{\varrho}$, the identity would at this point be  =  for some purification $\tilde{\psi}$ of $\tilde{\varrho}$.

Visibility of convex decompositions is linked to causal simulation as follows:

Lemma 5.12. *Let S and $\tilde{S}$ be strategies with common behaviour P. If $S \leq_{c.s.} \tilde{S}$, then any convex decomposition of P visible from S is also visible from $\tilde{S}$.*

Proof. Let $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ and $\tilde{S} = (\tilde{\varrho}, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$. The main observation to make is that for any dilation ξ of ϱ with environment $\mathcal{E}_0$, there exists a dilation $\tilde{\xi}$ of $\tilde{\varrho}$, with environment $\mathcal{E}_0$, such that



(76)

By completeness of purifications (Theorem 2.22), it suffices to show this statement when $\xi = \psi$ is a purification of ϱ. This case however follows from the condition which defines the relation $S \leq_{c.s.} \tilde{S}$ in terms of Stinespring implementations (Eq. (46), with $\tilde{S}$ in place of S and S in place of S'), by tracing out the environments corresponding to the Stinespring dilations Σ_{P} and $\tilde{\Sigma}_{\mathsf{P}}$ of the measurements ensembles.

Having established this fact, let $\big((p_k, P_k)\big)_{k=1,\dots,n}$ be a convex decomposition visible from S. Pick states $\varrho_1, \dots, \varrho_n$ such that $\varrho = \sum_{k=1}^n p_k \varrho_k$ and $(\varrho_k, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ has behaviour P_k. Now, consider the dilation $\xi := \sum_{k=1}^n p_k \varrho_k \otimes |k\rangle\langle k|$ of ϱ, with environment $\mathcal{E}_0 = \mathbb{C}^n$. By the introductory observation, we find a dilation $\tilde{\xi}$ and $\tilde{\varrho}$ such that Eq. (76) holds. The channel on the right hand side of Eq. (76) is given by $\sum_{k=1}^n p_k P_k \otimes |k\rangle\langle k|$. Since the left hand side must be identical, the $\mathcal{E}_0$-marginal of $\tilde{\xi}$ is $\sum_{k=1}^n p_k |k\rangle\langle k|$, so $\tilde{\xi}$ itself must be of the form $\sum_{k=1}^n p_k \tilde{\varrho}_k \otimes |k\rangle\langle k|$ for some states $\tilde{\varrho}_1, \dots, \tilde{\varrho}_n$ with $\sum_{k=1}^n p_k \tilde{\varrho}_k = \tilde{\varrho}$. Letting $\tilde{P}_k$ denote the behaviour of $(\tilde{\varrho}_k, \tilde{\Lambda}_{\mathsf{A}}, \tilde{\Lambda}_{\mathsf{B}})$, the decomposition $\big((\tilde{P}_k, p_k)\big)_{k=1,\dots,n}$ is obviously visible from $\tilde{S}$. But since the left hand side of Eq. (76) is given by $\sum_{k=1}^n p_k \tilde{P}_k \otimes |k\rangle\langle k|$ and must coincide with the right hand side, we have $\tilde{P}_k = P_k$ for all $k = 1, \dots, n$ since $p_1, \dots, p_n \neq 0$. Altogether we conclude that the decomposition $\big((p_k, P_k)\big)_{k=1,\dots,n}$ is visible from $\tilde{S}$, as desired. □

Remark 5.13 (Ekert-Implementations)**.** The implementations appearing in Eq. (76) have a natural operational interpretation, especially in light of Section 4.C. Indeed, they correspond to causal dilations in which the side-information in the environment requires no causes thus thus exists in advance of seeing the inputs x_{A} and x_{B} to the Bell-scenario. Let us call an implementation of $S = (\varrho, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$ an *Ekert-implementation* if it corresponds to a component-wise dilation of the form $(\xi, \Lambda_{\mathsf{A}}, \Lambda_{\mathsf{B}})$, where the environment of ξ is a classical embedded system on which the state ξ is classical and has full rank. (Essentially, Ekert's insight [Eke91] concerns the leak of information to the environment in Ekert-implementations.) It is an easy exercise to show that Ekert-implementations of S correspond to convex decompositions of P visible from S, in the sense that they as channels are precisely given by $\sum_{k=1}^n p_k |k\rangle\langle k| \otimes P_k$ with $\big((p_k, P_k)\big)_{k=1,\dots,n}$ a convex decomposition visible from S. The essence of Lemma 5.12 consists in observing that if $S \leq_{c.s.} \tilde{S}$ then every Ekert-implementation of S is also an Ekert-implementation of $\tilde{S}$. ✠

Lemma 5.12 can be used to show the following:

