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*

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.

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

---

*kj264@cornell.edu

†echitamb@illinois.edu

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 $\{\Lambda_{x_1}\}_{x_1}$ such that $\sum_{x_1} \Lambda_{x_1}$ is trace preserving (TP). When performing an instrument on a quantum state ρ, a classical outcome x_1 is observed with probability $p_{x_1} = \text{Tr}[\Lambda_{x_1}[\rho]]$, and the postmeasurement state is then given by $\Lambda_{x_1}[\rho]/p_{x_1}$. Note that POVMs are a special type of instrument for which $\Lambda_{x_1}[\rho] = \text{Tr}[M_{x_1}\rho]$ for some collection of positive semidefinite operators $\{M_{x_1}\}_{x_1}$ with $\sum_{x_1} M_{x_1} = \mathbb{1}$. 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 $\{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ 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 monotones. In many cases, these games define a complete set of monotones 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 $\{\Lambda_{x_1|x_0}\}_{x_0,x_1}$. The classical program is the input value x_0, which dictates that the instrument
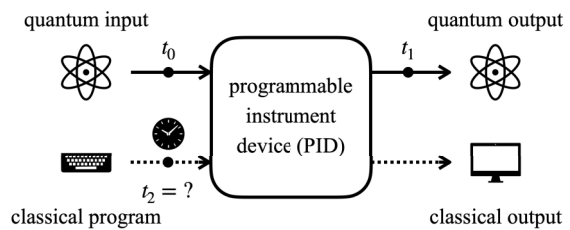

FIG. 1. A general controllable quantum device applies an instrument $\{\Lambda_{x_1|x_0}\}_{x_1}$ to the quantum input whenever a particular program x_0 is chosen. The characteristic time $\Delta t_D := t_1 - t_0$ is known as the *quantum delay time* of the device, and it measures how quickly the device 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."

$\{\Lambda_{x_1|x_0}\}_{x_1}$ 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 $\Delta t_D := t_1 - t_0$ 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 Δt_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

$$\sum_{x_1} \Lambda_{x_1|x_0} = \sum_{x_1} \Lambda_{x_1|x_0'} =: \Lambda \quad \forall x_0, x_0'. \tag{1}$$

In other words, all the instruments in the family $\{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ sum up to the same channel Λ. 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)*.

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 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 the 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
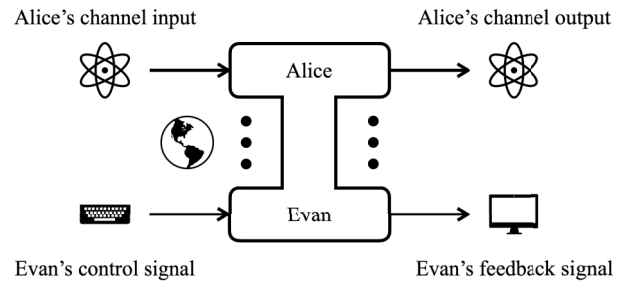

FIG. 2. The control device could also be split between spatially separated parties Alice and Evan. The nonsignaling constraint from Evan's control signal to Alice's channel output naturally arises if the spatial separation is so large that Evan's signal propagation time exceeds the anticipated quantum delay time between Alice's local input and output.

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-non-increasing 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 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 $\mathbb{M}$ is mathematically represented by a collection of positive semidefinite operators $\mathbb{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ such that $\sum_{x_1} M_{x_1|x_0} = \mathbb{1}$ 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 $\mathbb{M}$ 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:

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

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 PMD 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 $\mathbb{A}$ is mathematically represented by a collection of completely positive (CP) maps $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ such that $\sum_{x_1} \Lambda_{x_1|x_0}$ is trace preserving (TP) for all x_0. To make the quantum input and output systems A_0 and A_1 explicit, we sometimes write $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0,x_1}$.

Multi-instruments 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
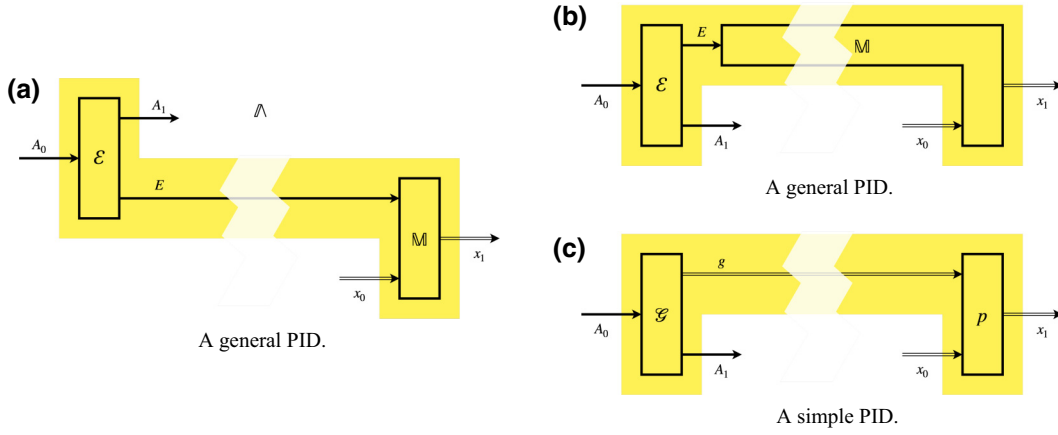
FIG. 3. Decomposition of a general PID [(a) and (b)] and that of a simple PID [(c)]. Solid arrows stand for quantum systems and hollow arrows for classical systems. Time flows from left to right. The opaque rips indicate that the quantum (left) and classical (right) parts of the devices are temporally separated under the late-program assumption. (a) A general PID $\mathbb{A}$ can be realized by connecting one output system E of a broadcast channel $\mathcal{E}^{A_0 \to A_1 E}$ to a PMD $\mathbb{M}$ using a quantum memory channel id^E. As such, the inner working of the PID can be understood as a process of channel steering (see Sec. IV A). (b) The general PID can be represented in a different configuration, with the system A_1 displaced downwards and the quantum memory channel subsumed within the PMD $\mathbb{M}$. (c) A simple PID can be realized with a "mother" instrument $\mathcal{G}$ and a classical channel p. In "steering" terms, the simple PID implements an unsteerable channel assemblage. Note that the simple PID has the same quantum delay time $\Delta t_{\mathrm{D}} \approx 0$ between A_0 and A_1 as the general PID does. However, the quantum memory of the general PID needs a much longer lifetime $\Delta t_{\mathrm{QM}} \gg \Delta t_{\mathrm{D}}$ in general to accept a late arriving classical program x_0, unlike the simple PID.

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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0, x_1}$ is called a *programmable instrument device* whenever it is nonsignaling from the classical input to the quantum output; namely, there exists a quantum channel Λ such that

$$\sum_{x_1} \Lambda_{x_1|x_0} = \Lambda \quad \forall x_0. \tag{3}$$

We say that the PID $\mathbb{A}$ is *simple* (and otherwise *nonsimple*) whenever there exists a quantum instrument $\mathcal{G} \equiv \{\mathcal{G}_g\}_g$ and a classical channel $p \equiv \{p_{x_1|x_0,g}\}_{x_0, x_1, g}$ such that

$$\Lambda_{x_1|x_0} = \sum_g p_{x_1|x_0,g} \mathcal{G}_g \quad \forall x_0, x_1. \tag{4}$$

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 $\Delta t_{\mathrm{QM}} \gg \Delta t_{\mathrm{D}} \approx 0$ 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) condition 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. 3(a). 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. 3(c). 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.
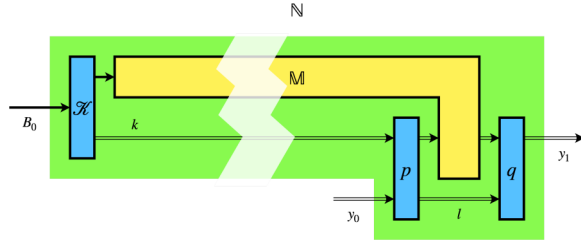
FIG. 4. The free simulation (blue) of a PMD $\mathbb{N}$ (green) using another PMD $\mathbb{M}$ (yellow) according to Eq. (5). It has been shown that nonsimple PMDs cannot be freely simulated by simple PMDs [20]. Simulations in this form are precisely those that can be realized without additional quantum memories under the late-program assumption.

FIG. 5. The free simulation (blue) of a PID $\mathbb{\Gamma}$ (green) using another PID $\mathbb{\Lambda}$ (yellow) according to Eq. (6). The simulation is composed of (i) pre-, post-, and side processing of the quantum part, (ii) pre-, post-, and side processing of the classical part, and (iii) an external classical memory connecting the quantum and classical parts. Under the late-program assumption, free simulations of PIDs represent the most general transformations given that using any quantum memory with a lifetime exceeding $\Delta t_D \approx 0$ is forbidden. This figure reduces to Fig. 4 when the quantum output systems A_1 and B_1 are trivial.

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 $\mathbb{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ and $\mathbb{N} \equiv \{N_{y_1|y_0}\}_{y_0,y_1}$ be two PMDs. We say that $\mathbb{M}$ can *freely simulate* $\mathbb{N}$, denoted by $\mathbb{M} \succcurlyeq_M \mathbb{N}$, whenever there exists a quantum instrument $\mathscr{K} \equiv \{\mathcal{K}_k\}_k$ 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

$$N_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}_k^\dagger \left[M_{x_1|x_0}\right] \quad \forall y_0, y_1, \quad (5)$$

where $(\cdot)^\dagger$ denotes adjunction. We call the transformation $\mathbb{M} \mapsto \mathbb{N}$ a *free simulation* of PMDs.

The operational significance of Definition 2 is demonstrated by the fact that $\mathbb{M} \succcurlyeq_M \mathbb{N}$ if and only if $\mathbb{M}$ can be physically transformed (i.e., via a quantum superchannel [49,50]) into $\mathbb{N}$ using no additional quantum memory [20], as represented in Fig. 4. Meanwhile, free simulations of PMDs have been shown to possess essential resource-theoretic properties, including preserving PMD simplicity and being able to generate the entire set of simple PMDs [20]. Thus they are formally qualified as the free transformations in a resource theory of PMD nonsimplicity, and the relation $\succcurlyeq_M$ is a legitimate incompatibility preorder on the set of PMDs.

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
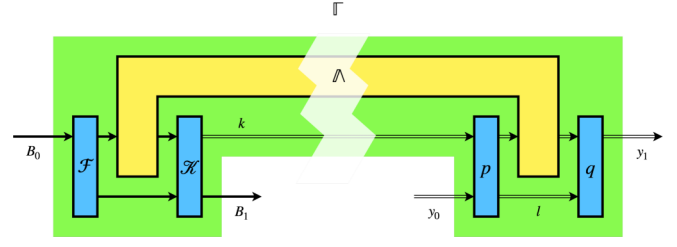
are semicausal quantum channels (i.e., quantum 2-combs), transformations between them are supposed to be quantum 4-combs [51].

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 Δt_D 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 Δt_{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.

Definition 3.—Let $\mathbb{\Lambda} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ and $\mathbb{\Gamma} \equiv \{\Gamma_{y_1|y_0}\}_{y_0,y_1}$ be two PIDs. We say that $\mathbb{\Lambda}$ can *freely simulate* $\mathbb{\Gamma}$, denoted by $\mathbb{\Lambda} \succcurlyeq_I \mathbb{\Gamma}$, whenever there exists a quantum channel $\mathcal{F}$, a quantum instrument $\mathscr{K} \equiv \{\mathcal{K}_k\}_k$, 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

$$\Gamma_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}_k \circ \left(\Lambda_{x_1|x_0} \otimes \mathrm{id}\right) \circ \mathcal{F}$$

$$\forall y_0, y_1. \quad (6)$$

We call the transformation $\mathbb{\Lambda} \mapsto \mathbb{\Gamma}$ 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 PMDs) when the quantum output systems of both $\mathbb{A}$ and Γ 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 $\succcurlyeq_I$ is an incompatibility 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 $\Delta t_D \approx 0$. This fact can be demonstrated by invoking the theory of quantum networks [51], combined with the observation that all input and output systems of $\mathbb{A}$ and Γ 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 $\mathcal{E}^{A_0 \to A_1 E}$ 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 $\Lambda^{A_0 \to A_1} := \mathrm{Tr}_E \circ \mathcal{E}^{A_0 \to A_1 E}$ by directing his system E to a PMD $\mathbb{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$. This steering process would lead to a family of instruments $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ on Alice's side, typically known as a *channel assemblage* [45], defined by

$$\Lambda_{x_1|x_0}^{A_0 \to A_1}[\cdot] := \mathrm{Tr}_E\left[\left(\mathbb{1}^{A_1} \otimes M_{x_1|x_0}^E\right) \mathcal{E}^{A_0 \to A_1 E}[\cdot]\right] \quad \forall x_0, x_1. \tag{7}$$

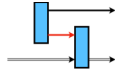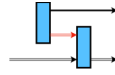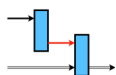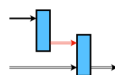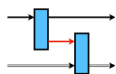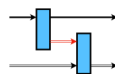The assemblage $\mathbb{A}$ is said to be *unsteerable* whenever it can be realized following Eq. (7) with $\mathcal{E}^{A_0 \to A_1 E}$ being an *incoherent extension* of its marginal channel $\Lambda^{A_0 \to A_1}$ [45], namely, whenever there exists a quantum instrument $\mathscr{G} \equiv \{\mathcal{G}_g\}_g$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

$$\mathcal{E}^{A_0 \to A_1 E} = \sum_g \mathcal{G}_g^{A_0 \to A_1} \otimes \epsilon_g^E. \tag{8}$$

It follows as an observation that if the PMD $\mathbb{M}$ is simple, then the channel assemblage $\mathbb{A}$ induced by Eq. (7) must be unsteerable regardless of $\mathcal{E}^{A_0 \to A_1 E}$ [45]. We introduce the following shorthand to denote a channel assemblage generated via steering.

Definition 4.—Let $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ be a channel assemblage, $\mathcal{E}^{A_0 \to A_1 E}$ a broadcast channel, and $\mathbb{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ a PMD. We say that $(\mathcal{E}, \mathbb{M})$ is a *steering decomposition* of $\mathbb{A}$, denoted by $\mathbb{A} \longleftarrow (\mathcal{E}, \mathbb{M})$, whenever Eq. (7) is satisfied, i.e., whenever $\mathbb{A}$ can be induced by a process of channel steering featuring $\mathcal{E}$ and $\mathbb{M}$.

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).

