Incompatibility as a Resource for Programmable Quantum Instruments

Kaiyuan Ji$^{1,2,*}$ and Eric Chitambar$^{3,†}$

$^1$Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China
$^2$School of Electrical and Computer Engineering, Cornell University, Ithaca, New York 14850, USA
$^3$Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

(Dated: March 11, 2024)

Quantum instruments represent the most general type of quantum measurement, as they incorporate processes with both classical and quantum outputs. In many scenarios, it may be desirable to have some “on-demand” device that is capable of implementing one of many possible instruments whenever the experimenter desires. We refer to such objects as programmable instrument devices (PIDs), and this paper studies PIDs from a resource-theoretic perspective. A physically important class of PIDs are those that do not require quantum memories to implement, and these are naturally “free” in this resource theory. Additionally, these free objects correspond precisely to the class of unsteerable channel assemblages in the study of channel steering. The traditional notion of measurement incompatibility emerges as a resource in this theory since any PID controlling an incompatible family of instruments requires a quantum memory to build. We identify an incompatibility preorder between PIDs based on whether one can be transformed into another using processes that do not require additional quantum memories. Necessary and sufficient conditions are derived for when such transformations are possible based on how well certain guessing games can be played using a given PID. Ultimately our results provide an operational characterization of incompatibility, and they offer semi-device-independent tests for incompatibility in the most general types of quantum instruments.

CONTENTS

I. Introduction 1
A. From programmability to nonsignaling 2
B. Organization of the paper 3

II. Programmable Quantum Devices 4
A. Programmable measurement devices 4
B. Programmable instrument devices 4

III. Free Simulations of Programmable Devices 5
A. Free simulations of PMDs 5
B. Free simulations of PIDs 6

IV. Relationship with Steering and Measurement Incompatibility 7
A. PIDs as assemblages in channel steering 7
B. Steering-equivalence mapping 8

V. Semi-Device-Independent Characterization with Nontransient Guessing Games 9
A. A complete set of incompatibility monotones 10
B. Robustness of incompatibility as the supremum of game advantage 11
C. Constructing experimentally friendly incompatibility tests 11

VI. Conclusions and Discussions 12
A. Summary of results 12
B. Outlook 13

Acknowledgements 13

A. Properties of Free Simulations 13
1. Proof of Theorem 1(1) 13
2. Proof of Theorem 1(2) 14
3. Proof of Theorem 1(3) 15
4. Proof of Theorem 1(4) 16
5. Proof of Theorem 1(5) 17

B. Properties of the Steering-Equivalence Mapping 18
1. Proof of Proposition 1 18
2. Proof of Proposition 2 19
3. Proof of Theorem 2(1) 20
4. Proof of Theorem 2(2) 20

C. Semi-Device-Independent Characterization 21
1. Proof of Theorem 3 21
2. Proof of Theorem 4 23
3. Proof of Proposition 3 26

References 27

I. INTRODUCTION

Incompatibility is a quintessential feature of quantum mechanics. Unlike classical systems in which conjugate variables have definite values at each moment in time, quantum systems are dictated by celebrated uncertainty relations, which place sharp restrictions on how well the measurement outcomes of two (or more) noncommuting observables can be predicted [1]. The incompatibility of noncommuting observables has wide-ranging applications in quantum information science from quantum cryptography [2, 3] to entanglement detection [4] to quantum error correction [5]. For more general types of
measurements beyond textbook observables, commutation relations are no longer sufficient to characterize measurement incompatibility. One instead considers the property of joint measurability, which means that a joint probability distribution can be defined for the given collection of measurement devices, each being described by a positive operator-valued measure (POVM) [6–8]. Incompatibility between POVMs in this sense means that such joint measurability is not possible.

Whereas POVMs characterize the classical output of a quantum measurement, a more general description of the measurement process also includes the quantum output. Here, one typically invokes the theory of quantum instruments [9], with an instrument formally being defined as a family of completely positive (CP) maps \( \{ \Lambda_{x_i} \}_{x_i} \) such that \( \sum_{x_i} \Lambda_{x_i} \) is trace preserving (TP). When performing an instrument on a quantum state \( \rho \), a classical outcome \( x_i \) is observed with probability \( p_{x_i} = \text{Tr}[\Lambda_{x_i} [\rho]] \), and the postmeasurement state is then given by \( \Lambda_{x_i} [\rho] / p_{x_i} \). Note that POVMs are a special type of instrument for which \( \Lambda_{x_i} [\rho] = \text{Tr}[M_{x_i} \rho] \) for some collection of positive semidefinite operators \( \{ M_{x_i} \}_{x_i} \) with \( \sum_{x_i} M_{x_i} = 1 \). Likewise, a quantum channel (i.e., a CPTP map) is also a type of quantum instrument, having just a single classical outcome. The notion of incompatibility can also be extended into the domain of channels and instruments [8, 10–12]. Similar to the case of POVMs, a family of instruments \( \{ \Lambda_{x_i|x_0} \}_{x_0,x_i} \) is compatible if all the constituent instruments can be simulated using a single instrument combined with classical postprocessing: incompatible instruments lack this property.

Extensive work has recently been conducted to capture incompatibility as a physical resource in quantum information processing [13–21]. This can be accomplished using the formal structure of a resource theory [22–26], in which objects are characterized as being either free or resourceful. Additionally, only a restricted set of physical operations can be performed by the experimenter, and these are unable to create resourceful objects from free ones. In the case of quantum incompatibility, the free objects are compatible families of POVMs or instruments, and the incompatible ones are resourceful.

By adopting a resource-theoretic perspective, one can establish operationally meaningful measures of incompatibility such as its robustness to noise [14, 18, 27–30]. The incompatibility in one family of POVMs or instruments can then be quantitatively compared to another. Resource theories also provide tools for detecting or “witnessing” the incompatibility present in general measurement devices [29, 31–34]. This certification can also be done in a semi-device-independent way [13, 18–20, 25, 35–37]. In other words, by attaining a certain score on some type of quantum measurement game, the experimenter can rest assured that he or she is controlling some family of incompatible POVMs or instruments without having full trust in the inner workings of the device. Crucially, the largest achievable score using some device cannot be increased using the allowed operations of the resource theory, and the scores therefore represent resource monotonies. In many cases, these games define a complete set of monotonies whose values provide necessary and sufficient conditions for convertibility of one object to another by the allowed operations [20, 34, 38–42]. We show in Sec. V that the same holds true for the guessing games considered in this paper, and furthermore, the advantage of using an incompatible device in these games can be quantitatively characterized by the aforementioned robustness measure. However, the general idea of relating convertibility to guessing games can be traced back to the original work of Blackwell on statistical comparisons [43] (see Ref. [38] for more discussions).

A. From programmability to nonsignaling

Our analysis of quantum incompatibility is motivated by the idea of “programmable” quantum instruments. Consider a generic controllable measurement device as depicted in Fig. 1, which is capable of implementing some family of instruments \( \{ \Lambda_{x_i|x_0} \}_{x_0,x_i} \). The classical program is the input value \( x_0 \), which dictates that the instrument \( \{ \Lambda_{x_i|x_0} \}_{x_i} \) be performed on the quantum input. We consider these devices to be modules in nature so that the classical or quantum output from one device can be connected to a classical or quantum input of another. This introduces a critical consideration of time: for the devices to function together properly, the outputs of one device must arrive at a time when the next device is ready to receive them. In practice, every physical device will have a characteristic quantum delay time [44], which measures how fast the device generates a quantum output when given a quantum input, and it corresponds to \( \Delta_D := t_1 - t_0 \) in Fig. 1. How about the timing of the classical program? As typically demanded by devices with multiple inputs, one would expect that the program be synchronized with the quantum input, or at least be within the finite window \( [t_0, t_1] \). However, if the quantum delay time \( \Delta_D \) appears to be short, such a hard constraint on timing can be unrealistic in practice, and therefore we have a particular interest in devices that give the experimenter full temporal freedom over when he or she can submit the program, a capability called programmability in Ref. [20]. As a consequence of programmability, the timing of the classical and quantum inputs need not be synchronized, and the classical program can arrive significantly before \( t_0 \) or after \( t_1 \). Without loss of generality, we need only to assume that the program arrives after the quantum output time \( t_1 \), which we refer to as the late-program assumption, and this is because an early arriving program can always be buffered in a classical memory before the quantum input arrives. Clearly, not every controllable quantum device can work through the late-program assumption, and a device can do so if and only if the quantum output is independent of the classical input after coarse-graining the classical output (see Sec. II B). Formally, this constraint is known as nonsignaling, which requires that

\[
\sum_{x_i} \Lambda_{x_i|x_0} = \sum_{x_i} \Lambda_{x_i|x_0'} =: \Lambda \quad \forall x_0, x_0'.
\]  

In other words, all the instruments in the family \( \{ \Lambda_{x_i|x_0} \}_{x_0,x_i} \) sum up to the same channel \( \Lambda \). Since the nonsignaling constraint in Eq. (1) is necessary and sufficient for programmability, we naturally refer to devices satisfying this constraint as programmable instrument devices (PIDs).

\[
\sum_{x_i} \Lambda_{x_i|x_0} = \sum_{x_i} \Lambda_{x_i|x_0'} =: \Lambda \quad \forall x_0, x_0'.
\]
While the quantum output at $t_1$ for a PID is independent of the classical program, the classical output will generally depend on the quantum input at $t_0$ (and certainly also on the classical input at $t_2$). Hence under the late-program assumption, the internal quantum memory of the PID might need to store quantum information for an indefinite amount of time until the experimenter chooses to issue a program. However, there is a special class of PIDs for which the quantum memory can be perfectly substituted with a classical memory. These are called simple PIDs, and they represent the “free” objects in the resource theory. Remarkably, simple PIDs are precisely those in which the family of instruments being implemented is compatible. In contrast, nonsimple PIDs, i.e., PIDs that are not simple, require quantum memories with an indefinite lifetime to support full programmability, and thus they are resources, demonstrating incompatibility. Of course, an indefinite lifetime of a quantum memory is an idealization and hence so is full programmability. In practice, every realizable PID will have a quantum memory with some finite lifetime $\Delta t_{QM} < \infty$. To pinpoint the differing demands on quantum memories for programmability, throughout this paper, we may as well assume that every PID satisfies $\Delta t_D \approx 0$ and that the internal quantum memory of any nonsimple PID satisfies $\Delta t_{QM} \gg \Delta t_D$. These assumptions are in line with our identification of only simple PIDs as being free.

There is an alternative justification for imposing the nonsignaling constraint in Eq. (1) not directly related to programmability. One could imagine that the device in Fig. 1 is a bipartite channel shared between spatially separated parties, with Evan controlling the classical input and output and Alice controlling the quantum input and output, as depicted in Fig. 2. Alice may be unaware of the existence of the eavesdropping party Evan, and she thinks of her quantum input and output as being connected by a local channel. She would then expect that the quantum delay time of her channel should be extremely short, limited only by the local inner workings of her device, which implies that the information from Evan’s control signal should not have enough time to propagate across spacetime and influence her channel output. Thus Alice’s output would be spacelike separated from the choice of Evan’s control signal, and the nonsignaling condition (from Evan to Alice) would hold, as in Eq. (1). Note that this reflects a scenario known as channel steering [45], as Evan is remotely manipulating Alice’s channel with his control signal without letting her detect him due to the nonsignaling constraint. Apart from being useful for better understanding quantum incompatibility, the scenario of channel steering is also relevant for cryptographic applications and one-sided device-independent testing of coherent channel extensions [45]. In fact, steerability and incompatibility are equivalent concepts when defined on nonsignaling families of instruments (see Sec. IV A), and so the resource theory of PID nonsimplicity that we develop in this paper is equivalently a resource theory of channel steering.

**B. Organization of the paper**

With this background and motivation in hand, we now conduct our resource-theoretic analysis of instrument incompatibility in terms of nonsimplicity of PIDs in more detail. The rest of the paper is organized as follows.

In Secs. II and III, we establish the basic pieces of our resource theory. We first review the traditional concept of measurement incompatibility and its resource theory in Secs. II A and III A. Then we extend the present theory to incorporate the more general concept of incompatibility between quantum instruments. Specifically, we formally introduce simple (i.e., free) versus nonsimple (i.e., resourceful) PIDs and the physical distinction between them in Sec. II B. We then propose a class of free transformations between PIDs, which is incompatibility-nonincreasing and has a clear operational meaning, in Sec. III B.

In Sec. IV, we focus on discussing and discovering the relationships between PID nonsimplicity, steering, and traditional measurement incompatibility. We explicate the spatiotemporal correspondence between channel steering and PIDs.

---

**Fig. 1.** A general controllable quantum device applies an instrument $\{\Lambda_{x_i|x_0}\}_{x_1}$ to the quantum input whenever a particular program $x_0$ is chosen. The characteristic time $\Delta t_D := t_1 - t_0$ is known as the quantum delay time of the device, and it measures how quickly the device functions as a q-to-q channel. The device is said to be fully programmable if the program is free to arrive at any time $t_2$, even outside the interval $[t_0, t_1]$, and we refer to such a device as a programmable instrument device. Given that classical memories are freely available, the full programmability of a PID is essentially its ability to withstand an arbitrarily late arrival of the program, termed the late-program assumption. In the framework of Ref. [44], a PID represents a so-called “multiprocess.”

**Fig. 2.** The control device could also be split between spatially separated parties Alice and Evan. The nonsignaling constraint from Evan’s control signal to Alice’s channel output naturally arises if the spatial separation is so large that Evan’s signal propagation time exceeds the anticipated quantum delay time between Alice’s local input and output.
in Sec. IV A. Then we generalize the concept of steering-equivalent observables to the scenario of channel steering in Sec. IV B, and in doing so we derive a monotonicity theorem that signifies a fundamental connection between the resource theory of PID nonsimplicity and that of measurement incompatibility.

Finally, in Sec. V, we provide a semi-device-independent characterization for PID nonsimplicity by designing a class of so-called “nontransient” guessing games as a temporal analog to the well-studied nonlocal games. In particular, nontransient guessing games can be used to characterize the incompatibility preorder between PIDs by providing necessary and sufficient conditions for convertibility from one PID to another under the aforementioned free transformations, as shown in Sec. V A. We further show in Sec. V B that the operational advantage of a given PID over simple PIDs in these games is tightly bounded from above by the PID’s robustness of incompatibility against noise. In Sec. V C, we discuss the experimental setup of nontransient guessing games and put forward a variant class of games that lowers the experimental requirement while also faithfully characterizing the incompatibility preorder between PIDs.

II. PROGRAMMABLE QUANTUM DEVICES

In this section, we first review the traditional concept of measurement incompatibility formulated as a resource for programmable measurement devices [20]. Then we extend the traditional framework by incorporating programmable instrument devices and a generalized concept of incompatibility for such devices.

A. Programmable measurement devices

A programmable measurement device (PMD) [20], alias a multimeter [46], is a quantum measurement device capable of implementing a family of POVMs conditioned on a classical control signal. A PMD $\Lambda$ is mathematically represented by a collection of positive semidefinite operators $\Lambda = \{ M_{x_1|x_0} \}_{x_0,x_1}$ such that $\sum_{x_1} M_{x_1|x_0} = 1$ for all $x_0$. The classical input $x_0$ is known as the program, which indicates the particular POVM $\{ M_{x_1|x_0} \}_{x_1}$ to be performed on the quantum input. The classical output $x_1$ labels the measurement outcome of the POVM. PMDs represent the most general qc-to-c CPTP maps.

A PMD $\Lambda$ is said to be simple whenever it can be simulated with a “mother” POVM followed by some (controllable) classical postprocessing; namely, it implements a compatible family of POVMs admitting the following decomposition:

$$M_{x_1|x_0} = \sum_g p_{x_1|x_0,g} G_g \quad \forall x_0, x_1, \quad (2)$$

where $\{ G_g \}_g$ is a POVM and $\{ p_{x_1|x_0,g} \}_{x_0,x_1,g}$ is a classical channel (i.e., a conditional probability distribution). A PMD not decomposable in the form of Eq. (2) is nonsimple, as the family of POVM it implements is incompatible. An advantage of studying quantum measurements in terms of PMDs is that it links measurement incompatibility and quantum memories in a physically motivated way. As discussed in Sec. I A, a practical conception of programmability should take into account the unavoidable asynchronicity between the quantum input and the classical program, and so programmable devices should (ideally) allow the experimenter to issue the program at any desirable time [20]. This temporal freedom can be simply captured by the late-program assumption, which we regard as a basic principle in the programmability context. For a PMD to function as a qc-to-c box under this assumption, an internal quantum memory is generally needed to store the quantum input until the program is submitted to the classical system $X_0$. However, if the PMD is controlling a compatible family of measurements, i.e., the PMD being simple, then no quantum memory is needed. Instead, the “mother” POVM $\{ G_g \}_g$ can be performed as soon as the quantum input arrives, and the outcome $g$ is stored in a classical memory until the program arrives. Thus, the requirement of a quantum memory to implement a PMD is another way of characterizing measurement incompatibility. It is then natural to identify simple PMDs as being “free” objects since they do not require quantum memories to implement, and a resource theory of PID nonsimplicity is physically well justified.

B. Programmable instrument devices

Next we extend the theory of programmability to quantum instruments. A quantum instrument is a generalized version of measurement that incorporates a quantum output representing the postmeasurement state. The most straightforward generalization of a PMD is a multi-instrument [46], a device capable of implementing a collection of quantum instruments conditioned on a classical control signal. A multi-instrument $\Lambda$ is mathematically represented by a collection of completely positive (CP) maps $\Lambda \equiv \{ \Lambda_{x_1|x_0} \}_{x_0,x_1}$ such that $\sum_{x_1} \Lambda_{x_1|x_0} 1 = \text{trace preserving (TP)}$ for all $x_0$. To make the quantum input and output systems $A_0$ and $A_1$ explicit, we sometimes write $\Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \}_{x_0,x_1}$.