Proposition 5.14. *Suppose that P self-tests $\tilde{S}$ a.t.c.s. Then, every convex decomposition of P is visible from $\tilde{S}$.*

Proof. By Lemma 5.12, it suffices to argue that every possible convex decomposition of P is visible from some strategy S with behaviour P. But this is easy to see: Given a convex decomposition $\left((p_k, P_k)\right)_{k=1,\dots,n}$, choose strategies $S_1 = (\varrho_1, \Lambda_{\mathsf{A},1}, \Lambda_{\mathsf{B},1}), \dots, S_n = (\varrho_n, \Lambda_{\mathsf{A},n}, \Lambda_{\mathsf{B},n})$ such that S_k has behaviour P_k. We may assume without loss of generality that the systems $\mathcal{H}_{\mathsf{A},k}$ are all equal to a fixed system, say $\mathcal{H}_{\mathsf{A}}$, by embedding them into a common system if necessary. Likewise, we may assume all the systems $\mathcal{H}_{\mathsf{B},k}$ equal to a fixed system $\mathcal{H}_{\mathsf{B}}$. Now, let S be the strategy whose state is $\sum_{k=1}^n p_k \, |k\rangle\langle k|_{\mathsf{A}} \otimes |k\rangle\langle k|_{\mathsf{B}} \otimes \varrho_k$ on the bipartite system $(\mathbb{C}^n \otimes \mathcal{H}_{\mathsf{A}}) \otimes (\mathbb{C}^n \otimes \mathcal{H}_{\mathsf{B}})$, and whose measurement ensemble Λ_{P} is defined by reading off the classical value k on the subsystem $\mathbb{C}^n$ and applying the according ensemble $\Lambda_{\mathsf{P},k}$. The behaviour of the strategy S is evidently $\sum_{k=1}^n p_k P_k = P$, and the convex decomposition $\left((p_k, P_k)\right)_{k=1,\dots,n}$ is visible from S by construction. □

Proposition 5.14 constitutes our key result about self-testing and convex decompositions: it expresses that a strategy which is self-tested according to causal simulation fully determines all possible decompositions of its behaviour.

In quantum theory, we may use the result to infer extremality of behaviours which self-test strategies (even according to local assisted simulation):

Corollary 5.15. *If P self-tests some quantum strategy a.t.l.a.s., then P is extremal in the convex set of quantum realisable behaviours.*

Proof. If P self-tests some quantum strategy a.t.l.a.s., then all quantum strategies with behaviour P are equivalent under local assisted simulation by Theorem 4.23. Therefore, P self-tests any strategy with behaviour P a.t.l.a.s. By the purification argument in the proof of Corollary 5.4, there is some pure-state strategy $\tilde{S}$ with behaviour P, and it holds in particular that P self-tests $\tilde{S}$ a.t.l.a.s. By Proposition 5.14, any convex decomposition of P is visible from $\tilde{S}$, and as $\tilde{S}$ is pure-state every convex decomposition visible from $\tilde{S}$ is trivial. Thus, P is extremal. □

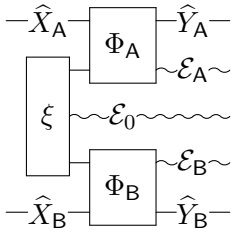
We note that there are known examples of extremal behaviours which do not self-test any strategy according to reducibility [TFR+21]. However, some or all of these might self-test a strategy according to local assisted simulation (a.t.l.a.s.); indeed the converse to Corollary 5.15 might be true. We leave this question open for future investigations.

5.D Exhausted Environments

In this final subsection, we point out a feature of self-testing which we believe has so far gone unnoticed. We do not currently know any concrete applications of this feature, but it shows up naturally from our operational reformulation and we suspect that it might be useful.

Suppose that parties A and B manage to produce a behaviour P which we know to self-test some strategy $\tilde{S}$ according to causal simulation (in particular, this holds if P self-tests $\tilde{S}$ in the conventional sense, i.e. according to reducibility). From the perspective of a referee who probes the behaviour by sending inputs $x_{\mathsf{A}}, x_{\mathsf{B}}$ and receiving outputs $y_{\mathsf{A}}, y_{\mathsf{B}}$, the information-processing is described by the channel P. From a global perspective including the environment, however, the information-processing is governed by the various

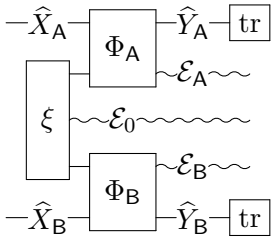
implementations



(77)

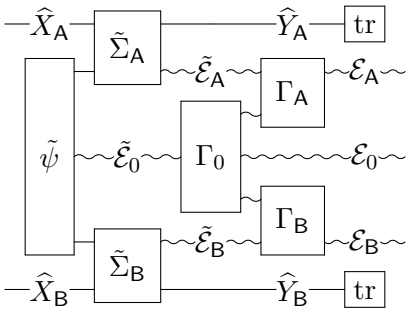
of a strategy $(\varrho, \Lambda_\mathsf{A}, \Lambda_\mathsf{B})$. (By the discussion in Section 4.C.4, this is even true when strategies corresponding to general circuits are employed; any possible information-processing is derivable from an implementation of the form (77).)

