HAL open science

Network-Device-Independent Certification of Causal Nonseparability

Hippolyte Dourdent, Alastair A. Abbott, Ivan Šupić, Cyril Branciard

▶ **To cite this version:**

Hippolyte Dourdent, Alastair A. Abbott, Ivan Šupić, Cyril Branciard. Network-Device-Independent Certification of Causal Nonseparability. 2023. hal-04193227

HAL Id: hal-04193227

https://inria.hal.science/hal-04193227

Preprint submitted on 1 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Network-Device-Independent Certification of Causal Nonseparability

Hippolyte Dourdent,[1] Alastair A. Abbott,[2] Ivan Šupić,[3] and Cyril Branciard[4]

[1] *ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels, Barcelona, Spain*

[2] *Univ. Grenoble Alpes, Inria, 38000 Grenoble, France*

[3] *LIP6, Sorbonne Université, CNRS, 4 Place Jussieu, 75005 Paris, France*

[4] *Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

(Dated: August 24, 2023)

Causal nonseparability is the property underlying quantum processes incompatible with a definite causal order. So far it has remained a central open question as to whether any process with a clear physical realisation can violate a causal inequality, so that its causal nonseparability can be certified in a device-independent way, as originally conceived. Here we present a method solely based on the observed correlations, which certifies the causal nonseparability of all the processes that can induce a causally nonseparable distributed measurement in a scenario with trusted quantum input states, as defined in [Dourdent *et al.*, Phys. Rev. Lett. **129**, 090402 (2022)]. This notably includes the celebrated quantum switch. This device-independent certification is achieved by introducing a network of untrusted operations, allowing one to self-test the quantum inputs on which the effective distributed measurement induced by the process is performed.

arXiv:2308.12760v1 [quant-ph] 24 Aug 2023

Introduction.—For both quantum and classical processes, it is usually assumed that if two parties interact only once with a physical medium, then only one-way influences are possible. Remarkably, this turns out to be an unnecessary restriction on the possible observable correlations. Formalised in the process matrix framework, where quantum theory is taken to hold locally but no global causal structure is assumed, some processes allow parties to establish correlations incompatible with a definite causal order [1].

Causal nonseparability, the property underlying such processes [1, 2], bears a foundational significance and is also the basis for advantages in a number of different tasks [3–8]. It can be certified in a *device-independent* (DI) way, i.e. requiring knowledge of the observed statistics only, based on the violation of "causal inequalities" by noncausal correlations [1, 9]. However, not all causally nonseparable process matrices are noncausal in this strong sense [10–12]. Indeed, a large class of quantum-realisable processes cannot violate causal inequalities [13, 14], and it is still unclear whether any process with a clear physical realisation is able to. This includes in particular the canonical "quantum switch" [3], the resource behind most known advantages arising from causal indefiniteness [3–8].

In a previous work [15], we showed (with another colleague) that the quantum switch, as well as a large class of bipartite causally nonseparable processes (and even all of them under a reasonable additional assumption), can display a new form of noncausality in a semi-DI with quantum inputs (SDI-QI) scenario. In the present work, we show how the gap between this SDI-QI certification and a fully DI certification can be bridged by self-testing the quantum inputs in a network scenario, where each party that receives a quantum input in the SDI-QI scenario shares an additional uncharacterised bipartite state with an extra untrusted party. We call this certification *network-device-independent (NDI)* to differentiate it from the standard DI scenario based on causal inequalities. This approach is motivated by the ability to exploit self-testing to transform a universal certification of entanglement in a SDI-QI scenario into a NDI certification [16, 17]. The apparent analogy between entanglement and causal nonseparability, however, was already seen to have severe limitations in the SDI-QI scenario [15], making it unclear whether NDI certification could be applied to this latter resource.

Causal (non)separability in the SDI-QI scenario.—The approach we present is a general method for transforming SDI-QI certification of causal nonseparability into NDI certification, applicable to a range of scenarios. For concreteness and clarity, however, we focus our initial presentation here on the bipartite scenario, before making some comments about its applicability to other scenarios and processes later.

Two parties, Alice and Bob, control separate quantum labs with input and output Hilbert spaces $\mathcal{H}^{A_I}$ and $\mathcal{H}^{A_O}$ for Alice, and $\mathcal{H}^{B_I}$ and $\mathcal{H}^{B_O}$ for Bob. They may also receive some auxiliary quantum states in Hilbert spaces $\mathcal{H}^{A'}, \mathcal{H}^{B'}$.[1] Let us consider the SDI-QI scenario introduced in [15] where Alice and Bob are provided with quantum input states $\rho_z^{A'} \in \mathcal{L}(\mathcal{H}^{A'})$ and $\rho_w^{B'} \in \mathcal{L}(\mathcal{H}^{B'})$, respectively, indexed by the labels z and w. They each perform some fixed quantum operations described

[1] All Hilbert spaces we consider are taken to be finite-dimensional. Notations: $\mathcal{L}(\mathcal{H}^X)$ denotes the space of linear operators on $\mathcal{H}^X$; superscripts are used to indicate on what spaces operators act; we shall write concisely $\mathcal{H}^{XY} = \mathcal{H}^X \otimes \mathcal{H}^Y$, $\mathcal{H}^A = \mathcal{H}^{A_I A_O}$, etc. Note that in [15], the auxiliary spaces in which the quantum inputs live were denoted with tildes (e.g., $\mathcal{H}^{\tilde{A}}$); here we switch to primes: the tildes will be used for the untrusted physical systems in the self-testing procedure below. Also, in [15] Alice and Bob's quantum input states were labeled by subscripts x and y; here we change these to z and w, as x and y will be introduced later as classical inputs for the self-testing procedure.

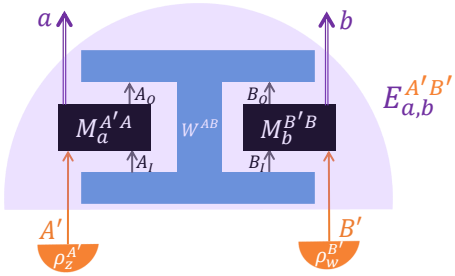

FIG. 1. The SDI-QI scenario: A process matrix W^{AB} connects two parties who receive trusted (orange) quantum inputs $\rho_z^{A'}$ and $\rho_w^{B'}$, resp. They each perform an untrusted (black) joint operation ($(M_a^{A'A})_a$ and $(M_b^{B'B})_b$, resp.), producing the classical outcomes a and b. The purple semicircle represents the D-POVM $(E_{a,b}^{A'B'})_{a,b}$ induced by these instruments and the process matrix.

as quantum instruments [18], i.e., sets of completely positive (CP) maps $\mathcal{M}_a : \mathcal{L}(\mathcal{H}^{A'A_I}) \to \mathcal{L}(\mathcal{H}^{A_O})$ and $\mathcal{M}_b : \mathcal{L}(\mathcal{H}^{B'B_I}) \to \mathcal{L}(\mathcal{H}^{B_O})$, whose indices a, b refer to some (classical) outcomes for Alice and Bob, and whose sums $\sum_a \mathcal{M}_a$ and $\sum_b \mathcal{M}_b$ are trace-preserving (TP). Using the Choi isomorphism [19] (see the Supplemental Material (SM) of [15], Sec. A), the CP maps $\mathcal{M}_a$, $\mathcal{M}_b$ can be represented as positive semidefinite matrices $M_a^{A'A}$ and $M_b^{B'B}$.

Within the process matrix framework, the correlations established by Alice and Bob in such a scenario with quantum inputs are given by the probabilities

$$\begin{aligned} P(a,b|\rho_z^{A'},\rho_w^{B'}) &= \left(M_a^{A'A}\otimes M_b^{B'B}\right) * \left(\rho_z^{A'}\otimes\rho_w^{B'}\otimes W^{AB}\right) \\ &= E_{a,b}^{A'B'} * \left(\rho_z^{A'}\otimes\rho_w^{B'}\right) = \mathrm{Tr}\left[\left(E_{a,b}^{A'B'}\right)^T\left(\rho_z^{A'}\otimes\rho_w^{B'}\right)\right] \end{aligned} \tag{1}$$

with

$$E_{a,b}^{A'B'} = \left(M_a^{A'A}\otimes M_b^{B'B}\right) * W^{AB}. \tag{2}$$

Here, $W^{AB} \in \mathcal{L}(\mathcal{H}^{AB})$ is the so-called "process matrix", a positive semidefinite matrix satisfying nontrivial linear constraints in order to always generate valid probabilities [1] (see e.g. [15], SM, Sec. B); and $*$ is the "link product" [20, 21], a convenient tool for calculations defined for any matrices $M^{XY} \in \mathcal{L}(\mathcal{H}^{XY})$, $N^{YZ} \in \mathcal{L}(\mathcal{H}^{YZ})$ as $M^{XY} * N^{YZ} = \mathrm{Tr}_Y[(M^{XY}\otimes \mathbb{1}^Z)^{T_Y}(\mathbb{1}^X\otimes N^{YZ})] \in \mathcal{L}(\mathcal{H}^{XZ})$ (where T_Y is the partial transpose over $\mathcal{H}^Y$; see also [15], SM, Sec. A).[2] The family $\mathbb{E}^{A'B'} := (E_{a,b}^{A'B'})_{a,b}$ defines an effective distributed measurement [22, 23] on the quantum inputs, termed a "distributed positive-operator-valued measure" (D-POVM); see Fig. 1. Note that by choosing $\{\rho_z^{A'}\}_z$ and $\{\rho_w^{B'}\}_w$ to be tomographically complete sets of states, one can completely reconstruct $\mathbb{E}^{A'B'}$ from the statistics obtained through Eq. (1).

The process matrix formalism makes no *a priori* assumption of a global causal structure relating Alice and Bob. In fact, assuming such a structure imposes further constraints, due to the inability for a party to signal to the causal past. Process matrices compatible, for example, with Alice acting before Bob (denoted $A \prec B$) are of the form $W^{A\prec B} = W^{A\prec B_I}\otimes \mathbb{1}^{B_O}$, and similarly $W^{B\prec A} = W^{B\prec A_I}\otimes \mathbb{1}^{A_O}$ for Bob before Alice ($B \prec A$), with $W^{A\prec B_I} \in \mathcal{L}(\mathcal{H}^{AB_I})$ and $W^{B\prec A_I} \in \mathcal{L}(\mathcal{H}^{A_IB})$ being themselves valid process matrices [1]. Process matrices that can be written as a convex mixture of matrices compatible with $A \prec B$ and $B \prec A$, i.e., of the form

$$W^{AB} = q\, W^{A\prec B_I}\otimes \mathbb{1}^{B_O} + (1-q)\, W^{B\prec A_I}\otimes \mathbb{1}^{A_O}, \tag{3}$$

with $q \in [0,1]$, are said to be "causally separable". They can be interpreted as being compatible with a definite (although perhaps probabilistic) causal order. Remarkably, there exist "causally nonseparable" process matrices that cannot be decomposed as in Eq. (3), and are thus incompatible with any definite causal order [1].

Causal nonseparability can always be certified in a device-dependent manner using a "causal witness" [10, 24], while only some processes can be certified in a DI way through the violation of a causal inequality [1, 10–12]. Here, we start by considering a relaxation of this DI scenario, by providing the parties with quantum inputs instead of classical ones. In [15], we showed that causally separable processes necessarily generate D-POVMs that are themselves causally separable in the sense that they can be written

$$\mathbb{E}^{A'B'} = q\,\mathbb{E}^{A'\prec B'} + (1-q)\,\mathbb{E}^{B'\prec A'} \tag{4}$$

with $q \in [0,1]$, where $\mathbb{E}^{A'\prec B'} = (E_{a,b}^{A'\prec B'})_{a,b}$ and $\mathbb{E}^{B'\prec A'} = (E_{a,b}^{B'\prec A'})_{a,b}$ are D-POVMs satisfying

$$\sum_b E_{a,b}^{A'\prec B'} = E_a^{A'}\otimes \mathbb{1}^{B'} \quad \forall a, \quad \sum_a E_a^{A'} = \mathbb{1}^{A'}, \tag{5}$$

$$\sum_a E_{a,b}^{B'\prec A'} = E_b^{B'}\otimes \mathbb{1}^{A'} \quad \forall b, \quad \sum_b E_b^{B'} = \mathbb{1}^{B'}, \tag{6}$$

and are respectively compatible with the causal order where Alice receives her quantum input and acts before Bob ($A' \prec B'$) and *vice versa* ($B' \prec A'$). Crucially, some causally nonseparable processes which cannot violate any causal inequality in the DI setting can still generate causally nonseparable D-POVMs. This causal nonseparability—and hence that of the process matrix inducing it through Eq. (2)—can then be certified by reconstructing the D-POVM (as mentioned above) and verifying that it cannot be decomposed as in Eq. (4) or, more directly, using these quantum inputs to violate a witness

[2] A full trace $\mathrm{Tr}[(M^Y)^T N^Y]$ and a tensor product $M^X\otimes N^Z$ can both be written as a link product. Moreover, the link product is commutative and associative.

inequality of the form $\mathcal{J} \geq 0$, where $\mathcal{J}$ denotes some linear combination of the probabilities $P(a,b|\rho_z^{A'}, \rho_w^{B'})$ [15].