Resource theories of nonsimple programmability	Programmable devices (nonsignaling assemblages)	Simple devices (compatible assemblages)
PSD nonsimplicity (EPR steering [52])	$\{\rho_{x_1\|x_0}\}_{x_0,x_1}$: $\sum_{x_1}\rho_{x_1\|x_0} = \rho\ \forall x_0$	$\{\rho_{x_1\|x_0}\}_{x_0,x_1}$: $\rho_{x_1\|x_0} = \sum_g p_{x_1\|x_0,g}\eta_g\ \forall x_0,x_1$
PMD nonsimplicity [20] (measurement incompatibility [13])	$\{M_{x_1\|x_0}\}_{x_0,x_1}$: $\sum_{x_1} M_{x_1\|x_0} = \mathbb{1}\ \forall x_0$	$\{M_{x_1\|x_0}\}_{x_0,x_1}$: $M_{x_1\|x_0} = \sum_g p_{x_1\|x_0,g}G_g\ \forall x_0,x_1$
PID nonsimplicity (channel steering) [this paper]	$\{\Lambda_{x_1\|x_0}\}_{x_0,x_1}$: $\sum_{x_1}\Lambda_{x_1\|x_0} = \Lambda\ \forall x_0$	$\{\Lambda_{x_1\|x_0}\}_{x_0,x_1}$: $\Lambda_{x_1\|x_0} = \sum_g p_{x_1\|x_0,g}\mathcal{G}_g\ \forall x_0,x_1$

Clearly, each channel assemblage $\mathbb{\Lambda} \hookleftarrow (\mathcal{E},\mathbb{M})$ 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 semilocalizability [48], each PID $\mathbb{\Lambda}$ admits a steering decomposition $\mathbb{\Lambda} \hookleftarrow (\mathcal{E},\mathbb{M})$, and therefore we can envision a process of channel steering going on within each PID, as shown in Fig. 3(a). Then we can observe from Fig. 3(c) 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 $\xi^{A_1 E}$ into a PMD $\mathbb{M}$, a state assemblage $\{\rho_{x_1|x_0}\}_{x_0,x_1}$ is generated on Alice's side:

$$\rho^{A_1}_{x_1|x_0} := \mathrm{Tr}_E\left[\left(\mathbb{1}^{A_1} \otimes M^E_{x_1|x_0}\right)\xi^{A_1 E}\right] \quad \forall x_0,x_1. \quad (9)$$

Accordingly, the assemblage $\{\rho_{x_1|x_0}\}_{x_0,x_1}$ is *unsteerable* when it can be induced by a bipartite state $\xi^{A_1 E}$ being separable:

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

where $\{\eta_g\}_g$ is a state ensemble and ϵ_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]:

$$\rho^{A_1}_{x_1|x_0} = \sum_g p_{x_1|x_0,g}\eta_g^{A_1} \quad \forall x_0,x_1, \quad (11)$$

where $\{p_{x_1|x_0,g}\}_{x_0,x_1,g}$ is a classical channel. As in channel steering, if the PMD $\mathbb{M}$ is simple, then the induced state assemblage must be unsteerable. Conversely, it has also been proved that all nonsimple PMDs are useful to generate *steerable* state assemblages when $\xi^{A_1 E}$ 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 non-simplicity (i.e., channel steering) with PMD nonsimplicity (i.e., measurement incompatibility). Given a PID $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0,x_1}$, the nonsignaling constraint guarantees the existence of the following channel:

$$\Lambda^{A_0 \to A_1} := \sum_{x_1} \Lambda_{x_1|x_0}^{A_0 \to A_1} \quad \forall x_0. \tag{12}$$

The Choi operator of $\Lambda_{x_1|x_0}^{A_0 \to A_1}$ is given by

$$J_{\Lambda_{x_1|x_0}}^{A_0 A_1} := \left(\mathrm{id}^{A_0} \otimes \Lambda_{x_1|x_0}^{\widetilde{A}_0 \to A_1} \right) \left[\phi_+^{A_0 \widetilde{A}_0} \right] \quad \forall x_0, x_1, \tag{13}$$

where $\widetilde{A}_0$ is a system isomorphic to A_0 and $\phi_+^{A_0 \widetilde{A}_0} \equiv \sum_{i,j} |i\rangle\langle j|^{A_0} \otimes |i\rangle\langle j|^{\widetilde{A}_0}$. The Choi operator of $\Lambda^{A_0 \to A_1}$ is given by

$$J_\Lambda^{A_0 A_1} := \sum_{x_1} J_{\Lambda_{x_1|x_0}}^{A_0 A_1} \quad \forall x_0. \tag{14}$$

Let A^* be a quantum system such that its associated Hilbert space, denoted by $\mathbb{H}^{A^*}$, is isomorphic to the support of $J_\Lambda^{A_0 A_1}$. In other words, it satisfies $\mathbb{H}^{A^*} \cong \mathrm{supp}(J_\Lambda^{A_0 A_1}) \subseteq \mathbb{H}^{A_0 A_1}$, where $\mathbb{H}^{A_0 A_1} \equiv \mathbb{H}^{A_0} \otimes \mathbb{H}^{A_1}$ is the composite Hilbert space associated with the systems A_0 and A_1. By Choi's theorem, $J_{\Lambda_{x_1|x_0}}^{A_0 A_1}$ is positive semidefinite since $\Lambda_{x_1|x_0}^{A_0 \to A_1}$ is CP for all x_0, x_1. By Eq. (14), this implies $\mathrm{supp}(J_{\Lambda_{x_1|x_0}}^{A_0 A_1}) \subseteq \mathrm{supp}(J_\Lambda^{A_0 A_1}) \cong \mathbb{H}^{A^*}$, and therefore the image of $J_{\Lambda_{x_1|x_0}}^{A_0 A_1}$ in the system A^*, denoted by $J_{\Lambda_{x_1|x_0}}^{A^*}$, is well defined for all x_0, x_1. The following definition generalizes the steering-equivalent observes (SEO) defined on state assemblages to PIDs.

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

$$S_{x_1|x_0}^{A^*} := (J_\Lambda^{A^*})^{-\frac{1}{2}} J_{\Lambda_{x_1|x_0}}^{A^*} (J_\Lambda^{A^*})^{-\frac{1}{2}} \quad \forall x_0, x_1. \tag{15}$$

It can be conveniently verified that $\mathsf{SEM}(\mathbb{A})$ is a valid PMD given any PID $\mathbb{A}$. It is also apparent from Definition 5 that, by the Choi–Jamiołkowski isomorphism, a PID $\mathbb{A}$ is nonsimple if and only if the PMD $\mathsf{SEM}(\mathbb{A})$ 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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0,x_1}$ be a PID, and let $\mathbb{S} \equiv \{S_{x_1|x_0}^{A^*}\}_{x_0,x_1}$ be a PMD such that $\mathbb{S} = \mathsf{SEM}(\mathbb{A})$. Then there exists an isometric channel $\mathcal{V}^{A_0 \to A_1 A^*}$ such that

$$\Lambda_{x_1|x_0}^{A_0 \to A_1} [\cdot] = \mathrm{Tr}_{A^*} \left[\left(\mathbb{1}^{A_1} \otimes (S_{x_1|x_0}^\top)^{A^*} \right) \mathcal{V}^{A_0 \to A_1 A^*} [\cdot] \right]$$
$$\forall x_0, x_1, \tag{16}$$

where $(\cdot)^\top$ denotes transposition under a fixed orthonormal basis.

The proof of Proposition 1 is in Appendix B 1.

Proposition 1 indicates that given any PID $\mathbb{A}$, there exists a broadcast channel $\mathcal{V}^{A_0 \to A_1 A^*}$ such that the steering decomposition $\mathbb{A} \hookleftarrow (\mathcal{V}, \mathsf{SEM}(\mathbb{A})^\top)$ holds, where $\mathcal{V}^{A_0 \to A_1 A^*}$ is an isometric dilation of $\Lambda^{A_0 \to A_1}$ and $\mathsf{SEM}(\mathbb{A})^\top$ is the elementwise transpose of the PMD $\mathsf{SEM}(\mathbb{A})$. In particular, this provides a formal and independent demonstration for the internal structure of a general PID as depicted in Fig. 3(a), which we previously argued by invoking the equivalence between semicausality and semilocalizability.

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 $\mathbb{A}$ be a PID, $\mathcal{E}^{A_0 \to A_1 E}$ a broadcast channel, and $\mathbb{M} \equiv \{M_{x_1|x_0}\}_{x_0,x_1}$ a PMD such that $\mathbb{A} \hookleftarrow (\mathcal{E}, \mathbb{M})$. Then we have $\mathbb{M}^\top \succcurlyeq_\mathrm{M} \mathsf{SEM}(\mathbb{A})$, where $\mathbb{M}^\top \equiv \{M_{x_1|x_0}^\top\}_{x_0,x_1}$.

The proof of Proposition 2 is in Appendix B 2.

Proposition 2 indicates that $\mathsf{SEM}(\mathbb{A})^\top$ is the least resourceful PMD that can be used to compose a given PID $\mathbb{A}$, because every other PMD $\mathbb{M}$ that suffices to do

so must be convertible to $\mathsf{SEM}(\mathbb{A})^\top$ via free simulations. This is the reason why we refer to the decomposition $\mathbb{A} \leftarrowtail (\mathcal{V}, \mathsf{SEM}(\mathbb{A})^\top)$ in Proposition 1 as the "canonical" or "minimal" decomposition for $\mathbb{A}$. To understand this physically, we can think of any PID decomposition $\mathbb{A} \leftarrowtail (\mathcal{E}, \mathbb{M})$ as having a configuration as depicted in Fig. 3(b), and the nonsimplicity of $\mathbb{A}$ is essentially attributed to the PMD $\mathbb{M}$, since that is where the quantum memory with a lifetime $\Delta t_{\mathrm{QM}} \gg \Delta t_{\mathrm{D}} \approx 0$ resides. Then Proposition 2 implies that when the decomposition $\mathbb{A} \leftarrowtail (\mathcal{E}, \mathbb{M})$ is canonical, namely, when $\mathcal{E}^{A_0 \to A_1 E} = \mathcal{V}^{A_0 \to A_1 A^*}$ is the minimal isometric dilation (i.e., the maximally coherent channel extension [45]), the physical resource within $\mathbb{M}$ 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 $\mathsf{SEM}(\mathbb{A})$ is not only an indicator, but also a *measure* of the nonsimplicity of $\mathbb{A}$.

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

(1) Faithfulness: $\mathbb{A}$ is a simple PID if and only if $\mathsf{SEM}(\mathbb{A})$ is a simple PMD.
(2) Monotonicity: if $\mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{\Gamma}$, then $\mathsf{SEM}(\mathbb{A}) \succcurlyeq_{\mathrm{M}} \mathsf{SEM}(\mathbb{\Gamma})$.

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 $f \circ \mathsf{SEM}$ 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. 3(b)].

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 $\mathscr{M} \equiv \{M_{m,n}\}_{m,n}$, and it has two stages separated by a time interval $\Delta t \gg \Delta t_{\mathrm{D}} \approx 0$.

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

(2) In the second stage (after Δt 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 $\Delta t_D \approx 0$.

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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ 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 $\mathbb{A}$ is no greater than the Δt_D 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 $\Delta t \gg \Delta t_D \approx 0$ after her quantum output is released. Therefore, the most general strategy for Alice to follow is to use her PID $\mathbb{A}$ to simulate whatever PID $\mathbb{A}'$ she can and to insert $\mathbb{A}'$ into the open slots of the game, as illustrated in Fig. 6. In addition, given that the quantum delay time of $\mathbb{A}$ is within Δt_D and that no quantum memory with a lifetime exceeding Δt_D is accessible (see Definition 6), we can rest assured that Alice's simulation of $\mathbb{A}'$ using $\mathbb{A}$ is a free simulation, i.e., $\mathbb{A} \succeq_I \mathbb{A}'$ (see Definition 3). As a result, Alice's maximum winning probability in the nontransient guessing game specified by the POVM $\mathscr{M}$ is given by

$$P_{\text{guess}}(\mathbb{A}; \mathscr{M}) := \max_{\mathbb{A}':\, \mathbb{A} \succeq_I \mathbb{A}'} \sum_{m,n} \text{Tr}\left[M_{m,n}\left(\text{id} \otimes \Lambda'_{n|m}\right)[\varphi_+]\right]. \tag{17}$$

By the transitivity of the incompatibility preorder $\succeq_I$ [i.e., the sequential composability of free simulations in
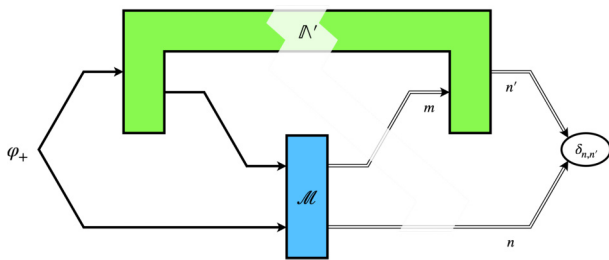

FIG. 6. A nontransient guessing game between Alice (the player) and Bob (the referee). The setting of the game is specified by Bob's POVM $\mathscr{M} \equiv \{M_{m,n}\}_{m,n}$. Alice's strategy to the game is represented by a PID $\mathbb{A}'$ (green), which is freely simulated by an actual PID $\mathbb{A}$ held in her hand. She wins the game whenever she makes a correct guess at one of Bob's outcome indices n given the other index m, i.e., whenever $n' = n$.

Theorem 1(3)], we can observe that Alice's winning probability $P_{\text{guess}}(\mathbb{A}; \mathscr{M})$ is a nonsimplicity monotone with respect to $\mathbb{A}$ given any POVM $\mathscr{M}$. The following theorem states that when considering all different POVMs, the winning probabilities $\{P_{\text{guess}}(\mathbb{A}; \mathscr{M})\}_{\mathscr{M}}$ compose a *complete* set of nonsimplicity monotones, which faithfully reflects convertibility between PIDs under free simulations.

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

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 $\mathscr{M}$, which equals

$$P_{\text{guess}}^{\text{simple}}(\mathscr{M}) := \max_{\Omega:\, \text{simple}} P_{\text{guess}}(\Omega; \mathscr{M}) \tag{18}$$

$$= \max_{\Omega:\, \text{simple}} \sum_{m,n} \text{Tr}\left[M_{m,n}\left(\text{id} \otimes \Omega_{n|m}\right)[\varphi_+]\right]. \tag{19}$$

Therefore, as long as Alice's winning probability in this game is observed to be greater than $P_{\text{guess}}^{\text{simple}}(\mathscr{M})$, we can conclude with certainty that Alice holds some device that functions as a nonsimple PID. Conversely, if Alice holds a nonsimple PID $\mathbb{A}$ 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 $\widehat{\mathscr{M}}$ such that $P_{\text{guess}}(\mathbb{A}; \widehat{\mathscr{M}}) > P_{\text{guess}}^{\text{simple}}(\widehat{\mathscr{M}})$.

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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$, denoted by $\mathsf{RoI}(\mathbb{A})$, is defined as

$$\mathsf{RoI}(\mathbb{A}) := \min_{r \geq 0} r \qquad (20a)$$

subject to: $\quad \left\{ \dfrac{\Lambda_{x_1|x_0} + r\Upsilon_{x_1|x_0}}{1+r} \right\}_{x_0,x_1} \quad$ is a simple PID, $\qquad (20b)$

$$\left\{\Upsilon_{x_1|x_0}\right\}_{x_0,x_1} \quad \text{is a PID.} \qquad (20c)$$

As we discussed before, Theorem 3 implies that every nonsimple PID can provide a nontrivial operational advantage over simple PIDs 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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ be a PID. Then

$$\sup_{\mathcal{M} \equiv \{M_{m,n}\}_{m,n}} \frac{P_{\text{guess}}(\mathbb{A}; \mathcal{M})}{P_{\text{guess}}^{\text{simple}}(\mathcal{M})} = 1 + \mathsf{RoI}(\mathbb{A}), \qquad (21)$$

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 $\mathbb{A}$ in any nontransient guessing game, in terms of the ratio $P_{\text{guess}}(\mathbb{A}; \mathcal{M})/P_{\text{guess}}^{\text{simple}}(\mathcal{M})$, can never exceed $1 + \mathsf{RoI}(\mathbb{A})$. 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 $\Delta t_{\text{D}} \approx 0$ 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 $\varsigma \equiv \{\sigma_{m,n,l}\}_{m,n,l}$ and a POVM $\mathcal{L} \equiv \{L_{l'}\}_{l'}$, and it has two stages separated by a time interval $\Delta t \gg \Delta t_{\text{D}} \approx 0$.

(1) In the first stage, Bob generates an index triple (m, n, l) with probability $\text{Tr}[\sigma_{m,n,l}]$ and sends the state $\sigma_{m,n,l}/\text{Tr}[\sigma_{m,n,l}]$ to Alice without announcing (m, n, l). 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 Δt 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, i.e., whenever $n' = n$ and $l' = l$.

Throughout the game, Alice has no access to quantum memories with a lifetime larger than $\Delta t_{\text{D}} \approx 0$.