Multinstruments represent the most general qc-to-qc CPTP maps. However, according to our conception of programmability, general multi-instruments are not suitable models for abstracting “programmable” quantum instruments. As discussed in Sec. I A generally and in Sec. II A for PMDs, practical programmable instruments should withstand the late-program assumption; they should function while the classical program is free to arrive at $X_0$ anytime after the quantum input arrives at $A_0$. By the same assumption, the program could even arrive after the device is scheduled to dispense some quantum output at $A_1$. For a device to be physically realizable in such a circumstance, there must be no signaling from the classical input $X_0$ to the quantum output $A_1$. This motivates the following definition.

Definition 1. A multi-instrument $\Lambda \equiv \{ \Lambda_{x_1|x_0} \}_{x_0,x_1}$ is called a programmable instrument device whenever it is nonsignaling from the classical input to the quantum output; namely, there
exists a quantum channel $\Lambda$ such that
\[ \sum_{x_1} \Lambda x_1 | x_0 = \Lambda \quad \forall x_0. \] (3)

We say that the PID $\Lambda$ is simple (and otherwise nonsimple) whenever there exists a quantum instrument $G \equiv \{G_g\}_g$ and a classical channel $p \equiv \{p_{x_1|x_0,g}\}_{x_0,x_1,g}$ such that
\[ \Lambda x_1 | x_0 = \sum_g p_{x_1|x_0,g} G_g \quad \forall x_0, x_1. \] (4)

The above definition of a programmable instrument device generalizes the definition of a PMD in a way that respects the late-program assumption. Likewise, the concept of incompatibility in terms of device nonsimplicity is extended from POVMs to instruments. Following an analogous argument that previously applies to PMDs, one can find that the difference between a nonsimple PID and a simple PID is precisely captured by whether the device needs a quantum memory with a non-negligible lifetime $\Delta_{QM} \gg \Delta_D \approx 0$ to implement. Accordingly, it is natural to identify nonsimple PIDs as resources in our theory of programmable instruments, whereas simple PIDs are free objects.

While our formulation of PIDs is motivated by the notion of programmability, the bipartite picture shown in Fig. 2 can be helpful in understanding the internal structure of such devices. We envision that Alice has the quantum input and output in her laboratory while Evan controls the classical input and output. The nonsignaling condition (from Evan to Alice) in the definition of a PID is also known as “semicausality” [47]. It has been proved that every semicausal map is “semilocalizable” [48], meaning that the map can be decomposed into local operations by Alice and Evan individually combined with one-way quantum communication from Alice to Evan, as shown in Fig. 3a.

Simple PIDs are then precisely those in which the one-way quantum communication can be replaced with one-way classical communication, as shown in Fig. 3c. In Secs. IV A and IV B, we provide a more formal statement (Proposition 1) and related discussions regarding the internal structure of a PID.

III. FREE SIMULATIONS OF PROGRAMMABLE DEVICES

In this section, we complete the construction of the resource theory of PID nonsimplicity by proposing a set of free transformations between PIDs. Hereafter, we will refer to free transformations applied to programmable devices (either PIDs or PMDs) as free simulations of the devices.

A. Free simulations of PMDs

Before we introduce what constitutes a free simulation of PIDs, we first recall the free simulations of PMDs in the resource theory of PMD nonsimplicity [20]. Later as we define the free simulations of PIDs, we must ensure that they reduce to the predefined simulations of PMDs when both the source PID and the target PID are PMDs.

**Definition 2** ([20]). Let $\mathcal{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ and $\mathcal{N} \equiv \{N_{y_1|y_0}\}_{y_0,y_1}$ be two PMDs. We say that $\mathcal{M}$ can freely simulate $\mathcal{N}$, denoted by $\mathcal{M} \gg_{SM} \mathcal{N}$, whenever there exists a quantum instrument $\mathcal{K} \equiv \{K_k\}_k$ and two classical channels $p \equiv \{p_{x_0,l|y_0,k}\}_{x_0,y_0,k,l}$ and $q \equiv \{q_{y_1|x_1,l}\}_{x_1,y_1,l}$ such that
\[ N_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} K_k \left[ M_{x_1|x_0} \right] \quad \forall y_0, y_1. \] (5)
The operational significance of Definition 2 is demonstrated by the fact that $M \succsim_{\mathcal{M}} N$ if and only if $M$ can be physically transformed (i.e., via a quantum superchannel [49, 50]) into $N$ using no additional quantum memory [20], as represented in Fig. 4. Meanwhile, free simulations of PDMs have been shown to possess essential resource-theoretic properties including preserving PDM simplicity and being able to generate the entire set of simple PDMs [20]. Thus they are formally qualified as the free transformations in a resource theory of PDM nonsimplicity, and the relation $\succsim_{\mathcal{M}}$ is a legitimate incompatibility preorder on the set of PDMs.

### B. Free simulations of PIDs

Now we are ready to propose the free simulations for the resource theory of PID nonsimplicity. In what follows, we first identify the complete class of PID transformations that do not require quantum memories to implement, and then we demonstrate its legitimacy as the set of free transformations from a resource-theoretic standpoint. Since PIDs are semicausal quantum channels (i.e., quantum 2-combs), transformations between them are supposed to be quantum 4-combs [51].

Programmability of PIDs highlights the temporal separation between its quantum systems and classical systems, as displayed in Fig. 3. So quantum memories across this separation are the only resource that should be forbidden when simulating PIDs. As a result, the experimenter should have the full ability to (i) have any physical process concatenated in sequence or appended in parallel to the quantum part or the classical part “locally,” and (ii) feed any side information generated by the quantum part into the classical part, as long as this information is stored in a classical memory before the classical program arrives, as represented in Fig. 5. Note that the quantum delay time $\Delta_{QM}$ between the quantum systems $A_0$ and $A_1$ is assumed negligible and not regarded as a resource compared to the internal quantum memory lifetime $\Delta_{QM}$. Hence, the side channel parallel to the quantum part need not be classical. The formal definition of a free simulation of PIDs is given as follows.

where $(\cdot)^\dagger$ denotes adjunction. We call the transformation $\Pi \mapsto N$ a free simulation of PDMs.

### Definition 3

Let $\Lambda \equiv \{K_{x_1|x_0} \}_{x_0,x_1}$ and $\Gamma \equiv \{\gamma_{y_1|y_0} \}_{y_0,y_1}$ be two PIDs. We say that $\Lambda$ can freely simulate $\Gamma$, denoted by $\Lambda \succsim \Gamma$, whenever there exists a quantum channel $\mathcal{F}$, a quantum instrument $\mathcal{K} \equiv \{K_k \}_{k}$, and two classical channels $p \equiv \{p_{x_0,l}|y_0,k \}_{x_0,y_0,k,l}$ and $q \equiv \{\gamma_{y_1|x_1,l} \}_{x_1,y_1,l}$ such that

$$\gamma_{y_1|y_0} = \sum_{x_0,x_1,l} q_{y_1|x_1,l} P_{x_0,l|y_0,k} K_k \circ (A_{x_1|x_0} \otimes \text{id}) \circ \mathcal{F} \quad \forall y_0,y_1.$$ (6)

We call the transformation $\Lambda \mapsto \Gamma$ a free simulation of PIDs.

As can be recognized from Eq. (6) or Fig. 5, free simulations of PIDs preserve the classical-to-quantum nonsignaling constraint, and thus they always map PIDs to PIDs. One can conveniently verify that Definition 3 reduces to Definition 2 (free simulations of PDMs) when the quantum output systems of both $\Lambda$ and $\Gamma$ are trivial (i.e., 1 dimensional).

The following theorem demonstrates the legitimacy of identifying the free simulations of PIDs as the free transformations for PID nonsimplicity. This implies that the relation $\succsim$ is an incomparability preorder on the set of PIDs.

**Theorem 1.** For fixed index sets where $x_0, x_1, y_0, y_1$ belong, the free simulations of PIDs have the following properties.

1. **Simplicity**: a simple PID cannot freely simulate any nonsimple PIDs.
2. **Reachability**: any PID can freely simulate any simple PID.
3. **Composability**: the sequential or parallel composition of free simulations is a free simulation.
4. **Closedness**: the limit of a sequence of free simulations (if exists) is a free simulation.
5. **Convexity**: the probabilistic mixture of free simulations is a free simulation.

The proof of Theorem 1 is in Appendix A.

Theorem 1(1) and (2) and the sequential composability in Theorem 1(3) guarantee that the free simulations meet the minimal requirements for qualifying as the free transformations
for PID nonsimplicity [24]. The parallel composability in Theorem 1(3) implies that the resulting resource theory admits a tensor-product structure [24]. Crucially, Theorem 1(4) and (5) indicate that the resource theory of PID nonsimplicity has the nice mathematical property of being closed and convex. Operationally, convexity means that the definition of free simulations of PIDs has implicitly included the use of shared randomness among the constituent physical units of a free simulation.

Before closing this section, we remark that our free simulations of PIDs constitute the complete set of physical transformations that do not exploit quantum memories whose lifetime exceeds $\Delta_D \approx 0$. This fact can be demonstrated by invoking the theory of quantum networks [51], combined with the observation that all input and output systems of $\Lambda$ and $\Gamma$ must be put in the present causal order in Fig. 5 (see Ref. [51, Theorem 8 and Fig. 11]). However, we leave it as an open question whether the set of free simulations considered here is the maximal set of transformations that do not generate PID nonsimplicity, or conversely, whether there exists a completely simplicity-preserving comb [24] that requires a quantum side memory with a non-negligible lifetime to implement.

IV. RELATIONSHIP WITH STEERING AND MEASUREMENT INCOMPATIBILITY

In this section, we expand on the relationship between PID nonsimplicity, steering, and PMD nonsimplicity. We first clarify that each PID can be implemented through a process of channel steering and vice versa, and so a resource theory of the former implies that of the latter. Persisting with the steering viewpoint, we then unfold some underlying connections between nonsimplicity of PIDs and of PMDs, or equivalently, between channel steering and measurement incompatibility, and ultimately we demonstrate that PMDs themselves can be cast as a measure of nonsimplicity for PIDs.

A. PIDs as assemblages in channel steering

As a dynamical generalization of the celebrated phenomenon of EPR steering [53–55], channel steering [45] provides a natural and effective way of understanding the internal structure of PIDs. Referencing the bipartite picture in Fig. 2, the scenario of channel steering is described as follows. Consider a broadcast channel $E_{A_0 \rightarrow A_1E}$ with the systems $A_0$ and $A_1$ held by Alice and the system $E$ leaked to Evan. Without any proactive interference, Evan can remotely “steer” the subchannel decomposition of Alice’s marginal channel $\Lambda_{A_0 \rightarrow A_1} := Tr_E \circ E_{A_0 \rightarrow A_1E}$ by directing his system $E$ to a PMD $\Pi \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$. This steering process would lead to a family of instruments $\Lambda \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ on Alice’s side, typically known as a channel assemblage [45], defined by

$$\Lambda_{A_0 \rightarrow A_1} [\cdot] := Tr_E \left( \left[ \pi^{A_1} \otimes M_{x_1|x_0}^{E} \right] E_{A_0 \rightarrow A_1E}[\cdot] \right) \quad \forall \chi_{0,1}. \quad (7)$$

The assemblage $\Lambda$ is said to be unsteerable whenever it can be realized following Eq. (7) with $E_{A_0 \rightarrow A_1E}$ being an incoherent extension of its marginal channel $\Lambda_{A_0 \rightarrow A_1}^{A_0 \rightarrow A_1}$ [45], namely, whenever there exists a quantum instrument $\mathcal{G} \equiv \{G_g\}_g$ and states $\epsilon_g$ for all $g$ such that

$$E_{A_0 \rightarrow A_1E} = \sum_g G_g^{A_0 \rightarrow A_1} \otimes \epsilon_g^E. \quad (8)$$

It follows as an observation that if the PMD $\Pi$ is simple, then the channel assemblage $\Lambda$ induced by Eq. (7) must be unsteerable regardless of $E_{A_0 \rightarrow A_1E}$ [45]. We introduce the following shorthand to denote a channel assemblage generated via steering.

Definition 4. Let $\Lambda \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ be a channel assemblage, $E_{A_0 \rightarrow A_1E}$ a broadcast channel, and $\Pi \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ a PMD. We say that $(E, \Pi)$ is a steering decomposition of $\Lambda$, denoted by $\Lambda \leftarrow (E, \Pi)$, whenever Eq. (7) is satisfied, i.e., whenever $\Lambda$ can be induced by a process of channel steering featuring $E$ and $\Pi$.

Clearly, each channel assemblage $\Lambda \leftarrow (E, \Pi)$ can be regarded as a PID, since by Eq. (7) the nonsignaling constraint in Eq. (3) is satisfied. Conversely, as we argued in Sec. II B, due to the equivalence between semicausality and semi-localizability [48], each PID $\Lambda$ admits a steering decomposition $\Lambda \leftarrow (E, \Pi)$, and therefore we can envision a process of channel steering going on within each PID, as shown in Fig. 3a. Then we can observe from Fig. 3c that simple PIDs are precisely those that implement unsteerable channel assemblages. Likewise, the free simulations of PIDs in Definition 3 correspond to transformations of channel assemblages that are realizable via Alice-to-Evan one-way local operations and classical communication (one-way LOCC). In this way, the resource theory of PID nonsimplicity that we developed in Secs. II B and III B can be equivalently interpreted as a resource theory of channel steering.

The scenario of channel steering reduces to that of EPR steering when the quantum input system $A_0$ is trivial [54, 55]. In this case, when Evan feeds his part of a bipartite state $\xi^{A_1E}$ into a PMD $\Pi$, a state assemblage $\{\rho_{x_1|x_0}\}_{x_0,x_1}$ is generated on Alice’s side:

$$\rho_{x_1|x_0}^{A_1} := Tr_E \left( \left[ \pi^{A_1} \otimes M_{x_1|x_0}^{E} \right] \xi^{A_1E} \right) \quad \forall \chi_{0,1}. \quad (9)$$

Accordingly, the assemblage $\{\rho_{x_1|x_0}\}_{x_0,x_1}$ is unsteerable when it can be induced by a bipartite state $\xi^{A_1E}$ being separable:

$$\xi^{A_1E} = \sum_g \eta_g^{A_1} \otimes \epsilon_g^E, \quad (10)$$

where $\{\eta_g\}_g$ is a state ensemble and $\epsilon_g$ is a state for all $g$. In parallel to Eq. (4), an unsteerable state assemblage $\{\rho_{x_1|x_0}\}_{x_0,x_1}$ demonstrates compatibility by admitting a so-called “local-hidden-state” model [54]:

$$\rho_{x_1|x_0}^{A_1} = \sum_g p_{x_1|x_0,g} \eta_g^{A_1} \quad \forall \chi_{0,1}, \quad (11)$$

where $\{p_{x_1|x_0,g}\}_{x_0,x_1,g}$ is a classical channel. As in channel steering, if the PMD $\Pi$ is simple, then the induced state assemblage must be unsteerable. Conversely, it has also been proved
Resource theories of nonsimple programmability

| Programmable devices (nonsignaling assemblages) | Simple devices (compatible assemblages) |
|------------------------------------------------|----------------------------------------|
| \( \{ p_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} p_{x_1|x_0} = \rho \ \forall x_0 \) | \( \{ p_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} p_{x_1|x_0} = \sum_{x_0,x_0'} \eta_{x_0} \ \forall x_0,x_1 \) |
| \( \{ M_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} M_{x_1|x_0} = \mathbb{I} \ \forall x_0 \) | \( \{ M_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} M_{x_1|x_0} = \sum_{x_0,x_0'} G_{x_0} \ \forall x_0,x_1 \) |
| \( \{ A_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} A_{x_1|x_0} = \Lambda \ \forall x_0 \) | \( \{ A_{x_1|x_0} \}_{x_0,x_1} : \sum_{x_1} A_{x_1|x_0} = \sum_{x_0,x_0'} G_{x_0} \ \forall x_0,x_1 \) |

**TABLE I.** A comparison between programmable devices (i.e., general objects) and simple devices (i.e., free objects) in the resource theories of nonsimplicity of programmable source, measurement, and instrument devices (PSDs, PMDs, and PIDs, respectively). Our resource theory of PID nonsimplicity, equivalently a resource theory of channel steering, is a generalized theory unifying both the resource theory of PSD nonsimplicity (i.e., EPR steering [52]) and that of PMD nonsimplicity [20] (i.e., measurement incompatibility [13]). Specifically, PID nonsimplicity reduces to PSD nonsimplicity when the quantum input is trivial, and it reduces to PMD nonsimplicity when the quantum output is trivial. Programmable devices in these theories are all subject to the nonsignaling constraint, and simple devices all implement compatible assemblages. The compatibility in these free objects can universally be viewed as a consequence of the classicality of the internal memory (hollow red arrows), which should in general be a quantum memory (solid red arrows).

that all nonsimple PMDs are useful to generate *steerable* state assemblages when \( \mathcal{E}^{A_1} \) is a pure state of maximum Schmidt rank [56, 57].

As a special case of channel steering, EPR steering can also be understood as a theory of nonsimplicity of programmable source devices (PSDs) (alias nonsignaling multisources [46]) in the programmability framework, and simple PSDs are those implementing unsteerable state assemblages. As a result, our resource theory of PID nonsimplicity subsumes the existing resource theory of EPR steering with one-way LOCC as free operations [52]. A comparison between resource theories of nonsimplicity of different types of programmable devices is provided in Table I.