NSDI-QI and NDI causal scenarios.—A bipartite SDI-QI scenario can be transformed into a 4-partite network scenario (NSDI-QI) where two additional separated parties—Charlie and Daisy—remotely prepare the quantum inputs. They control separate labs with input Hilbert spaces $\mathcal{H}^{C'}$ and $\mathcal{H}^{D'}$, and share maximally entangled pairs of d-dimensional states,[3] $|\phi^+\rangle^{C'A'}$ and $|\phi^+\rangle^{B'D'}$ (with $|\phi^+\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1}|j,j\rangle$) with Alice and Bob, respectively; see Fig. 2 (top). Charlie and Daisy can then remotely prepare the quantum inputs $\rho_z^{A'}$ and $\rho_w^{B'}$ by applying measurements with Choi representations $(M_{c|z}^{C'})_c = (\rho_z^{C'}, \mathbb{1}^{C'} - \rho_z^{C'})$ and $(M_{d|w}^{D'})_d = (\rho_w^{D'}, \mathbb{1}^{D'} - \rho_w^{D'})$ on their shares of entangled qudit pairs. When they observe the outcomes $c = 0$ and $d = 0$ corresponding to the POVM elements $\rho_z^{C'}$ and $\rho_w^{D'}$, respectively, then Alice and Bob receive precisely the desired quantum inputs since $\rho_z^{A'} \propto M_{0|z}^{C'} * |\phi^+\rangle\langle\phi^+|^{C'A'}$ and $\rho_w^{B'} \propto M_{0|w}^{D'} * |\phi^+\rangle\langle\phi^+|^{B'D'}$. Thereby, the observed correlations

$$P(c,a,b,d|z,w) = \left(M_{c|z}^{C'} \otimes E_{a,b}^{A'B'} \otimes M_{d|w}^{D'}\right) * \left(|\phi^+\rangle\langle\phi^+|^{C'A'} \otimes |\phi^+\rangle\langle\phi^+|^{B'D'}\right) \quad (7)$$

satisfy $P(0,a,b,0|z,w) = P(0|z)P(0|w)P(a,b|\rho_z^{A'}, \rho_w^{B'})$ and the causal nonseparability of $\mathbb{E}^{A'B'}$ can be certified by the violation of a witness inequality as in the SDI-QI scenario.

This NSDI-QI scenario so far assumes the sharing of maximally entangled states and trusts Charlie's and Daisy's quantum measurements in the "reference scenario" described above. However, these assumptions can be relaxed by exploiting self-testing techniques. Self-testing exploits the fact that quantum correlations which maximally violate a Bell inequality can sometimes be used to uniquely determine—up to local transformations—which state and measurements produced them [25, 26]. To self-test the maximally entangled states and Charlie and Daisy's measurements, Alice, Bob, Charlie and Daisy are all given (new) classical inputs x, y, z, w, resp., and perform untrusted operations $(\mathsf{M}_{a|x}^{\tilde{A}A})_a, (\mathsf{M}_{b|y}^{\tilde{B}B})_b, (\mathsf{M}_{c|z}^{\tilde{C}})_c, (\mathsf{M}_{d|w}^{\tilde{D}})_d$, producing outcomes a, b, c, d, resp. on the process matrix W^{AB} and the untrusted auxiliary states $\psi_1^{\tilde{C}\tilde{A}}$ and $\psi_2^{\tilde{B}\tilde{D}}$ to be self-tested, which are now in the (untrusted) "physical" Hilbert spaces $\mathcal{H}^{\tilde{A}}, \mathcal{H}^{\tilde{B}}, \mathcal{H}^{\tilde{C}}, \mathcal{H}^{\tilde{D}}$ (denoted with tildes)—that, *a priori*, may differ from those in the reference scenario $\mathcal{H}^{A'}, \mathcal{H}^{B'}, \mathcal{H}^{C'}, \mathcal{H}^{D'}$ (with primes); see Fig. 2

[3] Here d is the dimension of the quantum input spaces, which for simplicity we assume to be the same for both parties (although this can be easily relaxed without consequence).

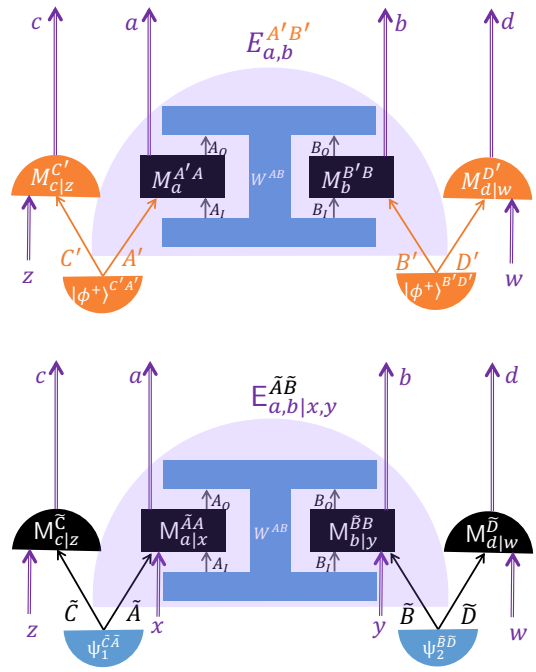

FIG. 2. From the NSDI-QI to the NDI scenario: In a NSDI-QI scenario (top), the quantum inputs of the SDI-QI scenario (Fig. 1) are prepared remotely by two other separated parties—Charlie and Daisy—performing suitable operations on their part of a maximally entangled state shared with Alice and Bob, so that $\rho_z^{A'} \propto M_{0|z}^{C'} * |\phi^+\rangle\langle\phi^+|^{C'A'}$ and $\rho_w^{B'} \propto M_{0|w}^{D'} * |\phi^+\rangle\langle\phi^+|^{B'D'}$ resp. (orange). These quantum inputs can be self-tested (bottom): Alice, Bob, Charlie and Daisy are given classical inputs x, y, z, w, and perform untrusted operations that produce outcomes a, b, c, d, resp. Verifying that the correlations $P(c,a|z,x)$ and $P(b,d|y,w)$ both maximally violate a Bell inequality allows one to self-test the preparation of the quantum inputs; from this, one can certify the causal nonseparability of the D-POVM $(\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}})_{a,b}$ (purple semicircle) for $x = y = \star$, and thus the causal nonseparability of the process matrix W^{AB}, in a NDI scenario.

(bottom). When applied to the process matrix, Alice and Bob's instruments generate a D-POVM $\mathbb{E}_{x,y}^{\tilde{A}\tilde{B}} = (\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}})_{a,b}$ for each x and y with elements $\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}} = \left(\mathsf{M}_{a|x}^{\tilde{A}A} \otimes \mathsf{M}_{b|y}^{\tilde{B}B}\right) * W^{AB}$, similarly to Eq. (2). We will show that in this scenario, the causal nonseparability of a D-POVM generated from a causally nonseparable process matrix W^{AB} can be certified solely based on the correlations $P(c,a,b,d|z,x,y,w)$, thus achieving a NDI certification of causal nonseparability.

NDI certification of causal nonseparability.—As suggested above, the idea to achieve this goal is to combine the SDI-QI certification with self-testing. For the protocol we present to function, it will be necessary for the quantum inputs in the initial SDI-QI setting to be chosen such that that the reference measurements $(M_{c|z}^{C'})_c = (\rho_z^{C'}, \mathbb{1}^{C'} - \rho_z^{C'})$ and $(M_{d|w}^{D'})_d = (\rho_w^{D'}, \mathbb{1}^{D'} - \rho_w^{D'})$

of the NSDI-QI setting are precisely those involved in the self-testing procedure we use. For simplicity, let us assume that $\mathsf{d} = 2^m$ for some $m \geq 1$. This can be done without loss of generality by embedding the reference experiment, i.e., the spaces $\mathcal{H}^{C'}$ and $\mathcal{H}^{D'}$, into larger 2^m-dimensional spaces if needed (see Appendix A). We then require not only that $\rho_z^{C'}$ and $\rho_w^{D'}$ form tomographically complete sets, but more precisely that they are (projectors onto the) eigenstates of orthogonal Pauli measurements (or, beyond the 2-dimensional case, tensor products thereof). We will assume henceforth that this is indeed the case.

When Alice and Bob receive generic classical inputs x, y (other than the specific $x, y = \star$ that we shall consider below), we interpret the experiment as a self-testing round, in which the correlations are used to test a Bell inequality: reaching its maximal quantum violation certifies that the untrusted states $\psi_1^{\tilde{C}\tilde{A}}, \psi_2^{\tilde{B}\tilde{D}}$ and measurements $(\mathsf{M}_{c|z}^{\tilde{C}})_c, (\mathsf{M}_{d|w}^{\tilde{D}})_d$ have the form of those used in the NSDI-QI setup. Such a statement can only be true up to local transformations, but these transformations relating the "physical", uncharacterised, states and measurements to their idealised "reference" counterparts can be explicitly incorporated into the analysis. When they both receive the special input $x = y = \star$, we instead interpret it as a certification round, in which one follows the same procedure as in a NSDI-QI certification of causal nonseparability using these self-tested quantum inputs. Crucially, the aforementioned local transformations preserve causal (non)separability, allowing the causal nonseparability of the physical process to be certified, despite lacking any knowledge of the Hilbert spaces upon which it acts.

The general protocol can thus be described as follows.

- (*Correlation generation*) Alice, Bob, Charlie and Daisy perform local operations and measurements on their respective subsystems to obtain the statistics $P(c, a, b, d|z, x, y, w)$.
- (*Self-testing*) One certifies that the auxiliary states each contain a maximally entangled state and that Charlie and Daisy each effectively perform the desired tomographically complete, orthogonal Pauli measurements on their subsystems by verifying that the distributions $P_y(c, a|z, x) := P(c, a|z, x, y)$ and $P_x(b, d|y, w) := P(b, d|x, y, w)$ both maximally violate a Bell inequality $\mathcal{I} \geq 0$ for some y and x, resp.[4]
- (*Causal nonseparability certification*) One verifies that the correlations $P_\star(a, b|z, w) := P(a, b|z, x{=}\star, y{=}\star, w, c{=}0, d{=}0)$ violate a witness inequality $\mathcal{J} \geq 0$, thus certifying that the process matrix is causally nonseparable.

Let us now explain these steps in more detail. We use the fact that, for $\mathsf{d} = 2^m$, one can self-test a d-dimensional maximally entangled state and orthogonal Pauli measurements (or tensor products thereof, for $\mathsf{d} > 2$) by the maximal quantum violation of a Bell inequality $\mathcal{I} \geq 0$ [17]. Further discussion of the inequality is provided in Appendix A, but its exact form (which depends on d) is not important for the argument here.

The first step of the protocol simply consists in generating the correlations $P(c, a, b, d|z, x, y, w)$, which can be written as

$$\begin{aligned} &P(c, a, b, d|z, x, y, w) \\ &= \left(\mathsf{M}_{c|z}^{\tilde{C}} \otimes \mathsf{M}_{a|x}^{\tilde{A}A} \otimes \mathsf{M}_{b|y}^{\tilde{B}B} \otimes \mathsf{M}_{d|w}^{\tilde{D}}\right) * \left(\psi_1^{\tilde{C}\tilde{A}} \otimes W^{AB} \otimes \psi_2^{\tilde{B}\tilde{D}}\right) \\ &= P(c|z)\, P(d|w)\, \mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}} * \left(\rho_{c|z}^{\tilde{A}} \otimes \rho_{d|w}^{\tilde{B}}\right) \end{aligned} \tag{8}$$

with the remotely prepared states satisfying $P(c|z)\, \rho_{c|z}^{\tilde{A}} = \mathsf{M}_{c|z}^{\tilde{C}} * \psi_1^{\tilde{C}\tilde{A}}$ and $P(d|w)\, \rho_{d|w}^{\tilde{B}} = \mathsf{M}_{d|w}^{\tilde{D}} * \psi_2^{\tilde{B}\tilde{D}}$, and with the D-POVM elements $\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}} = \left(\mathsf{M}_{a|x}^{\tilde{A}A} \otimes \mathsf{M}_{b|y}^{\tilde{B}B}\right) * W^{AB}$, as given above.

The second step is the self-testing procedure. The players verify that the correlations $P_y(c, a|z, x)$ and $P_x(b, d|y, w)$ reach the maximum quantum violation of the Bell inequality $\mathcal{I} \geq 0$. As we recall in Appendix A, it is known that if this is the case, then the physical states $\psi_1^{\tilde{C}\tilde{A}}$ and $\psi_2^{\tilde{B}\tilde{D}}$ and measurements $(\mathsf{M}_{c|z}^{\tilde{C}})_c$ and $(\mathsf{M}_{d|w}^{\tilde{D}})_d$ can be related to the ideal reference states and operations of the NSDI-QI scenario in the d-dimensional spaces $\mathcal{H}^{C'}, \mathcal{H}^{A'}, \mathcal{H}^{B'}, \mathcal{H}^{D'}$ via local isometries [16, 17]. We then show (see Appendix A) that these local isometries also relate the remotely prepared inputs[5] $\rho_{0|z}^{\tilde{A}}$ and $\rho_{0|w}^{\tilde{B}}$ to the reference quantum inputs $\rho_z^{A'} \propto M_{0|z}^{C'} * |\phi^+\rangle\langle\phi^+|^{C'A'}$ and $\rho_w^{B'} \propto M_{0|w}^{D'} * |\phi^+\rangle\langle\phi^+|^{B'D'}$. For example, for Alice's quantum inputs we prove that

$$\rho_{0|z}^{\tilde{A}} = V^{\mathbf{A}\dagger} \left(\rho_z^{A'} \otimes \xi_0^{A^{\mathrm{jf}}} + (\rho_z^{A'})^T \otimes \xi_1^{A^{\mathrm{jf}}}\right) V^{\mathbf{A}}, \tag{9}$$

where $V^{\mathbf{A}} : \mathcal{H}^{\tilde{A}} \to \mathcal{H}^{A'A^{\mathrm{jf}}}$ is an isometry and $\xi_0^{A^{\mathrm{jf}}}$, $\xi_1^{A^{\mathrm{jf}}}$ are (subnormalised) "junk and flag" states on a Hilbert space $\mathcal{H}^{A^{\mathrm{jf}}}$; a similar relation holds for Bob. This allows us to write Eq. (8), when $c = d = 0$, in terms of the reference scenario as

$$\begin{aligned} &P(0, a, b, 0|z, x, y, w) \\ &\qquad = P(0|z)\, P(0|w)\, E_{a,b|x,y}^{A'B'} * \left(\rho_z^{A'} \otimes \rho_w^{B'}\right), \end{aligned} \tag{10}$$

[4] Note that these distributions are independent of w and z, respectively, due to the network structure: e.g., Daisy only shares a (nonsignalling) quantum state with Alice, Bob and Charlie. However, contrary to the scenario of entanglement certification [16, 17], these may still depend on y and x, resp., as the process matrix may allow for signalling.

[5] In fact, we show that each $\rho_{c|z}^{\tilde{A}}$ and $\rho_{d|w}^{\tilde{B}}$ (for all c, d) can be related to a corresponding state in the reference scenario, but this stronger result is not needed here.

where the $E^{A'B'}_{a,b|x,y}$ are elements of an effective D-POVM (which $\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y}$ can likewise be related to through local isometries; see Eq. (B3) in Appendix B). Eq. (10) thus relates directly the observed correlations to those obtainable in the NSDI-QI setup.