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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}\}_{x_0,x_1}$ while playing the game, and that the quantum delay time of $\mathbb{A}$ is no greater than the Δt_{D} specified in Definition 8. Since Alice has no access to quantum memories with a lifetime exceeding Δt_{D}, her most general strategy is described by a PID $\mathbb{A}'$ such that $\mathbb{A}' \succcurlyeq_{\text{I}} \mathbb{A}$, as illustrated in Fig. 7. As a result, Alice's maximum winning probability in the postinformation guessing game specified by the state ensemble ς and the POVM $\mathcal{L}$ is given by

$$P'_{\text{guess}}(\mathbb{A}; \varsigma, \mathcal{L}) := \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\text{I}} \mathbb{A}'} \sum_{m,n,l} \text{Tr}\left[L_l \Lambda'_{n|m}\left[\sigma_{m,n,l}\right]\right]. \quad (22)$$

The following proposition states that for a certain POVM $\mathcal{L}$, the winning probabilities $\{P'_{\text{guess}}(\mathbb{A}; \varsigma, \mathcal{L})\}_\varsigma$
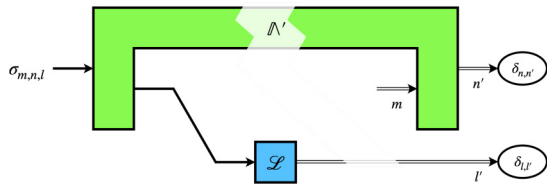
FIG. 7. A postinformation guessing game between Alice (the player) and Bob (the referee). The setting of the game is specified by Bob's state ensemble $\varsigma \equiv \{\sigma_{m,n,l}\}_{m,n,l}$ and POVM $\mathscr{L} \equiv \{L_{l'}\}_{l'}$. Alice's strategy to the game is represented by a PID $\mathbb{N}'$ (green), which is freely simulated by an actual PID $\mathbb{N}$ held in her hand. She wins the game whenever she makes a correct guess at one of Bob's indices n while not altering the index l given the index m, i.e., whenever $n' = n$ and $l' = l$.

compose a complete set of nonsimplicity monotones with respect to $\mathbb{N}$ when considering different sources ensembles.

Proposition 3.—Let $\mathbb{N} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathbb{X}_0, x_1 \in \mathbb{X}_1}$ and $\mathbb{\Gamma} \equiv \{\Gamma_{y_1|y_0}^{B_0 \to B_1}\}_{y_0 \in \mathbb{Y}_0, y_1 \in \mathbb{Y}_1}$ be two PIDs, and let $\mathscr{L} \equiv \{L_{l'}^{B_1}\}_{l' \in \mathbb{L}}$ be an informationally complete POVM. Then $\mathbb{N} \succeq_{\mathrm{I}} \mathbb{\Gamma}$ if and only if $P'_{\mathrm{guess}}(\mathbb{N}; \varsigma, \mathscr{L}) \geq P'_{\mathrm{guess}}(\mathbb{\Gamma}; \varsigma, \mathscr{L})$ for every state ensemble $\varsigma \equiv \{\sigma_{m,n,l}^{B_0}\}_{m \in \mathbb{Y}_0, n \in \mathbb{Y}_1, l \in \mathbb{L}}$.

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 3, 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 measurement 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 game parameters differ 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 Δt_{D}, which quantifies how quickly the device produces its quantum output, and (ii) the lifetime of the internal quantum memory Δt_{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 $\Delta t_{\mathrm{QM}}^{\mathrm{simple}} \leq \Delta t_{\mathrm{D}}$, whereas nonsimple PIDs must satisfy $\Delta t_{\mathrm{QM}} \gg \Delta t_{\mathrm{D}}$. Quantum memories are thus the resource that enables nonsimple programmability. To isolate the different memory demands between simple and nonsimple PIDs, we have assumed $\Delta t_{\mathrm{D}} \approx 0$ 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 (Propositions 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 monotonicity 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-nonbreaking 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 two 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 $\mathbb{N}'$ acts on one part of a bipartite POVM $\mathscr{M}$ (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.

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

$$\Lambda_{x_1|x_0}^{A_0\to A_1} = \sum_g p_{x_1|x_0,g}\mathcal{G}_g^{A_0\to A_1} \quad \forall x_0,x_1. \tag{A1}$$

Let $\Gamma \equiv \{\Gamma_{y_1|y_0}^{B_0\to B_1}\}_{y_0,y_1}$ be a PID such that $\Lambda \succcurlyeq_I \Gamma$. By Definition 3, there exists a quantum channel $\mathcal{F}'^{B_0\to A_0D}$, a quantum instrument $\{\mathcal{K}_k'^{A_1D\to B_1}\}_k$, 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

$$\Gamma_{y_1|y_0}^{B_0\to B_1} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l}'p_{x_0,l|y_0,k}'\mathcal{K}_k'^{A_1D\to B_1} \circ \left(\Lambda_{x_1|x_0}^{A_0\to A_1} \otimes \mathrm{id}^D\right) \circ \mathcal{F}'^{B_0\to A_0D} \tag{A2}$$

$$= \sum_{x_0,x_1,g,k,l} q_{y_1|x_1,l}'p_{x_1|x_0,g}p_{x_0,l|y_0,k}'\mathcal{K}_k'^{A_1D\to B_1} \circ \left(\mathcal{G}_g^{A_0\to A_1} \otimes \mathrm{id}^D\right) \circ \mathcal{F}'^{B_0\to A_0D} \quad \forall y_0,y_1. \tag{A3}$$

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

$$\mathcal{G}_{g,k}'^{B_0\to B_1} := \mathcal{K}_k'^{A_1D\to B_1} \circ \left(\mathcal{G}_g^{A_0\to A_1} \otimes \mathrm{id}^D\right) \circ \mathcal{F}'^{B_0\to A_0D} \quad \forall g,k, \tag{A4}$$

$$p_{y_1|y_0,g,k}'' := \sum_{x_0,x_1,l} q_{y_1|x_1,l}'p_{x_1|x_0,g}p_{x_0,l|y_0,k}' \quad \forall y_0,y_1,g,k. \tag{A5}$$

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

$$\sum_{g,k} p_{y_1|y_0,g,k}''\mathcal{G}_{g,k}'^{B_0\to B_1} = \sum_{x_0,x_1,g,k,l} q_{y_1|x_1,l}'p_{x_1|x_0,g}p_{x_0,l|y_0,k}'\mathcal{K}_k'^{A_1D\to B_1} \circ \left(\mathcal{G}_g^{A_0\to A_1} \otimes \mathrm{id}^D\right) \circ \mathcal{F}'^{B_0\to A_0D} \tag{A6}$$

$$= \Gamma_{y_1|y_0}^{B_0\to B_1} \quad \forall y_0,y_1. \tag{A7}$$

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

2. Proof of Theorem 1(2)

Let $\Lambda \equiv \{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0,x_1}$ be a PID, and let $\Gamma \equiv \{\Gamma_{y_1|y_0}^{B_0\to B_1}\}_{y_0,y_1}$ be a simple PID. By Definition 1, there exists a quantum instrument $\{\mathcal{G}_g^{B_0\to B_1}\}_g$ and a classical channel $\{p_{y_1|y_0,g}\}_{y_0,y_1,g}$ such that

$$\Gamma_{y_1|y_0}^{B_0\to B_1} = \sum_g p_{y_1|y_0,g}\mathcal{G}_g^{B_0\to B_1} \quad \forall y_0,y_1. \tag{A8}$$

Let δ 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 $\mathcal{F}'^{B_0\to A_0D}$, a quantum instrument $\{\mathcal{K}_g'^{A_1D\to B_1}\}_g$, and two classical channels $\{p_{x_0,l|y_0,g}'\}_{x_0,y_0,g,l}$ and $\{q_{y_1|x_1,l}'\}_{x_1,y_1,s}$ as follows:

$$\mathcal{F}'^{B_0\to A_0D} := |0\rangle\langle 0|^{A_0} \otimes \mathrm{id}^{B_0\to D}, \tag{A9}$$

$$\mathcal{K}_g'^{A_1D\to B_1} := \mathcal{G}_g^{D\to B_1} \circ \mathrm{Tr}_{A_1} \quad \forall g, \tag{A10}$$

$$p_{x_0,l|y_0,g}' := \delta_{x_0|0}p_{l|y_0,g} \quad \forall x_0,y_0,g,l, \tag{A11}$$

$$q_{y_1|x_1,l}' := \delta_{y_1|l} \quad \forall x_1,y_1,l. \tag{A12}$$

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

$$\sum_{x_0,x_1,g,l} q'_{y_1|x_1,l} p'_{x_0,l|y_0,g} \mathcal{K}'^{A_1 D \to B_1}_g \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}'^{B_0 \to A_0 D}$$

$$= \sum_{x_0,x_1,g,l} \delta_{y_1|l} \delta_{x_0|0} p_{l|y_0,g} \mathcal{G}^{D \to B_1}_g \circ \mathrm{Tr}_{A_1} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \left[|0\rangle\langle 0|^{A_0} \right] \otimes \mathrm{id}^{B_0 \to D} \right) \tag{A13}$$

$$= \sum_g p_{y_1|y_0,g} \mathcal{G}^{B_0 \to B_1}_g \tag{A14}$$

$$= \Gamma^{B_0 \to B_1}_{y_1|y_0} \quad \forall y_0, y_1. \tag{A15}$$

By Definition 3, this shows $\mathbb{\Lambda} \succeq_{\mathrm{I}} \Gamma$. This concludes the proof of Theorem 1(2).

3. Proof of Theorem 1(3)

First, we prove sequential composability. Consider a free simulation $\mathbb{\Lambda} \mapsto \Gamma$, where $\mathbb{\Lambda} \equiv \{\Lambda^{A_0 \to A_1}_{x_1|x_0}\}_{x_0,x_1}$ and $\Gamma \equiv \{\Gamma^{B_0 \to B_1}_{y_1|y_0}\}_{y_0,y_1}$ are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel $\mathcal{F}^{B_0 \to A_0 D}$, a quantum instrument $\{\mathcal{K}^{A_1 D \to B_1}_k\}_k$, 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

$$\Gamma^{B_0 \to B_1}_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}^{A_1 D \to B_1}_k \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \quad \forall y_0, y_1. \tag{A16}$$

Consider another free simulation $\Gamma \mapsto \Psi$, where $\Psi \equiv \{\Psi^{C_0 \to C_1}_{z_1|z_0}\}_{z_0,z_1}$ is a PID. By Definition 3, this free simulation can be represented by a quantum channel $\mathcal{F}'^{C_0 \to B_0 E}$, a quantum instrument $\{\mathcal{K}'^{B_1 E \to C_1}_{k'}\}_{k'}$, and two classical channels $\{p'_{y_0,l'|z_0,k'}\}_{y_0,z_0,k',l'}$ and $\{q'_{z_1|y_1,l'}\}_{y_1,z_1,l'}$ such that

$$\Psi^{C_0 \to C_1}_{z_1|z_0} = \sum_{y_0,y_1,k',l'} q'_{z_1|y_1,l'} p'_{y_0,l'|z_0,k'} \mathcal{K}'^{B_1 E \to C_1}_{k'} \circ \left(\Gamma^{B_0 \to B_1}_{y_1|y_0} \otimes \mathrm{id}^E \right) \circ \mathcal{F}'^{C_0 \to B_0 E} \quad \forall z_0, z_1. \tag{A17}$$

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

$$\Psi^{C_0 \to C_1}_{z_1|z_0} = \sum_{x_0,x_1,y_0,y_1,k,k',l,l'} q'_{z_1|y_1,l'} q_{y_1|x_1,l} p_{x_0,l|y_0,k} p'_{y_0,l'|z_0,k'} \mathcal{K}'^{B_1 E \to C_1}_{k'} \circ \left(\mathcal{K}^{A_1 D \to B_1}_k \otimes \mathrm{id}^E \right)$$

$$\circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{DE} \right) \circ \left(\mathcal{F}^{B_0 \to A_0 D} \otimes \mathrm{id}^E \right) \circ \mathcal{F}'^{C_0 \to B_0 E} \quad \forall z_0, z_1. \tag{A18}$$

Define a quantum channel $\mathcal{F}''^{C_0 \to A_0 DE}$, a quantum instrument $\{\mathcal{K}''^{A_1 DE \to C_1}_{k,k'}\}_{k,k'}$, and two classical channels $\{p''_{x_0,l,l'|z_0,k,k'}\}_{x_0,z_0,k,k',l,l'}$ and $\{q''_{z_1|x_1,l,l'}\}_{x_1,z_1,l,l'}$ as follows:

$$\mathcal{F}''^{C_0 \to A_0 DE} := \left(\mathcal{F}^{B_0 \to A_0 D} \otimes \mathrm{id}^E \right) \circ \mathcal{F}'^{C_0 \to B_0 E}, \tag{A19}$$

$$\mathcal{K}''^{A_1 DE \to C_1}_{k,k'} := \mathcal{K}'^{B_1 E \to C_1}_{k'} \circ \left(\mathcal{K}^{A_1 D \to B_1}_k \otimes \mathrm{id}^E \right) \quad \forall k, k', \tag{A20}$$

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

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

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

$$
\sum_{x_0,x_1,k,k',l,l'} q''_{z_1|x_1,l,l'} p''_{x_0,l,l'|z_0,k,k'} \mathcal{K}''^{A_1 DE \to C_1}_{k,k'} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{DE} \right) \circ \mathcal{F}''^{C_0 \to A_0 DE}
$$

$$
= \sum_{x_0,x_1,y_0,y_1,k,k',l,l'} q'_{z_1|y_1,l'} q_{y_1|x_1,l} p_{x_0,l|y_0,k} p'_{y_0,l'|z_0,k'} \mathcal{K}'^{B_1 E \to C_1}_{k'}
$$

$$
\circ \left(\mathcal{K}^{A_1 D \to B_1}_{k} \otimes \mathrm{id}^{E} \right) \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{DE} \right) \circ \left(\mathcal{F}^{B_0 \to A_0 D} \otimes \mathrm{id}^{E} \right) \circ \mathcal{F}'^{C_0 \to B_0 E} \tag{A23}
$$

$$
= \Psi^{C_0 \to C_1}_{z_1|z_0} \quad \forall z_0, z_1. \tag{A24}
$$

By Definition 3, this shows that the sequential composition $\mathbb{\Lambda} \mapsto \mathbb{\Psi}$ of the two free simulations is a free simulation.