### B. Steering-equivalence mapping

The concept of “steering-equivalent observables” (SEO) plays an essential role in studying the relationship between EPR steering and measurement incompatibility [58]. Now we generalize this concept from state assemblages to PIDs (i.e., to channel assemblages), so as to connect PID nonsimplicity (i.e., channel steering) with PMD nonsimplicity (i.e., measurement incompatibility). Given a PID \( \Lambda \equiv \{ \Lambda_{x_1|x_0} \}_{x_0,x_1} \), the nonsignaling constraint guarantees the existence of the following channel:

\[
\Lambda_{A_0 \rightarrow A_1} := \sum_{x_1} \Lambda_{A_0 \rightarrow A_1}^{x_1|x_0} \ \forall x_0.
\]  

(12)

The Choi operator of \( \Lambda_{A_0 \rightarrow A_1} \) is given by

\[
J_{\Lambda_{A_0 \rightarrow A_1}}^{A_0,A_1} := \left( \text{id}_{A_0} \otimes \Lambda_{A_0 \rightarrow A_1} \right) \left( \phi_{A_0}^{A_0,A_0} \right) \ \forall x_0,x_1.
\]  

(13)

where \( \phi_{A_0}^{A_0,A_0} \) is a system isomorphic to \( A_0 \) and \( \phi_{A_0}^{A_0,A_0} \equiv \sum_{i,j} |i \rangle \langle j |^{A_0} \otimes |i \rangle \langle j |^{A_0} \). The Choi operator of \( \Lambda_{A_0 \rightarrow A_1} \) is given by

\[
J_{\Lambda}^{A_0,A_1} := \sum_{x_1} J_{\Lambda_{A_0|A_1}}^{A_0,A_1} \ \forall x_0.
\]  

(14)

Let \( A^* \) be a quantum system such that its associated Hilbert space, denoted by \( \mathcal{H}^{A^*} \), is isomorphic to the support of \( \Lambda_{A_0|A_1} \). In other words, it satisfies \( \mathcal{H}^{A^*} \equiv \text{supp}(J_{\Lambda}^{A_0,A_1}) \subseteq \mathcal{H}^{A_0,A_1} \), where \( \mathcal{H}^{A_0,A_1} \equiv \mathcal{H}^{A_0} \otimes \mathcal{H}^{A_1} \) is the composite Hilbert space associated with the systems \( A_0 \) and \( A_1 \). By Choi’s theorem, \( J_{\Lambda_{A_0|A_1}}^{A_0,A_1} \) is positive semidefinite since \( \Lambda_{A_0 \rightarrow A_1} \) is CP for all \( x_0, x_1 \) .

By Eq. (14), this implies \( \text{supp}(J_{\Lambda_{A_0|A_1}}^{A_0,A_1}) \subseteq \text{supp}(J_{\Lambda}^{A_0,A_1}) \equiv \mathcal{H}^{A^*} \), and therefore the image of \( J_{\Lambda_{A_0|A_1}}^{A_0,A_1} \) in the system \( A^* \), denoted by \( J_{\Lambda_{A_0|A_1}}^{A_0,A_1} \), is well defined for all \( x_0, x_1 \). The following definition generalizes the steering-equivalent observables (SEO) defined on state assemblages to PIDs.

**Definition 5.** The steering-equivalence mapping, denoted by \( \text{SEM} \), is a mapping from the set of PIDs to the set of PMDs, and it sends a PID \( \Lambda \equiv \{ \Lambda_{A_0 \rightarrow A_1}^{x_1|x_0} \}_{x_0,x_1} \) to a PMD \( \Lambda \equiv \{ \Lambda_{A_0 \rightarrow A_1}^{x_1|x_0} \}_{x_0,x_1} \) where

\[
S_{1|x_0}^{A^*} := (J_{\Lambda}^{A_0,A_1})^{-\frac{1}{2}} J_{\Lambda_{A_0|A_1}}^{A_0,A_1} (J_{\Lambda}^{A_0,A_1})^{-\frac{1}{2}} \ \forall x_0,x_1.
\]  

(15)

It can be conveniently verified that \( \text{SEM}(\Lambda) \) is a valid PMD given any PID \( \Lambda \). It is also apparent from Definition 5 that, by the Choi–Jamiołkowski isomorphism, a PID \( \Lambda \) is nonsimple if and only if the PMD \( \text{SEM}(\Lambda) \) is nonsimple. That is to say, the membership problem of steerable channel assemblages can be reduced to the membership problem of incompatible...
families of POVMs through SEM. This generalizes Theorem 1 of Ref. [58], which addresses the membership problem of steerable state assemblages via the SEO. The SEO of a state assemblage has been shown to possess an operational interpretation as the transposed PMD that induces the state assemblage from a minimal state extension [58]. We show in the following proposition that this type of operational interpretation remains effective for SEM in the generalized scenario of channel steering.

**Proposition 1.** Let $\Lambda \equiv \{\Lambda_{x_0|\Lambda}^{A_0\rightarrow A_1}\}_{x_0, x_1}$ be a PID, and let $S \equiv \{S_{x_1|x_0}^{A_1\rightarrow A_0}\}_{x_0, x_1}$ be a PMD such that $S = \text{SEM}(\Lambda)$. Then there exists an isometric channel $\mathcal{V}_{A_0\rightarrow A_1A'}$ such that

$$\Lambda_{x_0|\Lambda}^{A_0\rightarrow A_1} [\cdot] = \text{Tr}_{A'} \left( \left[ \Lambda^{A_1} \otimes \left( S_{x_1|x_0}^{A_1} \right)^\dagger \right] \mathcal{V}_{A_0\rightarrow A_1A'} [\cdot] \right) \quad \forall x_0, x_1,$$

(16)

where $(-)^\dagger$ denotes transposition under a fixed orthonormal basis.

The proof of Proposition 1 is in Appendix B 1.

Proposition 1 indicates that given any PID $\Lambda$, there exists a broadcast channel $\mathcal{V}_{A_0\rightarrow A_1A'}$ such that the steering decomposition $\Lambda \leftarrow (\mathcal{V}, \text{SEM}(\Lambda))$ holds, where $\mathcal{V}_{A_0\rightarrow A_1A'}$ is an isometric dilation of $\Lambda_{x_0|\Lambda}^{A_0\rightarrow A_1}$ and $\text{SEM}(\Lambda)$ is the elementwise transpose of the PMD $\text{SEM}(\Lambda)$. In particular, this provides a formal and independent demonstration for the internal structure of a general PID as depicted in Fig. 3a, which we previously argued by invoking the equivalence between semicausality and semi-localizability.

On the other hand, the existence of a steering decomposition for any PID does not imply that such a decomposition is unique, and in fact it is not unique. Despite this, there is a close relation between any steering decomposition of a given PID and the “canonical” steering decomposition specified in Proposition 1, as given by the following proposition.

**Proposition 2.** Let $\Lambda$ be a PID, $\mathcal{E}_{A_0\rightarrow A_1E}$ a broadcast channel, and $\mathcal{M} \equiv \{M_{x_1|x_0}^{A_1\rightarrow A_0}\}_{x_0, x_1}$ a PMD such that $\Lambda \leftarrow (\mathcal{E}, \mathcal{M})$. Then we have $\mathcal{M}^\dagger \succeq \text{SEM}(\Lambda)$, where $\mathcal{M}^\dagger \equiv \{M_{x_1|x_0}^{A_1\rightarrow A_0}\}_{x_0, x_1}$.

The proof of Proposition 2 is in Appendix B 2.

Proposition 2 indicates that $\text{SEM}(\Lambda)^\dagger$ is the least resourceful PMD that can be used to compose a given PID $\Lambda$, because every other PMD $\mathcal{M}$ that suffices to do so must be convertible to $\text{SEM}(\Lambda)^\dagger$ via free simulations. This is the reason why we refer to the decomposition $\Lambda \leftarrow (\mathcal{E}, \mathcal{M})$ in Proposition 1 as the “canonical” or “minimal” decomposition for $\Lambda$. To understand this physically, we can think of any PID decomposition $\Lambda \leftarrow (\mathcal{E}, \mathcal{M})$ as having a configuration as depicted in Fig. 3b, and the nonsimplicity of $\Lambda$ is essentially attributed to the PMD $\mathcal{M}$, since that is where the quantum memory with a lifetime $\Delta_{QM} \gg \Delta_{TD} \approx 0$ resides. Then Proposition 2 implies that when the decomposition $\Lambda \leftarrow (\mathcal{E}, \mathcal{M})$ is canonical, namely, when $\mathcal{E}_{A_0\rightarrow A_1E} = \mathcal{V}_{A_0\rightarrow A_1A'}$ is the minimal isometric dilation (i.e., the maximally coherent channel extension [45]), the physical resource within $\mathcal{M}$ is best utilized.

Built on the aforementioned results, we conclude this section with the following theorem. We show that the mapping SEM behaves as a nonsimplicity monotone of PIDs (equivalently, a steering monotone of channel assemblages), in the sense that it preserves the incompatibility preorder specified by the free simulations of programmable devices. This means that the nonsimplicity of $\text{SEM}(\Lambda)$ is not only an indicator, but also a measure of the nonsimplicity of $\Lambda$.

**Theorem 2.** The mapping SEM is a faithful nonsimplicity monotone. Formally, it has the following properties.

1. Faithfulness: $\Lambda$ is a simple PID if and only if $\text{SEM}(\Lambda)$ is a simple PMD.
2. Monotonicity: if $\Lambda \succ \Gamma$, then $\text{SEM}(\Lambda) \succeq \text{SEM}(\Gamma)$.

The proof of Theorem 2 is detailed in Appendices B 3 and B 4. It is a proof by construction based on Propositions 1 and 2.

Conventionally, by “resource monotones” we allude to real-valued functions that are nonincreasing under free transformations [24]. In Theorem 2, however, the term has a broader meaning of order-preserving mappings under free transformations, and such mappings can in general be between objects, as SEM is. A direct application of such generalized monotones is to convert resource monotones in one resource theory to resource monotones in another resource theory. For instance, given any resource monotone $f$ for PMD nonsimplicity (i.e., measurement incompatibility), we immediately obtain an induced resource monotone $f \circ \text{SEM}$ for PID nonsimplicity (i.e., channel steering), regardless of what $f$ is and whether $f$ is real valued or not.

On the other hand, we remark that object-valued monotones like SEM are also interesting in their own rights, since they reveal fundamental connections between two different resource theories and may trigger insights into the physical nature of the resources involved. As for SEM, with its operational interpretation established in Proposition 1, Theorem 2 indicates that a more resourceful PID must have a more resourceful internal PMD under the “canonical” steering decomposition. This certainly supports our previous viewpoint of attributing the physical resource (i.e., a quantum memory with a non-negligible lifetime) in a PID to its internal PMD (as in Fig. 3b).

V. SEMI-DEVICE-INDEPENDENT CHARACTERIZATION WITH NONTRANSIENT GUESSING GAMES

In this section, we demonstrate the operational significance of PID nonsimplicity in the scenario of guessing games. We propose a class of guessing games with double temporal stages, and we call them nontransient guessing games. Just as nonlocal games feature spatial separation between different parties [39, 59], nontransient games feature temporal separation between different stages, and therefore they are a suitable setting for characterizing correlations that exist across time, such as memory effect [41] and programmability [20]. We show that the winning probabilities of PIDs in the nontransient guessing games compose a complete set of incompatibility monotones, fully characterizing the incompatibility preorder between PIDs. This also implies that every nonsimple PID provides a nontrivial advantage over simple PIDs in some guessing game.
Furthermore, we prove that this advantage is bounded from above by the robustness of incompatibility of the PID and that this bound is tight. Finally, we comment on the limitations of nontransient guessing games in terms of experimental difficulties, and we propose a variant class of guessing games that overcomes such difficulties while also giving rise to a complete set of incompatibility monotones.

A. A complete set of incompatibility monotones

A nontransient guessing game is a parametrized interactive protocol between a player and a referee. Although such a game will be utilized to test the player’s PID, the game itself is actually “semi-device-independent.” That is, it requires that the referee’s operations be faithfully executed, but does not make any assumptions about the player’s device or strategy.

Definition 6. A nontransient guessing game between Alice (the player) and Bob (the referee) is specified by a bipartite POVM \( \mathcal{M} \equiv \{M_{m,n}\}_{m,n} \), and it has two stages separated by a time interval \( \Delta t \gg \Delta t_0 \).

1. In the first stage, Bob sends Alice one half of a maximally entangled state \( \varphi_+ \equiv 1/d \sum_{i,j} |i\rangle \otimes |i\rangle \), and then he asks Alice to submit a quantum system back to him. Bob measures Alice’s submitted system and the other half of \( \varphi_+ \), jointly according to the POVM \( \mathcal{M} \) and obtains a tuple \((m,n)\) as the outcome.

2. In the second stage (after \( \Delta t \) has passed), Bob announces the index \( m \) to Alice, and then he asks Alice to make a guess \( n' \) at the other index \( n \). Alice wins the game whenever she guesses correctly, i.e., whenever \( n' = n \).

Throughout the game, Alice has no access to quantum memories with a lifetime larger than \( \Delta t_0 \approx 0 \).

We note that in Definition 6, Alice’s device and strategy are both uncharacterized. To serve the purpose of testing PIDs, we now assume that Alice holds a PID \( \Lambda \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1} \) in hand, which may count as an additional resource for her in the game. It is also convenient for us to assume that the quantum delay time of \( \Lambda \) is no greater than the \( \Delta t_0 \) specified in Definition 6.

We note that from Alice’s perspective, the setting she is dealing with perfectly satisfies the late-program assumption; namely, the classical signal \( m \) does not arrive until a significant time interval \( \Delta t \approx \Delta t_0 \approx 0 \) after her quantum output is released. Therefore, the most general strategy for Alice to follow is to use her PID \( \Lambda \) to simulate whatever PID \( \Lambda' \) she can and to insert \( \Lambda' \) into the open slots of the game, as illustrated in Fig. 6. In addition, given that the quantum delay time of \( \Lambda \) is within \( \Delta t_0 \) and that no quantum memory with a lifetime exceeding \( \Delta t_0 \) is accessible (see Definition 6), we can rest assured that Alice’s simulation of \( \Lambda' \) using \( \Lambda \) is a free simulation, i.e., \( \Lambda \gg \Lambda' \) (see Definition 3). As a result, Alice’s maximum winning probability in the nontransient guessing game specified by the POVM \( \mathcal{M} \) is given by

\[
P_{\text{guess}}(\Lambda; \mathcal{M}) := \max_{\Lambda' \gg \Lambda} \sum_{m,n} \text{Tr} \left[ M_{m,n} \left( \text{id} \otimes \Lambda'_{n|m} \right) \varphi_+ \right].
\]

(17)

By the transitivity of the incompatibility preorder \( \gg \) [i.e., the sequential composability of free simulations in Theorem 1(3)], we can observe that Alice’s winning probability \( P_{\text{guess}}(\Lambda; \mathcal{M}) \) is a nonsimplicity monotone with respect to \( \Lambda \) given any POVM \( \mathcal{M} \). The following theorem states that when considering all different POVMs, the winning probabilities \( \{P_{\text{guess}}(\Lambda; \mathcal{M})\}_\mathcal{M} \) compose a complete set of nonsimplicity monotones, which faithfully reflects convertibility between PIDs under free simulations.

Theorem 3. Let \( \Lambda \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1} \) and \( \Gamma \equiv \{\Gamma_{y_1|y_0}\}_{y_0,y_1} \) be two PIDs. Then \( \Lambda \gg \Gamma \) if and only if \( P_{\text{guess}}(\Lambda; \mathcal{M}) \geq P_{\text{guess}}(\Gamma; \mathcal{M}) \) for every bipartite POVM \( \mathcal{M} \equiv \{M_{m,n}\}_{m,n} \).

The “only if” part of Theorem 3 is evident from the monotonicity of winning probabilities as argued before. The proof of the “if” part is detailed in Appendix C 1. It is a proof by construction, utilizing the closedness and convexity of the resource theory [Theorem 1(4) and (5)], the hyperplane separation theorem [60], and a technique employed in Refs. [27, 61].

A prominent implication of Theorem 3 is that each nontransient guessing game can be used as a certification of nonsimplicity of a PID, and the combination of all such certifications is sufficient to compose a faithful criterion for deciding nonsimplicity. Specifically, since all simple PIDs are interconvertible via free simulations [Theorem 1(2)], Theorem 3 implies that all simple PIDs give rise to the same winning probability in any fixed nontransient guessing game specified by \( \mathcal{M} \), which equals

\[
P_{\text{simple guess}}(\mathcal{M}) := \max_{\Omega} \sum_{m,n} \text{Tr} \left[ M_{m,n} \left( \text{id} \otimes \Omega_{n|m} \right) \varphi_+ \right].
\]

(18)

Therefore, as long as Alice’s winning probability in this game is observed to be greater than \( P_{\text{simple guess}}(\mathcal{M}) \), we can conclude with certainty that Alice holds some device that functions as a nonsimple PID. Conversely, if Alice holds a nonsimple PID \( \Lambda \) and always follows an optimal strategy while playing the games, then there must exist a specific game that certifies
the nonsimplicity of her device, i.e., a POVM \( \mathcal{M} \) such that 
\[ P_{\text{guess}}(\Lambda; \mathcal{M}) > P_{\text{guess}}^\text{simple}(\mathcal{M}). \]

B. Robustness of incompatibility as the supremum of game advantage

The robustness of resource [62, 63] is a well-studied resource measure that reflects how tolerant a resource is to an admixture of generic noise. It is universally well defined in any resource theory, and it possesses many desirable properties as a resource measure, including faithfulness, monotonicity, and resource-dependent convexity (i.e., being a convex function when the set of free objects is convex) [24]. For the resource theory of PID nonsimplicity, the robustness measure can be defined as follows.