Now, let us consider what the information-processing looks like *from the perspective of the environment*. This amounts to tracing out the system $\widehat{Y}_\mathsf{A}$ and $\widehat{Y}_\mathsf{B}$ and considering the channel



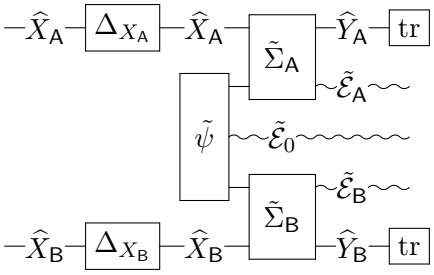
(78)

When the implementation is a Stinespring implementation, the channel (78) is a *complementary channel* [DS05] to the behaviour channel P, however as a causal channel (cf. Section 4.C) it has additionally a specific structure according to which the information in $\mathcal{E}_0$ is available from the beginning and the information in $\mathcal{E}_\mathsf{P}$ becomes available upon the input to $\widehat{X}_\mathsf{P}$. Since P self-tests $\tilde{S}$ a.t.c.s., the channel (78) must be of the form



(79)

for some channels Γ_0, Γ_A and Γ_B, with $(\tilde{\psi}, \tilde{\Sigma}_\mathsf{A}, \tilde{\Sigma}_\mathsf{B})$ a component-wise Stinespring dilation of $\tilde{S}$. As such, when the referee proves the behaviour by supplying only classical inputs to $\widehat{X}_\mathsf{P}$, the strongest possible leak of information to the environment is represented by the channel



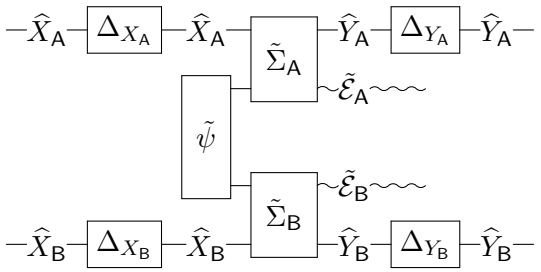
(80)

and this channel can be computed by the referee, because the strategy $\tilde{S}$ is known. Specifically, the information in $\tilde{\mathcal{E}}_0$ is the maximal information the environment may hold before the probing begins, and the information in $\tilde{\mathcal{E}}_{\mathsf{P}}$ is the maximal information which may leak to the environment after receiving the input to $\widehat{X}_{\mathsf{P}}$.

It is no surprise that, when computing the channel (80), the system $\tilde{\mathcal{E}}_{\mathsf{P}}$ will at least hold a copy of the input x_{P} and the output y_{P}, or information equivalent to these copies; indeed these two classical values may always be copied by the environment and kept in memory. It is however perfectly conceivable that the environment will after the interaction with the referee hold information additional to these memories, information which we may think of as undisclosed to the referee. (For example, one could imagine that $\mathcal{E}_{\mathsf{A}}$ would hold a copy of the output y_{B}, or that $\mathcal{E}_0$ would hold non-trivial information about either output.)

We will show now that this does not happen when $\tilde{S}$ is a pure-state projective strategy whose measurements are described by rank-one projections. This is the case e.g. in the optimal canonical strategy for the CHSH-game [CHSH69, ŠB20]. In other words, two parties A and B who play this game optimally essentially cannot know anything in advance of the game, and will be forced to reveal as output to the referee everything they learn upon receiving their inputs. Hence, they are effectively *exhausted* of information.

The proof of this is fairly simple. Suppose that the canonical strategy $\tilde{S} = (\tilde{\psi}, \tilde{\Pi}_{\mathsf{A}}, \tilde{\Pi}_{\mathsf{B}})$ is pure-state and projective, and that all of the projections $\tilde{\Pi}_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})$ have rank one (this implies, incidentally, than, $\dim \tilde{\mathcal{H}}_{\mathsf{P}} = |Y_{\mathsf{P}}|$). If we choose the Stinespring dilations $\tilde{\Sigma}_{\mathsf{A}}$ and $\tilde{\Sigma}_{\mathsf{B}}$ conveniently as in the proof of Proposition 4.14, then the channel