Next, we prove parallel composability. Consider a free simulation $\mathbb{\Lambda} \mapsto \mathbb{\Gamma}$, where $\mathbb{\Lambda} \equiv \{\Lambda^{A_0 \to A_1}_{x_1|x_0}\}_{x_0,x_1}$ and $\mathbb{\Gamma} \equiv \{\Gamma^{B_0 \to B_1}_{y_1|y_0}\}_{y_0,y_1}$ are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel $\mathcal{F}^{B_0 \to A_0 D}$, a quantum instrument $\{\mathcal{K}^{A_1 D \to B_1}_{k}\}_k$, 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

$$
\Gamma^{B_0 \to B_1}_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}^{A_1 D \to B_1}_{k} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{D} \right) \circ \mathcal{F}^{B_0 \to A_0 D} \quad \forall y_0, y_1. \tag{A25}
$$

Consider another free simulation $\mathbb{\Lambda}' \mapsto \mathbb{\Gamma}'$, where $\mathbb{\Lambda}' \equiv \{\Lambda'^{A'_0 \to A'_1}_{x'_1|x'_0}\}_{x'_0,x'_1}$ and $\mathbb{\Gamma}' \equiv \{\Gamma'^{B'_0 \to B'_1}_{y'_1|y'_0}\}_{y'_0,y'_1}$ are two PIDs. By Definition 3, this free simulation can be represented by a quantum channel $\mathcal{F}'^{B'_0 \to A'_0 D'}$, a quantum instrument $\{\mathcal{K}'^{A'_1 D' \to B'_1}_{k'}\}_{k'}$, 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

$$
\Gamma'^{B'_0 \to B'_1}_{y'_1|y'_0} = \sum_{x'_0,x'_1,k',l'} q'_{y'_1|x'_1,l'} p'_{x'_0,l'|y'_0,k'} \mathcal{K}'^{A'_1 D' \to B'_1}_{k'} \circ \left(\Lambda'^{A'_0 \to A'_1}_{x'_1|x'_0} \otimes \mathrm{id}^{D'} \right) \circ \mathcal{F}'^{B'_0 \to A'_0 D'} \quad \forall y'_0, y'_1. \tag{A26}
$$

Combining Eqs. (A25) and (A26), the parallel composition of the above two free simulations can be described as $\mathbb{\Lambda}'' \mapsto \mathbb{\Gamma}''$, where $\mathbb{\Lambda}'' \equiv \{\Lambda''^{A_0 A'_0 \to A_1 A'_1}_{x_1,x'_1|x_0,x'_0}\}_{x_0,x'_0,x_1,x'_1}$ and $\mathbb{\Gamma}'' \equiv \{\Gamma''^{B_0 B'_0 \to B_1 B'_1}_{y_1,y'_1|y_0,y'_0}\}_{y_0,y'_0,y_1,y'_1}$ are two PIDs, such that

$$
\Gamma''^{B_0 B'_0 \to B_1 B'_1}_{y_1,y'_1|y_0,y'_0} = \sum_{x_0,x'_0,x_1,x'_1,k,k',l,l'} q_{y_1|x_1,l} q'_{y'_1|x'_1,l'} p_{x_0,l|y_0,k} p'_{x'_0,l'|y'_0,k'} \left(\mathcal{K}^{A_1 D \to B_1}_{k} \otimes \mathcal{K}'^{A'_1 D' \to B'_1}_{k'} \right)
$$

$$
\circ \left(\Lambda''^{A_0 A'_0 \to A_1 A'_1}_{x_1,x'_1|x_0,x'_0} \otimes \mathrm{id}^{DD'} \right) \circ \left(\mathcal{F}^{B_0 \to A_0 D} \otimes \mathcal{F}'^{B'_0 \to A'_0 D'} \right) \quad \forall y_0, y'_0, y_1, y'_1. \tag{A27}
$$

Define a quantum channel $\mathcal{F}''^{B_0 B'_0 \to A_0 D A'_0 D'}$, a quantum instrument $\{\mathcal{K}''^{A_1 A'_1 DD' \to B_1 B'_1}_{k,k'}\}_{k,k'}$, and two classical channels $\{p''_{x_0,x'_0,l,l'|y_0,y'_0,k,k'}\}_{x_0,x'_0,y_0,y'_0,k,k',l,l'}$ and $\{q''_{y_1,y'_1|x_1,x'_1,l,l'}\}_{x_1,x'_1,y_1,y'_1,l,l'}$ as follows:

$$
\mathcal{F}''^{B_0 B'_0 \to A_0 A'_0 DD'} := \mathcal{F}^{B_0 \to A_0 D} \otimes \mathcal{F}'^{B'_0 \to A'_0 D'}, \tag{A28}
$$

$$
\mathcal{K}''^{A_1 A'_1 DD' \to B_1 B'_1}_{k,k'} := \mathcal{K}^{A_1 D \to B_1}_{k} \otimes \mathcal{K}'^{A'_1 D' \to B'_1}_{k'} \quad \forall k, k', \tag{A29}
$$

$$
p''_{x_0,x'_0,l,l'|y_0,y'_0,k,k'} := p_{x_0,l|y_0,k} p'_{x'_0,l'|y'_0,k'} \quad \forall x_0, x'_0, y_0, y'_0, k, k', l, l', \tag{A30}
$$

$$
q''_{y_1,y'_1|x_1,x'_1,l,l'} := q_{y_1|x_1,l} q'_{y'_1|x'_1,l'} \quad \forall x_1, x'_1, y_1, y'_1, l, l'. \tag{A31}
$$

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

$$
\sum_{x_0,x_1,k,k',l,l'} q''_{y_1,y'_1|x_1,x'_1,l,l'} p''_{x_0,x'_0,l,l'|y_0,y'_0,k,k'} \mathcal{K}''^{A_1A'_1DD'\to B_1B'_1}_{k,k'} \circ \left(\Lambda''^{A_0A'_0\to A_1A'_1}_{x_1,x'_1|x_0,x'_0} \otimes \mathrm{id}^{DD'} \right) \circ \mathcal{F}''^{B_0B'_0\to A_0A'_0DD'}
$$

$$
= \sum_{x_0,x'_0,x_1,x'_1,k,k',l,l'} q_{y_1|x_1,l} q'_{y'_1|x'_1,l'} p_{x_0,l|y_0,k} p'_{x'_0,l'|y'_0,k'} \left(\mathcal{K}^{A_1D\to B_1}_k \otimes \mathcal{K}'^{A'_1D'\to B'_1}_{k'} \right) \tag{A32}
$$

$$
\circ \left(\Lambda''^{A_0A'_0\to A_1A'_1}_{x_1,x'_1|x_0,x'_0} \otimes \mathrm{id}^{DD'} \right) \circ \left(\mathcal{F}^{B_0\to A_0D} \otimes \mathcal{F}'^{B'_0\to A'_0D'} \right)
$$

$$
= \Gamma''^{B_0B'_0\to B_1B'_1}_{y_1,y'_1|y_0,y'_0} \quad \forall y_0,y'_0,y_1,y'_1. \tag{A33}
$$

By Definition 3, this shows that the parallel composition $\mathbb{A}'' \mapsto \Gamma''$ 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 $\mathbb{A} \mapsto \Gamma$, where $\mathbb{A} \equiv \{\Lambda^{A_0\to A_1}_{x_1|x_0}\}_{x_0\in\mathrm{X}_0,x_1\in\mathrm{X}_1}$ and $\Gamma \equiv \{\Gamma^{B_0\to B_1}_{y_1|y_0}\}_{y_0\in\mathrm{Y}_0,y_1\in\mathrm{Y}_1}$ are two PIDs. By Definition 3, the free simulation can be represented by a quantum channel $\mathcal{F}^{B_0\to A_0D}$, a quantum instrument $\{\mathcal{K}^{A_1D\to B_1}_k\}_k$, 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

$$
\Gamma^{B_0\to B_1}_{y_1|y_0} = \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0,l|y_0,k} \mathcal{K}^{A_1D\to B_1}_k \circ \left(\Lambda^{A_0\to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0\to A_0D} \quad \forall y_0,y_1. \tag{A34}
$$

By Ref. [50, Theorem 1(4)], the dimensionality of the system D can be bounded by the product of the dimensionalities 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 $\{q'_{l'|l}\}_{l,l'}$ such that

$$
q_{y_1|x_1,l} = \sum_{l'\in\mathrm{L}'} \delta_{y_1|l'(x_1)} q'_{l'|l} \quad \forall x_1,y_1,l, \tag{A35}
$$

where $\mathrm{L}' := \mathrm{Y}_1^{\mathrm{X}_1}$ is the finite set of all functions from X_1 to Y_1. Define a classical channel $\{p'_{x_0,l'|y_0,k}\}_{x_0,y_0,k,l'}$ as follows:

$$
p'_{x_0,l'|y_0,k} := \sum_l q'_{l'|l} p_{x_0|y_0,k} \quad \forall x_0,y_0,k,l'. \tag{A36}
$$

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

$$
p'_{x_0,l'|y_0,k} = \sum_{k'\in\mathrm{K}'} \delta_{x_0,l'|k'(y_0)} p''_{k'|k} \quad \forall x_0,y_0,k,l', \tag{A37}
$$

where $\mathrm{K}' := (\mathrm{X}_0 \otimes \mathrm{L}')^{\mathrm{Y}_0}$ is the finite set of all functions from Y_0 to $\mathrm{X}_0 \otimes \mathrm{L}'$. Define a quantum instrument $\{\mathcal{K}'^{A_1D\to B_1}_{k'}\}_{k'}$ and two classical channels $\{p'''_{x_0,l'|y_0,k'}\}_{x_0,y_0,k',l'}$ and $\{q''_{y_1|x_1,l'}\}_{x_1,y_1,l'}$ as follows:

$$
\mathcal{K}'^{A_1D\to B_1}_{k'} := \sum_k p''_{k'|k} \mathcal{K}^{A_1D\to B_1}_k \quad \forall k', \tag{A38}
$$

$$
p'''_{x_0,l'|y_0,k'} := \delta_{x_0,l'|k'(y_0)} \quad \forall x_0,y_0,k',l', \tag{A39}
$$

$$
q''_{y_1|x_1,l'} := \delta_{y_1|l'(x_1)} \quad \forall x_1,y_1,l'. \tag{A40}
$$

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

$$\sum_{x_0,x_1,k',l'} q''_{y_1|x_1,l'} p'''_{x_0,l'|y_0,k'} \mathcal{K}'^{A_1 D \to B_1}_{k'} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A41}$$

$$= \sum_{x_0,x_1,k',l'} \delta_{y_1|l'(x_1)} \delta_{x_0,l'|k'(y_0)} \mathcal{K}'^{A_1 D \to B_1}_{k'} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A42}$$

$$= \sum_{x_0,x_1,k,k',l'} \delta_{y_1|l'(x_1)} \delta_{x_0,l'|k'(y_0)} p''_{k'|k} \mathcal{K}^{A_1 D \to B_1}_{k} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A43}$$

$$= \sum_{x_0,x_1,k,l'} \delta_{y_1|l'(x_1)} p'_{x_0,l'|y_0,k} \mathcal{K}^{A_1 D \to B_1}_{k} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A44}$$

$$= \sum_{x_0,x_1,k,l,l'} \delta_{y_1|l'(x_1)} q'_{l'|l} p_{x_0|y_0,k} \mathcal{K}^{A_1 D \to B_1}_{k} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A45}$$

$$= \sum_{x_0,x_1,k,l} q_{y_1|x_1,l} p_{x_0|y_0,k} \mathcal{K}^{A_1 D \to B_1}_{k} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \tag{A46}$$

$$= \Gamma^{B_0 \to B_1}_{y_1|y_0} \quad \forall y_0, y_1. \tag{A47}$$

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 ith free simulation is applied with probability $p_i \geq 0$ such that $\sum_i p_i = 1$. By Definition 3, the ith free simulation can be represented by a quantum channel $\mathcal{F}'^{B_0 \to A_0 D}_{(i)}$, a quantum instrument $\{\mathcal{K}'^{A_1 D \to B_1}_{k|i}\}_k$, and two classical channels $\{p'_{x_0,l|y_0,k,i}\}_{x_0,y_0,k,l}$ and $\{q'_{y_1|x_1,l,i}\}_{x_1,y_1,l}$. The probabilistic mixture of these free simulations is then described by $\mathbb{A} \mapsto \Gamma$, where $\mathbb{A} \equiv \{\Lambda^{A_0 \to A_1}_{x_1|x_0}\}_{x_0,x_1}$ and $\Gamma \equiv \{\Gamma^{B_0 \to B_1}_{y_1|y_0}\}_{y_0,y_1}$ are two PIDs, such that

$$\Gamma^{B_0 \to B_1}_{y_1|y_0} = \sum_i p_i \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l,i} p'_{x_0,l|y_0,k,i} \mathcal{K}'^{A_1 D \to B_1}_{k|i} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}'^{B_0 \to A_0 D}_{(i)} \quad \forall y_0, y_1. \tag{A48}$$

Define a quantum channel $\mathcal{F}''^{B_0 \to A_0 DK}$, a quantum instrument $\{\mathcal{K}''^{A_1 DK \to B_1}_{k,k'}\}_{k,k'}$, and a classical channel $\{p''_{x_0,l,l'|y_0,k,k'}\}_{x_0,y_0,k,l,k',l'}$ as follows:

$$\mathcal{F}''^{B_0 \to A_0 DK} := \sum_i p_i \mathcal{F}'^{B_0 \to A_0 D}_{(i)} \otimes |i\rangle\langle i|^K, \tag{A49}$$

$$\mathcal{K}''^{A_1 DK \to B_1}_{k,k'} [\cdot] := \mathcal{K}'^{A_1 D \to B_1}_{k|k'} \left[\left(\mathbb{1}^{A_1 D} \otimes \langle k'|^K \right) [\cdot] \left(\mathbb{1}^{A_1 D} \otimes |k'\rangle^K \right) \right] \quad \forall k, k', \tag{A50}$$

$$p''_{x_0,l,l'|y_0,k,k'} := p'_{x_0,l|y_0,k,k'} \delta_{l'|k'} \quad \forall x_0, y_0, k, l, k', l'. \tag{A51}$$

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

$$\sum_{x_0,x_1,k,l,k',l'} q'_{y_1|x_1,l,l'} p''_{x_0,l,l'|y_0,k,k'} \mathcal{K}''^{A_1 DK \to B_1}_{k,k'} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{DK} \right) \circ \mathcal{F}''^{B_0 \to A_0 DK}$$

$$= \sum_{x_0,x_1,k,l,k',l'} q'_{y_1|x_1,l,l'} p'_{x_0,s|y_0,k,k'} \delta_{l'|k'} \mathcal{K}''^{A_1 DK \to B_1}_{k|k'} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^{DK} \right) \circ \left(\sum_i p_i \mathcal{F}'^{B_0 \to A_0 D}_{(i)} \otimes \langle k'|i\rangle\langle i|k'\rangle^K \right) \tag{A52}$$

$$= \sum_i p_i \sum_{x_0,x_1,k,l} q'_{y_1|x_1,l,i} p'_{x_0,l|y_0,k,i} \mathcal{K}'^{A_1 D \to B_1}_{k|i} \circ \left(\Lambda^{A_0 \to A_1}_{x_1|x_0} \otimes \mathrm{id}^D \right) \circ \mathcal{F}'^{B_0 \to A_0 D}_{(i)} \tag{A53}$$

$$= \Gamma^{B_0 \to B_1}_{y_1|y_0} \quad \forall y_0, y_1. \tag{A54}$$

By Definition 3, this shows that the probabilistic mixture $\mathbb{A} \mapsto \Psi$ 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 $\mathbb{A} \equiv \{\Lambda^{A_0 \to A_1}_{x_1|x_0}\}_{x_0,x_1}$ be a PID. Let $\Lambda^{A_0 \to A_1} := \sum_{x_1} \Lambda^{A_0 \to A_1}_{x_1|x_0}$ be the marginal channel of $\mathbb{A}$ from A_0 to A_1. Let $r :=$ rank$(J^{A_0 A_1}_\Lambda)$. The Choi operator $J^{A_0 A_1}_\Lambda$ has a spectral decomposition as follows:

$$J^{A_0 A_1}_\Lambda = \sum_{i=0}^{r-1} a_i |\alpha_i\rangle\langle\alpha_i|^{A_0 A_1}, \tag{B1}$$

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

$$\left(\mathbb{1}^{A_0} \otimes V^{\widetilde{A}_0 \to A_1 A^*}\right) |\phi_+\rangle^{A_0 \widetilde{A}_0} = \sum_{i=0}^{r-1} \sqrt{a_i} |\alpha_i\rangle^{A_0 A_1} |\overline{\alpha_i}\rangle^{A^*}. \tag{B2}$$

Define a linear map $\mathcal{V}^{A_0 \to A_1 A^*}$ as follows:

$$\mathcal{V}^{A_0 \to A_1 A^*}[\cdot] := V^{A_0 \to A_1 A^*}[\cdot](V^\dagger)^{A_1 A^* \to A_0}. \tag{B3}$$

We note that $\mathcal{V}^{A_0 \to A_1 A^*}$ is an isometric dilation of $\Lambda^{A_0 \to A_1}$, as can be verified by marginalizing its Choi operator $J^{A_0 A_1 A^*}_{\mathcal{V}}$:

$$\text{Tr}_{A^*}\left[J^{A_0 A_1 A^*}_{\mathcal{V}}\right] = \text{Tr}_{A^*}\left[\left(\mathbb{1}^{A_0} \otimes V^{\widetilde{A}_0 \to A_1 A^*}\right) \phi^{A_0 \widetilde{A}_0}_+ \left(\mathbb{1}^{A_0} \otimes (V^\dagger)^{A_1 A^* \to \widetilde{A}_0}\right)\right] \tag{B4}$$

$$= \text{Tr}_{A^*}\left[\sum_{i,j=0}^{r-1} \sqrt{a_i a_j} |\alpha_i\rangle\langle\alpha_j|^{A_0 A_1} \otimes |\overline{\alpha_i}\rangle\langle\overline{\alpha_j}|^{A^*}\right] \tag{B5}$$

$$= \sum_{i=0}^{r-1} a_i |\alpha_i\rangle\langle\alpha_i|^{A_0 A_1} \tag{B6}$$

$$= J^{A_0 A_1}_\Lambda. \tag{B7}$$

Let $\mathbb{S} \equiv \{S^{A^*}_{x_1|x_0}\}_{x_0,x_1}$ be a PMD such that $\mathbb{S} = \text{SEM}(\mathbb{A})$. By Definition 5,

$$S^{A^*}_{x_1|x_0} = (J^{A^*}_\Lambda)^{-\frac{1}{2}} J^{A^*}_{\Lambda_{x_1|x_0}} (J^{A^*}_\Lambda)^{-\frac{1}{2}} \quad \forall x_0, x_1. \tag{B8}$$