**Definition 7.** The robustness of incompatibility of a PID \( \Lambda \equiv \{\Lambda_{x|y}\}_{x,y} \), denoted by \( \text{RoI}(\Lambda) \), is defined as

\[
\text{RoI}(\Lambda) := \min_{r \geq 0} \sup_{\mathcal{M} \equiv \{M_{m,n}\}_{m,n}} \left( \frac{P_{\text{guess}}(\Lambda; \mathcal{M})}{P_{\text{guess}}^\text{simple}(\mathcal{M})} + r \right),
\]

subject to:

\[
\left( \frac{\Lambda_{x|y} + r \mathcal{M}_{x|y}}{1 + r} \right)_{x,y} \quad \text{is a simple PID},
\]

\[
\{\mathcal{M}_{x|y}\}_{x,y} \quad \text{is a PID}.
\]

As we discussed before, Theorem 3 implies that every non-simple PID can provide a nontrivial operational advantage over simple PDIs in some nontransient guessing game. The following theorem shows that this advantage can be quantitatively characterized by the robustness of incompatibility of the PID.

**Theorem 4.** Let \( \Lambda \equiv \{\Lambda_{x|y}\}_{x,y} \) be a PID. Then

\[
\sup_{\mathcal{M} \equiv \{M_{m,n}\}_{m,n}} \frac{P_{\text{guess}}(\Lambda; \mathcal{M})}{P_{\text{guess}}^\text{simple}(\mathcal{M})} = 1 + \text{RoI}(\Lambda),
\]

where the supremum is over all bipartite POVMs.

The proof of Theorem 4 is detailed in Appendix C.2. It consists of two parts, following a similar structure to the proofs of comparable results in Refs. [18, 20, 25, 27]. The first part proves that the advantage provided by any PID \( \Lambda \) in any nontransient guessing game, in terms of the ratio \( \frac{P_{\text{guess}}(\Lambda; \mathcal{M})}{P_{\text{guess}}^\text{simple}(\mathcal{M})} \), can never exceed \( 1 + \text{RoI}(\Lambda) \). This is done via a slight reformulation of Eq. (20). The second part explicitly constructs a sequence of games (specified by a sequence of POVMs) that approaches the aforementioned robustness upper bound on the advantage arbitrarily close, thus showing that this upper bound is tight. This is done by utilizing the dual conic program of Eq. (20). It is worth mentioning that while the construction of the sequence of POVMs may require an unbounded number of measurement outcomes, the dimensionality of the quantum systems to be measured is bounded. We also note that the convexity of the resource theory of PID nonsimplicity [Theorem 1(5)] plays a crucial role here, as it guarantees strong duality between the conic programs for the robustness of incompatibility.

C. Constructing experimentally friendly incompatibility tests

We now take a closer look at the experimental setup of using nontransient guessing games to test PID nonsimplicity. According to Definition 6 and as illustrated in Fig. 6, throughout the game procedure, no quantum memory with a lifetime larger than \( \Delta T \approx 0 \) is ever needed by Alice or Bob. In this sense, nontransient guessing games are resource efficient, as they do not consume any physical resource they are actually testing. On the other hand, testing PID nonsimplicity or convertibility following the scheme proposed in Theorem 3 can still be costly to experiment. This is because the scheme requires Bob to be able to implement infinitely many different POVMs reliably, which is hard to achieve in realistic settings. Therefore, we are motivated to design nonsimplicity tests that are experimentally friendlier. In what follows, we propose a new class of guessing games that also gives rise to a complete set of incompatibility monotones. We call such games postinformation guessing games, as they generalize the games after the same name in Ref. [20] for testing incompatibility between POVMs. Compared to nontransient guessing games, postinformation guessing games are experimentally more convenient to realize.

**Definition 8.** A postinformation guessing game between Alice (the player) and Bob (the referee) is specified by a state ensemble \( \varsigma \equiv \{\sigma_{m,n,i}\}_{m,n,i} \) and a POVM \( \mathcal{L} \equiv \{L_i\}_1 \), and it has two stages separated by a time interval \( \Delta t \gg \Delta T \approx 0 \).

1. In the first stage, Bob generates an index triple \( (m, n, l) \) with probability \( \text{Tr}[\sigma_{m,n,i}] \) and sends the state \( \sigma_{m,n,i} \) to Alice without announcing \( (m, n) \). Then Bob asks Alice to submit a quantum system back to him. Bob measures Alice’s submitted system according to the POVM \( \mathcal{L} \) and obtains an outcome \( l' \).

2. In the second stage (after \( \Delta t \) has passed), Bob announces the index \( j \) to Alice, and then he asks Alice to make a guess \( n' \) at the index \( n \). Alice wins the game whenever she guesses correctly and in the meantime does not alter the index \( l' = l \), i.e., whenever \( n' = n \) and \( l' = l \).

Throughout the game, Alice has no access to quantum memories with a lifetime larger than \( \Delta T \approx 0 \).

We note that Definition 8 is also semi-device-independent in the sense that Alice’s device and strategy are uncharacterized. As before, we now assume that Alice is assisted by a PID \( \Lambda \equiv \{\Lambda_{x|y}\}_{x,y} \) while playing the game and that the quantum delay time of \( \Lambda \) is no greater than the \( \Delta T \) specified in Definition 8. Since Alice has no access to quantum memories with a lifetime exceeding \( \Delta T \), her most general strategy is described by a PID \( \Lambda' \) such that \( \Lambda' \succeq \Lambda \), as illustrated in Fig. 7. As a result, Alice’s maximum winning probability in the postinformation guessing game specified by the state ensemble \( \varsigma \) and the POVM \( \mathcal{L} \) is given by

\[
P_{\text{guess}}'(\Lambda'; \varsigma; \mathcal{L}) := \max_{\Lambda' \succeq \Lambda} \sum_{m,n,l} \text{Tr}[L_i \Lambda'_{n|m} (\sigma_{m,n,l})].
\]

The following proposition states that for a certain POVM \( \mathcal{L} \), the winning probabilities \( \{P_{\text{guess}}'(\Lambda'; \varsigma; \mathcal{L})\}_\varsigma \) compose a...
complete set of nonsimplicity monotones with respect to $\Lambda$ when considering different source ensembles.

**Proposition 3.** Let $\Lambda \equiv \{A_{x}^{a_{m,n,l}}\}_{m,n,l}^{1}$ and $\Gamma \equiv \{B_{y_{0}}^{b_{Y_{0}}}\}_{y_{0} \in Y_{0}}^{1}$ be two PIDs, and let $\mathcal{L} \equiv \{L_{n}^{p_{i}}\}_{P_{i} \in \mathcal{L}}^{1}$ be an informationally complete POM. Then $\Lambda \preceq_{\Gamma} \Gamma$ if and only if $P'_{\text{guess}}(\Gamma; \mathcal{L}) \geq P'_{\text{guess}}(\Lambda; \mathcal{L})$ for every state ensemble $\zeta \equiv \{\sigma_{m,n,l}^{a_{m,n,l}}\}_{m \in Y_{0}, n \in Y_{1}, l \in L}^{1}$.

The proof of Proposition 3 is in Appendix C 3.

Proposition 3 generalizes Ref. [20, Theorem 1] in the sense that it reduces to the latter when the index $l$ and Alice’s quantum output are trivial. Compared to Theorem 2, Proposition 3 enables faithful tests of convertibility between PIDs under free simulations (and thus of nonsimplicity of a PID) with a much lower experimental requirement. First, unlike in Theorem 3, the tests in Proposition 3 do not demand infinitely many different POVMs on Bob’s side. Besides, only single-system operations are involved to carry out such tests, while entanglement distribution and bipartite measurements are avoided, reducing the experimental difficulty to a further extent. Although Proposition 3 still requires an infinitude of different source ensembles, it is also true that one can bypass this challenge by using a single tomographically complete state ensemble and classical postprocessing to simulate all different source ensembles in these tests. As a result, while implementing the tests in Proposition 3, Bob can reuse his quantum hardware over and over again and only vary his classical postprocessing when the parameter differences between tests.

VI. CONCLUSIONS AND DISCUSSIONS

A. Summary of results

In this paper, we have conducted a resource-theoretic analysis of incompatibility between quantum instruments in terms of nonsimple programmability of quantum devices. We have been physically motivated by the notion of programmability, which envisions certain quantum devices as objects that can be programmed at any time, i.e., regardless of when the quantum input arrives. This naturally restricts the investigation to programmable instrument devices, which are classically controlled mechanisms that implement nonsignaling multi-instruments (Definition 1). Every PID possesses two characteristic time intervals: (i) the quantum delay time $\Delta_{D}$, which quantifies how quickly the device produces its quantum output, and (ii) the lifetime of the internal quantum memory $\Delta_{QM}$, which quantifies how long the device is able to store some form of quantum information to influence the classical output. To provide the experimenter with full temporal freedom on when the program can be issued, simple PIDs can have $\Delta_{D}^{\text{simple}} \leq \Delta_{D}$, whereas nonsimple PIDs must satisfy $\Delta_{QM} \gg \Delta_{D}$. Quantum memories are thus the resource that enables nonsimple programmability. To isolate the different memory demands between simple and nonsimple PIDs, we have assumed $\Delta_{D} \approx 0$ for all PIDs so that only nonsimple PIDs require a built-in quantum memory with a non-negligible lifetime to implement.

In the resource theory of PID nonsimplicity, the experimenter is allowed to perform arbitrary auxiliary processing that does not depend on quantum memories with a non-negligible lifetime. This restricts the allowable transformations between PIDs to a set of quantum combs called free simulations (Definition 3). As nonsimplicity of PIDs is mathematically captured by incompatibility of the family of instruments being implemented, the ability of one PID to freely simulate another identifies an incompatibility preorder between these devices (Theorem 1).

Every PID can be understood as a channel assemblage produced through a process of channel steering, and simple PIDs correspond to unsteerable channel assemblages. So yet another way to frame this work is as a resource theory of channel steering. From a practical point of view, channel steering offers a way of investigating properties of a given broadcast channel when the measurement device of one receiver is untrusted [13, 27]. We have deepened the connections between PID nonsimplicity and channel steering by deriving for every PID a unique steering decomposition (Proposition 1) and showing that this decomposition is “canonical” (Proposition 2). An essential ingredient of this decomposition, called the steering-equivalence mapping (Definition 5), has subsequently been identified as an object-valued incompatibility monotone (Theorem 2). This monotonicty result reflects a fundamental connection between the resource theory of PID nonsimplicity and that of measurement incompatibility, and consequently, any measure of incompatibility between POVMs previously studied in the literature [14, 16–18] can now be used to quantify incompatibility between quantum instruments.

We have also proposed operational schemes for measuring and benchmarking nonsimplicity of PIDs by designing a class of games, called nontransient guessing games (Definition 6). These games have temporally separated stages in a way that resembles the spatially separated parties in nonlocal games, and therefore they are adept at characterizing correlations that exist across time. We have shown that the maximum winning probability in any nontransient guessing game is a nonsimplicity monotone with respect to the player’s PID, and the collection of all such winning probabilities under different game settings provides a complete criterion for judging whether a given PID
can freely simulate another (Theorem 3). Since no assumption needs to be made about the player’s device or strategy, nontransient guessing games also provide semi-device-independent certifications for PID nonsimplicity. We have also established a tight upper bound on the operational advantage of a given PID over simple PIDs in nontransient guessing games in terms of a well-studied resource measure, namely the robustness against noise (Theorem 4). This result endows the robustness of incompatibility with a clear operational meaning. Considering the fact that testing PID convertibility using nontransient guessing games can be experimentally costly to implement, we have provided an alternative but experimentally friendlier scheme for such tests based on a class of so-called postinformation guessing games (Proposition 3).

B. Outlook

Our work leads to several directions for future research. First, we have treated PID nonsimplicity as a dynamical resource distributed over quantum networks [51] rather than carried by quantum channels [64, 65]. The difference here is that, PIDs are quantum 2-combs (i.e., quantum networks with two vertices), and so they can be manipulated by quantum 4-combs [51], whereas generic quantum channels can only be manipulated by quantum superchannels [49, 50]. A potential direction is to generalize the concept of incompatibility and its resource theory to more complex network layouts. It would also be interesting to study resources other than incompatibility in the network setting [66, 67], and a generic framework for studying resources in networks has been lacking.

Second, we have investigated the relationships between a number of quantum correlations (see Table I), and we have clarified that the presence of quantum memories (i.e., entanglement-breaking channels) is a precondition for any of these correlations. This is a qualitative remark, and one can continue to conduct a quantitative analysis on the pivotal role of quantum memories by asking what is the limit of nonclassicality in quantum correlations generated by unideal quantum memories (i.e., nonidentity channels). Following this line, one may further expect that a universal framework for studying various quantum correlations can be established based on the resource theory of quantum memories [41].

Finally, as we mentioned before, PIDs are quantum 2-combs with the second vertex being classical, and thus they can also be interpreted as quantum superchannels transforming one POVM to another. Interestingly, despite simple PIDs being a proper subset of general PIDs, any conversion between two single-party POVMs via general PIDs can always be realized via simple PIDs. This indicates a vanishing operational distinction between general and simple PIDs in terms of converting single-party POVMs. However, the distinction between general and simple PIDs becomes conspicuous when we consider convertibility between bipartite POVMs via partial action of these PIDs. This is well demonstrated by the nontransient guessing games, where a PID \( \Lambda \) acts on one part of a bipartite POVM. \( \rhd \) (see Fig. 6), and the performance gap between general and simple PIDs is nontrivial (Theorem 3). This kind of interplay between simple PIDs and (single-party and bipartite) POVMs is somehow reminiscent of that between entanglement-breaking channels and (single-party and bipartite) states. As for the latter, convertibility between single-party states via general channels is equivalent to convertibility via entanglement-breaking channels, whereas convertibility between bipartite states via partial action of general channels does not imply convertibility via partial action of entanglement-breaking channels. Partial action of PIDs on bipartite POVMs is still not fully understood, and we leave the exploration for future work.

ACKNOWLEDGEMENTS

We thank Vishal Singh and Mark M. Wilde for helpful comments on the presentation of the paper. This work was supported by the Department of Energy (DOE) Q-NEXT: National Quantum Information Science Research Center.

Appendix A: Properties of Free Simulations

In this Appendix, we demonstrate resource-theoretic properties possessed by free simulations of PIDs. These properties include simplicity [Theorem 1(1)], reachability [Theorem 1(2)], composability [Theorem 1(3)], closedness [Theorem 1(4)], and convexity [Theorem 1(5)].

1. Proof of Theorem 1(1)

Let \( \Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \}_{x_0,x_1} \) be a simple PID. By Definition 1, there exists a quantum instrument \( \{ \mathcal{G}_{g}^{A_0\rightarrow A_1} \}_{g} \) and a classical channel (i.e., a conditional probability distribution) \( \{ p_{x_1|x_0,g} \}_{x_0,x_1,g} \) such that

\[
\Lambda_{x_1|x_0}^{A_0\rightarrow A_1} = \sum_{g} p_{x_1|x_0,g} \mathcal{G}_{g}^{A_0\rightarrow A_1} \quad \forall x_0,x_1.
\]
Let $\mathbb{I} \equiv \{\Gamma_{y_0,y_1}^{B_0 \rightarrow B_1}\}_{y_0,y_1}$ be a PID such that $\mathbb{I} \succeq \mathbb{I}$. By Definition 3, there exists a quantum channel $\mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}}$, a quantum instrument $\{\mathcal{K}^{A_1 \rightarrow B_1}_{k}\}_k$, and two classical channels $\{p'_{x_0,l|y_0,k}\}_{x_0,y_0,k,l}$ and $\{q'_{y_1,x_1,l}\}_{x_1,y_1,l}$ such that

$$
\Gamma_{y_1|y_0}^{B_0 \rightarrow B_1} = \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l} P'_{x_0,l|y_0,k} \mathcal{K}^{A_1 \rightarrow B_1}_k \circ \left(\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \otimes \text{id}^D\right) \circ \mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}}
$$

$$
= \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l} P_{x_1|y_0,g} P'_{x_0,l|y_0,k} \mathcal{K}^{A_1 \rightarrow B_1}_k \circ \left(\mathcal{G}_g^{A_0 \rightarrow A_1} \otimes \text{id}^D\right) \circ \mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}} \forall y_0,y_1.
$$

(A3)

Define a quantum instrument $\{\mathcal{G}^{\gamma^{B_0 \rightarrow B_1}}_{g,k}\}_{g,k}$ and a classical channel $\{p''_{y_1|y_0,g,k}\}_{y_0,g,k}$ as follows:

$$
\mathcal{G}^{\gamma^{B_0 \rightarrow B_1}}_{g,k} := \mathcal{K}^{A_1 \rightarrow B_1}_k \circ \left(\mathcal{G}_g^{A_0 \rightarrow A_1} \otimes \text{id}^D\right) \circ \mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}} \forall g,k.
$$

(A4)

$$
p''_{y_1|y_0,g,k} := \sum_{x_0,x_1,l} q'_{y_1|x_1,l} P_{x_1|y_0,g} P'_{x_0,l|y_0,k} \forall y_0,y_1,g,k.
$$

(A5)

Then it follows from Eqs. (A3)–(A5) that

$$
\sum_{g,k} p''_{y_1|y_0,g,k} \mathcal{G}^{\gamma^{B_0 \rightarrow B_1}}_{g,k} = \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l} P_{x_1|y_0,g} P'_{x_0,l|y_0,k} \mathcal{K}^{A_1 \rightarrow B_1}_k \circ \left(\mathcal{G}_g^{A_0 \rightarrow A_1} \otimes \text{id}^D\right) \circ \mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}}
$$

$$
= \Gamma_{y_1|y_0}^{B_0 \rightarrow B_1} \forall y_0,y_1.
$$

(A6)

(A7)

By Definition 1, this shows that the simulated PID $\Gamma$ is simple. This concludes the proof of Theorem 1(1).