, (81)

where $\tilde{\mathcal{E}}_{\mathsf{A}} = \widehat{X}_{\mathsf{A}} \otimes \tilde{\mathcal{H}}_{\mathsf{A}}$ and $\tilde{\mathcal{E}}_{\mathsf{B}} = \widehat{X}_{\mathsf{B}} \otimes \tilde{\mathcal{H}}_{\mathsf{B}}$ (and where $\tilde{\mathcal{E}}_0$ is trivial by purity of the state), is given by

$$F \mapsto \sum_{x \in X, y \in Y} \Big(\langle x| F |x\rangle \, |x\rangle\langle x| \otimes \tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle\langle\tilde{\psi}| \, \tilde{\Pi}^x(y) \Big) \otimes |y\rangle\langle y| \tag{82}$$

for $F \in \mathrm{End}(\widehat{X}_{\mathsf{A}} \otimes \widehat{X}_{\mathsf{B}})$, parenthesising the parts which belong to the environment. Now, if we let $|\phi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})\rangle$ denote a unit vector corresponding to the rank-one projection $\tilde{\Pi}_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})$, and if $|\phi^x(y)\rangle = |\phi_{\mathsf{A}}^{x_{\mathsf{A}}}(y_{\mathsf{A}})\rangle \otimes |\phi_{\mathsf{B}}^{x_{\mathsf{B}}}(y_{\mathsf{B}})\rangle$ for $x = (x_{\mathsf{A}}, x_{\mathsf{B}})$ and $y = (y_{\mathsf{A}}, y_{\mathsf{B}})$, then we have the identity $\tilde{\Pi}^x(y) \, |\tilde{\psi}\rangle = \langle \phi^x(y) | \tilde{\psi} \rangle \, |\phi^x(y)\rangle$, with $|\langle \phi^x(y) | \tilde{\psi} \rangle|^2$ equal to the probability of output y on input x, namely $P^x(y)$. Hence, the channel (82) can be re-expressed as

$$F \mapsto \sum_{x \in X, y \in Y} \langle x| F |x\rangle \, P^x(y) \Big(\, |x\rangle\langle x| \otimes |\phi^x(y)\rangle\langle\phi^x(y)| \, \Big) \otimes |y\rangle\langle y| \, . \tag{83}$$

For any $\mathsf{P} \in \{\mathsf{A}, \mathsf{B}\}$ and any $x_{\mathsf{P}} \in X_{\mathsf{P}}$, the pure states $(\phi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}}))_{y_{\mathsf{P}} \in Y_{\mathsf{P}}}$ are perfectly distinguishable, indeed their vector representatives form an orthonormal basis of $\tilde{\mathcal{H}}_{\mathsf{P}}$. They can therefore be viewed as an encoding of $y_{\mathsf{P}} \in Y_{\mathsf{P}}$ in terms of pure states on $\tilde{\mathcal{H}}_{\mathsf{P}} \cong \widehat{Y}_{\mathsf{P}}$. More precisely, there exist unitaries $U_{\mathsf{P}} : \widehat{X}_{\mathsf{P}} \otimes \tilde{\mathcal{H}}_{\mathsf{P}} \to \widehat{X}_{\mathsf{P}} \otimes \widehat{Y}_{\mathsf{P}}$ which map $|x_{\mathsf{P}}\rangle \otimes |\phi_{\mathsf{P}}^{x_{\mathsf{P}}}(y_{\mathsf{P}})\rangle$

to $|x_{\mathsf{P}}\rangle \otimes |y_{\mathsf{P}}\rangle$. As such, the channel (81) is equivalent (under application of unitaries U_{A} and U_{B} to the systems $\tilde{\mathcal{E}}_{\mathsf{A}}$ and $\tilde{\mathcal{E}}_{\mathsf{B}}$) to the channel given by

$$F \mapsto \sum_{x\in X, y\in Y} \langle x| F |x\rangle \, P^x(y) \Big(|x\rangle\langle x| \otimes |y\rangle\langle y| \Big) \otimes |y\rangle\langle y| \,. \tag{84}$$

In other words, when executing the behaviour P the only possible information which can possibly reside in the environment is, upon input x_{P}, a copy of x_{P} and the corresponding output y_{P}.

6 Conclusions

Quantum self-testing is an important phenomenon in quantum cryptography which arose from Bell's considerations about experimental scenarios with separated parties. Today, it has many theoretical and practical ramifications in quantum information theory. However, the conventional mathematical formalisation of self-testing is phrased in the operator-algebraic formalism for quantum theory, and therefore resists an operational interpretation and immediate generalisation e.g. to alternative theories, formalisms or causal scenarios.

After discussing operational structure in physical theories (Section 2) and reviewing the conventional operator-algebraic narrative of quantum self-testing (Section 3), we have turned in Section 4 to rework the perception of self-testing and cast it anew as an operational phenomenon.