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

$$\mathcal{W}^{A^* \to A_0 A_1}[\cdot] := \sum_{i,j=0}^{r-1} \langle\alpha_i|[\cdot]|\alpha_j\rangle^{A^*} |\alpha_i\rangle\langle\alpha_j|^{A_0 A_1}. \tag{B9}$$

It follows that

$$\mathrm{Tr}_{A^*}\left[\left(\mathbb{1}^{A_0A_1}\otimes(S_{x_1|x_0}^\top)^{A^*}\right)J_\mathcal{V}^{A_0A_1A^*}\right]=\sum_{i,j=0}^{r-1}\sqrt{a_ia_j}\langle\overline{\alpha_j}|S_{x_1|x_0}^\top|\overline{\alpha_i}\rangle^{A^*}|\alpha_i\rangle\langle\alpha_j|^{A_0A_1} \tag{B10}$$

$$=\sum_{i,j=0}^{r-1}\sqrt{a_ia_j}\langle\alpha_i|S_{x_1|x_0}|\alpha_j\rangle^{A^*}|\alpha_i\rangle\langle\alpha_j|^{A_0A_1} \tag{B11}$$

$$=\left(\sum_{i=0}^{r-1}\sqrt{a_i}|\alpha_i\rangle^{A_0A_1}\langle\alpha_i|^{A^*}\right)S_{x_1|x_0}^{A^*}\left(\sum_{j=0}^{r-1}\sqrt{a_j}|\alpha_j\rangle^{A^*}\langle\alpha_j|^{A_0A_1}\right) \tag{B12}$$

$$=\mathcal{W}^{A^*\to A_0A_1}\left[(J_\Lambda^{A^*})^{\frac{1}{2}}S_{x_1|x_0}^{A^*}(J_\Lambda^{A^*})^{\frac{1}{2}}\right] \tag{B13}$$

$$=\mathcal{W}^{A^*\to A_0A_1}\left[J_{\Lambda_{x_1|x_0}}^{A^*}\right] \tag{B14}$$

$$=J_{\Lambda_{x_1|x_0}}^{A_0A_1}\quad\forall x_0,x_1. \tag{B15}$$

By the Choi–Jamiołkowski isomorphism, we can conclude that

$$\Lambda_{x_1|x_0}^{A_0\to A_1}[\cdot]=\mathrm{Tr}_{A^*}\left[\left(\mathbb{1}^{A_1}\otimes(S_{x_1|x_0}^\top)^{A^*}\right)\mathcal{V}^{A_0\to A_1A^*}[\cdot]\right]\quad\forall x_0,x_1. \tag{B16}$$

This concludes the proof of Proposition 1.

2. Proof of Proposition 2

Let $\mathbb{A}\equiv\{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0,x_1}$ be a PID. Let $\Lambda^{A_0\to A_1}:=\sum_{x_1}\Lambda_{x_1|x_0}^{A_0\to A_1}$ be the marginal channel of $\mathbb{A}$ from A_0 to A_1. Let $\mathcal{E}^{A_0\to A_1E}$ be a broadcast channel, and let $\mathbb{M}\equiv\{M_{x_1|x_0}^E\}_{x_0,x_1}$ be a PMD such that $\mathbb{A}\hookleftarrow(\mathcal{E},\mathbb{M})$. By Definition 4,

$$\Lambda_{x_1|x_0}^{A_0\to A_1}[\cdot]=\mathrm{Tr}_E\left[\left(\mathbb{1}^{A_1}\otimes M_{x_1|x_0}^E\right)\mathcal{E}^{A_0\to A_1E}[\cdot]\right]\quad\forall x_0,x_1. \tag{B17}$$

Let $\mathcal{V}^{A_0\to A_1EF}$ be an isometric dilation of $\mathcal{E}^{A_0\to A_1E}$. Let $V^{A_0\to A_1EF}$ be the isometry operator such that

$$\mathcal{V}^{A_0\to A_1EF}[\cdot]:=V^{A_0\to A_1EF}[\cdot](V^\dagger)^{A_1EF\to A_0}. \tag{B18}$$

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

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

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

$$\mathrm{Tr}_{EF}\left[\left(\mathbb{1}^{A_1}\otimes N_{x_1|x_0}^{EF}\right)\mathcal{V}^{A_0\to A_1EF}[\cdot]\right]=\mathrm{Tr}_E\left[\left(\mathbb{1}^{A_1}\otimes M_{x_1|x_0}^E\right)\mathrm{Tr}_F\circ\mathcal{V}^{A_0\to A_1EF}[\cdot]\right] \tag{B20}$$

$$=\mathrm{Tr}_E\left[\left(\mathbb{1}^{A_1}\otimes M_{x_1|x_0}^E\right)\mathcal{E}^{A_0\to A_1E}[\cdot]\right] \tag{B21}$$

$$=\Lambda_{x_1|x_0}^{A_0\to A_1}[\cdot]\quad\forall x_0,x_1. \tag{B22}$$

The vector $(\mathbb{1}^{A_0}\otimes V^{\tilde{A}_0\to A_1EF})|\phi_+\rangle^{A_0\tilde{A}_0}$ has a Schmidt decomposition as follows:

$$\left(\mathbb{1}^{A_0}\otimes V^{\tilde{A}_0\to A_1EF}\right)|\phi_+\rangle^{A_0\tilde{A}_0}=\sum_{i=0}^{r-1}\sqrt{a_i}|\alpha_i\rangle^{A_0A_1}|\beta_i\rangle^{EF}, \tag{B23}$$

where $a_i>0$ is a positive real number for all i and $\{|\alpha_i\rangle^{A_0A_1}\}_i$ and $\{|\beta_i\rangle^{EF}\}_i$ are two orthonormal sets of vectors. Let A^* be a system such that $\mathbb{H}^{A^*}\cong\mathrm{supp}(J_\Lambda^{A_0A_1})\subseteq\mathbb{H}^{A_0A_1}$. Let $|\alpha_i\rangle^{A^*}$ be the image of $|\alpha_i\rangle^{A_0A_1}$ in A^*. Then $\{|\alpha_i\rangle^{A^*}\}_i$ is an orthonormal

basis of $\mathbb{H}^{A^*}$. Let $|\overline{\beta_i}\rangle^{EF}$ be the complex conjugate of $|\beta_i\rangle^{EF}$ under a fixed orthonormal basis. Define an isometric channel $\mathcal{W}^{A^*\to EF}$ as follows:

$$\mathcal{W}^{A^*\to EF}[\cdot] = \sum_{i,j=0}^{r-1} \langle\alpha_i|[\cdot]|\alpha_j\rangle^{A^*}|\overline{\beta_i}\rangle\langle\overline{\beta_j}|^{EF}. \tag{B24}$$

Let $J_\mathcal{V}^{A^*EF}$ denote the image of the Choi operator $J_\mathcal{V}^{A_0A_1EF}$ of $\mathcal{V}^{A_0\to A_1EF}$ in the composite system A^*EF. Then

$$(J_\Lambda^{A^*})^{\frac{1}{2}}(\mathcal{W}^\dagger)^{EF\to A^*}\left[(N_{x_1|x_0}^\top)^{EF}\right](J_\Lambda^{A^*})^{\frac{1}{2}} = \sum_{i,j=0}^{r-1}\sqrt{a_ia_j}|\alpha_i\rangle\langle\alpha_i|^{A^*}(\mathcal{W}^\dagger)^{EF\to A^*}\left[(N_{x_1|x_0}^\top)^{EF}\right]|\alpha_j\rangle\langle\alpha_j|^{A^*} \tag{B25}$$

$$= \sum_{i,j=0}^{r-1}\sqrt{a_ia_j}\langle\overline{\beta_i}|N_{x_1|x_0}^\top|\overline{\beta_j}\rangle^{EF}|\alpha_i\rangle\langle\alpha_j|^{A^*} \tag{B26}$$

$$= \sum_{i,j=0}^{r-1}\sqrt{a_ia_j}\langle\beta_j|N_{x_1|x_0}|\beta_i\rangle^{EF}|\alpha_i\rangle\langle\alpha_j|^{A^*} \tag{B27}$$

$$= \mathrm{Tr}_{EF}\left[\left(\mathbb{1}^{A^*}\otimes(N_{x_1|x_0})^{EF}\right)\left(\sum_{i,j=0}^{r-1}\sqrt{a_ia_j}|\alpha_i\rangle\langle\alpha_j|^{A^*}\otimes|\beta_i\rangle\langle\beta_j|^{EF}\right)\right] \tag{B28}$$

$$= \mathrm{Tr}_{EF}\left[\left(\mathbb{1}^{A^*}\otimes N_{x_1|x_0}^{EF}\right)J_\mathcal{V}^{A^*EF}\right] \tag{B29}$$

$$= J_{\Lambda_{x_1|x_0}}^{A^*}\quad\forall x_0,x_1. \tag{B30}$$

Here Eq. (B29) follows from Eq. (B23) and the isomorphism between $\mathbb{H}^{A^*}$ and $\mathrm{supp}(J_\Lambda^{A_0A_1})$; Eq. (B30) follows from Eq. (B22) and the same isomorphism. It follows that

$$(J_\Lambda^{A^*})^{-\frac{1}{2}}J_{\Lambda_{x_1|x_0}}^{A^*}(J_\Lambda^{A^*})^{-\frac{1}{2}} = (\mathcal{W}^\dagger)^{EF\to A^*}\left[(N_{x_1|x_0}^\top)^{EF}\right]\quad\forall x_0,x_1. \tag{B31}$$

By Definitions 2 and 5, this implies $\mathbb{N}^\top \succcurlyeq_\mathrm{M} \mathsf{SEM}(\mathbb{\Lambda})$. Since Eq. (B19) implies $\mathbb{M}^\top \succcurlyeq_\mathrm{M} \mathbb{N}^\top$, by the transitivity of the preorder $\succcurlyeq_\mathrm{M}$ [20], we have $\mathbb{M}^\top \succcurlyeq_\mathrm{M} \mathsf{SEM}(\mathbb{\Lambda})$. This concludes the proof of Proposition 2.

3. Proof of Theorem 2(1)

By the Choi–Jamiołkowski isomorphism, a PID $\mathbb{\Lambda}\equiv\{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0,x_1}$ is simple if and only if its corresponding assemblage of Choi states $\{J_{\Lambda_{x_1|x_0}}^{A_0A_1}/d_{A_0}\}_{x_0,x_1}$ is an unsteerable state assemblage [45]. The state assemblage $\{J_{\Lambda_{x_1|x_0}}^{A_0A_1}/d_{A_0}\}_{x_0,x_1}$ is unsteerable if and only if its steering-equivalent observables are compatible [58, Theorem 1]. By Definition 5 and Ref. [58, Eq. (5)], $\mathsf{SEM}(\mathbb{\Lambda})$ is the steering-equivalent observables of the state assemblage $\{J_{\Lambda_{x_1|x_0}}^{A_0A_1}/d_{A_0}\}_{x_0,x_1}$. Therefore, $\mathbb{\Lambda}$ is simple if and only if $\mathsf{SEM}(\mathbb{\Lambda})$ is simple. This concludes the proof of Theorem 2(1).

4. Proof of Theorem 2(2)

Let $\mathbb{\Lambda}\equiv\{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0,x_1}$ and $\mathbb{\Gamma}\equiv\{\Gamma_{y_1|y_0}^{B_0\to B_1}\}_{y_0,y_1}$ be two PIDs such that $\mathbb{\Lambda}\succcurlyeq_\mathrm{I}\mathbb{\Gamma}$. By Definition 3, there exists a quantum channel $\mathcal{F}^{B_0\to A_0D}$, a quantum instrument $\{\mathcal{K}_k^{A_1D\to B_1}\}_k$, 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

$$\Gamma_{y_1|y_0}^{B_0\to B_1} = \sum_{x_0,x_1,k,l}q_{y_1|x_1,l}p_{x_0,l|y_0,k}\mathcal{K}_k^{A_1D\to B_1}\circ\left(\Lambda_{x_1|x_0}^{A_0\to A_1}\otimes\mathrm{id}^C\right)\circ\mathcal{F}^{B_0\to A_0D}\quad\forall y_0,y_1. \tag{B32}$$

Let $\Lambda^{A_0\to A_1}:=\sum_{x_1}\Lambda_{x_1|x_0}^{A_0\to A_1}$ be the marginal channel of $\mathbb{\Lambda}$ from A_0 to A_1. Let A^* be a system such that $\mathbb{H}^{A^*}\cong\mathrm{supp}(J_\Lambda^{A_0A_1})\subseteq\mathbb{H}^{A_0A_1}$. Let $\mathbb{S}\equiv\{S_{x_1|x_0}^{A^*}\}$ be a PMD such that $\mathbb{S}=\mathsf{SEM}(\mathbb{\Lambda})$. By Proposition 1, there exists an isometric

channel $\mathcal{V}^{A_0 \to A_1 A^*}$ such that

$$\Lambda_{x_1|x_0}^{A_0 \to A_1} [\cdot] = \mathrm{Tr}_{A^*} \left[\left(\mathbb{1}^{A_1} \otimes (S_{x_1|x_0}^\top)^{A^*} \right) \mathcal{V}^{A_0 \to A_1 A^*} [\cdot] \right] \quad \forall x_0, x_1. \tag{B33}$$

Define a quantum channel $\mathcal{E}^{B_0 \to B_1 A^* K}$ as follows:

$$\mathcal{E}^{B_0 \to B_1 A^* K} = \sum_k \left(\mathcal{K}_k^{A_1 D \to B_1} \otimes \mathrm{id}^{A^*} \right) \circ \left(\mathcal{V}^{A_0 \to A_1 A^*} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} \otimes |k\rangle\langle k|^K. \tag{B34}$$

Define a quantum instrument $\mathscr{K}' \equiv \{\mathcal{K}_{k'}^{\prime A^* K \to A^*}\}_{k'}$ as follows:

$$\mathcal{K}_{k'}^{\prime A^* K \to A^*} [\cdot] = \left(\mathbb{1}^{A^*} \otimes \langle k'|^K \right) [\cdot] \left(\mathbb{1}^{A^*} \otimes |k'\rangle^K \right) \quad \forall k'. \tag{B35}$$

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

$$N_{y_1|y_0}^{A^* K} := \sum_{x_0, x_1, k', l} q_{y_1|x_1, l} p_{x_0, l|y_0, k'} (\mathcal{K}_{k'}^{\prime\dagger})^{A^* \to A^* K} \left[S_{x_1|x_0}^{A^*} \right] \tag{B36}$$

$$= \sum_{x_0, x_1, k', l} q_{y_1|x_1, l} p_{x_0, l|y_0, k'} S_{x_1|x_0}^{A^*} \otimes |k'\rangle\langle k'|^K \quad \forall y_0, y_1. \tag{B37}$$