The last step consists in certifying causal nonseparability. The self-testing step ensures that, when Charlie and Daisy get the outputs $c = d = 0$, Alice and Bob receive, up to local isometries, the desired quantum inputs $\rho^{A'}_z$ and $\rho^{B'}_w$ from the SDI-QI scenario. We thus focus on the correlations observed when $x = y = \star$ which, in the reference experiment, should correspond to the measurements used to obtain a causally non-separable D-POVM. These correlations, following Eq. (10), are given by

$$\begin{aligned} P_\star(a,b|z,w) &= P(a,b|z,x{=}\star,y{=}\star,w,c{=}0,d{=}0) \\ &= E^{A'B'}_{a,b|\star,\star} * \left(\rho^{A'}_z \otimes \rho^{B'}_w\right), \end{aligned} \tag{11}$$

which precisely defines, following Eq. (1), a distribution $P'(a,b|\rho^{A'}_z,\rho^{B'}_w)$ in a SDI-QI setting with the reference quantum inputs for the certification of the effective D-POVM $\mathbb{E}^{A'B'}_{\text{eff}} := (E^{A'B'}_{a,b|\star,\star})_{a,b}$. In Appendix B we show that if W^{AB} is causally separable then $\mathbb{E}^{A'B'}_{\text{eff}}$ is a causally separable D-POVM. Hence, if $P'(a,b|\rho^{A'}_z,\rho^{B'}_w)$ violates the witness inequality, yielding $\mathcal{J} < 0$, this certifies the causally nonseparability of W^{AB}. By following the ideal reference protocol of the NSDI-QI scenario, one can thus obtain such a NDI certification for any process that can be certified in the bipartite SDI-QI scenario, i.e., which can generate a causally nonseparable D-POVM.

The NDI certification of causal nonseparability described above assumes that the self-testing is perfect: the maximal quantum violation of the Bell inequality $\mathcal{I} \geq 0$ is obtained, meaning that noiseless maximally entangled states and perfect measurements are realised (up to local isometries). Such requirements are experimentally unobtainable. However, the self-testing procedure and subsequent certification of causal nonseparability can be rendered robust to noise in a similar way to that of [17]; see Appendix C for further explanation.

Applications and scope.—The protocol described in the previous section exploits a general approach to transforming a SDI-QI scenario into a NDI one. Indeed, it can be readily generalised to provide NDI certification in settings beyond the basic bipartite one we described. For instance, in [15] we showed that all bipartite causally nonseparable process matrices can be certified in a SDI-QI way if we assume an additional natural assumption on the form of Alice's and Bob's instruments: that they factorise as a joint measurement on one subsystem of their quantum input and the system they receive from the process matrix, and a channel taking the rest of their quantum input into the process matrix. By assuming the same "measurement device and channel independent" (MDCI) structure for the physical instruments $(\mathsf{M}^{\tilde{A}A}_{a|\star})_a$ and $(\mathsf{M}^{\tilde{B}B}_{b|\star})_b$, and using independent auxiliary sources to self-test Alice and Bob's quantum input subsystems, our method here can readily be used to show that all bipartite causally nonseparable processes can also be certified in a NDI way (modulo these assumptions); see Appendix D for details on the soundness of the NDI certification in this MDCI scenario.

A process of particular interest is the quantum switch [3], which can be described in either a "(2+F)-partite" or "(P+2+F)-partite scenario" in which, in addition to Alice and Bob, a third party Fiona with no output Hilbert space performs a measurement in their causal future, and (in the (P+2+F)-partite scenario), a fourth party Phil with no input Hilbert space acts in their causal past. In the quantum switch, Phil prepares a target system and a control qubit, the latter of which coherently controls the order that Alice and Bob then act on the target system, before both the target and control are received by Fiona. (In the (2+F)-partite scenario, Phil's operations are fixed and absorbed into the process matrix.) The quantum switch is the canonical example of a physically realisable causally nonseparable process [13], and the resource behind most known advantages arising from causal indefiniteness [3–8]. Certifying its causal nonseparability is thus a key problem, and several experiments have done this in a device-dependent way using causal witnesses [27, 28]. In [15], the SDI-QI scenario was extended to the (2+F)-scenario to certify the causal nonseparability of the quantum switch, and the generality of our NDI approach can again be illustrated in these scenarios. In Appendix E we further extend the SDI-QI framework to the (P+2+F)-partite scenario—in which the SDI-QI approach can then be transformed into a NDI one by having Phil and Fiona also share maximally entangled states with additional parties, and following the natural generalisation of our bipartite protocol. In fact, for the quantum switch, in the SDI-QI setting it is not necessary to provide all parties with a quantum input: some parties may do with classical inputs, or no input at all. For example, in [15] the case where only Alice and Bob have quantum inputs was considered; in Appendix E we show that one can also certify the causal nonseparability of the quantum switch in a SDI-QI manner while providing only Phil with quantum inputs—a scenario that can then be turned into a NDI one by introducing just one additional party. We thus see that, despite its inability to violate any causal inequality [10, 11], the quantum switch can be certified in a particular type of DI scenario; see also the discussion below.

Discussion.—As we emphasised, our certification of causal nonseparability is *device-independent*, in the sense that it does not assume any knowledge about the underlying process or operations involved. Every element involved in the protocol is treated as a (quantum) black box with classical inputs and outputs, and all conclusions are drawn from the observed correlations. However, we called it *network-device-independent* to differentiate it from the standard approach to DI certification, which requires the violation of a causal inequality [1].

Indeed, our approach differs from the standard one as it assumes a network structure involving two separated untrusted parties, each independently performing operations on a system that is independently shared with a party involved in the process being certified. Moreover, unlike the standard approach which certifies causal indefiniteness in a theory-independent manner, our method certifies the theory-dependent notion of causal nonseparability. Indeed, it is based on the violation of a witness inequality by a causally nonseparable process matrix and on the self-testing of quantum input states through the maximal quantum violation of Bell inequalities. It thus assumes the validity of quantum theory and of its extension through the process matrix formalism. This, and the simple *network* structural assumption, is what allows one to certify—without trusting any device—the causal nonseparability of process matrices which cannot violate a causal inequality, such as the quantum switch [10, 11].

Finally, we note that some recent works have also proposed another approach to certify the indefinite causal order in the quantum switch in a device-independent way [29, 30]. In the scenario presented in [30] (also considered in [29]), an additional space-like separated party is introduced along with Alice, Bob and Fiona, and an inequality is derived based on the assumptions of "Definite causal order, Relativistic causality, and Free interventions" (DRF). It is then shown that the correlations generated from a quantum switch in which the control system is entangled with the newly introduced party can violate this "DRF inequality". Notably, our approach allows us to consider a scenario very close to that of [30] by transforming a SDI-QI test of a $(P+2+F)$-partite quantum switch, with only Phil (in the causal past) being given a quantum input, into a NDI test (as briefly considered above, and further discussed in Appendix E). It is then interesting to compare the assumptions and conclusions one may draw from these two approaches in which the causal indefiniteness of the quantum switch can be certified in a DI manner (see Appendix F for further discussion). A first important difference is that the approach of [30], and the inequality they derive, is not only device-independent, but also theory-independent—like standard causal inequalities, but contrary to our approach. Furthermore, while [30] explicitly introduced some form of no-signalling assumption at the ontological level ("Relativistic causality") between the new party and those involved in the switch, we did not need such an assumption. Instead, we make the simple network structure assumption which implies just the weaker operational notion of no-signalling. With a violation of a "DRF inequality", the conclusion that there is indefinite causal order can hence be escaped by rejecting this relativistic causality assumption; in our case, one would need to reject entirely the network structure assumption to salvage definite causal order (or, in both cases, one could reject the "Free interventions assumption", which we indeed also made implicitly). It will be insightful to further investigate the connections between the two approaches, possibly in different scenarios. Recall indeed that our approach allows one to turn any SDI-QI certification of causal nonseparability into a NDI one; it would be interesting to see if the approach of [30] can similarly be made more systematic.

Acknowledgements.—We thank Jean-Daniel Bancal, Nicolas Brunner and Tein van der Lugt for enlightening discussions, and acknowledge financial support from the EU NextGen Funds, the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA program), the PEPR integrated project EPiQ ANR-22-PETQ-0007 part of Plan France 2030, the ERC Starting grant QUSCO, and the Agence Nationale de la Recherche (project ANR-22-CE47-0012).

For the purpose of open access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

[1] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].

[2] J. Wechs, A. A. Abbott, and C. Branciard, On the definition and characterisation of multipartite causal (non)separability, New J. Phys. **21**, 013027 (2019), arXiv:1807.10557 [quant-ph].

[3] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013), arXiv:0912.0195 [quant-ph].

[4] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301 (2012), arXiv:1109.5154 [quant-ph].

[5] M. Araújo, F. Costa, and Č. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, Phys. Rev. Lett. **113**, 250402 (2014), arXiv:1401.8127 [quant-ph].

[6] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, Exponential communication complexity advantage from quantum superposition of the direction of communication, Phys. Rev. Lett. **117**, 100502 (2016), arXiv:1605.07372 [quant-ph].

[7] D. Ebler, S. Salek, and G. Chiribella, Enhanced communication with the assistance of indefinite causal order, Phys. Rev. Lett. **120**, 120502 (2018), arXiv:1711.10165 [quant-ph].

[8] X. Zhao, Y. Yang, and G. Chiribella, Quantum metrology with indefinite causal order, Phys. Rev. Lett. **124**, 190503 (2020), arXiv:1912.02449 [quant-ph].

[9] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, The simplest causal inequalities and their violation, New J. Phys. **18**, 013008 (2016),

arXiv:1508.01704 [quant-ph].

[10] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New J. Phys. **17**, 102001 (2015), arXiv:1506.03776 [quant-ph].

[11] O. Oreshkov and C. Giarmatzi, Causal and causally separable processes, New J. Phys. **18**, 093020 (2016), arXiv:1506.05449 [quant-ph].

[12] A. Feix, M. Araújo, and Č. Brukner, Causally nonseparable processes admitting a causal model, New. J. Phys. **18**, 083040 (2016), arXiv:1604.03391 [quant-ph].

[13] J. Wechs, H. Dourdent, A. A. Abbott, and C. Branciard, Quantum circuits with classical versus quantum control of causal order, PRX Quantum **2**, 030335 (2021), arXiv:2101.08796 [quant-ph].

[14] T. Purves and A. J. Short, Quantum theory cannot violate a causal inequality, Phys. Rev. Lett. **127**, 110402 (2021), arXiv:2101.09107 [quant-ph].

[15] H. Dourdent, A. A. Abbott, N. Brunner, I. Šupić, and C. Branciard, Semi-device-independent certification of causal nonseparability with trusted quantum inputs, Phys. Rev. Lett. **129**, 090402 (2022), arXiv:2107.10877 [quant-ph].

[16] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, Device-independent entanglement certification of all entangled states, Phys. Rev. Lett. **121**, 180503 (2018), arXiv:1801.10444 [quant-ph].

[17] J. Bowles, I. Šupić, D. Cavalcanti, and A. Acín, Self-testing of Pauli observables for device-independent entanglement certification, Phys. Rev. A **98**, 042336 (2018), arXiv:1801.10446 [quant-ph].

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

[19] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).

[20] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, Phys. Rev. Lett. **101**, 060401 (2008), arXiv:0712.1325 [quant-ph].

[21] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009), arXiv:0904.4483 [quant-ph].

[22] I. Šupić, P. Skrzypczyk, and D. Cavalcanti, Measurement-device-independent entanglement and randomness estimation in quantum networks, Phys. Rev. A **95**, 042340 (2017), arXiv:1702.04752 [quant-ph].

[23] M. J. Hoban and A. B. Sainz, A channel-based framework for steering, non-locality and beyond, New J. Phys. **20**, 053048 (2018), arXiv:1708.00750 [quant-ph].

[24] C. Branciard, Witnesses of causal nonseparability: an introduction and a few case studies, Sci. Rep. **6**, 26018 (2016), arXiv:1603.00043 [quant-ph].

[25] D. Mayers and A. Yao, Self testing quantum apparatus, Quantum Inf. Comput. **4**, 273 (2004), arXiv:quant-ph/0307205.

[26] I. Šupić and J. Bowles, Self-testing of quantum systems: a review, Quantum **4**, 337 (2020), arXiv:1904.10042 [quant-ph].

[27] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Sci. Adv. **3**, e1602589 (2017), arXiv:1608.01683 [quant-ph].

[28] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite causal order in a quantum switch, Phys. Rev. Lett. **121**, 090503 (2018), arXiv:1803.04302 [quant-ph].

[29] S. Gogioso and N. Pinzani, The geometry of causality, arXiv:2303.09017 [quant-ph] (2023).

[30] T. van der Lugt, J. Barrett, and G. Chiribella, Device-independent certification of indefinite causal order in the quantum switch, arXiv:2208.00719 [quant-ph] (2022).

[31] J. Kaniewski, Self-testing of binary observables based on commutation, Phys. Rev. A **95**, 062323 (2017), arXiv:1702.06845 [quant-ph].

[32] J. Bavaresco, M. Araújo, Č. Brukner, and M. T. Quintino, Semi-device-independent certification of indefinite causal order, Quantum **3**, 176 (2019), arXiv:1903.10526 [quant-ph].

APPENDIX

Appendix A: Self-testing the quantum inputs

In this appendix we begin by recalling briefly the framework of self-testing [25, 26], before considering in more detail the self-testing of maximally entangled states and tomographically complete sets of orthogonal Pauli measurements, and how this can be used to self-test the remotely prepared quantum inputs in the NDI scenario.

We will focus here on Charlie and Alice, who share a quantum state $\psi_1^{\tilde{C}\tilde{A}}$ and perform local measurements on their share of the state described by POVMs $(\mathsf{M}^{\tilde{C}}_{c|z})_c$ and $(\mathsf{M}^{\tilde{A}}_{a|x})_a$, respectively. The self-testing results for Bob and Daisy can be derived in an entirely analogous way.