In Section 4.A we have proposed a new definition of self-testing (Definition 4.2) according to the concept of local simulation (Definition 4.7). It specialises in the case of quantum theory to the standard operator-algebraic definition (Theorem 4.9), it answers an unresolved issue regarding a restriction to projective strategies (Proposition 4.13), and it naturally lends itself to definitions of approximate self-testing (Remark 4.18).

Our reformulation shows how to think of self-testing operationally, in terms of side-information which gradually leaks to an environment. By meditating on this idea, we have proposed two relaxations of local simulation, namely local assisted simulation (Definition 4.19) and causal simulation (Definition 4.28). In Section 4.B, we have studied local assisted simulation and seen that it is in quantum theory the equivalence relation generated by local simulation (Theorem 4.23 and Remark 4.24). This sheds new light on quantum self-testing by uncovering a two-fold character of the phenomenon. In particular, it suggests a notion of self-testing with emphasis on upper bounding the abilities of adversaries (Definition 4.25). Causal simulation yields a further relaxation of self-testing, and we have explained in Section 4.C how this ultimately realises quantum self-testing within a larger framework of causally structured channels and dilations (most important in this regard are Theorem 4.44, Theorem 4.46 and Theorem 4.49). It is our hope that this realisation may eventually be beneficial in cryptographic contexts and point to modular, composable features of self-testing (see also Section 4.3 in Ref. [HL21]).

In Section 5, we have shown that our operational reformulation of self-testing may be used to give conceptually simple proofs of several features of self-testing. Specifically, we have considered extractability of canonical states (Proposition 5.1) and canonical measurements (Proposition 5.6), limitations on convex decompositions of behaviours (Proposition 5.14 and Corollary 5.15), and a novel feature of information exhaustion (Section 5.D). The fact that we are able to recover well-known features from our reformulation indicates that it is at least as potent as the conventional formulation — we hope, of course, that our reformulation will eventually prove itself even more potent by easing more arguments

which involve self-testing.

The current work opens up several avenues for future investigation. A rather concrete problem is the further study of *approximate* simulation and *robust* self-testing. As mentioned in Remark 4.18, the Bures distance β [KSW08a, KSW08b] and purified diamond distance $P_\diamond$ [HL21] have properties which render them suitable for quantifying approximate simulation, but we have not determined the distances in concrete cases. In particular, if $\tilde{S}$ is the canonical optimal strategy for the CHSH-game [CHSH69, ŠB20], we do not know, given $\delta > 0$, what is the optimal value of $\varepsilon > 0$ such that any strategy S whose behaviour is δ-close to the behaviour of $\tilde{S}$ locally ε-simulates $\tilde{S}$ (with closeness measured either by β or $P_\diamond$).[37] An important insight from our work is that the local isometries which relate one strategy to another according to local simulation can be chosen dependently on the local inputs x_A and x_B (cf. Proposition 4.14); this turns out to be invisible in the case of exact local simulation (Lemma 4.17), but we suspect that it will result in differences in the approximate case.

Another open problem is the following. Usually, a proof that a concrete behaviour P self-tests a concrete strategy $\tilde{S}$ (e.g. the optimal strategy for the CHSH-game) is obtained by deriving certain operator-algebraic identities which the constituents of a general strategy S must satisfy if its behaviour is P, and then from these identities concluding the existence of local isometries and a residual state witnessing reducibility to the canonical strategy $\tilde{S}$. Given that we have been able to reformulate self-testing operationally, the following question is natural: *Is it possible to find an operational proof that $\tilde{S}$ is self-tested by its behaviour?* Such a proof would not directly rely on operator-algebra, but would rather use operational features of the theory **QIT**. Even for well-studied cases such as the canonical optimal strategy in the CHSH-game, an operational proof of self-testing (according to any mode of simulation) would yield new insights. It would even be interesting to find an operational proof of the fact that the CHSH-game has the property of exhausting quantum players of information, as described in Section 5.D.[38]

Finally, it would be natural to investigate self-testing statements in other causal scenarios and other operational theories.[39] A more systematic study of this might clarify the special case of quantum self-testing and eventually bring new insights to the field.

Acknowledgements

We thank Jędrzej Kaniewski for comments on an earlier version of this manuscript. We acknowledge financial support from the European Research Council (ERC Grant Agreement No. 81876), VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059), the QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme (QuantAlgo project) via the Innovation Fund Denmark and the Novo Nordisk Foundation (Grant NNF20OC0059939 'Quantum for Life'). LM is also supported by VILLUM FONDEN (Young Investigator Grant No. 37532).

[37] Likewise, we do not know the optimal value of ε when instead of assuming the behaviours to be δ-close it is merely assumed that the winning probability of S is δ-close to the optimal winning probability.