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

$$\mathrm{Tr}_{A^* K} \left[\left(\mathbb{1}^{B_1} \otimes (N_{y_1|y_0}^\top)^{A^* K} \right) \mathcal{E}^{B_0 \to B_1 A^* K} [\cdot] \right]$$

$$= \sum_{x_0, x_1, k, l, k'} q_{y_1|x_1, l} p_{x_0, l|y_0, k'} \mathrm{Tr}_{A^*} \left[\left(\mathbb{1}^{B_1} \otimes (S_{x_1|x_0}^\top)^{A^*} \right) \left(\mathcal{K}_k^{A_1 D \to B_1} \otimes \mathrm{id}^{A^*} \right) \right.$$

$$\left. \circ \left(\mathcal{V}^{A_0 \to A_1 A^*} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} [\cdot] \otimes \langle k'|k\rangle\langle k|k'\rangle^K \right] \tag{B38}$$

$$= \sum_{x_0, x_1, k, l} q_{y_1|x_1, l} p_{x_0, l|y_0, k} \mathcal{K}_k^{A_1 D \to B_1} \circ \mathrm{Tr}_{A^*} \left[\left(\mathbb{1}^{A_1 D} \otimes (S_{x_1|x_0}^\top)^{A^*} \right) \left(\mathcal{V}^{A_0 \to A_1 A^*} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} [\cdot] \right] \tag{B39}$$

$$= \sum_{x_0, x_1, k, l} q_{y_1|x_1, l} p_{x_0, l|y_0, k} \mathcal{K}_k^{A_1 D \to B_1} \circ \left(\Lambda_{x_1|x_0}^{A_0 \to A_1} \otimes \mathrm{id}^D \right) \circ \mathcal{F}^{B_0 \to A_0 D} [\cdot] \tag{B40}$$

$$= \Gamma_{y_1|y_0}^{B_0 \to B_1} [\cdot] \quad \forall y_0, y_1. \tag{B41}$$

By Definition 4, this implies $\mathbb{\Gamma} \hookleftarrow (\mathcal{E}, \mathbb{N}^\top)$, where $\mathbb{N}^\top \equiv \{(N_{y_1|y_0}^\top)^{A^* K}\}_{y_0, y_1}$. Then by Proposition 2, we have $\mathbb{N} \succeq_M$ SEM($\mathbb{\Gamma}$). Since SEM($\mathbb{\Lambda}$) = $\mathbb{S} \succeq_M \mathbb{N}$, by the transitivity of the preorder $\succeq_M$ [20], we have SEM($\mathbb{\Lambda}$) $\succeq_M$ SEM($\mathbb{\Gamma}$). 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 $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathrm{X}_0, x_1 \in \mathrm{X}_1}$ be a PID. Let $\mathscr{M} \equiv \{M_{m,n}^{C_0 C_1}\}_{m \in \mathrm{M}, n \in \mathrm{N}}$ be a bipartite POVM. By Eq. (17), Alice's maximum winning probability in the nontransient guessing game specified by $\mathscr{M}$ equals

$$P_{\mathrm{guess}}(\mathbb{A}; \mathscr{M}) := \frac{1}{d_{C_0}} \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \sum_{m \in \mathrm{M}, n \in \mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} \left(\mathrm{id}^{C_0} \otimes \Lambda_{n|m}^{\prime \widetilde{C}_0 \to C_1} \right) \left[\phi_+^{C_0 \widetilde{C}_0} \right] \right] \tag{C1}$$

$$= \frac{1}{d_{C_0}} \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \sum_{m \in \mathrm{M}, n \in \mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\Lambda_{n|m}'}^{C_0 C_1} \right]. \tag{C2}$$

First, we prove the necessity of the convertibility conditions in Theorem 3. Let $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathrm{X}_0, x_1 \in \mathrm{X}_1}$ and $\Gamma \equiv \{\Gamma_{y_1|y_0}^{B_0 \to B_1}\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1}$ be two PIDs such that $\mathbb{A} \succcurlyeq_{\mathrm{I}} \Gamma$. Let $\mathscr{M} \equiv \{M_{m,n}^{C_0 C_1}\}_{m \in \mathrm{M}, n \in \mathrm{N}}$ be a bipartite POVM. By Theorem 1(3), the preorder $\succcurlyeq_{\mathrm{I}}$ is transitive, thus

$$P_{\mathrm{guess}}(\mathbb{A}; \mathscr{M}) = \frac{1}{d_{C_0}} \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \sum_{m \in \mathrm{M}, n \in \mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\Lambda_{n|m}'}^{C_0 C_1} \right] \tag{C3}$$

$$\geq \frac{1}{d_{C_0}} \max_{\mathbb{A}': \Gamma \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \sum_{m \in \mathrm{M}, n \in \mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\Lambda_{n|m}'}^{C_0 C_1} \right] \tag{C4}$$

$$= P_{\mathrm{guess}}(\Gamma; \mathscr{M}). \tag{C5}$$

Next, we prove the sufficiency of the convertibility conditions in Theorem 3 by contradiction. Let $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathrm{X}_0, x_1 \in \mathrm{X}_1}$ and $\Gamma \equiv \{\Gamma_{y_1|y_0}^{B_0 \to B_1}\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1}$ be two PIDs such that $P_{\mathrm{guess}}(\mathbb{A}; \mathscr{M}) \geq P_{\mathrm{guess}}(\Gamma; \mathscr{M})$ for every bipartite POVM $\mathscr{M} \equiv \{M_{m,n}^{C_0 C_1}\}_{m \in \mathrm{M}, n \in \mathrm{N}}$. We assume $\mathbb{A} \not\succcurlyeq_{\mathrm{I}} \Gamma$. This means that

$$\Gamma \notin \left\{ \mathbb{A}' \equiv \left\{ \Lambda_{y_1|y_0}^{\prime B_0 \to B_1} \right\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} : \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}' \right\}. \tag{C6}$$

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 $\{\mathcal{O}_{y_0, y_1}^{B_0 \to B_1}\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1}$ and a positive real number $\varepsilon > 0$ such that

$$\left\langle \left\{ \mathcal{O}_{y_0, y_1}^{B_0 \to B_1} \right\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1}, \left\{ \Gamma_{y_1|y_0}^{B_0 \to B_1} \right\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \right\rangle > \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \left\langle \left\{ \mathcal{O}_{y_0, y_1}^{B_0 \to B_1} \right\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1}, \left\{ \Lambda_{y_1|y_0}^{\prime B_0 \to B_1} \right\}_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \right\rangle + \varepsilon. \tag{C7}$$

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

$$\sum_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \mathrm{Tr}\left[J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} J_{\Gamma_{y_1|y_0}}^{B_0 B_1} \right] > \max_{\mathbb{A}': \mathbb{A} \succcurlyeq_{\mathrm{I}} \mathbb{A}'} \sum_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \mathrm{Tr}\left[J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} J_{\Lambda_{y_1|y_0}'}^{B_0 B_1} \right] + \varepsilon. \tag{C8}$$

Let $c := \max_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \| J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} \|_\infty$. Define a bipartite POVM $\widehat{\mathscr{M}} \equiv \{\widehat{M}_{m,n}^{B_0 B_1}\}_{m \in \mathrm{Y}_0, n \in \mathrm{N}}$ such that $\mathrm{N} \supset \mathrm{Y}_1$ as follows:

$$\widehat{M}_{m,n}^{B_0 B_1} := \begin{cases} \dfrac{1}{2c \,|\mathrm{Y}_0| \,|\mathrm{Y}_1|} \left(J_{\mathcal{O}_{m,n}}^{B_0 B_1} + c \mathbb{1}^{B_0 B_1} \right) & \forall m \in \mathrm{Y}_0, n \in \mathrm{Y}_1, \\[2ex] \dfrac{1}{|\mathrm{Y}_0| \,(|\mathrm{N}| - |\mathrm{Y}_1|)} \left(\mathbb{1}^{B_0 B_1} - \sum_{y_0 \in \mathrm{Y}_0, y_1 \in \mathrm{Y}_1} \widehat{M}_{y_0, y_1}^{B_0 B_1} \right) & \forall m \in \mathrm{Y}_0, n \in \mathrm{N} \setminus \mathrm{Y}_1. \end{cases} \tag{C9}$$

It can be verified that $\widehat{\mathscr{M}}$ is a valid POVM, as $\widehat{M}_{m,n}^{B_0B_1} \geq 0$ for all $m \in \mathtt{Y}_0, n \in \mathtt{N}$ and $\sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \widehat{M}_{m,n}^{B_0B_1} = \mathbb{1}^{B_0B_1}$. Then

$$P_{\text{guess}}(\mathbb{T}; \widehat{\mathscr{M}}) = \frac{1}{d_{B_0}} \max_{\mathbb{T}': \mathbb{T} \succeq_{\mathtt{I}} \mathbb{T}'} \sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Gamma'_{n|m}}^{B_0B_1}\right] \tag{C10}$$

$$\geq \frac{1}{d_{B_0}} \sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[\widehat{M}_{y_0,y_1}^{B_0B_1} J_{\Gamma_{y_1|y_0}}^{B_0B_1}\right] \tag{C11}$$

$$= \frac{1}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|}\left(\sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1} J_{\Gamma_{y_1|y_0}}^{B_0B_1}\right] + c \sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\Gamma_{y_1|y_0}}^{B_0B_1}\right]\right) \tag{C12}$$

$$= \frac{1}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|}\sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1} J_{\Gamma_{y_1|y_0}}^{B_0B_1}\right] + \frac{1}{2|\mathtt{Y}_1|} \tag{C13}$$

$$> \frac{1}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|}\left(\max_{\mathbb{A}': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}'} \sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1} J_{\Lambda'_{y_1|y_0}}^{B_0B_1}\right] + \varepsilon\right) + \frac{1}{2|\mathtt{Y}_1|} \tag{C14}$$

$$= \frac{1}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \max_{\mathbb{A}': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}'}\left(\sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1} J_{\Lambda'_{y_1|y_0}}^{B_0B_1}\right] + c \sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[J_{\Lambda'_{y_1|y_0}}^{B_0B_1}\right]\right) + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C15}$$

$$= \frac{1}{d_{B_0}} \max_{\mathbb{A}': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}'} \sum_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1} \text{Tr}\left[\widehat{M}_{y_0,y_1}^{B_0B_1} J_{\Lambda'_{y_1|y_0}}^{B_0B_1}\right] + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C16}$$

$$\geq \frac{1}{d_{B_0}} \max_{\mathbb{A}'': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}''} \sum_{m \in \mathtt{Y}_0, n \in \mathtt{Y}_1} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right] + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C17}$$

$$= \frac{1}{d_{B_0}} \max_{\mathbb{A}'': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}''}\left(\sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right] - \sum_{m \in \mathtt{Y}_0, n \in \mathtt{N} \setminus \mathtt{Y}_1} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right]\right) + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C18}$$

$$\geq \frac{1}{d_{B_0}} \max_{\mathbb{A}'': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}''}\left(\sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right] - \frac{1}{|\mathtt{Y}_0|(|\mathtt{N}| - |\mathtt{Y}_1|)} \sum_{m \in \mathtt{Y}_0, n \in \mathtt{N} \setminus \mathtt{Y}_1} \text{Tr}\left[J_{\Lambda''_{n|m}}^{B_0B_1}\right]\right) + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C19}$$

$$\geq \frac{1}{d_{B_0}} \max_{\mathbb{A}'': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}''}\left(\sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right] - \frac{1}{|\mathtt{Y}_0|(|\mathtt{N}| - |\mathtt{Y}_1|)} \sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[J_{\Lambda''_{n|m}}^{B_0B_1}\right]\right) + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C20}$$

$$= \frac{1}{d_{B_0}} \max_{\mathbb{A}'': \mathbb{A} \succeq_{\mathtt{I}} \mathbb{A}''} \sum_{m \in \mathtt{Y}_0, n \in \mathtt{N}} \text{Tr}\left[\widehat{M}_{m,n}^{B_0B_1} J_{\Lambda''_{n|m}}^{B_0B_1}\right] - \frac{1}{|\mathtt{N}| - |\mathtt{Y}_1|} + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|} \tag{C21}$$

$$= P_{\text{guess}}(\mathbb{A}; \widehat{\mathscr{M}}) - \frac{1}{|\mathtt{N}| - |\mathtt{Y}_1|} + \frac{\varepsilon}{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|}. \tag{C22}$$

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 $\mathbb{A}'' \equiv \{\Lambda''^{B_0 \to B_1}_{n|m}\}_{m \in \mathtt{Y}_0, n \in \mathtt{N}}$ in contrast to the PID $\mathbb{A}' \equiv \{\Lambda'^{B_0 \to B_1}_{y_1|y_0}\}_{y_0 \in \mathtt{Y}_0, y_1 \in \mathtt{Y}_1}$ in Eq. (C16), and Ineq. (C17) follows from the fact that the freedom to make a guess n outside the set of target outcomes $\mathtt{Y}_1$ does not increase the maximum winning probability. Choose $\mathtt{N} \supseteq \mathtt{Y}_1$ to be an arbitrary index set such that

$$|\mathtt{N}| > |\mathtt{Y}_1| + \frac{2cd_{B_0}|\mathtt{Y}_0||\mathtt{Y}_1|}{\varepsilon}. \tag{C23}$$

Then Eq. (C22) implies $P_{\text{guess}}(\mathbb{A}; \widehat{\mathscr{M}}) < P_{\text{guess}}(\mathbb{\Gamma}; \widehat{\mathscr{M}})$, which contradicts the assumption that $P_{\text{guess}}(\mathbb{A}; \mathscr{M}) \geq P_{\text{guess}}(\mathbb{\Gamma}; \mathscr{M})$ for every $\mathscr{M}$. Therefore, we must have $\mathbb{A} \succcurlyeq_{\text{I}} \mathbb{\Gamma}$. This concludes the proof of Theorem 3.

2. Proof of Theorem 4

Let $\mathbb{A} \equiv \{\Lambda_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathsf{X}_0, x_1 \in \mathsf{X}_1}$ be a PID. By Definition 7, the robustness of incompatibility of $\mathbb{A}$ equals

$$\mathsf{RoI}(\mathbb{A}) := \min_{r \geq 0} r \tag{C24a}$$

$$\text{subject to:} \quad \left\{ \frac{\Lambda_{x_1|x_0} + r \Upsilon_{x_1|x_0}}{1 + r} \right\}_{x_0, x_1} \quad \text{is a simple PID,} \tag{C24b}$$

$$\left\{ \Upsilon_{x_1|x_0} \right\}_{x_0, x_1} \quad \text{is a PID.} \tag{C24c}$$

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 $\widehat{\mathbb{\Upsilon}} \equiv \{\widehat{\Upsilon}_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathsf{X}_0, x_1 \in \mathsf{X}_1}$ be an optimal solution to Program (C24), which is a PID. Define a PID $\widehat{\mathbb{\Omega}} \equiv \{\widehat{\Omega}_{x_1|x_0}^{A_0 \to A_1}\}_{x_0 \in \mathsf{X}_0, x_1 \in \mathsf{X}_1}$ as follows:

$$\widehat{\Omega}_{x_1|x_0}^{A_0 \to A_1} := \frac{\Lambda_{x_1|x_0}^{A_0 \to A_1} + \mathsf{RoI}(\mathbb{A}) \widehat{\Upsilon}_{x_1|x_0}^{A_0 \to A_1}}{1 + \mathsf{RoI}(\mathbb{A})} \quad \forall x_0, x_1. \tag{C25}$$

By Eq. (C24b), $\widehat{\mathbb{\Omega}}$ is simple. Let $\mathfrak{C}$ denote the cone generated by the set of unsteerable state assemblages in $A_0 A_1$:

$$\mathfrak{C} := \left\{ \left\{ \omega_{x_1|x_0}^{A_0 A_1} \right\}_{x_0 \in \mathsf{X}_0, x_1 \in \mathsf{X}_1} : \omega_{x_1|x_0}^{A_0 A_1} = \sum_g p_{x_1|x_0, g} \eta_g^{A_0 A_1}, \; \eta_g^{A_0 A_1} \geq 0, \; p_{x_1|x_0, g} \geq 0, \; \sum_{x_1} p_{x_1|x_0, g} = 1 \quad \forall x_0, x_1, g \right\}. \tag{C26}$$

We note that the cone $\mathfrak{C}$ is generating with respect to the space of state ensembles in $A_0 A_1$. Let $\mathfrak{C}^*$ denote the dual cone of $\mathfrak{C}$:

$$\mathfrak{C}^* := \left\{ \left\{ \kappa_{x_1|x_0}^{A_0 A_1} \right\}_{x_0 \in \mathsf{X}_0, x_1 \in \mathsf{X}_1} : \sum_{x_0, x_1} \text{Tr}\left[\kappa_{x_1|x_0}^{A_0 A_1} \omega_{x_1|x_0}^{A_0 A_1} \right] \geq 0 \quad \forall \left\{ \omega_{x_1|x_0}^{A_0 A_1} \right\}_{x_0, x_1} \in \text{cone}(\mathfrak{U}^{A_0 A_1}) \right\}. \tag{C27}$$