The general goal of self-testing is to use the observed correlations $P(c,a|z,x)$ to show that the state $\psi_1^{\tilde{C}\tilde{A}}$ and/or the measurements $(\mathsf{M}^{\tilde{C}}_{c|z})_c$, $(\mathsf{M}^{\tilde{A}}_{a|x})_a$ must be of a specific, desired form. More precisely, the observed correlations are given by the Born rule as[6]

$$P(c,a|z,x) = (\mathsf{M}^{\tilde{C}}_{c|z} \otimes \mathsf{M}^{\tilde{A}}_{a|x}) * \psi_1^{\tilde{C}\tilde{A}} = \mathrm{Tr}\left[(\mathsf{M}^{\tilde{C}}_{c|z} \otimes \mathsf{M}^{\tilde{A}}_{a|x})^T \psi_1^{\tilde{C}\tilde{A}}\right], \qquad \forall z,x,c,a, \tag{A1}$$

in a "physical" experiment described by $(\psi_1^{\tilde{C}\tilde{A}}, (\mathsf{M}^{\tilde{C}}_{c|z})_c, (\mathsf{M}^{\tilde{A}}_{a|x})_a)$, which we would typically like to relate (in part or wholly) to a "reference" experiment $(|\psi_1\rangle\langle\psi_1|^{C'A'}, (M^{C'}_{c|z})_c, (M^{A'}_{a|x})_a)$. In general, one cannot identify a unique reference experiment compatible with the observed correlations, but only an equivalence class of such experiments. Indeed, there are two classes of transformations of quantum states and measurements which do not change the observed correlations. The first are local isometries, encompassing local rotations and embeddings in Hilbert spaces of higher dimension. The other is a simultaneous transposition of the underlying quantum state and measurements. We thus seek to show a relation between the physical and reference experiments up to such local isometries and transpositions.

Self-testing the maximally entangled state and Pauli measurements

For simplicity, let us focus here on the case of $d = 2$ where the quantum inputs are two-dimensional. We will discuss at the end of the section how the self-testing procedure can be generalised to arbitrary $d > 2$.

In order to self-test the remotely prepared quantum inputs, we will first show how Alice and Charlie can self-test the fact that they effectively share a maximally entangled state $|\phi^+\rangle^{C'A'} = (|00\rangle + |11\rangle)/\sqrt{2}$ and that Charlie's reference measurements $(M^{C'}_{c|z})_c$ effectively correspond to the tomographically complete Pauli measurements $\sigma_x, \sigma_y, \sigma_z$.

Although the physical state $\psi_1^{\tilde{C}\tilde{A}}$ that Charlie and Alice share may not be pure, it will be convenient to work with a purification $|\psi_1\rangle^{\tilde{C}\tilde{A}R}$ of $\psi_1^{\tilde{C}\tilde{A}}$, where $\mathcal{H}^R$ is a purification system that the measurements $(\mathsf{M}^{\tilde{C}}_{c|z})_c$ and $(\mathsf{M}^{\tilde{A}}_{a|x})_a$ act, by definition, trivially upon. Since no operations act nontrivially on this space, we will generally, in a slight abuse of notation, omit writing explicitly this space, writing simply $|\psi_1\rangle^{\tilde{C}\tilde{A}}$ and leave identities acting on $\mathcal{H}^R$ implicit. For simplicity, we will also assume in what follows that the physical measurements $(\mathsf{M}^{\tilde{C}}_{c|z})_c$ and $(\mathsf{M}^{\tilde{A}}_{a|x})_a$ are projective. Although common in self-testing proofs (and justifiable on reasonable grounds [26]), this assumption is not a fundamental one in our NDI protocol, and could be removed using the techniques of [31] to *derive* this projectivity from the maximal violation of a Bell inequality.

The self-testing of a maximally entangled state and the three mutually orthogonal Pauli measurements can be obtained using an extended form of the CHSH inequality, as shown in Lemma 1 of [16, 17]. Here we simply restate the elements of the results necessary for what follows, directing the interested reader to the proofs in [16, 17].

Let us consider an experiment involving three measurements for Charlie and six for Alice, labelled $z = 1, \ldots, 3$ and $x = 1, \ldots, 6$ respectively, all of which are binary valued, and denote $\mathcal{E}_{z,x} = P(c = a|z,x) - P(c \neq a|z,x)$ the expectation value of the correlation between the outputs c and a for inputs z and x. Starting from the extended CHSH inequality

$$\mathcal{E} := \mathcal{E}_{1,1} + \mathcal{E}_{1,2} + \mathcal{E}_{2,1} - \mathcal{E}_{2,2} + \mathcal{E}_{1,3} + \mathcal{E}_{1,4} - \mathcal{E}_{3,3} + \mathcal{E}_{3,4} + \mathcal{E}_{2,5} + \mathcal{E}_{2,6} - \mathcal{E}_{3,5} + \mathcal{E}_{3,6} \leq 6 \tag{A2}$$

---

[6] Note that as we work here with the Choi representation, the measurement operators are transposed with respect to those that one would usually write (e.g., the Choi representation for a projector onto a state $|\phi\rangle$ is $(|\phi\rangle\langle\phi|)^T$). This also explains the presence of transposes on $(\mathsf{M}^{\tilde{C}}_{c|z})^T$ and $(M^{C'}_{c|z})^T$ in the self-testing statement of Eq. (A5) and in the subsequent equations below, compared to the notations used in the literature [16, 17].

obtained by the sum of three CHSH inequalities, each with a local bound of 2, let us consider the Bell inequality $\mathcal{I} = 6 - \mathcal{E} \geq 0$. In the experiment we consider (described by quantum theory), the correlators $\mathcal{E}_{z,x}$ are obtained as

$$\mathcal{E}_{z,x} = (\mathsf{M}^{\tilde{C}}_{0|z} \otimes \mathsf{M}^{\tilde{A}}_{0|x}) * \psi_1^{\tilde{C}\tilde{A}} + (\mathsf{M}^{\tilde{C}}_{1|z} \otimes \mathsf{M}^{\tilde{A}}_{1|x}) * \psi_1^{\tilde{C}\tilde{A}} - (\mathsf{M}^{\tilde{C}}_{0|z} \otimes \mathsf{M}^{\tilde{A}}_{1|x}) * \psi_1^{\tilde{C}\tilde{A}} - (\mathsf{M}^{\tilde{C}}_{1|z} \otimes \mathsf{M}^{\tilde{A}}_{0|x}) * \psi_1^{\tilde{C}\tilde{A}}. \tag{A3}$$

Since $\mathcal{E}$ is the sum of three CHSH inequalities, its maximum quantum value is $\mathcal{E}_Q = 6\sqrt{2}$, and the maximum quantum violation of $\mathcal{I} \geq 0$ is $\mathcal{I}_Q = 6 - 6\sqrt{2} < 0$. If this value is reached by Alice and Charlie's correlations $P_y(c, a|z, x)$ (for some input y of Bob) in their physical experiment, then by Lemma 1 of Refs. [16, 17] there exist local isometries $V^{\mathbf{C}} : \mathcal{H}^{\tilde{C}} \to \mathcal{H}^{C'C^{\mathrm{j}}C^{\mathrm{f}}}$ and $V^{\mathbf{A}} : \mathcal{H}^{\tilde{A}} \to \mathcal{H}^{A'A^{\mathrm{j}}A^{\mathrm{f}}}$, where C^{j} and A^{j} are "junk" systems and C^{f} and A^{f} are (2-dimensional) "flag" systems, such that

$$V^{\mathbf{C}} \otimes V^{\mathbf{A}} \left|\psi_1\right\rangle^{\tilde{C}\tilde{A}} = \left|\phi^+\right\rangle^{C'A'} \otimes \left|\xi_0\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|00\right\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}} + \left|\phi^+\right\rangle^{C'A'} \otimes \left|\xi_1\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|11\right\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}} \tag{A4}$$

and

$$V^{\mathbf{C}} \otimes V^{\mathbf{A}} \left((\mathsf{M}^{\tilde{C}}_{c|z})^T \left|\psi_1\right\rangle^{\tilde{C}\tilde{A}}\right) = \left((M^{C'}_{c|z})^T \left|\phi^+\right\rangle^{C'A'}\right) \otimes \left|\xi_0\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|00\right\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}} + \left(M^{C'}_{c|z} \left|\phi^+\right\rangle^{C'A'}\right) \otimes \left|\xi_1\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|11\right\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}}, \tag{A5}$$

where we omit identity operators acting on Alice's side, the junk states $\left|\xi_0\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}}$ and $\left|\xi_1\right\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}}$ satisfy $\| \left|\xi_0\right\rangle \|^2 + \| \left|\xi_1\right\rangle \|^2 = 1$, and Charlie's reference measurements $M^{C'}_{c|z}$ are precisely the tomographically complete set of projectors

$$(M^{C'}_{c|1})^T = \tfrac{1}{2}(\mathbb{1} + (-1)^c \sigma_Z), \quad (M^{C'}_{c|2})^T = \tfrac{1}{2}(\mathbb{1} + (-1)^c \sigma_X), \quad (M^{C'}_{c|3})^T = \tfrac{1}{2}(\mathbb{1} + (-1)^c \sigma_Y). \tag{A6}$$

From this self-testing result, we can determine how the remotely prepared inputs $\rho^{\tilde{A}}_{c|z} \propto \mathsf{M}^{\tilde{A}}_{c|z} * \psi_1^{\tilde{C}\tilde{A}}$ relate to the corresponding reference quantum inputs $\rho^{A'}_{c|z} \propto M^{C'}_{c|z} * \left|\phi^+\right\rangle\!\left\langle\phi^+\right|^{C'A'}$, which here are precisely the tomographically complete states $\rho^{A'}_{c|z} = (M^{C'}_{c|z})^T$ specified in Eq. (A6). Multiplying Eqs. (A4) and (A5) by $V^{\mathbf{C}\dagger} \otimes V^{\mathbf{A}\dagger}$, and recalling that $\psi_1^{\tilde{C}\tilde{A}} = \mathrm{Tr}_R \left|\psi_1\right\rangle\!\left\langle\psi_1\right|^{\tilde{C}\tilde{A}R}$ (with the implicit purification space made explicit here, but left implicit below on both the state and in the partial trace), it follows that (again with some implicit identity operators)

$$\begin{aligned}
\mathsf{M}^{\tilde{C}}_{c|z} * \psi_1^{\tilde{C}\tilde{A}} &= \mathrm{Tr}_{\tilde{C}} \left((\mathsf{M}^{\tilde{C}}_{c|z})^T \left|\psi_1\right\rangle\!\left\langle\psi_1\right|^{\tilde{C}\tilde{A}}\right) \\
&= V^{\mathbf{A}\dagger} \left(\mathrm{Tr}_{\tilde{C}} \left[V^{\mathbf{C}\dagger} \left(\sum_{k,l=0,1} \left((M^{C'}_{c|z})^{T^{\bar{k}}} \left|\phi^+\right\rangle\!\left\langle\phi^+\right|^{C'A'}\right) \otimes \left|\xi_k\right\rangle\!\left\langle\xi_l\right|^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|kk\right\rangle\!\left\langle ll\right|^{C^{\mathrm{f}}A^{\mathrm{f}}}\right) V^{\mathbf{C}}\right]\right) V^{\mathbf{A}} \\
&= V^{\mathbf{A}\dagger} \left(\mathrm{Tr}_{C'C^{\mathrm{j}}C^{\mathrm{f}}} \left[V^{\mathbf{C}} V^{\mathbf{C}\dagger} \left(\sum_{k,l=0,1} \left((M^{C'}_{c|z})^{T^{\bar{k}}} \left|\phi^+\right\rangle\!\left\langle\phi^+\right|^{C'A'}\right) \otimes \left|\xi_k\right\rangle\!\left\langle\xi_l\right|^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|kk\right\rangle\!\left\langle ll\right|^{C^{\mathrm{f}}A^{\mathrm{f}}}\right)\right]\right) V^{\mathbf{A}},
\end{aligned} \tag{A7}$$

where $(\cdot)^{T^{\bar{k}}}$ denotes the "transpose to the power $\bar{k} := 1 - k$". Noting now that $V^{\mathbf{C}} V^{\mathbf{C}\dagger}$ effectively acts as the identity on the state of Eq. (A5), it can be removed from the previous equation. The partial trace then kills the cross terms with $k \neq l$ and we obtain

$$P(c|z)\, \rho^{\tilde{A}}_{c|z} = \mathsf{M}^{\tilde{C}}_{c|z} * \psi_1^{\tilde{C}\tilde{A}} = P(c|z)\, V^{\mathbf{A}\dagger} \left(\sum_{k=0,1} (\rho^{A'}_{c|z})^{T^k} \otimes \xi_k^{A^{\mathrm{jf}}}\right) V^{\mathbf{A}}, \tag{A8}$$

with $A^{\mathrm{jf}} = A^{\mathrm{j}}A^{\mathrm{f}}$, $\xi_k^{A^{\mathrm{jf}}} = \mathrm{Tr}_{C^{\mathrm{j}}} \left|\xi_k\right\rangle\!\left\langle\xi_k\right|^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes \left|k\right\rangle\!\left\langle k\right|^{A^{\mathrm{f}}}$ such that $\sum_k \mathrm{Tr}(\xi_k^{A^{\mathrm{jf}}}) = 1$, and $P(c|z)\, \rho^{A'}_{c|z} = \mathrm{Tr}_{C'} \left((M^{C'}_{c|z})^T \left|\phi^+\right\rangle\!\left\langle\phi^+\right|^{C'A'}\right) = M^{C'}_{c|z} * \left|\phi^+\right\rangle\!\left\langle\phi^+\right|^{C'A'}$.[7]

By the symmetry of the considered network scenario (cf. Eq. (8) and Fig. 2), we obtain a similar self-testing statement for Bob and Daisy: provided that their correlations $P_x(b, d|y, w)$ also violate the Bell inequality Eq. (A2)

[7] One may note that up to a proportionality constant, $\rho^{A'}_{c|z}$ is formally the same matrix as $M^{C'}_{c|z}$, in the space A' rather than C'. (This can indeed be interpreted as some form of quantum teleportation from system C' to system A'.)

up to its quantum maximum (for some input x of Alice), there exists a local isometry of the form $V^{\mathbf{B}} : \mathcal{H}^{\tilde{B}} \to \mathcal{H}^{B'B^{\mathrm{jf}}}$ such that we can write

$$P(d|w)\,\rho^{\tilde{B}}_{d|w} = \mathsf{M}^{\tilde{D}}_{d|w} * \psi_2^{\tilde{B}\tilde{D}} = P(d|w)\,V^{\mathbf{B}^\dagger}\left(\sum_{l=0,1}(\rho^{B'}_{d|w})^{T^l}\otimes\zeta_l^{B^{\mathrm{jf}}}\right)V^{\mathbf{B}}, \tag{A9}$$

with $\sum_l \mathrm{Tr}(\zeta_l^{B^{\mathrm{jf}}}) = 1$ and $P(d|w)\,\rho^{B'}_{d|w} = \mathrm{Tr}_{D'}\left((M^{D'}_{d|w})^T\,|\phi^+\rangle\!\langle\phi^+|^{B'D'}\right) = M^{D'}_{d|w} * |\phi^+\rangle\!\langle\phi^+|^{B'D'}$.