[38] It is worth observing that this proof would necessarily involve aspects of quantum theory and cannot be a feature of the CHSH-game by itself, since the game is not exhaustive for classical adversaries who play optimally — for example, in the optimal classical strategy which is the equal probabilistic mixture of optimal deterministic strategies, both parties A and B know the function which the other party uses to determine an output, while this function need not be disclosed to the referee.

[39] A few ideas in this direction are presented in Ref. [HL21], in particular as regards the theory **CIT**.

Authors' Contributions

This paper is based on the PhD thesis of NGHL [HL21]. NGHL is the main author of the paper, including conceptualisation and methodology of the work and the formal analysis and writing. MC and LM contributed to the conceptualisation and methodology of the work and reviewed and edited the draft. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

References

[AC04] Samson Abramsky and Bob Coecke, *A Categorical Semantics of Quantum Protocols*, Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, 2004, IEEE, 2004, https://doi.org/10.1109/LICS.2004.1319636.

[Bae06] John C. Baez, *Quantum Quandaries: A Category-Theoretic Perspective*, The structural foundations of quantum gravity (2006), 240–265, https://doi.org/10.1093/acprof:oso/9780199269693.003.0008.

[Bar07] Jonathan Barrett, *Information Processing in Generalized Probabilistic Theories*, Physical Review A **75** (2007), no. 3, 032304, https://doi.org/10.1103/PhysRevA.75.032304.

[Bel64] John S. Bell, *On the Einstein Podolsky Rosen paradox*, Physics Physique Fizika **1** (1964), no. 3, 195–200, https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195.

[BGP10] Cyril Branciard, Nicolas Gisin, and Stefano Pironio, *Characterizing the nonlocal correlations created via entanglement swapping*, Physical review letters **104** (2010), no. 17, 170401, https://doi.org/10.1103/PhysRevLett.104.170401.

[BMP18] Cédric Bamps, Serge Massar, and Stefano Pironio, *Device-independent randomness generation with sublinear shared quantum resources*, Quantum **2** (2018), 86, https://doi.org/10.22331/q-2018-08-22-86.

[BRGP12] Cyril Branciard, Denis Rosset, Nicolas Gisin, and Stefano Pironio, *Bilocal versus nonbilocal correlations in entanglement-swapping experiments*, Physical Review A **85** (2012), no. 3, 032119, https://doi.org/10.1103/PhysRevA.85.032119.

[BW11] Howard Barnum and Alexander Wilce, *Information Processing in Convex Operational Theories*, Electronic Notes in Theoretical Computer Science **270** (2011), no. 1, 3–15, https://doi.org/10.1016/j.entcs.2011.01.002.

[BW16] ———, *Post-Classical Probability Theory*, Quantum Theory: Informational Foundations and Foils, Springer, 2016, https://doi.org/10.1007/978-94-017-7303-4_11.

[CDP09] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, *Theoretical Framework for Quantum Networks*, Physical Review A **80** (2009), no. 2, 022339, https://doi.org/10.1103/PhysRevA.80.022339.

[CDP10] ———, *Probabilistic Theories with Purification*, Physical Review A **81** (2010), no. 6, 062348, https://doi.org/10.1103/PhysRevA.81.062348.

[CDP11] ———, *Informational derivation of quantum theory*, Physical Review A **84** (2011), no. 1, 012311, https://doi.org/10.1103/PhysRevA.84.012311.

[CFH21] Matthias Christandl, Roberto Ferrara, and Karol Horodecki, *Upper bounds on device-independent quantum key distribution*, Physical Review Letters **126** (2021), no. 16, 160501, https://doi.org/10.1103/PhysRevLett.126.160501.

[CFS16] Bob Coecke, Tobias Fritz, and Robert W. Spekkens, *A mathematical theory of resources*, Information and Computation **250** (2016), 59–86, https://doi.org/10.1016/j.ic.2016.02.008.

[CGJV19] Andrea Coladangelo, Alex B Grilo, Stacey Jeffery, and Thomas Vidick, *Verifier-on-a-leash: new schemes for verifiable delegated quantum computation, with quasilinear resources*, Annual International Conference on the The-

ory and Applications of Cryptographic Techniques, Springer, 2019, https://doi.org/10.1007/978-3-030-17659-4_9.

[Chi14] Giulio Chiribella, *Distinguishability and copiability of programs in general process theories*, Int J Software Informatics **1** (2014), no. 2, 209–223, https://doi.org/10.48550/arXiv.1411.3035.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, *Proposed experiment to test local hidden-variable theories*, Physical review letters **23** (1969), no. 15, 880–884, https://doi.org/10.1103/PhysRevLett.23.880.

[Cir93] Boris S. Cirel'son, *Some results and problems on quantum Bell-type inequalities*, Hadronic Journal Supplement **8** (1993), no. 4, 329–345.