Substituting $\omega_{x_1|x_0}^{A_0 A_1}$ for $J_{\Lambda_{x_1|x_0}}^{A_0 A_1} + r J_{\Upsilon_{x_1|x_0}}^{A_0 A_1}$, Program (C24) can be reformulated as a conic program as follows:

$$\mathsf{RoI}(\mathbb{A}) = \frac{1}{d_{A_0} |\mathsf{X}_0|} \min_{\left\{ \omega_{x_1|x_0}^{A_0 A_1} \right\}_{x_0, x_1}} \sum_{x_0, x_1} \text{Tr}\left[\omega_{x_1|x_0}^{A_0 A_1} \right] - 1 \tag{C28a}$$

$$\text{subject to:} \quad \omega_{x_1|x_0}^{A_0 A_1} - J_{\Lambda_{x_1|x_0}}^{A_0 A_1} \geq 0 \quad \forall x_0, x_1, \tag{C28b}$$

$$d_{A_0} \sum_{x_1} \text{Tr}_{A_1}\left[\omega_{x_1|x_0}^{A_0 A_1} \right] = \sum_{x_1} \text{Tr}\left[\omega_{x_1|x_0}^{A_0 A_1} \right] \mathbb{1}^{A_0} \quad \forall x_0, \tag{C28c}$$

$$\left\{ \omega_{x_1|x_0}^{A_0 A_1} \right\}_{x_0, x_1} \in \mathfrak{C}. \tag{C28d}$$

Invoking the theory of conic programming duality [60], since $\omega_{x_1|x_0}^{A_0 A_1} = 2 d_{A_0} \mathbb{1}^{A_0 A_1}$ 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

$$
\mathrm{RoI}(\mathbb{\Lambda}) = \frac{1}{d_{A_0} |\mathrm{X}_0|} \max_{\{\alpha_{x_1|x_0}^{A_0 A_1}\}_{x_0,x_1}} \sum_{x_0,x_1} \mathrm{Tr}\left[\alpha_{x_1|x_0}^{A_0 A_1} J_{\Lambda_{x_1|x_0}}^{A_0 A_1}\right] - 1 \tag{C29a}
$$

$$
\text{subject to: } \quad \alpha_{x_1|x_0}^{A_0 A_1} \geq 0 \quad \forall x_0, x_1, \tag{C29b}
$$

$$
\sum_{x_0} \mathrm{Tr}\left[\beta_{x_0}^{A_0}\right] = d_{A_0} |\mathrm{X}_0|, \tag{C29c}
$$

$$
\left\{\beta_{x_0}^{A_0} \otimes \mathbb{1}^{A_1} - \alpha_{x_1|x_0}^{A_0 A_1}\right\}_{x_0,x_1} \in \mathfrak{C}^*. \tag{C29d}
$$

Let $\{\widehat{\alpha}_{x_1|x_0}^{A_0 A_1}\}_{x_0,x_1}$, $\{\widehat{\beta}_{x_1|x_0}^{A_0}\}_{x_0,x_1}$ 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 $\mathscr{M} \equiv \{M_{m,n}^{C_0 C_1}\}_{m\in\mathrm{M},n\in\mathrm{N}}$ be a bipartite POVM. Let $\widehat{\mathbb{\Lambda}}' \equiv \{\widehat{\Lambda}_{n|m}'^{C_0 \to C_1}\}_{m\in\mathrm{M},n\in\mathrm{N}}$ be an optimal solution to the maximization in Eq. (C2), namely, the PID simulated by $\mathbb{\Lambda}$ under Alice's optimal strategy. Let $\widehat{\mathbb{\Upsilon}}' \equiv \{\widehat{\Upsilon}_{n|m}'^{C_0 \to C_1}\}_{m\in\mathrm{M},n\in\mathrm{N}}$ and $\widehat{\mathbb{\Omega}}' \equiv \{\widehat{\Omega}_{n|m}'^{C_0 \to C_1}\}_{m\in\mathrm{M},n\in\mathrm{N}}$ denote the PIDs obtained by applying the same simulation strategy to $\widehat{\mathbb{\Upsilon}}$ and $\widehat{\mathbb{\Omega}}$, respectively. Then

$$
P_{\mathrm{guess}}(\mathbb{\Lambda}; \mathscr{M}) = \frac{1}{d_{C_0}} \sum_{m\in\mathrm{M},n\in\mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\widehat{\Lambda}_{n|m}'}^{C_0 C_1}\right] \tag{C30}
$$

$$
= \frac{1}{d_{C_0}} \sum_{m\in\mathrm{M},n\in\mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1}\left((1 + \mathrm{RoI}(\mathbb{\Lambda})) J_{\widehat{\Omega}_{n|m}'}^{C_0 C_1} - \mathrm{RoI}(\mathbb{\Lambda}) J_{\widehat{\Upsilon}_{n|m}'}^{C_0 C_1}\right)\right] \tag{C31}
$$

$$
\leq \frac{1 + \mathrm{RoI}(\mathbb{\Lambda})}{d_{C_0}} \sum_{m\in\mathrm{M},n\in\mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\widehat{\Omega}_{n|m}'}^{C_0 C_1}\right] \tag{C32}
$$

$$
\leq \frac{1 + \mathrm{RoI}(\mathbb{\Lambda})}{d_{C_0}} \max_{\mathbb{\Omega}:\, \widehat{\mathbb{\Omega}}' \succcurlyeq_I \mathbb{\Omega}} \sum_{m\in\mathrm{M},n\in\mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\Omega_{n|m}}^{C_0 C_1}\right] \tag{C33}
$$

$$
= \frac{1 + \mathrm{RoI}(\mathbb{\Lambda})}{d_{C_0}} \max_{\mathbb{\Omega}:\, \mathrm{simple}} \sum_{m\in\mathrm{M},n\in\mathrm{N}} \mathrm{Tr}\left[M_{m,n}^{C_0 C_1} J_{\Omega_{n|m}}^{C_0 C_1}\right] \tag{C34}
$$

$$
= (1 + \mathrm{RoI}(\mathbb{\Lambda})) P_{\mathrm{guess}}^{\mathrm{simple}}(\mathscr{M}). \tag{C35}
$$

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

$$
\sup_{\mathscr{M}} \frac{P_{\mathrm{guess}}(\mathbb{\Lambda}; \mathscr{M})}{P_{\mathrm{guess}}^{\mathrm{simple}}(\mathscr{M})} \leq 1 + \mathrm{RoI}(\mathbb{\Lambda}). \tag{C36}
$$

Next, we show that Ineq. (C36) can be equalized by an infinite sequence of POVMs on $A_0 A_1$. Let $c := \|\sum_{x_0\in\mathrm{X}_0, x_1\in\mathrm{X}_1} \widehat{\alpha}_{x_1|x_0}^{A_0 A_1}\|_\infty$. Define a bipartite POVM $\widehat{\mathscr{M}} \equiv \{\widehat{M}_{m,n}^{A_0 A_1}\}_{m\in\mathrm{X}_0, n\in\mathrm{N}}$ such that $\mathrm{N} \supset \mathrm{X}_1$ as follows:

$$
\widehat{M}_{m,n}^{A_0 A_1} := \begin{cases} \dfrac{1}{c} \widehat{\alpha}_{n|m}^{A_0 A_1} & \forall m \in \mathrm{X}_0, n \in \mathrm{X}_1, \\[2ex] \dfrac{1}{|\mathrm{X}_0|\,(|\mathrm{N}| - |\mathrm{X}_1|)}\left(\mathbb{1}^{A_0 A_1} - \sum_{x_0\in\mathrm{X}_0, x_1\in\mathrm{X}_1} \widehat{M}_{x_0,x_1}^{A_0 A_1}\right) & \forall m \in \mathrm{X}_0, n \in \mathrm{N} \setminus \mathrm{X}_1. \end{cases} \tag{C37}
$$

It can be verified that $\widehat{\mathscr{M}}$ is a valid POVM, as $\widehat{M}_{m,n}^{A_0 A_1} \geq 0$ for all $m \in X_0, n \in N$ and $\sum_{m \in X_0, n \in N} \widehat{M}_{m,n}^{A_0 A_1} = \mathbb{1}^{A_0 A_1}$. Then

$$P_{\text{guess}}(\mathbb{A}, \widehat{\mathscr{M}}) = \frac{1}{d_{A_0}} \max_{\mathbb{A}' : \mathbb{A} \succcurlyeq_1 \mathbb{A}'} \sum_{m \in X_0, n \in N} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Lambda'_{n|m}}^{A_0 A_1}\right] \tag{C38}$$

$$\geq \frac{1}{d_{A_0}} \sum_{x_0 \in X_0, x_1 \in X_1} \text{Tr}\left[\widehat{M}_{x_0, x_1}^{A_0 A_1} J_{\Lambda_{x_1|x_0}}^{A_0 A_1}\right] \tag{C39}$$

$$= \frac{1}{d_{A_0} c} \sum_{x_0 \in X_0, x_1 \in X_1} \text{Tr}\left[\widehat{\alpha}_{x_1|x_0}^{A_0 A_1} J_{\Lambda_{x_1|x_0}}^{A_0 A_1}\right] \tag{C40}$$

$$= \frac{|X_0| \, (1 + \text{Rol}(\mathbb{A}))}{c}. \tag{C41}$$

Here Eq. (C41) is by the definition of $\{\widehat{\alpha}_{x_1|x_0}^{A_0 A_1}\}_{x_0, x_1}$ and follows from Eq. (C29a). In addition,

$$P_{\text{guess}}^{\text{simple}}(\widehat{\mathscr{M}}) = \frac{1}{d_{A_0}} \max_{\Omega : \text{simple}} \sum_{m \in X_0, n \in N} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] \tag{C42}$$

$$= \frac{1}{d_{A_0}} \max_{\Omega : \text{simple}} \left(\sum_{m \in X_0, n \in X_1} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] + \sum_{m \in X_0, n \in N \setminus X_1} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] \right) \tag{C43}$$

$$\leq \frac{1}{d_{A_0}} \max_{\Omega : \text{simple}} \left(\sum_{m \in X_0, n \in X_1} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] + \frac{1}{|X_0| \, (|N| - |X_1|)} \sum_{m \in X_0, n \in N \setminus X_1} \text{Tr}\left[J_{\Omega_{n|m}}^{A_0 A_1}\right] \right) \tag{C44}$$

$$\leq \frac{1}{d_{A_0}} \max_{\Omega : \text{simple}} \left(\sum_{m \in X_0, n \in X_1} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] + \frac{1}{|X_0| \, (|N| - |X_1|)} \sum_{m \in X_0, n \in N} \text{Tr}\left[J_{\Omega_{n|m}}^{A_0 A_1}\right] \right) \tag{C45}$$

$$= \frac{1}{d_{A_0}} \max_{\Omega : \text{simple}} \sum_{m \in X_0, n \in X_1} \text{Tr}\left[\widehat{M}_{m,n}^{A_0 A_1} J_{\Omega_{n|m}}^{A_0 A_1}\right] + \frac{1}{|N| - |X_1|} \tag{C46}$$

$$\leq \frac{1}{d_{A_0}} \max_{\Omega' : \text{simple}} \sum_{x_0 \in X_0, x_1 \in X_1} \text{Tr}\left[\widehat{M}_{x_0, x_1}^{A_0 A_1} J_{\Omega'_{x_1|x_0}}^{A_0 A_1}\right] + \frac{1}{|N| - |X_1|} \tag{C47}$$

$$= \frac{1}{d_{A_0} c} \max_{\Omega' : \text{simple}} \sum_{x_0 \in X_0, x_1 \in X_1} \text{Tr}\left[\widehat{\alpha}_{x_1|x_0}^{A_0 A_1} J_{\Omega'_{x_1|x_0}}^{A_0 A_1}\right] + \frac{1}{|N| - |X_1|} \tag{C48}$$

$$= \frac{1}{d_{A_0} c} \max_{\Omega' : \text{simple}} \sum_{x_0 \in X_0, x_1 \in X_1} \left(\text{Tr}\left[\left(\widehat{\beta}_{x_0}^{A_0} \otimes \mathbb{1}^{A_1}\right) J_{\Omega'_{x_1|x_0}}^{A_0 A_1}\right] - \text{Tr}\left[\left(\widehat{\beta}_{x_0}^{A_0} \otimes \mathbb{1}^{A_1} - \widehat{\alpha}_{x_1|x_0}^{A_0 A_1}\right) J_{\Omega'_{x_1|x_0}}^{A_0 A_1}\right] \right) + \frac{1}{|N| - |X_1|} \tag{C49}$$

$$\leq \frac{1}{d_{A_0} c} \max_{\Omega' : \text{simple}} \sum_{x_0 \in X_0, x_1 \in X_1} \text{Tr}\left[\left(\widehat{\beta}_{x_0}^{A_0} \otimes \mathbb{1}^{A_1}\right) J_{\Omega'_{x_1|x_0}}^{A_0 A_1}\right] + \frac{1}{|N| - |X_1|} \tag{C50}$$

$$= \frac{1}{d_{A_0} c} \sum_{x_0 \in X_0} \text{Tr}\left[\widehat{\beta}_{x_0}^{A_0}\right] + \frac{1}{|N| - |X_1|} \tag{C51}$$

$$= \frac{|X_0|}{c} + \frac{1}{|N| - |X_1|}. \tag{C52}$$

Here Eqs. (C44) and (C48) follow from Eq. (C37); Ineq. (C50) follows from Eq. (C29c); Eq. (C52) follows from Eq. (C29d). Note that the PID in Ineq. (C47) has the form $\Omega' \equiv \{\Omega_{x_1|x_0}'^{A_0 \to A_1}\}_{x_0 \in X_0, x_1 \in X_1}$ in contrast to the PID $\Omega \equiv$

$\{\Omega_{n|m}^{A_0\to A_1}\}_{m\in\mathbb{X}_0,n\in\mathbb{N}}$ 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 $\mathbb{Y}_1$. Then it follows from Eqs. (C41) and (C52) that

$$\sup_{\mathscr{M}}\frac{P_{\text{guess}}(\mathbb{A};\mathscr{M})}{P_{\text{guess}}^{\text{simple}}(\mathscr{M})}\geq\lim_{|\mathbb{N}|\to\infty}\frac{P_{\text{guess}}(\mathbb{A};\widehat{\mathscr{M}})}{P_{\text{guess}}^{\text{simple}}(\widehat{\mathscr{M}})}\tag{C53}$$

$$=\lim_{|\mathbb{N}|\to\infty}\frac{1+\text{Rol}(\mathbb{A})}{1+\frac{c}{|\mathbb{X}_0|(|\mathbb{N}|-|\mathbb{X}_1|)}}\tag{C54}$$

$$=1+\text{Rol}(\mathbb{A}).\tag{C55}$$

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

$$\sup_{\mathscr{M}}\frac{P_{\text{guess}}(\mathbb{A};\mathscr{M})}{P_{\text{guess}}^{\text{simple}}(\mathscr{M})}=1+\text{Rol}(\mathbb{A}).\tag{C56}$$