Thus, we have shown that both remotely prepared inputs $\rho^{\tilde{A}}_{c|z} \propto \mathsf{M}^{\tilde{C}}_{c|z} * \psi_1^{\tilde{C}\tilde{A}}$ and $\rho^{\tilde{B}}_{d|w} \propto \mathsf{M}^{\tilde{D}}_{d|w} * \psi_2^{\tilde{B}\tilde{D}}$ can be related to the reference quantum inputs $\rho^{A'}_{c|z} \propto M^{C'}_{c|z} * |\phi^+\rangle\!\langle\phi^+|^{C'A'}$ and $\rho^{B'}_{d|w} \propto M^{D'}_{d|w} * |\phi^+\rangle\!\langle\phi^+|^{B'D'}$ via local isometries. That is, we are able to turn the self-test of the maximally entangled state that Charlie and Alice share, along with Charlie's tomographically complete set of orthogonal Pauli measurements, into a self-test of the quantum inputs used in the reference NSDI-QI certification protocol (and likewise for Bob and Daisy). Note that in the NDI protocol presented in the main text, we assumed that Charlie's (tomographically complete) quantum inputs in the reference scenario are obtained as the remotely prepared states when $c = 0$ for the different inputs z, whereas here the $c = 0$ outcomes only give one projector for each measurement, which is not quite a tomographically complete set. This detail can be overcome by considering some additional measurements in the NDI protocol obtained by relabelling some of the above measurements' outcomes, so as to complete the set of reference quantum inputs (and likewise for Daisy's measurements).

Finally, while the above approach, presented explicitly for the case of two-dimensional quantum inputs (i.e., $\mathsf{d} = 2$), it can be readily generalised to $\mathsf{d} = 2^m$, for $m > 1$. Indeed, the maximally entangled state $|\phi^+_{\mathsf{d}}\rangle^{C'A'} = \frac{1}{\sqrt{\mathsf{d}}}\sum_{j=0}^{\mathsf{d}-1}|j,j\rangle$, for $\mathsf{d} = 2^m$, can be self-tested through the parallel repetition of the extended CHSH inequality (A2) [17]. In this case, Charlie has 3^m measurement settings corresponding to the nontrivial tensor products of orthogonal single-qubit Pauli measurements, each of which has 2^m outcomes so that the $M^{C'}_{c|z}$ one self-tests are precisely the projectors onto the 6^m eigenstates of m-fold Pauli measurements. As in the $\mathsf{d} = 2$ case, the quantum inputs prepared on the $c = 0$ outcomes can be extended to a tomographically complete set by adding some additional measurements with relabelled outcomes.

For $\mathsf{d} \neq 2^m$, the above procedure can again be utilised by first embedding the reference scenario into some d'-dimensional spaces for some $\mathsf{d}' = 2^m > \mathsf{d}$. By writing the D-POVM $\mathbb{E}^{A'B'}$ in these d'-dimensional spaces (and taking $\mathcal{H}^{C'}$ and $\mathcal{H}^{D'}$ to hence also be d'-dimensional) one can then proceed using the general approach described above for $\mathsf{d}' = 2^m$ before restricting ourselves to quantum inputs in the d-dimensional subspace of interest.

Appendix B: Certifying causal nonseparability in a network-device-independent manner

The self-testing statements above allow us to directly relate the observed correlations to those of the ideal reference systems in the NSDI-QI setup. To see this, let us define for convenience $|V^{\mathbf{A}^\dagger}\rangle\!\rangle^{A'A^{\mathrm{jf}}\tilde{A}} := \sum_i |i\rangle^{A'A^{\mathrm{jf}}} \otimes V^{\mathbf{A}^\dagger}|i\rangle^{A'A^{\mathrm{jf}}}$, where $\{|i\rangle^{A'A^{\mathrm{jf}}}\}_i$ is the computational basis of $\mathcal{H}^{A'A^{\mathrm{jf}}}$. With these notations one can write, using the link product,

$$V^{\mathbf{A}^\dagger}\big[(\rho^{A'}_{c|z})^{T^k}\otimes\xi_k^{A^{\mathrm{jf}}}\big]V^{\mathbf{A}} = |V^{\mathbf{A}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}^\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \big[(\rho^{A'}_{c|z})^{T^k}\otimes\xi_k^{A^{\mathrm{jf}}}\big] = \left(|V^{\mathbf{A}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}^\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi_k^{A^{\mathrm{jf}}}\right)^{T^k_{A'}} * \rho^{A'}_{c|z}, \tag{B1}$$

and similarly for $V^{\mathbf{B}^\dagger}\big[(\rho^{B'}_{d|w})^{T^l}\otimes\zeta_l^{B^{\mathrm{jf}}}\big]V^{\mathbf{B}}$, introducing $|V^{\mathbf{B}^\dagger}\rangle\!\rangle^{B'B^{\mathrm{jf}}\tilde{B}}$ in a similar way to $|V^{\mathbf{A}^\dagger}\rangle\!\rangle^{A'A^{\mathrm{jf}}\tilde{A}}$ above. From Eqs. (A8) and (A9), the correlations $P(c,a,b,d|z,x,y,w)$ can then be written as

$$\begin{aligned}
P(c,a,b,d|z,x,y,w) &= \left(\mathsf{M}^{\tilde{C}}_{c|z}\otimes\mathsf{M}^{\tilde{A}A}_{a|x}\otimes\mathsf{M}^{\tilde{B}B}_{b|y}\otimes\mathsf{M}^{\tilde{D}}_{d|w}\right) * \left(\psi_1^{\tilde{C}\tilde{A}}\otimes W^{AB}\otimes\psi_2^{\tilde{B}\tilde{D}}\right)\\
&= P(c|z)\,P(d|w)\left(\mathsf{M}^{\tilde{A}A}_{a|x}\otimes\mathsf{M}^{\tilde{B}B}_{b|y}\right) * W^{AB} * \left[\left(\sum_{k=0,1}V^{\mathbf{A}^\dagger}\big[(\rho^{A'}_{c|z})^{T^k}\otimes\xi_k^{A^{\mathrm{jf}}}\big]V^{\mathbf{A}}\right)\otimes\left(\sum_{l=0,1}V^{\mathbf{B}^\dagger}\big[(\rho^{B'}_{d|w})^{T^l}\otimes\zeta_l^{B^{\mathrm{jf}}}\big]V^{\mathbf{B}}\right)\right]\\
&= P(c|z)\,P(d|w)\,\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \left[\left(\sum_{k=0,1}\left(|V^{\mathbf{A}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}^\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi_k^{A^{\mathrm{jf}}}\right)^{T^k_{A'}} * \rho^{A'}_{c|z}\right)\otimes\left(\sum_{l=0,1}\left(|V^{\mathbf{B}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}^\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta_l^{B^{\mathrm{jf}}}\right)^{T^l_{B'}} * \rho^{B'}_{d|w}\right)\right]\\
&= P(c|z)\,P(d|w)\;E^{A'B'}_{a,b|x,y} * \left(\rho^{A'}_{c|z}\otimes\rho^{B'}_{d|w}\right)
\end{aligned} \tag{B2}$$

with the "physical" D-POVM elements $\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} = \left(\mathsf{M}^{\tilde{A}A}_{a|x} \otimes \mathsf{M}^{\tilde{B}B}_{b|y}\right) * W^{AB}$ and the new D-POVM elements effectively acting on the "reference" quantum input states

$$E^{A'B'}_{a,b|x,y} := \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \left[\sum_{k=0,1} \left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi^{A^{\mathrm{jf}}}_k \right)^{T^k_{A'}} \otimes \sum_{l=0,1} \left(|V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l \right)^{T^l_{B'}} \right]. \tag{B3}$$

The correlations (cf. Eq. (11))

$$P_\star(a,b|z,w) = P(a,b|z,x{=}\star,y{=}\star,w,c{=}0,d{=}0) = E^{A'B'}_{a,b|\star,\star} * \left(\rho^{A'}_z \otimes \rho^{B'}_w\right) = P'(a,b|\rho^{A'}_z,\rho^{B'}_w) \tag{B4}$$

thus describe precisely the correlations in a SDI-QI setting with the reference quantum inputs $\rho^{A'}_z$ and $\rho^{B'}_w$ and the effective D-POVM $\mathbb{E}^{A'B'}_{\mathrm{eff}} = (E^{A'B'}_{a,b|*,*})_{a,b}$ involving the D-POVM elements defined in Eq. (B3).

We now show that the local transformations used in this self-testing procedure preserve causal separability, thus allowing us to certify the causal nonseparability of the physical D-POVM $\mathbb{E}^{\tilde{A}\tilde{B}} = (\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|*,*})_{a,b}$—and hence the underlying physical process W^{AB}—despite lacking any knowledge of the Hilbert spaces upon which it acts.

Let us for that assume that the process matrix is compatible with the causal order $A \prec B$. Such a process matrix takes the form $W = W^{A\prec B_I} \otimes \mathbb{1}^{B_O}$, with $\mathrm{Tr}_{B_I} W^{A\prec B_I} = W^{A_I} \otimes \mathbb{1}^{A_O}$ and $\mathrm{Tr}\, W^{A_I} = 1$. Using the fact that $\sum_b \mathsf{M}^{\tilde{B}B}_{b|y} * \mathbb{1}^{B_O} = \sum_b \mathrm{Tr}_{B_O} \mathsf{M}^{\tilde{B}B}_{b|y} = \mathbb{1}^{\tilde{B}B_I}$ according to the overall trace-preserving condition for a quantum instrument, we have

$$\sum_b \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} = \sum_b \left(\mathsf{M}^{\tilde{A}A}_{a|x} \otimes \mathsf{M}^{\tilde{B}B}_{b|y}\right) * W^{A\prec B_I} \otimes \mathbb{1}^{B_O} = \left(\mathsf{M}^{\tilde{A}A}_{a|x} \otimes \mathbb{1}^{\tilde{B}B_I}\right) * W^{A\prec B_I} = \mathsf{E}^{\tilde{A}}_{a|x} \otimes \mathbb{1}^{\tilde{B}} \tag{B5}$$

with $\mathsf{E}^{\tilde{A}}_{a|x} := \mathsf{M}^{\tilde{A}A}_{a|x} * \mathrm{Tr}_{B_I} W^{A\prec B_I} = \mathrm{Tr}_{A_O} \mathsf{M}^{\tilde{A}A}_{a|x} * W^{A_I}$ (such that $(\mathsf{E}^{\tilde{A}}_{a|x})_a$ forms Alice's marginal POVM applied to her physical quantum input). From Eq. (B3), we then obtain[8]

$$\begin{aligned}
\sum_b E^{A'B'}_{a,b|x,y} &= \mathsf{E}^{\tilde{A}}_{a|x} \otimes \mathbb{1}^{\tilde{B}} * \left[\sum_{k=0,1} \left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi^{A^{\mathrm{jf}}}_k \right)^{T^k_{A'}} \otimes \sum_{l=0,1} \left(|V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l \right)^{T^l_{B'}} \right] \\
&= \left[\mathsf{E}^{\tilde{A}}_{a|x} * \sum_{k=0,1} \left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi^{A^{\mathrm{jf}}}_k \right)^{T^k_{A'}} \right] \otimes \sum_{l=0,1} \left(\mathrm{Tr}_{\tilde{B}}\, |V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l \right)^{T^l} \\
&= E^{A'}_{a|x} \otimes \mathbb{1}^{B'}.
\end{aligned} \tag{B6}$$

with

$$E^{A'}_{a|x} := \mathsf{E}^{\tilde{A}}_{a|x} * \sum_{k=0,1} \left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi^{A^{\mathrm{jf}}}_k \right)^{T^k_{A'}}, \tag{B7}$$

which now forms a single-partite POVM acting on Alice's reference quantum input (one can in particular check that $\sum_a \mathsf{E}^{\tilde{A}}_{a|x} = \mathbb{1}^{\tilde{A}}$ and $\sum_a E^{A'}_{a|x} = \mathbb{1}^{A'}$).

Thus, from Eq. (B6) we can conclude that, starting from a causally ordered $W = W^{A\prec B_I} \otimes \mathbb{1}^{B_O}$ compatible with $A \prec B$, we obtain a D-POVM $(E^{A'B'}_{a,b|x,y})_{a,b}$ compatible with the order $A' \prec B'$. Similarly, from a process matrix compatible with the causal order $B \prec A$ we obtain a D-POVM $(E^{A'B'}_{a,b|x,y})_{a,b}$ compatible with the order $B' \prec A'$. We can then conclude that any D-POVM $(E^{A'B'}_{a,b|x,y})_{a,b}$ induced by a causally separable process matrix is causally separable, as defined in [15].