[CK11] Roger Colbeck and Adrian Kent, *Private randomness expansion with untrusted devices*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 9, 095305, https://doi.org/10.1088/1751-8113/44/9/095305.

[CL13] Bob Coecke and Raymond Lal, *Causal Categories: Relativistically interacting Processes*, Foundations of Physics **43** (2013), no. 4, 458–501, https://doi.org/10.1007/s10701-012-9646-8.

[Coe14] Bob Coecke, *Terminality implies non-signalling*, arXiv preprint (2014), https://doi.org/10.48550/arXiv.1405.3681.

[Col06] Roger Colbeck, *Quantum and relativistic protocols for secure multi-party computation*, Ph.D. thesis, University of Cambridge, 2006, https://doi.org/10.48550/arXiv.0911.3814.

[Col20] Andrea Coladangelo, *A two-player dimension witness based on embezzlement, and an elementary proof of the non-closure of the set of quantum correlations*, Quantum **4** (2020), 282, https://doi.org/10.22331/q-2020-06-18-282.

[CS18] Andrea Coladangelo and Jalex Stark, *Unconditional separation of finite and infinite-dimensional quantum correlations*, arXiv preprint (2018), https://doi.org/10.48550/arXiv.1804.05116.

[CY14] Matthew Coudron and Henry Yuen, *Infinite randomness expansion with a constant number of devices*, Proceedings of the forty-sixth annual ACM symposium on Theory of computing, 2014, https://doi.org/10.1145/2591796.2591873.

[DS05] Igor Devetak and Peter W. Shor, *The capacity of a quantum channel for simultaneous transmission of classical and quantum information*, Communications in Mathematical Physics **256** (2005), no. 2, 287–303, https://doi.org/10.1007/s00220-005-1317-6.

[Eke91] Artur K. Ekert, *Quantum Cryptography based on Bell's Theorem*, Physical Review Letters **67** (1991), no. 6, 661–663, https://doi.org/10.1103/PhysRevLett.67.661.

[Fri12] Tobias Fritz, *Beyond bell's theorem: correlation scenarios*, New Journal of Physics **14** (2012), no. 10, 103001, https://doi.org/10.1088/1367-2630/14/10/103001.

[Fri20] ———, *A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics*, Advances in Mathematics **370** (2020), 107239, https://doi.org/10.1016/j.aim.2020.107239.

[GKW+18] Koon Tong Goh, Jędrzej Kaniewski, Elie Wolfe, Tamás Vértesi, Xingyao Wu, Yu Cai, Yeong-Cherng Liang, and Valerio Scarani, *Geometry of the set of quantum correlations*, Physical Review A **97** (2018), no. 2, 022104, https://doi.org/10.1103/PhysRevA.97.022104.

[Har01] Lucien Hardy, *Quantum theory from five reasonable axioms*, arXiv preprint (2001), https://doi.org/10.48550/arXiv.quant-ph/0101012.

[Har13] Lucien Hardy, *A formalism-local framework for general probabilistic theories, including quantum theory*, Mathematical Structures in Computer Science **23** (2013), no. 2, 399–440, https://doi.org/10.1017/S0960129512000163.

[HL21] Nicholas Gauguin Houghton-Larsen, *A Mathematical Framework for Causally Structured Dilations and its Relation to Quantum Self-Testing*, Ph.D. thesis, University of Copenhagen, 2021, https://doi.org/10.48550/arXiv.2103.02302.

[HM15] Karol Horodecki and Gláucia Murta, *Bounds on quantum nonlocality via partial transposition*, Physical Review A **92** (2015), no. 1, 010301, https://doi.org/10.1103/PhysRevA.92.010301.

[JMS20] Rahul Jain, Carl A. Miller, and Yaoyun Shi, *Parallel device-independent quantum key distribution*, IEEE transactions on information theory **66** (2020), no. 9, 5567–5584, https://doi.org/10.1109/TIT.2020.2986740.

[JNV+21] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen, *MIP*=RE*, Commun. ACM **64** (2021), no. 11, 131–138, https://doi.org/10.1145/3485628.

[Kan17] Jędrzej Kaniewski, *Self-testing of binary observables based on commutation*, Physical Review A **95** (2017), no. 6, 062323, https://doi.org/10.1103/PhysRevA.95.062323.

[KSW08a] Dennis Kretschmann, Dirk Schlingemann, and Reinhard F. Werner, *A continuity theorem for Stinespring's dilation*, Journal of Functional Analysis **255** (2008), no. 8, 1889–1904, https://doi.org/10.1016/j.jfa.2008.07.023.

[KSW08b] ———, *The information-disturbance tradeoff and the continuity of Stinespring's representation*, IEEE transactions on information theory **54** (2008), no. 4, 1708–1717, https://doi.org/10.1109/TIT.2008.917696.