2. Proof of Theorem 1(2)

Let $\mathbb{A} \equiv \{\mathcal{A}_{x_0}^{A_0 \rightarrow A_1}\}_{x_0}$ be a PID, and let $\mathbb{I} \equiv \{\Gamma_{y_0,y_1}^{B_0 \rightarrow B_1}\}_{y_0,y_1}$ be a simple PID. By Definition 1, there exists a quantum instrument $\{\mathcal{G}_g^{B_0 \rightarrow B_1}\}_g$ and a classical channel $\{p_{y_1|y_0,g}\}_{y_0,y_1,g}$ such that

$$
\Gamma_{y_1|y_0}^{B_0 \rightarrow B_1} = \sum_{g} p_{y_1|y_0,g} \mathcal{G}_g^{B_0 \rightarrow B_1} \forall y_0,y_1.
$$

(A8)

Let $\delta$ denote the classical identity channel, which satisfies $\delta_{b|a} := 1$ if $a = b$ and $\delta_{b|a} := 0$ if $a \neq b$. Define a quantum channel $\mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}}$, a quantum instrument $\{\mathcal{K}^{A_1 \rightarrow B_1}_g\}_g$, and two classical channels $\{p'_{x_0,l|y_0,g}\}_{x_0,y_0,g,l}$ and $\{q'_{y_1,x_1,l}\}_{x_1,y_1,l}$ as follows:

$$
\mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}} := |0\rangle \langle 0|^{A_0} \otimes \text{id}^{B_0 \rightarrow D},
$$

(A9)

$$
\mathcal{K}^{A_1 \rightarrow B_1}_g := \mathcal{G}_g^{B_0 \rightarrow B_1} \circ \text{Tr}_{A_1} \forall g.
$$

(A10)

$$
p'_{x_0,l|y_0,g} := \delta_{x_0|0} P_{l|y_0,g} \forall x_0,y_0,g,l.
$$

(A11)

$$
q'_{y_1,x_1,l} := \delta_{y_1,l} \forall x_1,y_1,l.
$$

(A12)

Then it follows from Eqs. (A8)–(A12) that

$$
\sum_{x_0,x_1,g,l} q'_{y_1|x_1,l} P'_{x_0,l|y_0,g} \mathcal{K}^{A_1 \rightarrow B_1}_g \circ \left(\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \otimes \text{id}^D\right) \circ \mathcal{T}^{\Gamma_{B_0 \rightarrow A_0 D}}
$$

$$
= \sum_{x_0,x_1,g,l} \delta_{y_1,l} \delta_{x_0|0} P_{l|y_0,g} \mathcal{G}_g^{B_0 \rightarrow B_1} \circ \left(\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \left[|0\rangle \langle 0|^{A_0}\right] \otimes \text{id}^{B_0 \rightarrow D}\right)
$$

$$
= \sum_{g} p_{y_1|y_0,g} \mathcal{G}_g^{B_0 \rightarrow B_1} \forall y_0,y_1.
$$

(A13)

(A14)

(A15)

By Definition 3, this shows $\mathbb{A} \succeq \mathbb{I}$. This concludes the proof of Theorem 1(2).
3. Proof of Theorem 1(3)

First, we prove sequential composable. Consider a free simulation \( \Lambda \mapsto \Gamma \), where \( \Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \}_{x_0,x_1} \) and \( \Gamma \equiv \{ \Gamma_{y_1|y_0}^{B_0\rightarrow B_1} \}_{y_0,y_1} \) are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel \( \mathcal{F}_{B_0\rightarrow A_0}^D \), a quantum instrument \( \{ \mathcal{K}_k^{A_0\rightarrow B_1} \}_{k} \), and two classical channels \( \{ p_{x_0,l|y_0,k} \}_{x_0,y_0,k,l} \) and \( \{ q_{y_1|x_1,l} \}_{x_1,y_1,l} \) such that

\[
\Gamma_{y_1|y_0}^{B_0\rightarrow B_1} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}_k^{A_0\rightarrow B_1} = \left( \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \otimes \text{id}_D \right) \circ \mathcal{F}_{B_0\rightarrow A_0}^D \quad \forall y_0,y_1. \tag{A16}
\]

Consider another free simulation \( \Lambda \mapsto \Psi \), where \( \Psi \equiv \{ \Psi_{z_1|z_0}^{C_0\rightarrow C_1} \}_{z_0,z_1} \) is a PID. By Definition 3, this free simulation can be represented by a quantum channel \( \mathcal{F}_{C_0\rightarrow B_0}^E \), a quantum instrument \( \{ \mathcal{K}_k^{B_0\rightarrow E_1} \}_{k} \), and two classical channels \( \{ p'_{y_0,l|z_0,k,k'} \}_{y_0,z_0,k,k',l'} \) and \( \{ q'_{z_1|y_1,l,l'} \}_{y_1,z_1,l,l'} \) such that

\[
\Psi_{z_1|z_0}^{C_0\rightarrow C_1} = \sum_{y_0,y_1,k,k',l,l'} q'_{z_1|y_1,l,l'} p'_{y_0,l|z_0,k,k'} \mathcal{K}_k^{B_0\rightarrow E_1} = \left( \Gamma_{y_1|y_0}^{B_0\rightarrow B_1} \otimes \text{id}_E \right) \circ \mathcal{F}_{C_0\rightarrow B_0}^E \quad \forall z_0,z_1. \tag{A17}
\]

Combining Eqs. (A16) and (A17), the sequential composition of the above two free simulations can be described as \( \Lambda \mapsto \Psi \) such that

\[
\Psi_{z_1|z_0}^{C_0\rightarrow C_1} = \sum_{x_0,x_1,y_0,y_1,k,k',l,l'} q'_{z_1|y_1,l,l'} q_{y_1|x_1,l} p_{x_0,l|y_0,k} p'_{y_0,l|z_0,k,k'} \mathcal{K}_k^{B_0\rightarrow E_1} = \left( \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \otimes \text{id}_{DE} \right) \circ \left( \mathcal{F}_{B_0\rightarrow A_0}^D \otimes \text{id}_E \right) \circ \mathcal{F}_{C_0\rightarrow B_0}^E \quad \forall z_0,z_1. \tag{A18}
\]

Define a quantum channel \( \mathcal{F}_{C_0\rightarrow A_0}^{DE} \), a quantum instrument \( \{ \mathcal{K}_k^{A_0\rightarrow DE} \}_{k,k'} \), and two classical channels \( \{ p''_{x_0,l|y_0,k,k',l,l'} \}_{x_0,z_0,k,k',l,l'} \) and \( \{ q''_{z_1|y_1,l,l'} \}_{y_1,z_1,l,l'} \) as follows:

\[
\mathcal{F}_{C_0\rightarrow A_0}^{DE} := \left( \mathcal{F}_{B_0\rightarrow A_0}^D \otimes \text{id}_E \right) \circ \mathcal{F}_{C_0\rightarrow B_0}^E, \tag{A19}
\]

\[
\mathcal{K}_k^{A_0\rightarrow DE} := \mathcal{K}_k^{B_0\rightarrow E_1} \circ \left( \mathcal{K}_k^{A_0\rightarrow B_1} \otimes \text{id}_E \right) \quad \forall k,k', \tag{A20}
\]

\[
p''_{x_0,l|y_0,k,k',l,l'} := \sum_{y_0} p_{x_0,l|y_0,k} p'_{y_0,l|z_0,k,k'} \quad \forall x_0,z_0,k,k',l,l', \tag{A21}
\]

\[
q''_{z_1|y_1,l,l'} := \sum_{y_1} q'_{z_1|y_1,l,l'} q_{y_1|x_1,l} \quad \forall x_1,z_1,l,l'. \tag{A22}
\]

Then it follows from Eqs. (A18)–(A22) that

\[
\Psi_{z_1|z_0}^{C_0\rightarrow C_1} = \sum_{x_0,x_1,k,k',l,l'} q''_{z_1|x_1,l,l'} p''_{x_0,l|y_0,k,k'} \mathcal{K}_k^{A_0\rightarrow DE} = \left( \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \otimes \text{id}_{DE} \right) \circ \mathcal{F}_{\mathcal{F}_{C_0\rightarrow A_0}^{DE}} \quad \forall x_0,x_1,k,k',l,l'. \tag{A23}
\]

\[
= \sum_{x_0,x_1,y_0,y_1,k,k',l,l'} q'_{z_1|y_1,l,l'} q_{y_1|x_1,l} p_{x_0,l|y_0,k} p'_{y_0,l|z_0,k,k'} \mathcal{K}_k^{B_0\rightarrow E_1} = \left( \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \otimes \text{id}_D \right) \circ \mathcal{F}_{B_0\rightarrow A_0}^D \quad \forall y_0,y_1. \tag{A24}
\]

By Definition 3, this shows that the sequential composition \( \Lambda \mapsto \Psi \) of the two free simulations is a free simulation.

Next, we prove parallel composable. Consider a free simulation \( \Lambda \mapsto \Gamma \), where \( \Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \}_{x_0,x_1} \) and \( \Gamma \equiv \{ \Gamma_{y_1|y_0}^{B_0\rightarrow B_1} \}_{y_0,y_1} \) are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel \( \mathcal{F}_{B_0\rightarrow A_0}^D \), a quantum instrument \( \{ \mathcal{K}_k^{A_0\rightarrow B_1} \}_{k} \), and two classical channels \( \{ p_{x_0,l|y_0,k} \}_{x_0,y_0,k,l} \) and \( \{ q_{y_1|x_1,l} \}_{x_1,y_1,l} \) such that

\[
\Gamma_{y_1|y_0}^{B_0\rightarrow B_1} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}_k^{A_0\rightarrow B_1} = \left( \Lambda_{x_1|x_0}^{A_0\rightarrow A_1} \otimes \text{id}_D \right) \circ \mathcal{F}_{B_0\rightarrow A_0}^D \quad \forall y_0,y_1. \tag{A25}
\]
Consider another free simulation $\Lambda' \mapsto \Gamma'$, where $\Lambda' \equiv \{ \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1'} \}_{x_0} \times \{ \Lambda_{x_1'|x_0'}^{A_1 \rightarrow A_1'} \}_{x_0'}$ are two PIDs. By Definition 3, this free simulation can be represented by a quantum channel $\mathcal{F}^{B_0' \rightarrow A_0 A_1', D'}$, a quantum instrument $\{ \mathcal{K}_{k'}^{A_1 D' \rightarrow B_1} \}_{k'}$, and two classical channels $\{ p_{x_0', l'} \}_{l'} \times \{ q_{x_1', l'} \}_{l'}$ such that

$$
\Gamma^{B_0' \rightarrow B_1'}_{x_1|x_0} = \sum_{x_0', x_1', l'} q_{x_1', l'}^{(x_1, l)'} p_{x_0', l'}^{(x_0, l)} |y'_0 \rangle \langle k' |^{D'} \circ \mathcal{F}^{B_0' \rightarrow A_0'D'} \quad \forall y_0', y_1'.
$$

(A26)

Combining Eqs. (A25) and (A26), the parallel composition of the above two free simulations can be described as $\Lambda'' \mapsto \Gamma''$, where $\Lambda'' \equiv \{ \Lambda_{x_1,x_1'|x_0,x_0'}^{A_0 A_1' \rightarrow A_1 A_1'} \}_{x_0,x_0',x_1,x_1'}$ and $\Gamma'' \equiv \{ \Gamma^{B_0' \rightarrow B_1'}_{x_1|x_0} \}_{x_0,y_0,y_1}$ are two PIDs, such that

$$
\Gamma^{B_0 B_0' \rightarrow B_1 B_1'}_{y_1,y_1'|y_0,y_0'} = \sum_{x_0,x_0',x_1,x_1', y_0,y_0'} q_{y_1|x_1}^{(A_1 D')^{(A_1 D')^{(A_1 D')}}} p_{x_0,x_0',l,l'} |y_0\rangle \langle k |^{D'} \circ \mathcal{F}^{B_0 \rightarrow A_0 D} \circ \mathcal{F}^{B_0' \rightarrow A_1 D'}
$$

(A27)

Define a quantum channel $\mathcal{F}^{B_0 B_0' \rightarrow A_0 A_1' D'}$, a quantum instrument $\{ \mathcal{K}_{k,k'}^{A_1 D \rightarrow B_1} \}_{k,k'}$, and two classical channels $\{ p_{x_0,x_0',l,l'} \}_{x_0,x_0',y_0',y_0',y_0,y_0'}$ and $\{ q_{y_1|x_1} \}_{y_1,y_1'^{D'}}$ as follows:

$$
\mathcal{F}^{B_0 B_0' \rightarrow A_0 A_1' D'} := \mathcal{F}^{B_0 \rightarrow A_0 D} \circ \mathcal{F}^{B_0' \rightarrow A_0 D'}
$$

(A28)

$$
\mathcal{K}_{k,k'}^{A_1 D \rightarrow B_1} := \mathcal{K}_{k}^{A_1 D \rightarrow B_1} \circ \mathcal{K}_{k'}^{A_1 D \rightarrow B_1} \quad \forall k, k'.
$$

(A29)

$$
p_{x_0,x_0',l,l'}^{A_0 A_1'} := p_{x_0,x_0',l,l'}^{A_0 A_1'} |y_0\rangle \langle k |^{D'} \quad \forall x_0, x_0', y_0, y_0', y_0, y_0, y_0', y_0, y_0', l, l'.
$$

(A30)

$$
q_{y_1|x_1}^{A_1 D'} := q_{y_1|x_1}^{A_1 D'} |y_1\rangle \langle y_1 |^{D'} \quad \forall x_1, x_1', y_1, y_1', l, l'.
$$

(A31)

Then it follows from Eqs. (A27)-(A31) that

$$
\sum_{x_0,x_0',x_1,x_1',y_0,y_0'} q_{y_1|x_1}^{A_1 D'} p_{x_0,x_0',l,l'}^{A_0 A_1'} |y_0\rangle \langle k |^{D'} \circ \mathcal{F}^{B_0 \rightarrow A_0 D} \circ \mathcal{F}^{B_0' \rightarrow A_0 D'}
$$

$$
= \sum_{x_0,x_0',x_1,x_1',l,l'} q_{y_1|x_1}^{A_1 D'} p_{x_0,x_0',l,l'}^{A_0 A_1'} |y_0\rangle \langle k |^{D'} \circ \mathcal{F}^{B_0 \rightarrow A_0 D} \circ \mathcal{F}^{B_0' \rightarrow A_0 D'}
$$

$$
= \mathcal{F}^{B_0 B_0' \rightarrow A_0 A_1' D'}
$$

(A32)

By Definition 3, this shows that the parallel composition $\Lambda'' \mapsto \Gamma''$ of the two free simulations is a free simulation. This concludes the proof of Theorem 1(3).

4. Proof of Theorem 1(4)

Consider a free simulation $\Lambda \mapsto \Gamma$, where $\Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \}_{x_0 \in X_0, x_1 \in X_1}$ and $\Gamma \equiv \{ \Gamma^{B_0 \rightarrow B_1}_{y_1|y_0} \}_{y_0 \in Y_0, y_1 \in Y_1}$ are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel $\mathcal{F}^{B_0 \rightarrow A_0 D}$, a quantum instrument $\{ \mathcal{K}_{k}^{A_1 D \rightarrow B_1} \}_{k}$, and two classical channels $\{ p_{x_0,l|y_0,k} \}_{x_0,y_0,k,l}$ and $\{ q_{y_1|x_1,l} \}_{y_1,x_1,l}$ such that

$$
\Gamma^{B_0 \rightarrow B_1}_{y_1|y_0} = \sum_{x_0,x_1,l,k} q_{y_1|x_1,l,p_{x_0,l|y_0,k}}^{A_0 \rightarrow A_1} \circ \mathcal{F}^{B_0 \rightarrow A_0 D} \quad \forall y_0, y_1.
$$

(A34)

By Ref. [50, Theorem 1(4)], the dimensionality of the system $D$ can be bounded by the product of the dimensionality of the systems $A_0$ and $B_0$. Since every classical channel can be decomposed into a probabilistic mixture of deterministic classical channels, there exists a conditional probability distribution $\{ q_{y_1|x_1,l} \}_{y_1,x_1,l}$ such that

$$
q_{y_1|x_1,l} = \sum_{l \in L} \delta_{y_1|p_{y_1|l}} q_{y_1|x_1,l}^{(x_1,l)} \quad \forall x_1, y_1, l.
$$

(A35)
where \( \mathcal{Y} \) is the finite set of all functions from \( X_0 \) to \( Y_0 \). Define a classical channel \( \{p_{x_0,l'|y_0,k} \}_{x_0,y_0,k,l'} \) as follows:

\[
p'_{x_0,l'|y_0,k} := \sum_i q_{l'|i} p_{x_0|y_0,k} \quad \forall x_0, y_0, k, l'.
\]

(A36)

Likewise, there exists a conditional probability distribution \( \{p'_{k'|k} \}_{k,k'} \) such that

\[
p'_{x_0,l'|y_0,k} = \sum_{k'\in K'} \delta_{x_0,l'|k'} p''_{k'|k} \quad \forall x_0, y_0, k, l'.
\]

(A37)

where \( K' := (X_0 \otimes \mathcal{Y}) \) is the finite set of all functions from \( Y_0 \) to \( X_0 \otimes L' \). Define a quantum instrument \( \{\mathcal{K}'_{A|D\rightarrow B_1} \}_{k'} \) and two classical channels \( \{p''_{x_0,l'|y_0,k} \}_{x_0,y_0,k,l'} \) and \( \{q_{y_1|x_1,l} \}_{x_1,y_1,l'} \) as follows:

\[
\mathcal{K}'_{A|D\rightarrow B_1} := \sum_k p''_{k'|k} \mathcal{K}'_{A|D\rightarrow B_1} \quad \forall k',
\]

(A38)

\[
p''_{x_0,l'|y_0,k} := \delta_{x_0,l'|k} p''_{k'|k} \quad \forall x_0, y_0, k', l',
\]

(A39)

\[
q_{y_1|x_1,l} := \delta_{y_1,l'} \quad \forall x_1, y_1, l'.
\]

(A40)

Then it follows from Eqs. (A34)–(A40) that

\[
\sum_{x_0,x_1,k,l'} q_{y_1|x_1,l} p''_{x_0,l'|y_0,k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D} = \sum_{x_0,x_1,k,l'} \delta_{y_1,l'(x_1)} \delta_{x_0,l'|k} p''_{k'|k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A41)

\[
= \sum_{x_0,x_1,k,l'} \delta_{y_1,l'(x_1)} \delta_{x_0,l'|k} p''_{k'|k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A42)

\[
= \sum_{x_0,x_1,k,l'} \delta_{y_1,l'(x_1)} \delta_{x_0,l'|k} p''_{k'|k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A43)

\[
= \sum_{x_0,x_1,k,l'} \delta_{y_1,l'(x_1)} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A44)

\[
= \sum_{x_0,x_1,k,l'} \delta_{y_1,l'(x_1)} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A45)

\[
= \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0|y_0,k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D}
\]

(A46)

\[
= \mathcal{T}_{B_0\rightarrow A_0D} \quad \forall y_0, y_1.
\]

(A47)

By Definition 3, this shows that every free simulation can be realized using side channels of bounded size, and therefore the set of free simulations is closed. This concludes the proof of Theorem 1(4).