---

[8] Here we use the fact that $\mathrm{Tr}_{\tilde{B}}\, |V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} = (V^{\mathbf{B}}V^{\mathbf{B}\dagger})^T$, so that $\mathrm{Tr}_{\tilde{B}}\, |V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l = \mathrm{Tr}_{B^{\mathrm{jf}}} \left(V^{\mathbf{B}}V^{\mathbf{B}\dagger}(\mathbb{1}^{B'} \otimes \zeta^{B^{\mathrm{jf}}}_l)\right)^T$. To evaluate this further, note that $V^{\mathbf{A}}V^{\mathbf{A}\dagger}$ acts as the identity on the state of Eq. (A5), and therefore on the state obtained after tracing out over $C'C^{\mathrm{jf}}$—that is, on $\sum_k (\rho^{A'}_{c|z})^{T^k} \otimes \xi^{A^{\mathrm{jf}}}_k$, for some tomographically complete set of states $\rho^{A'}_{c|z}$. The same holds on Bob's side, so that $V^{\mathbf{B}}V^{\mathbf{B}\dagger}[\sum_l (\rho^{B'})^{T^{\bar{l}}} \otimes \zeta^{B^{\mathrm{jf}}}_l] = \sum_l (\rho^{B'})^{T^{\bar{l}}} \otimes \zeta^{B^{\mathrm{jf}}}_l$ for all $\rho^{B'}$. It then follows that $\mathrm{Tr}\left[\left(\sum_l \left(\mathrm{Tr}_{\tilde{B}}\, |V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l\right)^{T^l}\right)\rho^{B'}\right] = \mathrm{Tr}\left[\sum_l \mathrm{Tr}_{B^{\mathrm{jf}}} \left(V^{\mathbf{B}}V^{\mathbf{B}\dagger}(\mathbb{1}^{B'} \otimes \zeta^{B^{\mathrm{jf}}}_l)\right)^{T^{\bar{l}}} \rho^{B'}\right] = \mathrm{Tr}\left[V^{\mathbf{B}}V^{\mathbf{B}\dagger}[\sum_l (\rho^{B'})^{T^{\bar{l}}} \otimes \zeta^{B^{\mathrm{jf}}}_l]\right] = \mathrm{Tr}\left[\sum_l (\rho^{B'})^{T^{\bar{l}}} \otimes \zeta^{B^{\mathrm{jf}}}_l\right] = (\mathrm{Tr}\,\rho^{B'})(\sum_l \mathrm{Tr}\,\zeta^{B^{\mathrm{jf}}}_l) = \mathrm{Tr}\,\rho^{B'}$ for all $\rho^{B'}$, which implies that $\sum_l \left(\mathrm{Tr}_{\tilde{B}}\, |V^{\mathbf{B}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta^{B^{\mathrm{jf}}}_l\right)^{T^l} = \mathbb{1}^{B'}$.

This is the case, in particular, when $x = y = \star$, so that the effective D-POVM $\mathbb{E}_{\text{eff}}^{A'B'} = (E_{a,b|*,*}^{A'B'})_{a,b}$ is causally separable whenever W^{AB} is causally separable. As a result, the correlations $P'(a,b|\rho_z^{A'}, \rho_w^{B'})$ defined in Eq. (B4) obtained from a causally separable process W^{AB} must satisfy any witness inequality $\mathcal{J} \geq 0$. If these correlations violate such an inequality, giving $\mathcal{J} < 0$, then this therefore proves the causal nonseparability of the process in a NDI manner.

Such a violation—and hence NDI certification—can moreover be obtained for any process W^{AB} that can (for appropriately chosen reference instruments $(M_a^{A'A})_a$ and $(M_b^{B'B})_b$) induce a causally nonseparable D-POVM $\mathbb{E}^{A'B'} = (E_{a,b}^{A'B'})_{a,b}$ in the reference scenario, i.e., which cannot be decomposed according to Eqs. (4)–(6). Indeed, as outlined in [15], the causal nonseparability of $\mathbb{E}^{A'B'}$ can always be witnessed by a family $\mathbb{S}^{A'B'} = (S_{a,b}^{A'B'})_{a,b}$ of operators (a "witness") such that $\sum_{a,b} S_{a,b}^{A'B'} * E_{a,b}^{A'B'} < 0$ only if $\mathbb{E}^{A'B'}$ is causally nonseparable. Concretely, writing $S_{a,b}^{A'B'} = \sum_{z,w} s_{a,b}^{(z,w)} \rho_z^{A'} \otimes \rho_w^{B'}$ using the tomographically complete sets of states from the NDI-protocol, one obtains the witness inequality

$$\mathcal{J} := \sum_{a,b,z,w} s_{a,b}^{(z,w)} P'(a,b|\rho_z^{A'}, \rho_w^{B'}) \geq 0. \tag{B8}$$

By implementing precisely the reference NSDI-QI protocol, one simply has $E_{a,b}^{A'B'} = E_{a,b|*,*}^{A'B'} = \mathsf{E}_{a,b|*,*}^{\tilde{A}\tilde{B}}$ for all a, b so that the reference, effective, and physical D-POVMs all coincide, and the witness inequality can indeed, as expected, be violated.

Appendix C: Robustness of the procedure

The exact self-testing conditions (i.e., strictly reaching the maximal quantum violation of the Bell inequality under consideration) are never met in any realistic scenario. Thus, to make self-testing applicable in practice one needs to ensure that it is "robust". In particular, the reference and physical experiments should be sufficiently close if the self-testing correlations are approximately reproduced by the physical experiment. In this section we discuss how our NDI certification procedure is affected if the states shared by Alice and Charlie and by Bob and Daisy approximately reproduce the self-testing correlations.

The robust self-testing of the maximally entangled pair of qubits (resp., qudits) is discussed in Appendix B (resp., E) of [17]. Let us restrict ourselves again here, for simplicity, to the case of qubits (for qudits one can adapt the approach of [17] in a similar manner). The robust self-testing statements—i.e., the robust version of Eq. (A5)—then take the form[9]

$$\Big\| V^{\mathbf{C}} \otimes V^{\mathbf{A}} \left((\mathsf{M}_{c|z}^{\tilde{C}})^T |\psi_1\rangle^{\tilde{C}\tilde{A}} \right) - \left(\left((M_{c|z}^{T})^{C'} |\phi^+\rangle^{C'A'} \right) \otimes |\xi_0\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes |00\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}} + \left(M_{c|z}^{C'} |\phi^+\rangle^{C'A'} \right) \otimes |\xi_1\rangle^{C^{\mathrm{j}}A^{\mathrm{j}}} \otimes |11\rangle^{C^{\mathrm{f}}A^{\mathrm{f}}} \right) \Big\| \leq \theta, \tag{C1}$$

where θ depends on the distance between observed and reference correlations (e.g., in the $d = 2$ case, if the violation of the Bell inequality (A2) reaches some value $6\sqrt{2} - \epsilon$, one can take $\theta = \kappa\sqrt{\epsilon}$, for some multiplicative factor κ [17]). Denoting by $|\Psi_{c|z}\rangle$ the state in the right-hand side of Eq. (A5) (in the noiseless case), this equation can be rewritten as

$$V^{\mathbf{C}} \otimes V^{\mathbf{A}} \left((\mathsf{M}_{c|z}^{\tilde{C}})^T |\psi_1\rangle^{\tilde{C}\tilde{A}} \right) = |\Psi_{c|z}\rangle^{C'C^{\mathrm{jf}}A'A^{\mathrm{jf}}} + |\omega_{c|z}\rangle^{C'C^{\mathrm{jf}}A'A^{\mathrm{jf}}}, \tag{C2}$$

where $|\omega_{c|z}\rangle$ is an unknown state with vector norm smaller than or equal to θ. Eq. (A8) then becomes, in the noisy case,

$$\begin{aligned} P(c|z)\, \rho_{c|z}^{\tilde{A}} = \mathsf{M}_{c|z}^{\tilde{C}} * \psi_1^{\tilde{C}\tilde{A}} &= V^{\mathbf{A}\dagger} \operatorname{Tr}_{C'C^{\mathrm{jf}}} \left((|\Psi_{c|z}\rangle + |\omega_{c|z}\rangle)(\langle\Psi_{c|z}| + \langle\omega_{c|z}|) \right) V^{\mathbf{A}} \\ &= P(c|z)\, V^{\mathbf{A}\dagger} \left(\sum_{k=0,1} (\rho_{c|z}^{A'})^{T^k} \otimes \xi_k^{A^{\mathrm{jf}}} \right) V^{\mathbf{A}} + \delta_{c|z}^{\tilde{A}}, \end{aligned} \tag{C3}$$

[9] Recall that we assume, for simplicity, that the physical measurements $(\mathsf{M}_{c|z}^{\tilde{C}})_c$ and $(\mathsf{M}_{a|x}^{\tilde{A}})_a$ are projective. In the noisy case this projectivity property cannot be derived directly by using the techniques of [31], which require the maximal violation of a Bell inequality, although we expect that the noise-robust statements of [31] should still allow one to evade the strict requirement for the projectivity assumption (working out the details is left as an open question).

with

$$\delta^{\tilde{A}}_{c|z} = V^{\mathbf{A}^\dagger} \operatorname{Tr}_{C'C^{\mathrm{jf}}} \left(|\Psi_{c|z}\rangle\langle\omega_{c|z}| + |\omega_{c|z}\rangle\langle\Psi_{c|z}| + |\omega_{c|z}\rangle\langle\omega_{c|z}| \right) V^{\mathbf{A}}. \tag{C4}$$

The norm of this correction term can be bounded in the same way as it is done in Appendix G of [17]. Namely, using the triangle inequality and noting that $\| \, |\Psi_{c|z}\rangle \, \| = \frac{1}{\sqrt{2}}$, we get

$$\left\| \delta^{\tilde{A}}_{c|z} \right\|_1 \leq \sqrt{2}\,\theta + \theta^2, \tag{C5}$$

where $\| \cdot \|_1$ is the trace norm.

Similar statements can be made on Bob and Daisy's side, allowing one to write $\mathsf{M}^{\tilde{D}}_{d|w} * \psi_2^{\tilde{B}\tilde{D}}$ in a similar way to Eq. (C3), in terms of a correction term $\delta^{\tilde{B}}_{d|w}$ whose trace norm is also upper-bounded as in Eq. (C5) (for simplicity we take the same value of θ on both sides). The observed correlations $P(c,a,b,d|z,x,y,w)$, following Eq. (B2), then become

$$\begin{aligned}
P(c,a,b,d|z,x,y,w) &= \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \left(\left(\mathsf{M}^{\tilde{C}}_{c|z} * \psi_1^{\tilde{C}\tilde{A}} \right) \otimes \left(\mathsf{M}^{\tilde{D}}_{d|w} * \psi_2^{\tilde{B}\tilde{D}} \right) \right) \\
&= \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \left(P(c|z) \sum_{k=0,1} \left(|V^{\mathbf{A}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}^\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi_k^{A^{\mathrm{jf}}} \right)^{T^k_{A'}} * \rho^{A'}_{c|z} + \delta^{\tilde{A}}_{c|z} \right) \\
&\qquad \otimes \left(P(d|w) \sum_{l=0,1} \left(|V^{\mathbf{B}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}^\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta_l^{B^{\mathrm{jf}}} \right)^{T^l_{B'}} * \rho^{B'}_{d|w} + \delta^{\tilde{B}}_{d|w} \right) \\
&= P(c|z)\, P(d|w)\; E^{A'B'}_{a,b|x,y} * \left(\rho^{A'}_{c|z} \otimes \rho^{B'}_{d|w} \right) + P(c|z)\, \mathsf{E}^{A'\tilde{B}}_{a,b|x,y} * \left(\rho^{A'}_{c|z} \otimes \delta^{\tilde{B}}_{d|w} \right) \\
&\qquad + P(d|w)\, \mathsf{E}^{\tilde{A}B'}_{a,b|x,y} * \left(\delta^{\tilde{A}}_{c|z} \otimes \rho^{B'}_{d|w} \right) + \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \left(\delta^{\tilde{A}}_{c|z} \otimes \delta^{\tilde{B}}_{d|w} \right)
\end{aligned} \tag{C6}$$

with the "physical" D-POVM elements defined as $\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} = \left(\mathsf{M}^{\tilde{A}A}_{a|x} \otimes \mathsf{M}^{\tilde{B}B}_{b|y} \right) * W^{AB}$, the D-POVM elements $E^{A'B'}_{a,b|x,y}$ acting on the "reference" input states as given in Eq. (B3), and with the new D-POVM elements effectively acting on the "reference" and the "physical" quantum input states

$$\begin{aligned}
\mathsf{E}^{A'\tilde{B}}_{a,b|x,y} &\coloneqq \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \sum_{k=0,1} \left(|V^{\mathbf{A}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}^\dagger}|^{A'A^{\mathrm{jf}}\tilde{A}} * \xi_k^{A^{\mathrm{jf}}} \right)^{T^k_{A'}}, \\
\mathsf{E}^{\tilde{A}B'}_{a,b|x,y} &\coloneqq \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} * \sum_{l=0,1} \left(|V^{\mathbf{B}^\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}^\dagger}|^{B'B^{\mathrm{jf}}\tilde{B}} * \zeta_l^{B^{\mathrm{jf}}} \right)^{T^l_{B'}}.
\end{aligned} \tag{C7}$$