[KU17] Aleks Kissinger and Sander Uijlen, *A Categorical Semantics for Causal Structure*, 2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, 2017, https://doi.org/10.23638/LMCS-15(3:15)2019.

[McK16] Matthew McKague, *Interactive Proofs for BQP via Self-Tested Graph States*, Theory of Computing **12** (2016), no. 3, 1–42, https://doi.org/10.4086/toc.2016.v012a003.

[ML13] Saunders Mac Lane, *Categories for the Working Mathematician*, vol. 5, Springer Science & Business Media, 2013, https://doi.org/10.1007/978-1-4757-4721-8.

[MS16] Carl A. Miller and Yaoyun Shi, *Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices*, Journal of the ACM (JACM) **63** (2016), no. 4, 1–63, https://doi.org/10.1145/2885493.

[MY98] Dominic Mayers and Andrew Yao, *Quantum Cryptography with Imperfect Apparatus*, Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No. 98CB36280), IEEE, 1998, https://doi.org/10.1109/sfcs.1998.743501.

[MY04] Dominic Mayers and Andrew Yao, *Self testing quantum apparatus*, Quantum Info. Comput. **4** (2004), no. 4, 273–286, https://doi.org/10.48550/arXiv.quant-ph/0307205.

[MYS12] Matthew McKague, Tzyh Haur Yang, and Valerio Scarani, *Robust self-testing of the singlet*, Journal of Physics A: Mathematical and Theoretical **45** (2012), no. 45, 455304, https://doi.org/10.1088/1751-8113/45/45/455304.

[Nai40] Mark Naimark, *Spectral functions of a symmetric operator*, Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya **4** (1940), no. 3, 277–318.

[NV17] Anand Natarajan and Thomas Vidick, *A quantum linearity test for robustly verifying entanglement*, Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, 2017, https://doi.org/10.1145/3055399.3055468.

[NW19] Anand Natarajan and John Wright, *NEEXP is contained in MIP*, 2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2019, https://doi.org/10.1109/FOCS.2019.00039.

[Per17] Paolo Perinotti, *Causal Structures and the Classification of Higher Order Quantum Computations*, Time in physics, Springer, 2017, https://doi.org/10.1007/978-3-319-68655-4_7.

[RB22] Marc-Olivier Renou and Salman Beigi, *Nonlocality for generic networks*, Phys. Rev. Lett. **128** (2022), 060401, https://doi.org/10.1103/PhysRevLett.128.060401.

[RUV13] Ben W. Reichardt, Falk Unger, and Umesh Vazirani, *Classical Command of Quantum Systems*, Nature **496** (2013), no. 7446, 456–460, https://doi.org/10.1038/nature12035.

[ŠASA16] Ivan Šupić, Remigiusz Augusiak, Alexia Salavrakos, and Antonio Acín, *Self-testing protocols based on the chained Bell inequalities*, New Journal of Physics **18** (2016), no. 3, 035013, https://doi.org/10.1088/1367-2630/18/3/035013.

[ŠB20] Ivan Šupić and Joseph Bowles, *Self-testing of quantum systems: a review*, Quantum **4** (2020), 337, https://doi.org/10.22331/q-2020-09-30-337.

[Sel04] Peter Selinger, *Towards a Semantics for Higher-Order Quantum Computation*, Proceedings of the 2nd International Workshop on Quantum Programming Languages, TUCS General Publication, vol. 33, 2004, pp. 127–143.

[SFK+20] David Schmid, Thomas C. Fraser, Ravi Kunjwal, Ana Belen Sainz, Elie Wolfe, and Robert W. Spekkens, *Understanding the interplay of entanglement and nonlocality: motivating and developing a new branch of entanglement theory*, arXiv preprint (2020), https://doi.org/10.48550/arXiv.2004.09194.

[Sti55] William Forrest Stinespring, *Positive functions on C*-algebras*, Proceedings of the American Mathematical Society **6** (1955), no. 2, 211–216, https://doi.org/10.2307/2032342.

[TCR10] Marco Tomamichel, Roger Colbeck, and Renato Renner, *Duality between smooth min- and max-entropies*, IEEE Transactions on information theory **56** (2010), no. 9, 4674–4681, https://doi.org/10.1109/TIT.2010.2054130.

[TFR+21] Armin Tavakoli, Máté Farkas, Denis Rosset, Jean-Daniel Bancal, and Jedrzej Kaniewski, *Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments*, Science Advances **7** (2021), no. 7, https://doi.org/10.1126/sciadv.abc3847.

[Tom12] Marco Tomamichel, *A framework for non-asymptotic quantum information theory*, Ph.D. thesis, ETH Zürich, 2012, https://doi.org/10.48550/arXiv.1203.2142.