5. Proof of Theorem 1(5)

Consider an ensemble of free simulations labeled by a random index \( i \). The \( i \)th free simulation is applied with probability \( p_i \geq 0 \) such that \( \sum_i p_i = 1 \). By Definition 3, the \( i \)th free simulation can be represented by a quantum channel \( \mathcal{T}_{B_0\rightarrow A_0D} \), a quantum instrument \( \{\mathcal{K}'_{A|D\rightarrow B_1} \}_{k} \), and two classical channels \( \{p_{x_0,l'|y_0,k} \}_{x_0,y_0,k,l'} \) and \( \{q_{y_1|x_1,l} \}_{x_1,y_1,l'} \). The probabilistic mixture of these free simulations is then described by \( \Lambda \mapsto \Gamma \), where \( \Lambda \equiv \{\Lambda_{x_1|x_0} \}_{x_0,x_1} \) and \( \Gamma \equiv \{\mathcal{T}_{B_0\rightarrow B_1} \}_{y_0,y_1} \) are two PIDs, such that

\[
\Gamma_{y_0\rightarrow B_1} = \sum_i p_i \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l} p'_{x_0,l'|y_0,k} \mathcal{K}'_{A|D\rightarrow B_1} \circ (\Lambda_{x_1|x_0} \otimes \text{id}^D) \circ \mathcal{T}_{B_0\rightarrow A_0D} \quad \forall y_0, y_1.
\]

(A48)

Define a quantum channel \( \mathcal{T}'_{B_0\rightarrow A_0DK} \), a quantum instrument \( \{\mathcal{K}'_{A|DK\rightarrow B_1} \}_{k,k'} \), and a classical channel \( \{p_{x_0,l'|y_0,k,k'} \}_{x_0,y_0,k,k',l'} \) as follows:

\[
\mathcal{T}'_{B_0\rightarrow A_0DK} := \sum_i p_i \mathcal{T}_{B_0\rightarrow A_0D} \otimes |i\rangle \langle i|_K.
\]

(A49)
\[ \mathcal{K}_{k,k'}^{A_0|D_K \rightarrow B_i} [\cdot] := \mathcal{K}_{k,k'}^{A_0|D} [\cdot] \left( \mathbb{1}^{A_i} \otimes \langle k' | K \right) \right \} \forall k,k'. \] (A50)

\[ p'_{X_0,i'|y_0,k,k'} := p'_{X_0,i'|y_0,k,k'} \mathcal{K}_{k,k'}^{A_0|D_K \rightarrow B_i} \quad \forall X_0, y_0, k, i', l'. \] (A51)

Then it follows from Eqs. (A48)–(A51) that

\[ \sum_{x_0,x_1,k,k',l'} q'_{x_1,i,l,l'} p'_{x_0,i',y_0,k,k'} \mathcal{K}_{k,k'}^{A_0|D_K \rightarrow B_i} \left( \Lambda_{x_0|y_0} \otimes \text{id}^{D_K} \right) \circ \mathcal{T}^{x_0} B_0 \rightarrow A_0 D_K \]

\[ = \sum_{x_0,x_1,k,k',l'} q'_{x_1,i,l,l'} p'_{x_0,i',y_0,k,k'} \mathcal{K}_{k,k'}^{A_0|D_K \rightarrow B_i} \left( \Lambda_{x_0|y_0} \otimes \text{id}^{D_K} \right) \circ \mathcal{T}^{x_0} B_0 \rightarrow A_0 D_K \]

\[ = \Gamma_{B_0 \rightarrow A_0} \quad \forall y_0, y_1. \] (A54)

By Definition 3, this shows that the probabilistic mixture \( \Lambda \mapsto \psi \) of the ensemble of free simulations is a free simulation. This concludes the proof of Theorem 1(5).

### Appendix B: Properties of the Steering-Equivalence Mapping

In this Appendix, we demonstrate useful properties possessed by the steering-equivalence mapping. These properties include an operational interpretation (Proposition 1), canonicity (Proposition 2), faithfulness [Theorem 2(1)], and monotonicity [Theorem 2(2)].

#### 1. Proof of Proposition 1

Let \( \Lambda \equiv \{ A_{x_0|y_0} \}_{x_0,x_1} \) be a PID. Let \( \Lambda_{A_0 \rightarrow A_1} := \sum_{x_1} \Lambda_{x_0 \rightarrow A_1} \) be the marginal channel of \( \Lambda \) from \( A_0 \) to \( A_1 \). Let \( r := \text{rank}(J_{A_0 \rightarrow A_1}^A) \). The Choi operator \( J_{A_0 \rightarrow A_1}^A \) has a spectral decomposition as follows:

\[ J_{A_0 \rightarrow A_1}^A = \sum_{i=0}^{r-1} \alpha_i |\alpha_i^A \rangle \langle \alpha_i^A |, \] (B1)

where \( \alpha_i > 0 \) is a positive real number for all \( i \) and \( \{|\alpha_i^A \rangle \}_{i} \) is an orthonormal set of vectors. Let \( A^* \) be a system such that \( \mathcal{H}^A \equiv \text{supp}(J_{A_0 \rightarrow A_1}^A) \subseteq \mathcal{H}^A_{A_0 A_1} \). Let \( |\alpha_i^A \rangle \) be the image of \( |\alpha_i \rangle \) in \( A^* \). Let \( |\overline{\alpha_i}^A \rangle \) be the complex conjugate of \( |\alpha_i^A \rangle \) under a fixed orthonormal basis. Then \( \{|\overline{\alpha_i}^A \rangle \}_{i} \) is an orthonormal basis of \( \mathcal{H}^{A^*} \). By Ref. [68, Eq. (2.2.36)], there exists an operator \( V_{A_0 \rightarrow A_1}^{A^*} \) such that

\[ \left( \mathbb{1}^{A_0} \otimes V_{A_0 \rightarrow A_1}^{A^*} \right) |\phi_i^A \rangle A_0 \hat{A}_0 = \sum_{i=0}^{r-1} \sqrt{\alpha_i} |\alpha_i^A \rangle \langle A_0 A_1 |\overline{\alpha_i}^A \rangle^A. \] (B2)

Define a linear map \( V_{A_0 \rightarrow A_1}^{A^*} \) as follows:

\[ V_{A_0 \rightarrow A_1}^{A^*} [\cdot] := V_{A_0 \rightarrow A_1}^{A^*} [\cdot] (V^{A_1 A^* \rightarrow A_0}) \] (B3)

We note that \( V_{A_0 \rightarrow A_1}^{A^*} \) is an isometric dilation of \( \Lambda_{A_0 \rightarrow A_1} \), as can be verified by marginalizing its Choi operator \( J_{A_0 \rightarrow A_1}^A \):

\[ \text{Tr}_{A^*} \left[ J_{A_0 \rightarrow A_1}^A \right] = \text{Tr}_{A^*} \left[ \left( \mathbb{1}^{A_0} \otimes V_{A_0 \rightarrow A_1}^{A^*} \right) |\phi_i^A \rangle A_0 \hat{A}_0 \left( \mathbb{1}^{A_0} \otimes (V^{A_1 A^* \rightarrow A_0}) \right) \right] \]

\[ = \text{Tr}_{A^*} \left[ \sum_{i,j=0}^{r-1} \sqrt{\alpha_i} \alpha_j |\alpha_i^A \rangle \langle \alpha_j^A | \otimes |\overline{\alpha_i}^A \rangle^A \langle \alpha_j^A |^A \right] \]

\[ = \sum_{i=0}^{r-1} \alpha_i |\alpha_i^A \rangle \langle A_0 A_1 | \]

\[ = J_{A_0 \rightarrow A_1}^A. \] (B7)
Let $S \equiv \{ S_{x_1|x_0}^{A'_{x_1|x_0}} \}_{x_0,x_1}$ be a PMD such that $S = \text{SEM}(\Lambda)$. By Definition 5,

$$S_{x_1|x_0}^{A'} = (J_{A'}^{A})^{-\frac{1}{2}} J_{A_{x_1|x_0}}^{A'} (J_{A'}^{A})^{-\frac{1}{2}} \quad \forall x_0,x_1. \quad (B8)$$

Define an isometric channel $\mathcal{W}^{A_0 \rightarrow A_1}$ as follows:

$$\rho_{\mathcal{W}^{A_0 \rightarrow A_1} : [\cdot]} := \sum_{i,j=0}^{r-1} \langle \alpha_i | [\cdot] | \alpha_j \rangle A_i \langle \alpha_j | A_0 A_1. \quad (B9)$$

It follows that

$$\text{Tr}_{A'} \left[ \left( \mathbb{1}_{A_0 A_1} \otimes (S_{x_1|x_0}^{T})^{A'} \right) J_{A_0 A_1}^{A'} \right] = \sum_{i,j=0}^{r-1} \sqrt{\mu_i \mu_j} \langle \alpha_i | S_{x_1|x_0}^{T} | \alpha_j \rangle A_i \langle \alpha_j | A_0 A_1 \quad (B10)$$

$$= \sum_{i,j=0}^{r-1} \sqrt{\mu_i \mu_j} \langle \alpha_i | S_{x_1|x_0} \alpha_j | A_i \rangle \langle \alpha_j | A_0 A_1 \quad (B11)$$

$$= \sum_{i,j=0}^{r-1} \sqrt{\mu_i \mu_j} \langle \alpha_i | S_{x_1|x_0}^{T} | \alpha_j \rangle A_i \langle \alpha_j | A_0 A_1 \quad (B12)$$

$$= \mathcal{W}^{A_0 \rightarrow A_1} \left[ (J_{A}^{A})^{-\frac{1}{2}} S_{x_1|x_0}^{T} (J_{A}^{A})^{-\frac{1}{2}} \right] \quad (B13)$$

$$= \mathcal{W}^{A_0 \rightarrow A_1} \left[ J_{A_{x_1|x_0}}^{A'} \right] \quad (B14)$$

$$= J_{A_{x_1|x_0}}^{A_1} \quad \forall x_0,x_1. \quad (B15)$$

By the Choi–Jamiołkowski isomorphism, we can conclude that

$$\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} : [\cdot] = \text{Tr}_{A'} \left[ \left( \mathbb{1}_{A_1} \otimes (S_{x_1|x_0}^{T})^{A'} \right) \mathcal{W}^{A_0 \rightarrow A_1} : [\cdot] \right] \quad \forall x_0,x_1. \quad (B16)$$

This concludes the proof of Proposition 1.

2. Proof of Proposition 2

Let $\Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \}_{x_0,x_1}$ be a PID. Let $\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} := \sum_{x_1} \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1}$ be the marginal channel of $\Lambda$ from $A_0$ to $A_1$. Let $\mathcal{E}^{A_0 \rightarrow A_1 E}$ be a broadcast channel, and let $\mathbb{B} \equiv \{ M_{x_1|x_0}^{E} \}_{x_0,x_1}$ be a PMD such that $\Lambda \leftrightarrow (\mathcal{E}, \mathbb{B})$. By Definition 4,

$$\Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} : [\cdot] = \text{Tr}_{E} \left[ \left( \mathbb{1}_{A_1} \otimes M_{x_1|x_0}^{E} \right) \mathcal{E}^{A_0 \rightarrow A_1 E} : [\cdot] \right] \quad \forall x_0,x_1. \quad (B17)$$

Let $\mathcal{V}^{A_0 \rightarrow A_1 E}$ be an isometric dilation of $\mathcal{E}^{A_0 \rightarrow A_1 E}$. Let $V^{A_0 \rightarrow A_1 EF}$ be the isometry operator such that

$$\mathcal{V}^{A_0 \rightarrow A_1 EF} : [\cdot] := V^{A_0 \rightarrow A_1 EF} : (\mathbb{1}_{A_1} \otimes \mathcal{E}^{A_0 \rightarrow A_1 E} : [\cdot]) \quad (B18)$$

Define a PMD $\mathbb{N} \equiv \{ N_{x_1|x_0}^{EF} \}_{x_0,x_1}$ as follows:

$$N_{x_1|x_0}^{EF} = M_{x_1|x_0}^{E} \otimes \mathbb{1}_{F} \quad \forall x_0,x_1. \quad (B19)$$

It follows from Eqs. (B17) and (B19) that

$$\text{Tr}_{E} \left[ \left( \mathbb{1}_{A_1} \otimes N_{x_1|x_0}^{EF} \right) \mathcal{V}^{A_0 \rightarrow A_1 EF} : [\cdot] \right] = \text{Tr}_{E} \left[ \left( \mathbb{1}_{A_1} \otimes M_{x_1|x_0}^{E} \right) \text{Tr}_{F} \circ \mathcal{V}^{A_0 \rightarrow A_1 EF} : [\cdot] \right] \quad (B20)$$

$$= \text{Tr}_{E} \left[ \left( \mathbb{1}_{A_1} \otimes M_{x_1|x_0}^{E} \right) \mathcal{E}^{A_0 \rightarrow A_1 EF} : [\cdot] \right] \quad (B21)$$

$$= \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} : [\cdot] \quad \forall x_0,x_1. \quad (B22)$$
where the steering-equivalent observables of the state assemblage

By Definitions 2 and 5, this implies

and the same isomorphism. It follows that

\[ \mathcal{W}^{A^* \rightarrow \mathcal{E}} \] as follows:

\[ \mathcal{W}^{A^* \rightarrow \mathcal{E}} [\cdot] = \sum_{i,j=0}^{r-1} \langle \alpha_i | \cdot | \alpha_j \rangle^{A^*} \langle \beta_i | \beta_j \rangle^{\mathcal{E}}. \] (B24)

Let \( J^{A^*\mathcal{E}} \) denote the image of the Choi operator \( J^{A_{\mathcal{E}} \rightarrow A^*\mathcal{E}} \) of \( \mathcal{W}^{A_{\mathcal{E}} \rightarrow A^*\mathcal{E}} \) in the composite system \( A^*\mathcal{E} \). Then

\[ (J^{A^*\mathcal{E}})^\frac{1}{2} (\mathcal{W}^T)^{\mathcal{E} \rightarrow A^*} \left[ (N^T_{x_1|x_0})^{\mathcal{E}} \right] (J^{A^*\mathcal{E}})^\frac{1}{2} = \sum_{i,j=0}^{r-1} \sqrt{a_i a_j} \langle \alpha_i | \cdot | \alpha_j \rangle^{A^*} \langle \mathcal{W}^T \rangle^{\mathcal{E} \rightarrow A^*} \left[ (N^T_{x_1|x_0})^{\mathcal{E}} \right] |\alpha_j\rangle \langle \alpha_j |^{A^*} \] (B25)

\[ = \sum_{i,j=0}^{r-1} \sqrt{a_i a_j} \langle \beta_j | N^T_{x_1|x_0} | \beta_j \rangle^{\mathcal{E}} |\alpha_i\rangle \langle \alpha_i |^{A^*} \] (B26)

\[ = \sum_{i,j=0}^{r-1} \sqrt{a_i a_j} \langle \beta_i | N^T_{x_1|x_0} | \beta_j \rangle^{\mathcal{E}} |\alpha_i\rangle \langle \alpha_i |^{A^*} \] (B27)

\[ = \text{Tr}_{\mathcal{E}} \left[ (\mathbb{1}^{A^*} \otimes N^{\mathcal{E}}_{x_1|x_0}) \left( J^{A^*\mathcal{E}} \right) \right] \] (B28)

\[ \text{Tr}_{\mathcal{E}} \left[ (\mathbb{1}^{A^*} \otimes N^{\mathcal{E}}_{x_1|x_0}) \left( J^{A^*\mathcal{E}} \right) \right] = J^{A^*}_{\Lambda_{x_1|x_0}} \quad \forall x_0,x_1. \] (B29)

Here Eq. (B29) follows from Eq. (B23) and the isomorphism between \( H^A \) and \( \text{supp}(J^{A_{\mathcal{E}}\rightarrow A^*}) \); Eq. (B30) follows from Eq. (B22) and the same isomorphism. It follows that

\[ (J^{A^*\mathcal{E}})^{-\frac{1}{2}} J^{A^*}_{\Lambda_{x_1|x_0}} (J^{A^*\mathcal{E}})^{-\frac{1}{2}} = (\mathcal{W}^T)^{\mathcal{E} \rightarrow A^*} \left[ (N^T_{x_1|x_0})^{\mathcal{E}} \right] \quad \forall x_0,x_1. \] (B31)

By Definitions 2 and 5, this implies \( N^T \|^M \text{SEM}(\Lambda) \). Since Eq. (B19) implies \( M^T \|^M \text{SEM}(\Lambda) \), by the transitivity of the preorder \(\geq_M \) [20], we have \( M^T \|^M \text{SEM}(\Lambda) \). This concludes the proof of Proposition 2.