The correlations $P_\star(a,b|z,w)$ used in our protocol to certify causal nonseparability are then

$$\begin{aligned}
P_\star(a,b|z,w) &= P(a,b|z,x{=}\star,y{=}\star,w,c{=}0,d{=}0) \\
&= P'(a,b|\rho^{A'}_z,\rho^{B'}_w) + \mathsf{E}^{A'\tilde{B}}_{a,b|\star,\star} * \left(\rho^{A'}_z \otimes \frac{\delta^{\tilde{B}}_{0|w}}{P(0|w)} \right) \\
&\qquad + \mathsf{E}^{\tilde{A}B'}_{a,b|\star,\star} * \left(\frac{\delta^{\tilde{A}}_{0|z}}{P(0|z)} \otimes \rho^{B'}_w \right) + \mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|\star,\star} * \left(\frac{\delta^{\tilde{A}}_{0|z}}{P(0|z)} \otimes \frac{\delta^{\tilde{B}}_{0|w}}{P(0|w)} \right),
\end{aligned} \tag{C8}$$

with $P'(a,b|\rho^{A'}_z,\rho^{B'}_w) = E^{A'B'}_{a,b|\star,\star} * \left(\rho^{A'}_z \otimes \rho^{B'}_w \right)$. We can bound the second term in the right-hand side of Eq. (C8) as

$$\begin{aligned}
\left| \operatorname{Tr} \left[\left(\mathsf{E}^{A'\tilde{B}}_{a,b|\star,\star} \right)^T \left(\rho^{A'}_z \otimes \frac{\delta^{\tilde{B}}_{0|w}}{P(0|w)} \right) \right] \right| &\leq \left\| \left(\mathsf{E}^{A'\tilde{B}}_{a,b|\star,\star} \right)^T \left(\rho^{A'}_z \otimes \frac{\delta^{\tilde{B}}_{0|w}}{P(0|w)} \right) \right\|_1 \leq \left\| \mathsf{E}^{A'\tilde{B}}_{a,b|\star,\star} \right\|_\infty \left\| \rho^{A'}_z \otimes \frac{\delta^{\tilde{B}}_{0|w}}{P(0|w)} \right\|_1 \\
&\leq \frac{1}{P(0|w)} \left\| \rho^{A'}_z \right\|_1 \left\| \delta^{\tilde{B}}_{0|w} \right\|_1 \leq \frac{1}{P(0|w)} \left(\sqrt{2}\,\theta + \theta^2 \right),
\end{aligned} \tag{C9}$$

where in the first inequality we used the fact that $|\operatorname{Tr}(A)| \leq \operatorname{Tr}|A| = \|A\|_1$, the second one follows from Hölder's inequality, in the third inequality we used the fact that the maximal eigenvalue of $\mathsf{E}^{A'\tilde{B}}_{a,b|\star,\star}$ cannot be larger than 1,

and in the last inequality we used the normalisation of $\rho_z^{A'}$ together with the analogue of the bound (C5) above, for $\left\|\delta_{0|w}^{\tilde{B}}\right\|_1$. In a similar way, for the last two terms in Eq. (C8), we can write

$$\left|\mathrm{Tr}\left[\left(\mathsf{E}_{a,b|\star,\star}^{\tilde{A}B'}\right)^T\left(\frac{\delta_{0|z}^{\tilde{A}}}{P(0|z)}\otimes\rho_w^{B'}\right)\right]\right|\leq\frac{1}{P(0|z)}\left(\sqrt{2}\,\theta+\theta^2\right), \tag{C10}$$

and

$$\left|\mathrm{Tr}\left[\left(\mathsf{E}_{a,b|\star,\star}^{\tilde{A}\tilde{B}}\right)^T\left(\frac{\delta_{0|z}^{\tilde{A}}}{P(0|z)}\otimes\frac{\delta_{0|w}^{\tilde{B}}}{P(0|w)}\right)\right]\right|\leq\frac{1}{P(0|w)P(0|z)}\left(\sqrt{2}\,\theta+\theta^2\right)^2. \tag{C11}$$

Consider now a witness inequality of the form of Eq. (B8), $\mathcal{J}_{\text{ref.}}:=\sum_{a,b,z,w}s_{a,b}^{(z,w)}P'(a,b|\rho_z^{A'},\rho_w^{B'})\geq 0$. Evaluating the same witness expression with the experimentally obtained probabilities $P_\star(a,b|z,w)$ instead of $P'(a,b|\rho_z^{A'},\rho_w^{B'})$, so as to define $\mathcal{J}_{\text{exp.}}:=\sum_{a,b,z,w}s_{a,b}^{(z,w)}P_\star(a,b|z,w)$, we can write

$$\mathcal{J}_{\text{exp.}}=\mathcal{J}_{\text{ref.}}+\mathcal{J}_{\text{correction}} \tag{C12}$$

with

$$\mathcal{J}_{\text{correction}}=\sum_{a,b,z,w}s_{a,b}^{(z,w)}\left\{\mathsf{E}_{a,b|\star,\star}^{A'\tilde{B}}*\left(\rho_z^{A'}\otimes\frac{\delta_{0|w}^{\tilde{B}}}{P(0|w)}\right)+\mathsf{E}_{a,b|\star,\star}^{\tilde{A}B'}*\left(\frac{\delta_{0|z}^{\tilde{A}}}{P(0|z)}\otimes\rho_w^{B'}\right)+\mathsf{E}_{a,b|\star,\star}^{\tilde{A}\tilde{B}}*\left(\frac{\delta_{0|z}^{\tilde{A}}}{P(0|z)}\otimes\frac{\delta_{0|w}^{\tilde{B}}}{P(0|w)}\right)\right\}. \tag{C13}$$

The previous analysis allows us to bound $\mathcal{J}_{\text{correction}}$ as $|\mathcal{J}_{\text{correction}}|\leq\mathcal{J}_{\text{correction}}^{\max}$, with an upper bound $\mathcal{J}_{\text{correction}}^{\max}$ on the maximum correction that can be evaluated explicitly given $(s_{a,b}^{(z,w)})_{a,b,z,w}$ using Eqs. (C9)–(C11), and that is (for small θ) of order $\mathcal{O}(\theta)$—with $\mathcal{J}_{\text{correction}}=0$ when $\theta=0$. Given the witness inequality $\mathcal{J}_{\text{ref.}}\geq 0$, which implies

$$\mathcal{J}_{\text{exp.}}\geq-\mathcal{J}_{\text{correction}}^{\max}, \tag{C14}$$

we thus find that a violation of Eq. (C14) certifies causal nonseparability in a manner that is robust to noise.

Appendix D: "Network-MDCI" certification of all causally nonseparable process matrices

It was shown in [15] that all bipartite causally nonseparable process matrices can be certified with trusted quantum inputs, if we assume that Alice and Bob's (trusted) ancillary Hilbert spaces have a bipartite structure of the form $\mathcal{H}^{A'}=\mathcal{H}^{A'_IA'_O}$ and $\mathcal{H}^{B'}=\mathcal{H}^{B'_IB'_O}$, and that their instruments have a specific structure as well—namely, that they factorise as a joint measurement of system A'_I together with the system A_I received from the process matrix, and a channel taking system A'_O into the system A_O that is fed into the process matrix (and similarly on Bob's side). For this certification, it suffices to use separate (product) quantum inputs in systems A'_I and A'_O (and in systems B'_I and B'_O). This certification, called the measurement-device-and-channel-independent (MDCI) certification, can be transformed into a NDI certification by making similar assumptions on the (now untrusted) physical spaces and instruments.

Let us first recall that the above assumptions imply that the Choi matrices of Alice's instruments decompose as $\mathsf{M}_{a|x}^{\tilde{A}A}=\mathsf{M}_{a|x}^{\tilde{A}_IA_I}\otimes\mathsf{M}_{|x}^{\tilde{A}_OA_O}$, for some POVM $(\mathsf{M}_{a|x}^{\tilde{A}_IA_I})_a$ and a channel $\mathsf{M}_{|x}^{\tilde{A}_OA_O}$, and similarly for Bob. Given that the channels $\mathsf{M}_{|x}^{\tilde{A}_OA_O}$ and $\mathsf{M}_{|y}^{\tilde{B}_OB_O}$ must be trace-preserving, these decompositions imply that

$$\mathrm{Tr}_{A_O}\,\mathsf{M}_{a|x}^{\tilde{A}A}=\mathsf{M}_{a|x}^{\tilde{A}_IA_I}\otimes\mathbb{1}^{\tilde{A}_O},\quad\mathrm{Tr}_{B_O}\,\mathsf{M}_{b|y}^{\tilde{B}B}=\mathsf{M}_{b|y}^{\tilde{B}_IB_I}\otimes\mathbb{1}^{\tilde{B}_O}, \tag{D1}$$

for all a,b, with $\sum_a\mathsf{M}_{a|x}^{\tilde{A}_IA_I}=\mathbb{1}^{\tilde{A}_IA_I}$ and $\sum_b\mathsf{M}_{b|y}^{\tilde{B}_IB_I}=\mathbb{1}^{\tilde{B}_IB_I}$. To obtain our results below, it is in fact sufficient to assume that Eq. (D1) holds, for some particular values of a and b.

Consider a causally ordered process matrix $W=W^{A\prec B_I}\otimes\mathbb{1}^{B_O}$. The structure of Eq. (D1) implies that the physical D-POVM elements $\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}}=\left(\mathsf{M}_{a|x}^{\tilde{A}A}\otimes\mathsf{M}_{b|y}^{\tilde{B}B}\right)*W^{AB}$ satisfy

$$\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}}=\left(\mathsf{M}_{a|x}^{\tilde{A}A}\otimes\mathrm{Tr}_{B_O}\,\mathsf{M}_{b|y}^{\tilde{B}B}\right)*W^{A\prec B_I}=\left(\mathsf{M}_{a|x}^{\tilde{A}A}\otimes\mathsf{M}_{b|y}^{\tilde{B}_IB_I}\right)*W^{A\prec B_I}\otimes\mathbb{1}^{\tilde{B}_O}=\mathsf{E}_{a,b|x,y}^{\tilde{A}\tilde{B}_I}\otimes\mathbb{1}^{\tilde{B}_O} \tag{D2}$$

with $\mathsf{E}^{\tilde{A}\tilde{B}_I}_{a,b|x,y} = \left(\mathsf{M}^{\tilde{A}A}_{a|x} \otimes \mathsf{M}^{\tilde{B}_I B_I}_{b|y}\right) * W^{A\prec B_I}$.

Moving on to the NDI scenario, let us focus, for now, on the MDCI assumption on Bob's side. Instead of introducing just one source shared between Bob and Daisy, we now introduce two independent sources shared between Bob and two new parties, Daisy$_I$, with Hilbert space $\mathcal{H}^{\tilde{D}_I}$, and Daisy$_O$, with Hilbert space $\mathcal{H}^{\tilde{D}_O}$ (we use the same notations as before, just adding the subscripts $_I$ and $_O$ appropriately). Self-testing the quantum inputs in spaces $\mathcal{H}^{\tilde{B}_I}$ and $\mathcal{H}^{\tilde{B}_O}$ separately, we can write, similarly to Eq. (A9),

$$P(d_I|w_I)\,\rho^{\tilde{B}_I}_{d_I|w_I} = \mathsf{M}^{\tilde{D}_I}_{d_I|w_I} * \psi^{\tilde{B}_I\tilde{D}_I}_{2,I} = P(d_I|w_I)\,V^{\mathbf{B}_I\dagger}\left(\sum_{l=0,1}(\rho^{B'_I}_{d_I|w_I})^{T^l}\otimes\zeta^{B^{\rm jf}_I}_l\right)V^{\mathbf{B}_I},$$

$$P(d_O|w_O)\,\rho^{\tilde{B}_O}_{d_O|w_O} = \mathsf{M}^{\tilde{D}_O}_{d_O|w_O} * \psi^{\tilde{B}_O\tilde{D}_O}_{2,O} = P(d_O|w_O)\,V^{\mathbf{B}_O\dagger}\left(\sum_{l'=0,1}(\rho^{B'_O}_{d_O|w_O})^{T^{l'}}\otimes\zeta^{B^{\rm jf}_O}_{l'}\right)V^{\mathbf{B}_O}, \tag{D3}$$

for some reference quantum input states $\rho^{B'_I}_{d_I|w_I}\in\mathcal{L}(\mathcal{H}^{B'_I})$ and $\rho^{B'_O}_{d_O|w_O}\in\mathcal{L}(\mathcal{H}^{B'_O})$.

Then as in Eq. (B2), and still for a causally ordered $W = W^{A\prec B_I}\otimes\mathbb{1}^{B_O}$, the correlations $P(c,a,b,d_I,d_O|z,x,y,w_I,w_O)$ can then be written (using the above equations together with Eq. (A8)) as

$$\begin{aligned}
P(c,a,b,d_I,d_O|z,x,y,w_I,w_O) &= \left(\mathsf{M}^{\tilde{C}}_{c|z}\otimes\mathsf{M}^{\tilde{A}A}_{a|x}\otimes\mathsf{M}^{\tilde{B}B}_{b|y}\otimes\mathsf{M}^{\tilde{D}_I}_{d_I|w_I}\otimes\mathsf{M}^{\tilde{D}_O}_{d_O|w_O}\right)*\left(\psi^{\tilde{C}\tilde{A}}_1\otimes W^{A\prec B_I}\otimes\mathbb{1}^{B_O}\otimes\psi^{\tilde{B}_I\tilde{D}_I}_{2,I}\otimes\psi^{\tilde{B}_O\tilde{D}_O}_{2,O}\right)\\
&= P(c|z)\,P(d_I|w_I)\,P(d_O|w_O)\,\left(\mathsf{E}^{\tilde{A}\tilde{B}_I}_{a,b|x,y}\otimes\mathbb{1}^{\tilde{B}_O}\right)*\Bigg(\sum_{k=0,1}\left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\rm jf}\tilde{A}}*\xi^{A^{\rm jf}}_k\right)^{T^k_{A'}}*\rho^{A'}_{c|z}\\
&\qquad\otimes\sum_{l=0,1}\left(|V^{\mathbf{B}_I\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}_I\dagger}|^{B'_IB^{\rm jf}_I\tilde{B}_I}*\zeta^{B^{\rm jf}_I}_l\right)^{T^l_{B'_I}}*\rho^{B'_I}_{d_I|w_I}\\
&\qquad\otimes\sum_{l'=0,1}\left(|V^{\mathbf{B}_O\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}_O\dagger}|^{B'_OB^{\rm jf}_O\tilde{B}_O}*\zeta^{B^{\rm jf}_O}_{l'}\right)^{T^{l'}_{B'_O}}*\rho^{B'_O}_{d_O|w_O}\Bigg)\\
&= P(c|z)\,P(d_I|w_I)\,P(d_O|w_O)\,\left(E^{A'B'_I}_{a,b|x,y}\otimes\mathbb{1}^{B'_O}\right)*\left(\rho^{A'}_{c|z}\otimes\rho^{B'_I}_{d_I|w_I}\otimes\rho^{B'_O}_{d_O|w_O}\right)
\end{aligned}\tag{D4}$$

with

$$E^{A'B'_I}_{a,b|x,y} = \mathsf{E}^{\tilde{A}\tilde{B}_I}_{a,b|x,y}*\left[\sum_{k=0,1}\left(|V^{\mathbf{A}\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{A}\dagger}|^{A'A^{\rm jf}\tilde{A}}*\xi^{A^{\rm jf}}_k\right)^{T^k_{A'}}\otimes\sum_{l=0,1}\left(|V^{\mathbf{B}_I\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}_I\dagger}|^{B'_IB^{\rm jf}_I\tilde{B}_I}*\zeta^{B^{\rm jf}_I}_l\right)^{T^l_{B'_I}}\right], \tag{D5}$$

where we used the fact that $\sum_{l'}\left(\mathrm{Tr}_{\tilde{B}_O}|V^{\mathbf{B}_O\dagger}\rangle\!\rangle\langle\!\langle V^{\mathbf{B}_O\dagger}|^{B'_OB^{\rm jf}_O\tilde{B}_O}*\zeta^{B^{\rm jf}_O}_{l'}\right)^{T^{l'}_{B'_O}} = \mathbb{1}^{B'_O}$, cf. Footnote 8.