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 $\mathbb{A}\equiv\{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0\in\mathbb{X}_0,x_1\in\mathbb{X}_1}$ and $\Gamma\equiv\{\Gamma_{y_1|y_0}^{A_0\to A_1}\}_{y_0\in\mathbb{Y}_0,y_1\in\mathbb{Y}_1}$ be two PIDs such that $\mathbb{A}\succcurlyeq_I\Gamma$. Let $\mathscr{L}\equiv\{L_{l'}^{B_1}\}_{l'\in\mathbb{L}}$ be a POVM, and let $\varsigma\equiv\{\sigma_{m,n,l}^{B_0}\}_{m\in\mathbb{Y}_0,n\in\mathbb{Y}_1,l\in\mathbb{L}}$ be a state ensemble. By Theorem 1(3), the preorder $\succcurlyeq_I$ is transitive, thus

$$P'_{\text{guess}}(\mathbb{A};\varsigma,\mathscr{L})=\max_{\mathbb{A}':\mathbb{A}\succcurlyeq_I\mathbb{A}'}\sum_{m,n,l}\text{Tr}\left[L_l^{B_1}\Lambda_{n|m}'^{B_0\to B_1}\left[\sigma_{m,n,l}^{B_0}\right]\right]\tag{C57}$$

$$\geq\max_{\mathbb{A}':\Gamma\succcurlyeq_I\mathbb{A}'}\sum_{m,n,l}\text{Tr}\left[L_l^{B_1}\Lambda_{n|m}'^{B_0\to B_1}\left[\sigma_{m,n,l}^{B_0}\right]\right]\tag{C58}$$

$$=P'_{\text{guess}}(\Gamma;\varsigma,\mathscr{L}).\tag{C59}$$

Next, we prove the sufficiency of the convertibility conditions in Proposition 3 by contradiction. Let $\mathbb{A}\equiv\{\Lambda_{x_1|x_0}^{A_0\to A_1}\}_{x_0\in\mathbb{X}_0,x_1\in\mathbb{X}_1}$ be two PIDs, and let $\mathscr{L}\equiv\{L_{l'}^{B_1}\}_{l'\in\mathbb{L}}$ be an informationally complete POVM such that $P'_{\text{guess}}(\mathbb{A};\varsigma,\mathscr{L})\geq P'_{\text{guess}}(\mathbb{A};\varsigma,\mathscr{L})$ for every state ensemble $\varsigma\equiv\{\sigma_{m,n,l}^{B_0}\}_{m\in\mathbb{Y}_0,n\in\mathbb{Y}_1,l\in\mathbb{L}}$. We assume $\mathbb{A}\not\succcurlyeq_I\Gamma$. Following the argument that leads to Eq. (C8), there exists a set of Hermiticity-preserving linear maps $\{\mathcal{O}_{y_0,y_1}^{B_0\to B_1}\}_{y_0\in\mathbb{Y}_0,y_1\in\mathbb{Y}_1}$ such that

$$\sum_{y_0,y_1}\text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1}J_{\Gamma_{y_1|y_0}}^{B_0B_1}\right]>\max_{\mathbb{A}':\mathbb{A}\succcurlyeq_I\mathbb{A}'}\sum_{y_0,y_1}\text{Tr}\left[J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1}J_{\Lambda_{y_1|y_0}'}^{B_0B_1}\right].\tag{C60}$$

Since $\mathscr{L}$ is an informationally complete POVM, there exists a set of Hermitian operators $\{\mu_{y_0,y_1,l'}^{B_0}\}_{y_0\in\mathbb{Y}_0,y_1\in\mathbb{Y}_1,l'\in\mathbb{L}}$ such that

$$J_{\mathcal{O}_{y_0,y_1}}^{B_0B_1}=\sum_{l'}\mu_{y_0,y_1,l'}^{B_0}\otimes L_{l'}^{B_1}\quad\forall y_0,y_1.\tag{C61}$$

Let $c:=\max_{y_0\in\mathbb{Y}_0,y_1\in\mathbb{Y}_1}\|\mu_{y_0,y_1}^{B_0}\|_\infty$ and $c':=\sum_{m\in\mathbb{Y}_0,n\in\mathbb{Y}_1,l\in\mathbb{L}}\text{Tr}[\mu_{m,n,l}^{B_0}]$. Define a state ensemble $\widehat{\varsigma}\equiv\{\widehat{\sigma}_{m,n,l}^{B_0}\}_{m\in\mathbb{Y}_0,n\in\mathbb{Y}_1,l\in\mathbb{L}}$ as follows:

$$\widehat{\sigma}_{m,n,l}^{B_0}:=\frac{1}{c'+cd_{B_0}|\mathbb{Y}_0||\mathbb{Y}_1||\mathbb{L}|}\left((\mu_{m,n,l}^\top)^{B_0}+c\mathbb{1}^{B_0}\right)\quad\forall m,n,l.\tag{C62}$$

It can be verified that $\widehat{\varsigma}$ is a valid state ensemble, as $\widehat{\sigma}_{m,n,l}^{B_0} \geq 0$ for all $m \in \mathtt{Y}_0, n \in \mathtt{Y}_0, l \in \mathtt{L}$ and $\sum_{m,n,l} \mathrm{Tr}[\widehat{\sigma}_{m,n,l}^{B_0}] = 1$. Then

$$P'_{\mathrm{guess}}(\mathbb{A}; \widehat{\varsigma}, \mathscr{L}) = \max_{\mathbb{A}': \mathbb{A}' \succcurlyeq_I \mathbb{A}'} \sum_{m,n,l} \mathrm{Tr}\left[L_l^{B_1} \Gamma_{n|m}'^{B_0 \to B_1}\left[\widehat{\sigma}_{m,n,l}^{B_0}\right]\right] \tag{C63}$$

$$= \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \max_{\mathbb{A}': \mathbb{A}' \succcurlyeq_I \mathbb{A}'} \left(\sum_{m,n,l} \mathrm{Tr}\left[L_l^{B_1} \Lambda_{n|m}'^{B_0 \to B_1}\left[(\mu_{m,n,l}^\top)^{B_0}\right]\right] + c \sum_{m,n,l} \mathrm{Tr}\left[L_l^{B_1} \Lambda_{n|m}'^{B_0 \to B_1}\left[\mathbb{1}^{B_0}\right]\right]\right) \tag{C64}$$

$$= \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \max_{\mathbb{A}': \mathbb{A}' \succcurlyeq_I \mathbb{A}'} \left(\sum_{m,n,l} \mathrm{Tr}\left[\left(\mu_{m,n,l}^{B_0} \otimes L_l^{B_1}\right)\left(\mathrm{id}^{B_0} \otimes \Lambda_{n|m}'^{\widetilde{B}_0 \to B_1}\right)\left[\phi_+^{B_0 \widetilde{B}_0}\right]\right] + cd_{B_0} |\mathtt{Y}_0|\right) \tag{C65}$$

$$= \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \left(\max_{\mathbb{A}': \mathbb{A}' \succcurlyeq_I \mathbb{A}'} \sum_{y_0, y_1} \mathrm{Tr}\left[J_{\mathcal{O}_{m,n}}^{B_0 B_1} J_{\Lambda'_{y_1 | y_0}}^{B_0 B_1}\right] + cd_{B_0} |\mathtt{Y}_0|\right). \tag{C66}$$

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

$$P'_{\mathrm{guess}}(\mathbb{\Gamma}; \widehat{\varsigma}, \mathscr{L}) = \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \left(\max_{\mathbb{\Gamma}': \mathbb{\Gamma}' \succcurlyeq_I \mathbb{\Gamma}'} \sum_{y_0, y_1} \mathrm{Tr}\left[J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} J_{\Gamma'_{y_1 | y_0}}^{B_0 B_1}\right] + cd_{B_0} |\mathtt{Y}_0|\right) \tag{C67}$$

$$\geq \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \left(\sum_{y_0, y_1} \mathrm{Tr}\left[J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} J_{\Gamma_{y_1 | y_0}}^{B_0 B_1}\right] + cd_{B_0} |\mathtt{Y}_0|\right) \tag{C68}$$

$$> \frac{1}{c' + cd_{B_0} |\mathtt{Y}_0| |\mathtt{Y}_1| |\mathtt{L}|} \left(\max_{\mathbb{A}': \mathbb{A}' \succcurlyeq_I \mathbb{A}'} \sum_{y_0, y_1} \mathrm{Tr}\left[J_{\mathcal{O}_{y_0, y_1}}^{B_0 B_1} J_{\Lambda'_{y_1 | y_0}}^{B_0 B_1}\right] + cd_{B_0} |\mathtt{Y}_0|\right) \tag{C69}$$

$$= P'_{\mathrm{guess}}(\mathbb{A}; \widehat{\varsigma}, \mathscr{L}). \tag{C70}$$

This contradicts the assumption that $P'_{\mathrm{guess}}(\mathbb{A}; \widehat{\varsigma}, \mathscr{L}) \geq P'_{\mathrm{guess}}(\mathbb{\Gamma}; \widehat{\varsigma}, \mathscr{L})$ for every $\widehat{\varsigma}$. Therefore, we have $\mathbb{A} \succcurlyeq_I \mathbb{\Gamma}$. This concludes the proof of Proposition 3.

---

[1] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, Rev. Mod. Phys. **89**, 015002 (2017).

[2] M. Koashi, Unconditional security of quantum key distribution and the uncertainty principle, J. Phys.: Conf. Ser. **36**, 98 (2006).

[3] M. Tomamichel and R. Renner, Uncertainty relation for smooth entropies, Phys. Rev. Lett. **106**, 110506 (2011).

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[5] J. M. Renes and M. M. Wilde, Polar codes for private and quantum communication over arbitrary channels, IEEE Trans. Inf. Theory **60**, 3090 (2014).

[6] P. Lahti, Coexistence and joint measurability in quantum mechanics, Int. J. Theor. Phys. **42**, 893 (2003).

[7] T. Heinosaari, D. Reitzner, and P. Stano, Notes on joint measurability of quantum observables, Found. Phys. **38**, 1133 (2008).

[8] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A: Math. Theor. **49**, 123001 (2016).

[9] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).

[10] T. Heinosaari, T. Miyadera, and D. Reitzner, Strongly incompatible quantum devices, Found. Phys. **44**, 34 (2013).

[11] T. Heinosaari and T. Miyadera, Incompatibility of quantum channels, J. Phys. A: Math. Theor. **50**, 135302 (2017).

[12] G. M. D'Ariano, P. Perinotti, and A. Tosini, Incompatibility of observables, channels and instruments in information theories, J. Phys. A: Math. Theor. **55**, 394006 (2022).

[13] M. F. Pusey, Verifying the quantumness of a channel with an untrusted device, J. Opt. Soc. Am. B **32**, A56 (2015).

[14] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015).

[15] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz, Incompatibility breaking quantum channels, J. Phys. A: Math. Theor. **48**, 435301 (2015).

[16] L. Guerini, J. Bavaresco, M. T. Cunha, and A. Acín, Operational framework for quantum measurement simulability, J. Math. Phys. **58**, 092102 (2017).

[17] M. Oszmaniec and T. Biswas, Operational relevance of resource theories of quantum measurements, Quantum **3**, 133 (2019).

[18] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Quantifying quantum resources with conic programming, Phys. Rev. Lett. **122**, 130404 (2019).

[19] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, All sets of incompatible measurements give an advantage in quantum state discrimination, Phys. Rev. Lett. **122**, 130403 (2019).

[20] F. Buscemi, E. Chitambar, and W. Zhou, Complete resource theory of quantum incompatibility as quantum programmability, Phys. Rev. Lett. **124**, 120401 (2020).

[21] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini, Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication, Quantum **7**, 1035 (2023).

[22] M. Horodecki and J. Oppenheim, (Quantumness in the context of) resource theories, Int. J. Mod. Phys. B **27**, 1345019 (2013).

[23] B. Coecke, T. Fritz, and R. W. Spekkens, A mathematical theory of resources, Inf. Comput. **250**, 59 (2016).

[24] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[25] R. Takagi and B. Regula, General resource theories in quantum mechanics and beyond: Operational characterization via discrimination tasks, Phys. Rev. X **9**, 031053 (2019).

[26] Z.-W. Liu, K. Bu, and R. Takagi, One-shot operational quantum resource theory, Phys. Rev. Lett. **123**, 020401 (2019).

[27] M. Piani and J. Watrous, Necessary and sufficient quantum information characterization of Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **114**, 060404 (2015).

[28] E. Haapasalo, Robustness of incompatibility for quantum devices, J. Phys. A: Math. Theor. **48**, 255303 (2015).

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

[30] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: A unified framework, New J. Phys. **21**, 113053 (2019).

[31] C. Carmeli, T. Heinosaari, T. Miyadera, and A. Toigo, Witnessing incompatibility of quantum channels, J. Math. Phys. **60**, 122202 (2019).

[32] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum incompatibility witnesses, Phys. Rev. Lett. **122**, 130402 (2019).

[33] T. Heinosaari, T. Miyadera, and R. Takakura, Testing incompatibility of quantum devices with few states, Phys. Rev. A **104**, 032228 (2021).

[34] B. Regula and R. Takagi, One-shot manipulation of dynamical quantum resources, Phys. Rev. Lett. **127**, 060402 (2021).

[35] C. Carmeli, T. Heinosaari, and A. Toigo, State discrimination with postmeasurement information and incompatibility of quantum measurements, Phys. Rev. A **98**, 012126 (2018).

[36] R. Uola, T. Bullock, T. Kraft, J.-P. Pellonpää, and N. Brunner, All quantum resources provide an advantage in exclusion tasks, Phys. Rev. Lett. **125**, 110402 (2020).

[37] J. Mori, Operational characterization of incompatibility of quantum channels with quantum state discrimination, Phys. Rev. A **101**, 032331 (2020).

[38] F. Buscemi, Comparison of quantum statistical models: Equivalent conditions for sufficiency, Commun. Math. Phys. **310**, 625 (2012).

[39] F. Buscemi, All entangled quantum states are nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[40] F. Buscemi, Degradable channels, less noisy channels, and quantum statistical morphisms: An equivalence relation, Probl. Inf. Transm. **52**, 201 (2016).

[41] D. Rosset, F. Buscemi, and Y.-C. Liang, Resource theory of quantum memories and their faithful verification with minimal assumptions, Phys. Rev. X **8**, 021033 (2018).

[42] P. Skrzypczyk and N. Linden, Robustness of measurement, discrimination games, and accessible information, Phys. Rev. Lett. **122**, 140403 (2019).

[43] D. Blackwell, Equivalent comparisons of experiments, Ann. Math. Stat. **24**, 265 (1953).

[44] E. Chitambar, G. Gour, K. Sengupta, and R. Zibakhsh, Quantum Bell nonlocality as a form of entanglement, Phys. Rev. A **104**, 052208 (2021).

[45] M. Piani, Channel steering, J. Opt. Soc. Am. B **32**, A1 (2015).

[46] G. Gour, T. Heinosaari, and R. Spekkens, A resource theory of quantum incompatibility, APS March Meeting 2018, Session S26: Quantum Resource Theories I, Los Angeles, California (2018).

[47] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, Causal and localizable quantum operations, Phys. Rev. A **64**, 052309 (2001).

[48] T. Eggeling, D. Schlingemann, and R. F. Werner, Semicausal operations are semilocalizable, Europhys. Lett. **57**, 782 (2002).

[49] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Transforming quantum operations: Quantum supermaps, Europhys. Lett. **83**, 30004 (2008).

[50] G. Gour, Comparison of quantum channels by superchannels, IEEE Trans. Inf. Theory **65**, 5880 (2019).

[51] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[52] R. Gallego and L. Aolita, Resource theory of steering, Phys. Rev. X **5**, 041008 (2015).

[53] E. Schrödinger, Discussion of probability relations between separated systems, Math. Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[54] 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).

[55] S. J. Jones, H. M. Wiseman, and A. C. Doherty, Entanglement, Einstein-Podolsky-Rosen correlations, Bell nonlocality, and steering, Phys. Rev. A **76**, 052116 (2007).

[56] M. T. Quintino, T. Vértesi, and N. Brunner, Joint measurability, Einstein-Podolsky-Rosen steering, and Bell nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).

[57] R. Uola, T. Moroder, and O. Gühne, Joint measurability of generalized measurements implies classicality, Phys. Rev. Lett. **113**, 160403 (2014).

[58] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, One-to-one mapping between steering and joint measurability problems, Phys. Rev. Lett. **115**, 230402 (2015).

[59] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23**, 880 (1969).

[60] S. P. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK; New York, 2004).

[61] P. Lipka-Bartosik and P. Skrzypczyk, Operational advantages provided by nonclassical teleportation, Phys. Rev. Res. **2**, 023029 (2020).

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

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

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

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

[66] S. Milz and K. Modi, Quantum stochastic processes and quantum non-Markovian phenomena, PRX Quantum **2**, 030201 (2021).

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

[68] S. Khatri and M. M. Wilde, Principles of quantum communication theory: A modern approach, arXiv:2011.04672 (2020).