3. Proof of Theorem 2(1)

By the Choi–Jamiołkowski isomorphism, a PID \( \Lambda \equiv \{ A_{x_1|x_0} \}_{x_0,x_1} \) is simple if and only if its corresponding assemblage of Choi states \( \{ J_{\Lambda_{x_1|x_0}}(d_{A_0}) \}_{x_0,x_1} \) is an unsteerable state assemblage [45]. The state assemblage \( \{ J_{\Lambda_{x_1|x_0}}(d_{A_0}) \}_{x_0,x_1} \) is unsteerable if and only if its steering-equivalent observables are compatible [58, Theorem 1]. By Definition 5 and Ref. [58, Eq. (5)], \( \text{SEM}(\Lambda) \) is the steering-equivalent observables of the state assemblage \( \{ J_{\Lambda_{x_1|x_0}}(d_{A_0}) \}_{x_0,x_1} \). Therefore, \( \Lambda \) is simple if and only if \( \text{SEM}(\Lambda) \) is simple. This concludes the proof of Theorem 2(1).

4. Proof of Theorem 2(2)

Let \( \Lambda \equiv \{ A_{x_1|x_0} \}_{x_0,x_1} \) and \( \Gamma \equiv \{ \Gamma_{y_0,B_{0|1}} \}_{y_0,y_1} \) be two PIDs such that \( \Lambda \geq \Gamma \). By Definition 3, there exists a quantum channel \( \mathcal{F}^{B_0 \rightarrow A_0D} \), a quantum instrument \( \{ H_{k}^{A_1 \rightarrow D_{B_{1|0}}} \}_{k} \), and two classical channels \( \{ p_{x_0,y_0|k,l} \}_{x_0,y_0,k,l} \) and \( \{ q_{y_1|x_1,l} \}_{x_1,y_1,l} \) such that

\[ \Gamma_{y_0,B_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,y_0|k,l} K_{k}^{A_1 \rightarrow B_{0|1}} \circ \{ A_{x_1|x_0} \otimes \text{id}^C \} \circ \mathcal{F}^{B_0 \rightarrow A_0D} \quad \forall y_0,y_1. \] (B32)
Let $\Lambda_{0 \rightarrow A_1} \equiv \sum_{x_1} \Lambda_{0 \rightarrow A_1|x_0}$ be the marginal channel of $\Lambda$ from $A_0$ to $A_1$. Let $A^*$ be a system such that $H_{A^*} \equiv \text{supp}(\Lambda_{A_0A_1}) \subseteq H_{A_0A_1}$. Let $S \equiv \{S_{x_1|x_0}\}$ be a PMD such that $S = \text{SEM}(\Lambda)$. By Proposition 1, there exists an isometric channel $V_{A_0 \rightarrow A_1A^*}$ such that

$$\Lambda_{0 \rightarrow A_1} \{ x \} = \text{Tr}_{A^*} \left( \left( I_{A_1} \otimes (S_{x_1|x_0}^{T}) \right) V_{A_0 \rightarrow A_1A^*} \{ x \} \right) \quad \forall x_0, x_1. \quad (B33)$$

Define a quantum channel $\mathcal{E}_{B_0 \rightarrow B_1A^*K}$ as follows:

$$\mathcal{E}_{B_0 \rightarrow B_1A^*K} = \sum_k \left( \mathcal{K}_k^{A_1D \rightarrow B_1} \otimes \text{id}_{A^*} \right) \circ \left( V_{A_0 \rightarrow A_1A^*} \otimes \text{id}_D \right) \circ \mathcal{F}_{B_0 \rightarrow A_0D} \otimes |k\rangle \langle k'|_K. \quad (B34)$$

Define a quantum instrument $\mathcal{J}^{A^*} \equiv \{ \mathcal{K}_{k_i}^{A^* \rightarrow A^*} \}_{k_i}$ as follows:

$$\mathcal{K}_{k_i}^{A^* \rightarrow A^*} \{ ] \} = \left( \mathcal{P}_{A^*} \otimes |k'|_K \right) \left( \mathcal{P}_{A^*} \otimes |k'|_K \right) \quad \forall k_i. \quad (B35)$$

Define a PMD $\mathcal{N} \equiv \{ N_{y_1|y_0}^{A^*K} \}_{y_0,y_1}$ as follows:

$$N_{y_1|y_0}^{A^*K} := \sum_{x_0, x_1, k, l} q_{y_1|x_1,l} \mathcal{P}_{x_0,l|y_0,k_i} \mathcal{K}_{k_i}^{A_1D \rightarrow B_1} \left( S_{x_1|x_0}^{T} \right) \quad (B36)$$

By Definition 2, Eq. (B36) implies $\mathcal{S} \succ_M \mathcal{N}$. It follows from Eqs. (B32)–(B34), and (B37) that

$$\text{Tr}_{A^*K} \left( \left( \mathcal{P}_{B_1} \otimes (N_{y_1|y_0}^{A^*K})^{A^*K} \right) \mathcal{E}_{B_0 \rightarrow B_1A^*K} \{ ] \} \right)$$

By Definition 4, this implies $\mathcal{I} \leftrightarrow (\mathcal{E}, \mathcal{N}^T)$, where $\mathcal{N}^T \equiv \{ (N_{y_1|y_0}^{A^*K})_y \}_{y_0,y_1}$. Then by Proposition 2, we have $\mathcal{N} \succ_M \text{SEM}(\mathcal{I})$. Since $\text{SEM}(\Lambda) = \mathcal{S} \succ_M \mathcal{N}$, by the transitivity of the preorder $\succ_M [20]$, we have $\text{SEG}(\Lambda) \succ_M \text{SEM}(\mathcal{I})$. This concludes the proof of Theorem 2(2).

**Appendix C: Semi-Device-Independent Characterization**

In this Appendix, we demonstrate the semi-device-independent characterization of PID nonsimplicity with guessing games. This includes providing a complete set of incompatibility monotones (Theorem 3) and an operational interpretation of the robustness of incompatibility (Theorem 4) based on nontransient guessing games, as well as a complete set of incompatibility monotones (Proposition 3) based on postinformation guessing games.

**1. Proof of Theorem 3**

Let $\Lambda \equiv \{ \Lambda_{x_1|x_0}^{A_0 \rightarrow A_1} \}_{x_0 \in X_0, x_1 \in X_1}$ be a PID. Let $\mathcal{M} \equiv \{ M_{m,n}^{C_0C_1} \}_{m \in M, n \in N}$ be a bipartite POVM. By Eq. (17), Alice’s maximum winning probability in the nontransient guessing game specified by $\mathcal{M}$ equals

$$P_{\text{guess}}(\Lambda; \mathcal{M}) := \frac{1}{d_{C_0}} \max_{\Lambda', x_1 \in X_1} \sum_{m \in M, n \in N} \text{Tr} \left( M_{m,n}^{C_0C_1} \left( \text{id}_{C_0} \otimes \Lambda_{x_1|0}^{A_0 \rightarrow A_1} \right) \left( \phi_{C_0}^{C_0} \phi_{C_0}^{C_0} \right) \right) \quad (C1)$$
First, we prove the necessity of the convertibility conditions in Theorem 3. Let \( \Lambda \equiv \{ \Lambda_{x_{0}|x_{1}} \}_{x_{0} \in X_{0}, x_{1} \in X_{1}} \) and \( \Gamma \equiv \{ \Gamma_{y_{1}|y_{0}} \}_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \) be two PIDs such that \( \Lambda \not\succ_{1} \Gamma \). Let \( \mathcal{M} \equiv \{ M_{c_{m,n}} \} \) be a bipartite POVM. By Theorem 1(3), the preorder \( \succ_{1} \) is transitive, thus

\[
P_{\text{guess}}(\Lambda; \mathcal{M}) = \frac{1}{d_{c_{0}}} \max_{\Lambda' \succ_{1} \Lambda} \sum_{m \in R, n \in R} \text{Tr} \left[ M_{c_{m,n}} f_{c_{m,n}} \right].
\]

Next, we prove the sufficiency of the convertibility conditions in Theorem 3 by contradiction. Let \( \Lambda \equiv \{ \Lambda_{x_{0}|x_{1}} \}_{x_{0} \in X_{0}, x_{1} \in X_{1}} \) and \( \Gamma \equiv \{ \Gamma_{y_{1}|y_{0}} \}_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \) be two PIDs such that \( P_{\text{guess}}(\Lambda; \mathcal{M}) \geq P_{\text{guess}}(\Gamma; \mathcal{M}) \) for every bipartite POVM \( \mathcal{M} \equiv \{ M_{c_{m,n}} \} \). We assume \( \Lambda \not\succ_{1} \Gamma \). This means that

\[
\exists \; \Lambda' \equiv \{ \Lambda'_{y_{1}|y_{0}} \}_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \; : \; \Lambda \not\succ_{1} \Lambda'.
\]

By Theorem 1(4) and (5), the set of PIDs on the right-hand side of Eq. (C6) is closed and convex. By the hyperplane separation theorem \([60]\), there exists a set of Hermiticity-preserving linear maps \( \{ O_{y_{0}|y_{1}} \}_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \) and a positive real number \( \varepsilon > 0 \) such that

\[
\sum_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \text{Tr} \left[ J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} \right] > \max_{\Lambda' \succ_{1} \Lambda} \sum_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \text{Tr} \left[ J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} \right] + \varepsilon.
\]

The above equation can be interpreted based on the Hilbert–Schmidt inner product as follows:

\[
\sum_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \text{Tr} \left[ J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} \right] > \max_{\Lambda' \succ_{1} \Lambda} \sum_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \text{Tr} \left[ J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} \right] + \varepsilon.
\]

Let \( c := \max_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \| J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} \| \). Define a bipartite POVM \( \hat{\mathcal{M}} \equiv \{ \hat{M}_{m,n} \} \) such that \( \hat{\mathcal{M}} \) satisfies as follows:

\[
\hat{M}_{m,n} := \frac{1}{2 \| y_{0} \| \| y_{1} \|} \left( J_{O_{y_{0}|y_{1}}} B_{O_{y_{0}|y_{1}}} + c I_{y_{1}|y_{0}} \right) \quad \forall m \in Y_{0}, n \in Y_{1}.
\]

It can be verified that \( \hat{\mathcal{M}} \) is a valid POVM, as \( \hat{M}_{m,n} \geq 0 \) for all \( m \in Y_{0}, n \in Y_{1} \) and \( \sum_{m \in Y_{0}, n \in Y_{1}} \hat{M}_{m,n} = I_{y_{1}|y_{0}} \). Then

\[
P_{\text{guess}}(\Gamma; \hat{\mathcal{M}}) = \frac{1}{2 cd_{B_{1}}} \max_{\Lambda' \succ_{1} \Lambda} \sum_{y_{0} \in Y_{0}, y_{1} \in Y_{1}} \text{Tr} \left[ \hat{M}_{m,n} f_{c_{m,n}} \right].
\]
Then Eq. (C22) implies
\[ \equiv \{
\]
Program (C24) exists. Let
\[ \Omega \equiv \left\{ x_A \mid 1 \to 0 \right\}
\]
\[ = \left\{ \omega_{x_0} N_{n|m}, \omega_{x_0} N_{n|m} \right\} - \frac{1}{|N|} \sum_{m \in N} \max_{m \in N} \left\{ j_{m,n} N_{n|m} \right\} + \frac{1}{2c} d_{B_0} |Y_0| |Y_1|
\]
\[ \leq \frac{1}{d_{B_0}} \left( \lambda^{\prime} : \lambda \to \lambda^{\prime} \right) \sum_{m \in N} \max_{m \in N} \left\{ j_{m,n} N_{n|m} \right\} - \frac{1}{|N|} \sum_{m \in N} \max_{m \in N} \left\{ j_{m,n} N_{n|m} \right\} + \frac{1}{2c} d_{B_0} |Y_0| |Y_1|
\]
\[ = \frac{1}{d_{B_0}} \left( \lambda^{\prime} : \lambda \to \lambda^{\prime} \right) \sum_{m \in N} \max_{m \in N} \left\{ j_{m,n} N_{n|m} \right\} - \frac{1}{|N|} \sum_{m \in N} \max_{m \in N} \left\{ j_{m,n} N_{n|m} \right\} + \frac{1}{2c} d_{B_0} |Y_0| |Y_1|
\]
\[ = P_{	ext{guess}}(\lambda^{\prime} : \lambda^{\prime}) - \frac{1}{|N|} \frac{1}{|Y_1|} + \frac{1}{2c} d_{B_0} |Y_0| |Y_1|
\]
Here Eqs. (C12), (C16), and (C19) follow from Eq. (C9); Ineq. (C14) follows from Ineq. (C8). Note that the PID in Ineq. (C17) has the form \( \lambda^{\prime} = \left\{ \lambda_{n|m}^{B_0 \to B_1} \right\} \) in contrast to the PID \( \lambda = \left\{ \lambda_{x_0} \right\} \) in Eq. (C16), and Ineq. (C17) follows from the fact that the freedom to make a guess \( n \) outside the set of target outcomes \( Y_1 \). This does not increase the maximum winning probability. Choose \( N \geq Y_1 \) to be an arbitrary index set such that
\[ |N| > |Y_1| + \frac{2c} {d_{B_0}} |Y_0| |Y_1|.
\]
Then Eq. (C22) implies \( P_{	ext{guess}}(\lambda^{\prime} : \lambda^{\prime}) < P_{	ext{guess}}(\lambda : \lambda^{\prime}) \), which contradicts the assumption that \( P_{	ext{guess}}(\lambda : \lambda) \geq P_{	ext{guess}}(\lambda : \lambda) \) for every \( \lambda^{\prime} \). Therefore, we must have \( \lambda \geq \lambda^{\prime} \). This concludes the proof of Theorem 3.

2. Proof of Theorem 4

Let \( \Lambda \equiv \left\{ A_{x_1|x_0} \right\}_{x_0 \in x_0, x_1 \in x_1} \) be a PID. By Definition 7, the robustness of incompatibility of \( \Lambda \) equals
\[ \text{Rol}(\Lambda) := \min_{r \geq 0} \left\{ \frac{\Lambda_{x_1|x_0} + r Y_{x_1|x_0}}{1 + r} \right\}_{x_0, x_1} \quad \text{subject to:} \quad \left\{ \frac{\Lambda_{x_1|x_0} + r Y_{x_1|x_0}}{1 + r} \right\}_{x_0, x_1} \text{is a simple PID}, \]
\[ \left\{ Y_{x_1|x_0} \right\}_{x_0, x_1} \text{is a PID.} \]

Since the set of simple PIDs is closed and convex and has a nonzero volume in the space of PIDs, the optimal solution to Program (C24) exists. Let \( \tilde{\Omega} \equiv \left\{ \tilde{A}_{x_1|x_0} \right\}_{x_0 \in x_0, x_1 \in x_1} \) be an optimal solution to Program (C24), which is a PID. Define a PID
\[ \tilde{\Omega} \equiv \left\{ \tilde{A}_{x_1|x_0} \right\}_{x_0 \in x_0, x_1 \in x_1} \] as follows:
\[ \tilde{\Omega}_{x_1|x_0} := \frac{A_{x_1|x_0} + \text{Ral}(\Lambda) \tilde{A}_{x_1|x_0}}{1 + \text{Rol}(\Lambda)} \forall x_0, x_1.
\]
By Eq. (C24b), \( \tilde{\Omega} \) is simple. Let \( \mathcal{C} \) denote the cone generated by the set of unsteerable state assemblages in \( A_0 A_1 \):
\[ \mathcal{C} := \left\{ \omega_{x_1|x_0} : \omega_{x_1|x_0} + \sum_g p_{x_1|x_0,g} \eta_{g A_0 A_1}, \eta_{A_0 A_1} \geq 0, p_{x_1|x_0,g} \geq 0, \sum_{x_1} p_{x_1|x_0,g} = 1 \forall x_0, x_1, g \right\}.
\]
We note that the cone $\mathcal{C}$ is generating with respect to the space of state ensembles in $A_0A_1$. Let $\mathcal{C}^*$ denote the dual cone of $\mathcal{C}$:

$$
\mathcal{C}^* := \left\{ \{A_{x_1|x_0}\}_{x_0 \in \mathcal{X}_0, x_1 \in \mathcal{X}_1} : \sum_{x_0, x_1} \text{Tr} \left[ A_{x_1|x_0} A_{x_1|x_0} \right] \geq 0 \quad \forall \left\{ \omega_{x_1|x_0} \right\}_{x_0, x_1} \in \text{cone}(\mathcal{U} \{A_{x_1|x_0}\}) \right\}.
$$

(C27)

Substituting $\omega_{x_1|x_0}$ for $J^{A_0A_1}_{A_1|x_0} + rJ^{A_0A_1}_{A_1|x_0}$ Program (C24) can be reformulated as a conic program as follows:

$$
\text{Rol}(\lambda) = \frac{1}{dA_0} \min_{\omega_{x_1|x_0}} \left\{ \sum_{x_0, x_1} \text{Tr} \left[ \omega_{x_1|x_0} A_{x_1|x_0} \right] - 1 \right\}
$$

subject to:

$$
\omega_{x_1|x_0} - f_{A_0A_1} A_{x_1|x_0} \geq 0 \quad \forall x_0, x_1,
$$

$$
d_{A_0} \sum_{x_1} \text{Tr} \left[ \omega_{x_1|x_0} A_{x_1|x_0} \right] = \sum_{x_1} \text{Tr} \left[ \omega_{x_1|x_0} A_{x_1|x_0} \right] \leq A_0 \quad \forall x_0,
$$

$$
\left\{ \omega_{x_1|x_0} \right\}_{x_0, x_1} \in \mathcal{C}.
$$

(C28)

Invoking the theory of conic programming duality [60], since $\omega_{x_1|x_0} = 2d_{A_0} A_{x_1|x_0}$ for all $x_0, x_1$ is a strictly feasible solution to Program (C28), by Slater’s condition, strong duality holds, and the optimal solution to the dual program exists. The dual program of Program (C28) is given by

$$
\text{Rol}(\lambda) = \frac{1}{dA_0} \max_{\alpha_{x_1|x_0}} \left\{ \sum_{x_0, x_1} \text{Tr} \left[ \alpha_{x_1|x_0} A_{x_1|x_0} \right] - 1 \right\}
$$

subject to:

$$
\alpha_{x_1|x_0} \geq 0 \quad \forall x_0, x_1,
$$

$$
\sum_{x_0} \text{Tr} \left[ \beta_{x_0} \right] = d_{A_0} \left| \omega_{x_0} \right|,
$$

$$
\left\{ \beta_{x_0} \otimes A_{x_1|x_0} - \alpha_{x_1|x_0} \right\}_{x_0, x_1} \in \mathcal{C}^*.
$$

(C29)

Let $\{\tilde{\alpha}_{x_1|x_0}^{A_0A_1}\}_{x_0, x_1}$, $\{\tilde{\beta}_{x_1|x_0}^{A_0A_1}\}_{x_0, x_1}$ be an optimal solution to Program (C29). First, we prove an upper bound on the game advantage in terms of the robustness of incompatibility. Let $\mathcal{H} \equiv \{M_{m,n}^{C_0C_1}\}_{m \in \mathcal{M}, n \in \mathcal{N}}$ be a bipartite POVM. Let $\hat{\lambda} \equiv \{\lambda_{n|mi}^{C_0\rightarrow C_1}\}_{m \in \mathcal{M}, n \in \mathcal{N}}$ be an optimal solution to the maximization in Eq. (C2), namely, the PID simulated by $\Lambda$ under Alice’s optimal strategy. Let $\hat{\lambda}' \equiv \{\lambda_{n|mi}^{C_0\rightarrow C_1}\}_{m \in \mathcal{M}, n \in \mathcal{N}}$ and $\hat{\lambda}_{\text{simple}} \equiv \{\lambda_{n|mi}^{C_0\rightarrow C_1}\}_{m \in \mathcal{M}, n \in \mathcal{N}}$ denote the PIDs obtained by applying the same simulation strategy to $\hat{\lambda}$ and $\hat{\lambda}_0$, respectively. Then

$$
P_{\text{guess}}(\hat{\lambda}; \mathcal{H}) = \frac{1}{dC_0} \sum_{m \in \mathcal{M}, n \in \mathcal{N}} \text{Tr} \left[ M_{m,n}^{C_0C_1} \lambda_{n|mi}^{C_0\rightarrow C_1} \right]
$$

(C30)

$$
= \frac{1}{dC_0} \sum_{m \in \mathcal{M}, n \in \mathcal{N}} \text{Tr} \left[ M_{m,n}^{C_0C_1} \left( 1 + \text{Rol}(\lambda) \right) \lambda_{n|mi}^{C_0\rightarrow C_1} - \text{Rol}(\lambda) \lambda_{n|mi}^{C_0\rightarrow C_1} \right]
$$

(C31)

$$
\leq \frac{1 + \text{Rol}(\lambda)}{dC_0} \sum_{m \in \mathcal{M}, n \in \mathcal{N}} \text{Tr} \left[ M_{m,n}^{C_0C_1} \lambda_{n|mi}^{C_0\rightarrow C_1} \right]
$$

(C32)

$$
\leq \frac{1 + \text{Rol}(\lambda)}{dC_0} \max_{\Omega: \text{simple}} \sum_{m \in \mathcal{M}, n \in \mathcal{N}} \text{Tr} \left[ M_{m,n}^{C_0C_1} \Omega_{n|mi}^{C_0}\right]
$$

(C33)

$$
= \frac{1 + \text{Rol}(\lambda)}{dC_0} \lambda_{\text{simple}} \sum_{m \in \mathcal{M}, n \in \mathcal{N}} \text{Tr} \left[ M_{m,n}^{C_0C_1} \Omega_{n|mi}^{C_0} \right]
$$

(C34)

$$
= (1 + \text{Rol}(\lambda)) P_{\text{Guess}}(\cdot; \mathcal{H}) .
$$

(C35)

Here Eq. (C31) follows from Eq. (C25) and the linearity of free simulations; Eq. (C34) follows from $\hat{\lambda}'$ being simple and Theorem 1(1) and (2). This implies that

$$
\sup_{\mathcal{H}} \frac{P_{\text{guess}}(\hat{\lambda}; \mathcal{H})}{P_{\text{guess}}(\cdot; \mathcal{H})} \leq 1 + \text{Rol}(\lambda).
$$

(C36)
Next, we show that Ineq. (C36) can be equalized by an infinite sequence of POVMs on $A_0A_1$. Let $c := \| \sum_{x_0, x_1, x_i \in X_1} \alpha_{x_1|x_0}^{A_0A_1} \|_\infty$. Define a bipartite POVM $\tilde{M} = \{ \tilde{M}_{m,n} \}_{m \in X_0, n \in \mathbb{N}}$ such that $N \supset X_1$ as follows:

$$\tilde{M}_{m,n} := \left( \frac{1}{|N|} |m\rangle\langle m| \right) \cdot \sum_{x_0, x_1, x_i \in X_1} \alpha_{x_1|x_0}^{A_0A_1} - \sum_{x_0, x_1, x_i \in X_1} \tilde{M}_{x_0, x_1}^{A_0A_1} \right) \cdot \forall m \in X_0, n \in \mathbb{N} \setminus X_1. \tag{C37}$$

It can be verified that $\tilde{M}$ is a valid POVM, as $\tilde{M}_{m,n} \geq 0$ for all $m \in X_0, n \in \mathbb{N}$ and $\sum_{m \in X_0, n \in \mathbb{N}} \tilde{M}_{m,n}^{A_0A_1} = 1^{A_0A_1}$. Then

$$p_{\text{guess}}(\tilde{M}) = \frac{1}{d_{A_0} \cdot |\Omega|} \sum_{m \in X_0} |\Omega| \sum_{n \in \mathbb{N}} \text{Tr} \left( \tilde{M}_{m,n}^{A_0A_1} \cdot \Omega_{nm}^{A_0A_1} \right) \tag{C42}$$

Here Eqs. (C42) and (C48) follow from Eq. (C37); Ineq. (C50) follows from Eq. (C29a); Eq. (C52) follows from Eq. (C29d). Note that the PID in Ineq. (C47) has the form $\tilde{\Omega} = \{ \Omega^{A_0A_1}_{x_1|x_0} \}_{x_0, x_1, x_i \in X_1}$ in contrast to the PID $\Omega = \{ \Omega_{n|m}^{A_0A_1} \}_{m \in X_0, n \in \mathbb{N}}$ in Eq. (C46), and Ineq. (C47) follows from the fact that the optimal strategy in Eq. (C46) necessarily makes a guess $n$ within the set
of target outcomes $\gamma_1$. Then it follows from Eqs. (C41) and (C52) that

$$\sup_{\mathcal{M}} \frac{P_{\text{guess}}(\hat{\gamma}; \mathcal{M})}{P_{\text{simple guess}}(\cdot; \mathcal{M})} \geq \lim_{|b| \to \infty} \frac{P_{\text{guess}}(\hat{\gamma}; \mathcal{M})}{P_{\text{simple guess}}(\cdot; \mathcal{M})}$$

(C53)

$$= \lim_{|b| \to \infty} \frac{1 + \text{Rol}(\hat{\gamma})}{1 + \frac{c}{|\xi_0|(|b| - |\gamma_1|)}}$$

(C54)

$$= 1 + \text{Rol}(\hat{\gamma}).$$

(C55)

Combining Eqs. (C36) and (C55), we can conclude that

$$\sup_{\mathcal{M}} \frac{P_{\text{guess}}(\hat{\gamma}; \mathcal{M})}{P_{\text{simple guess}}(\cdot; \mathcal{M})} = 1 + \text{Rol}(\hat{\gamma}).$$

(C56)

This concludes the proof of Theorem 4.

3. Proof of Proposition 3

First, we prove the necessity of the convertibility conditions in Proposition 3. Let $\Lambda \equiv \{\Lambda^0_{x_1|x_0}\}_{x_0 \in X_0, x_1 \in X_1}$ and $\Gamma \equiv \{\Gamma^0_{y_1|y_0}\}_{y_0 \in Y_0, y_1 \in Y_1}$ be two PIDs such that $\Lambda \succeq \Gamma$. Let $\mathcal{L}' \equiv \{L^{B_i}_{l}\}_{l \in L}$ be a POVM, and let $\zeta \equiv \{\sigma^0_{m,n,l}\}_{m \in Y_0, n \in Y_1, l \in L}$ be a state ensemble. By Theorem 1(3), the preorder $\succeq$ is transitive, thus

$$P'_{\text{guess}}(\hat{\gamma}; \zeta, \mathcal{L}') = \max_{\Lambda' : \Gamma \succeq \Gamma'} \sum_{m, n, l} \text{Tr} \left[ L^{B_i}_{l} \Lambda'^{B_i \to B_i}_{n|m} \left[ \sigma^0_{m,n,l} \right] \right]$$

(C57)

$$\geq \max_{\Lambda' : \Gamma \succeq \Gamma'} \sum_{m, n, l} \text{Tr} \left[ L^{B_i}_{l} \Lambda'^{B_i \to B_i}_{n|m} \left[ \sigma^0_{m,n,l} \right] \right]$$

(C58)

$$= P'_{\text{guess}}(\Gamma; \zeta, \mathcal{L'}).$$

(C59)

Next, we prove the sufficiency of the convertibility conditions in Proposition 3 by contradiction. Let $\Lambda \equiv \{\Lambda^0_{x_1|x_0}\}_{x_0 \in X_0, x_1 \in X_1}$ be two PIDs, and let $\mathcal{L}' \equiv \{L^{B_i}_{l}\}_{l \in L}$ be an informationally complete POVM such that $P'_{\text{guess}}(\hat{\gamma}; \zeta, \mathcal{L}') \geq P'_{\text{guess}}(\hat{\gamma}; \zeta, \mathcal{L'})$ for every state ensemble $\zeta \equiv \{\sigma^0_{m,n,l}\}_{m \in Y_0, n \in Y_1, l \in L}$. We assume $\Lambda \succeq \Gamma$. Following the argument that leads to Eq. (C8), there exists a set of Hermiticity-preserving linear maps $\{O_{y_0,y_1}^{B_i}\}_{y_0 \in Y_0, y_1 \in Y_1}$ such that

$$\sum_{y_0, y_1} \text{Tr} \left[ O_{y_0,y_1}^{B_i} O_{y_0,y_1}^{B_i} \right] > \max_{\Lambda' : \Gamma \succeq \Gamma'} \sum_{y_0, y_1} \text{Tr} \left[ O_{y_0,y_1}^{B_i} O_{y_0,y_1}^{B_i} \right].$$

(C60)

Since $\mathcal{L}'$ is an informationally complete POVM, there exists a set of Hermitian operators $\{O_{y_0,y_1}^{B_i}\}_{y_0 \in Y_0, y_1 \in Y_1, l \in L}$ such that

$$O_{y_0,y_1}^{B_i} = \sum_{l'} \mu_{y_0,y_1,l'} \otimes L^{B_i}_{l'} \quad \forall y_0, y_1.$$ (C61)

Let $c := \max_{y_0 \in Y_0, y_1 \in Y_1} \|\mu^0_{y_0,y_1}\|_{\infty}$ and $c' := \sum_{m \in Y_0, n \in Y_1, l \in L} \text{Tr}[\mu^0_{m,n,l}]$. Define a state ensemble $\tilde{\zeta} \equiv \{\tilde{\sigma}^0_{m,n,l}\}_{m \in Y_0, n \in Y_1, l \in L}$ as follows:

$$\tilde{\sigma}^0_{m,n,l} := \frac{1}{c' + c d_{B_0} |Y_0||Y_1||L|} \left( (\mu^0_{m,n,l})_{B_0} + c' B_0 \right) \quad \forall m, n, l.$$ (C62)

It can be verified that $\tilde{\zeta}$ is a valid state ensemble, as $\tilde{\sigma}^0_{m,n,l} \geq 0$ for all $m \in Y_0, n \in Y_1, l \in L$ and $\sum_{m,n,l} \text{Tr}[\tilde{\sigma}^0_{m,n,l}] = 1$. Then

$$P'_{\text{guess}}(\hat{\gamma}; \tilde{\zeta}, \mathcal{L}') = \max_{\Lambda' : \Gamma \succeq \Gamma'} \sum_{m, n, l} \text{Tr} \left[ L^{B_i}_{l} \Gamma'^{B_i \to B_i}_{n|m} \left[ \tilde{\sigma}^0_{m,n,l} \right] \right]$$

(C63)

$$= \frac{1}{c' + c d_{B_0} |Y_0||Y_1||L|} \max_{\Lambda' : \Gamma \succeq \Gamma'} \left( \sum_{m, n, l} \text{Tr} \left[ L^{B_i}_{l} \Gamma'^{B_i \to B_i}_{n|m} \left[ (\mu^0_{m,n,l})_{B_0} \right] \right] + c \sum_{m, n, l} \text{Tr} \left[ L^{B_i}_{l} \Gamma'^{B_i \to B_i}_{n|m} \left[ B_0 \right] \right] \right)$$ (C64)
This contradicts the assumption that $P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\}$ follows Eq. (C62). It follows that

$$P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\) = \frac{1}{c' + c d B_0} |Y_0| |Y_1| |L| \left( \max_{B': A \nabla' B'} \sum_{y_0, y_1} \text{Tr} \left( \left( \mu_{B'_0, A'_0, y_1} \otimes L'_{B'} \right) \left( i d^{B'_0} \otimes \Lambda_{m,n,d}^{B'_0} \right) \left[ \phi_{+}^{B'_0} \right] + c d B_0 |Y_0| \right) \right) \quad (C65)$$

$$= \frac{1}{c' + c d B_0} |Y_0| |Y_1| |L| \left( \max_{B': A \nabla' B'} \sum_{y_0, y_1} \text{Tr} \left( f_{B'_0} B'_1 f_{B'_0} B'_1 \right) + c d B_0 |Y_0| \right). \quad (C66)$$

Here Eq. (C64) follows from Eq. (C62). It follows that

$$P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\) = \frac{1}{c' + c d B_0} |Y_0| |Y_1| |L| \left( \max_{B': A \nabla' B'} \sum_{y_0, y_1} \text{Tr} \left( f_{B'_0} B'_1 f_{B'_0} B'_1 \right) + c d B_0 |Y_0| \right) \quad (C67)$$

$$= P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\). \quad (C70)$$

This contradicts the assumption that $P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\) ≥ $P'_{\text{guess}}(\{; \vec{\zeta}, \mathcal{L}\) for every $\vec{\zeta}$. Therefore, we have $\Lambda \nabla 1 \Gamma$. This concludes the proof of Proposition 3.
M. Piani and J. Watrous, Necessary and sufficient quantum information characterization of Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. 114, 060404 (2015).

E. Haapasalo, Robustness of incompatibility for quantum devices, Journal of Physics A: Mathematical and Theoretical 48, 255303 (2015).

L. Guerini, M. T. Quintino, and L. Aolita, Distributed sampling, quantum communication witnesses, and measurement incompatibility, Phys. Rev. A 100, 042308 (2019).

S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New Journal of Physics 21, 113053 (2019).

A. van der Wolf, F. Buscemi, Degradable channels, less noisy channels, and quantum resources, Phys. Rev. Lett. 123, 012303 (2019).

E. Schrödinger, Discussion of probability relations between separated systems, Mathematical Proceedings of the Cambridge Philosophical Society 31, 555–563 (1935).

H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox, Phys. Rev. Lett. 98, 140402 (2007).

A. W. Harrow and M. A. Nielsen, Robustness of quantum gates in the presence of noise, Phys. Rev. A 68, 012308 (2003).

F. G. S. L. Brandão and G. Gour, Reversible framework for quantum resource theories, Phys. Rev. Lett. 115, 070503 (2015).

G. Gour and A. Winter, How to quantify a dynamical quantum resource, Phys. Rev. Lett. 123, 150401 (2019).

G. Gour and C. M. Scandolo, Dynamical resources (2020), arXiv:2101.01552.

S. Milz and K. Modi, Quantum stochastic processes and quantum non-markovian phenomena, PRX Quantum 2, 030201 (2021).

S. Milz, J. Bavaresco, and G. Chiribella, Resource theory of causal connection, Quantum 6, 788 (2022).

S. Khatri and M. M. Wilde, Principles of quantum communication theory: A modern approach (2020).