Hence, we see that for a causally ordered $W = W^{A\prec B_I}\otimes\mathbb{1}^{B_O}$, the structure $\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y} = \mathsf{E}^{\tilde{A}\tilde{B}_I}_{a,b|x,y}\otimes\mathbb{1}^{\tilde{B}_O}$ that the MDCI assumption (on Bob's side) implies on the physical D-POVM elements $\mathsf{E}^{\tilde{A}\tilde{B}}_{a,b|x,y}$ gets transferred to the effective D-POVM elements $E^{A'B'}_{a,b|x,y} = E^{A'B'_I}_{a,b|x,y}\otimes\mathbb{1}^{B'_O}$ that act on the reference input states. By symmetry, the same structure preservation applies for a causally ordered $W = W^{B\prec A_I}\otimes\mathbb{1}^{A_O}$ (using the MDCI assumption on Alice's side, introducing now two independent sources shared by two new parties Charlie$_I$ and Charlie$_O$). It follows that for a causally separable process matrix $W^{AB} = q\,W^{A\prec B_I}\otimes\mathbb{1}^{B_O} + (1-q)\,W^{B\prec A_I}\otimes\mathbb{1}^{A_O}$, the D-POVM elements acting on the reference input states (in some reference Hilbert spaces $\mathcal{H}^{A'_I}$, $\mathcal{H}^{A'_O}$, $\mathcal{H}^{B'_I}$ and $\mathcal{H}^{B'_O}$) decompose as

$$E^{A'_IA'_OB'_IB'_O}_{a,b|x,y} = q\,E^{A'_IA'_OB'_I}_{a,b|x,y}\otimes\mathbb{1}^{B'_O} + (1-q)\,E^{A'_IB'_IB'_O}_{a,b|x,y}\otimes\mathbb{1}^{A'_O}. \tag{D6}$$

In this scenario (involving all four additional parties Charlie$_I$, Charlie$_O$, Daisy$_I$ and Daisy$_O$ together with Alice and Bob), the correlations

$$\begin{aligned}
P_\star(a,b|z,w) &= P(a,b|z_I,z_O,x{=}\star,y{=}\star,w_I,w_O,c_I{=}0,c_O{=}0,d_I{=}0,d_O{=}0)\\
&= E^{A'_IA'_OB'_IB'_O}_{a,b|\star,\star}*\left(\rho^{A'_I}_{z_I}\otimes\rho^{A'_O}_{z_O}\otimes\rho^{B'_I}_{w_I}\otimes\rho^{B'_O}_{w_O}\right)
\end{aligned}\tag{D7}$$

then define a distribution $P'(a,b|\rho_{z_I}^{A'_I},\rho_{z_O}^{A'_O},\rho_{w_I}^{B'_I},\rho_{w_O}^{B'_O})$ in a SDI-QI setting with the reference quantum inputs, that must respect any witness inequality for correlations induced by D-POVM elements of the form of Eq. (D6). Now, we have shown in [15] that any bipartite causally nonseparable process matrix can violate such a witness inequality. We thus conclude here that any bipartite causally nonseparable process matrix can be certified in the scenario considered here, which combines the NDI approach with the MDCI structure assumption—i.e., in a "N-MDCI" ("Network-MDCI") manner.

Recall that in [15] we also studied MDCI certification in the "(2+F)-partite" scenario (which involves an additional party, Fiona, in the causal future of Alice and Bob and who has no output Hilbert space), showing that a similar result holds for all "TTU-" and "TUU-noncausal" processes as defined in [32]—that is, causally nonseparable processes that can be certified with trusted operations for Alice and Bob (but not for Fiona), or for Alice only. In particular, we showed that all "TTU-" and "TUU-noncausal" processes could be certified in a "MDCI-MDCI-DI" and "MDCI-DI-DI" manner, respectively, when imposing the extra assumption of Eq. (D1) to the parties with formally trusted operations (and still with trusted quantum inputs). Following the same arguments as above, we can conclude that all such processes—including the quantum switch–can also be certified in a N-MDCI manner for those parties, thus relaxing the need for trusted quantum inputs.

Although we leave it as an open problem, we expect it to be rather straightforward to further generalise all these results to the (P+2+F)-partite scenario, which also includes a party in the global past and that we will consider in Appendix E.

Appendix E: SDI-QI and NDI certification of causal nonseparability in the "(P+2+F)-partite" scenario

We consider here the "(P+2+F)-partite" scenario introduced in the main text, which extends the (2+F)-partite scenario (considered in particular in [15]) by adding a new party, Phil, in the causal past of all other parties. Phil is only connected to the process matrix through his output Hilbert space $\mathcal{H}^P$ (he has no input Hilbert space). We shall first recall the definition of causal nonseparability in this scenario, before considering its SDI-QI certification when all four parties receive quantum inputs, as well as when only Phil does, and applying this to the quantum switch. As in the bipartite scenario, this SDI-QI certification can be readily transformed into NDI certification, and we will briefly discuss this transformation, which is performed analogously to the bipartite case, at the end of the section.

1. Causal (non)separability in the "(P+2+F)-partite" scenario

Given that Phil is in the global past of the process, and that Fiona is in the global future, the only relevant orders in this scenario are $P\prec A\prec B\prec F$ and $P\prec B\prec A\prec F$. The causally separable process matrices are those that can be decomposed as [2]

$$W^{PABF}=W^{P\prec A\prec B\prec F}+W^{P\prec B\prec A\prec F} \tag{E1}$$

such that

$$\begin{aligned}\mathrm{Tr}_F\, W^{P\prec A\prec B\prec F}&=W^{P\prec A\prec B_I}\otimes\mathbb{1}^{B_O}, \qquad \mathrm{Tr}_F\, W^{P\prec B\prec A\prec F}=W^{P\prec B\prec A_I}\otimes\mathbb{1}^{A_O},\\ \mathrm{Tr}_{B_I}\, W^{P\prec A\prec B_I}&=W^{P\prec A_I}\otimes\mathbb{1}^{A_O}, \qquad \mathrm{Tr}_{A_I}\, W^{P\prec B\prec A_I}=W^{P\prec B_I}\otimes\mathbb{1}^{B_O},\\ \mathrm{Tr}_{A_I}\, W^{P\prec A_I}&+\mathrm{Tr}_{B_I}\, W^{P\prec B_I}=\mathbb{1}^P,\end{aligned} \tag{E2}$$

with positive semidefinite matrices $W^{P\prec A\prec B_I}\in\mathcal{L}(\mathcal{H}^{PAB_I})$, $W^{P\prec B\prec A_I}\in\mathcal{L}(\mathcal{H}^{PBA_I})$, $W^{P\prec A_I}\in\mathcal{L}(\mathcal{H}^{PA_I})$, $W^{P\prec B_I}\in\mathcal{L}(\mathcal{H}^{PB_I})$.

Note that in general $W^{P\prec A\prec B\prec F}$ and $W^{P\prec B\prec A\prec F}$ are not necessarily valid (subnormalised) process matrices themselves (unless $\mathrm{Tr}_{A_I}\, W^{P\prec A_I}\propto\mathbb{1}^P$ and $\mathrm{Tr}_{B_I}\, W^{P\prec B_I}\propto\mathbb{1}^P$). Indeed causally separable processes allow for the order between Alice and Bob to be influenced by Phil's action (in a classically controlled manner [13]), rather than involving only convex combinations of processes with the fixed orders $P\prec A\prec B\prec F$ and $P\prec B\prec A\prec F$. This explains why Eq. (E1), contrary to Eqs. (3) and (D6), does not include convex weights, which would otherwise suggest such a combination.

2. SDI-QI certification with quantum inputs for all 4 parties

In a general (P+2+F)-partite SDI-QI scenario, all four parties Phil, Alice, Bob and Fiona receive some quantum inputs $\rho_t^{P'}$, $\rho_z^{A'}$, $\rho_w^{B'}$ and $\rho_v^{F'}$ in some Hilbert spaces $\mathcal{L}(\mathcal{H}^{P'})$, $\mathcal{L}(\mathcal{H}^{A'})$, $\mathcal{L}(\mathcal{H}^{B'})$ and $\mathcal{L}(\mathcal{H}^{F'})$ respectively; see Fig. 3.

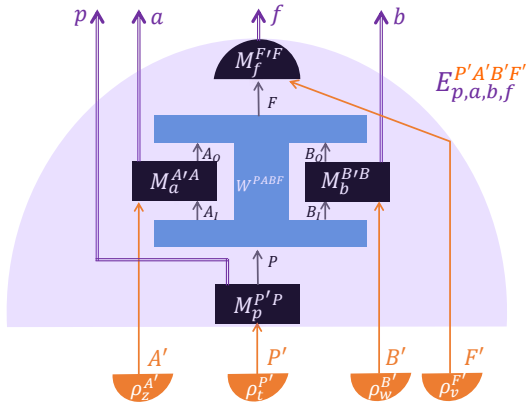

FIG. 3. $(P{+}2{+}F)$-partite SDI-QI scenario with quantum inputs for Phil, Alice, Bob and Fiona. The trusted quantum inputs (orange) $\rho_t^{P'} \in \mathcal{L}(\mathcal{H}^{P'}), \rho_z^{A'} \in \mathcal{L}(\mathcal{H}^{A'}), \rho_w^{B'} \in \mathcal{L}(\mathcal{H}^{B'}), \rho_v^{F'} \in \mathcal{L}(\mathcal{H}^{F'})$ are prepared and effectively measured by the D-POVM $(E_{p,a,b,f}^{P'A'B'F'})_{p,a,b,f}$ generated by Phil, Alice, Bob and Fiona's untrusted operations on the process matrix W^{PABF}.

Generalising the bipartite case introduced in the main text, the correlations established by the four parties are

$$\begin{aligned} P(p,a,b,f|\rho_t^{P'},\rho_z^{A'},\rho_w^{B'},\rho_v^{F'}) &= \left(M_p^{P'P}\otimes M_a^{A'A}\otimes M_b^{B'B}\otimes M_f^{F'F}\right) * \left(\rho_t^{P'}\otimes\rho_z^{A'}\otimes\rho_w^{B'}\otimes\rho_v^{F'}\otimes W^{PABF}\right) \\ &= E_{p,a,b,f}^{P'A'B'F'} * \left(\rho_t^{P'}\otimes\rho_z^{A'}\otimes\rho_w^{B'}\otimes\rho_v^{F'}\right), \end{aligned} \tag{E3}$$

where we introduced the D-POVM elements

$$E_{p,a,b,f}^{P'A'B'F'} = \left(M_p^{P'P}\otimes M_a^{A'A}\otimes M_b^{B'B}\otimes M_f^{F'F}\right) * W^{PABF}. \tag{E4}$$

Assuming that W^{PABF} is a causally separable process matrix satisfying Eqs. (E1)–(E2), the induced D-POVM $\mathbb{E}^{P'A'B'F'} = (E_{p,a,b,f}^{P'A'B'F'})_{p,a,b,f}$ also decomposes as

$$\mathbb{E}^{P'A'B'F'} = \mathbb{E}^{P'\prec A'\prec B'\prec F'} + \mathbb{E}^{P'\prec B'\prec A'\prec F'} \tag{E5}$$

with $\mathbb{E}^{P'\prec A'\prec B'\prec F'} = (E_{p,a,b,f}^{P'\prec A'\prec B'\prec F'})_{p,a,b,f}$ and $\mathbb{E}^{P'\prec B'\prec A'\prec F'} = (E_{p,a,b,f}^{P'\prec B'\prec A'\prec F'})_{p,a,b,f}$ such that

$$\begin{aligned} E_{p,a,b,f}^{P'\prec A'\prec B'\prec F'} &= (M_p^{P'P}\otimes M_a^{A'A}\otimes M_b^{B'B}\otimes M_f^{F'F}) * W^{P\prec A\prec B\prec F}, \\ \sum_f E_{p,a,b,f}^{P'\prec A'\prec B'\prec F'} &= (M_p^{P'P}\otimes M_a^{A'A}\otimes \mathrm{Tr}_{B_O} M_b^{B'B}) * W^{P\prec A\prec B_I}\otimes \mathbb{1}^{F'} = E_{p,a,b}^{P'\prec A'\prec B'}\otimes \mathbb{1}^{F'}, \\ \sum_b E_{p,a,b}^{P'\prec A'\prec B'} &= (M_p^{P'P}\otimes \mathrm{Tr}_{A_O} M_a^{A'A}) * W^{P\prec A_I}\otimes \mathbb{1}^{B'} = E_{p,a}^{P'\prec A'}\otimes \mathbb{1}^{B'}, \\ \sum_a E_{p,a}^{P'\prec A'} &= M_p^{P'P} * \mathrm{Tr}_{A_I} W^{P\prec A_I}\otimes \mathbb{1}^{A'} = E_p^{P'\,[A\prec B]}\otimes \mathbb{1}^{A'}, \end{aligned} \tag{E6}$$

analogously for $E_{p,a,b,f}^{P'\prec B'\prec A'\prec F'}$ and the order $P\prec B\prec A\prec F$, and with

$$E_p^{P'\,[A\prec B]} + E_p^{P'\,[B\prec A]} = M_p^{P'P} * \mathbb{1}^P = \mathrm{Tr}_P\, M_p^{P'P}, \tag{E7}$$

such that, since $\sum_p M_p^{P'P}$ is a CPTP map,

$$\sum_p (E_p^{P'\,[A\prec B]} + E_p^{P'\,[B\prec A]}) = \mathbb{1}^{P'}. \tag{E8}$$

Note that the individual sets of positive semidefinite matrices $\mathbb{E}^{P'\prec A'\prec B'\prec F'} = (E_{p,a,b,f}^{P'\prec A'\prec B'\prec F'})_{p,a,b,f}$, $\mathbb{E}^{P'\prec B'\prec A'\prec F'} = (E_{p,a,b,f}^{P'\prec B'\prec A'\prec F'})_{p,a,b,f}$, $\mathbb{E}^{P'\prec A'\prec B'} = (E_{p,a,b}^{P'\prec A'\prec B'})_{p,a,b}$, $\mathbb{E}^{P'\prec B'\prec A'} = (E_{p,a,b}^{P'\prec B'\prec A'})_{p,a,b}$, $\mathbb{E}^{P'\prec A'} = (E_{p,a}^{P'\prec A'})_{p,a}$,

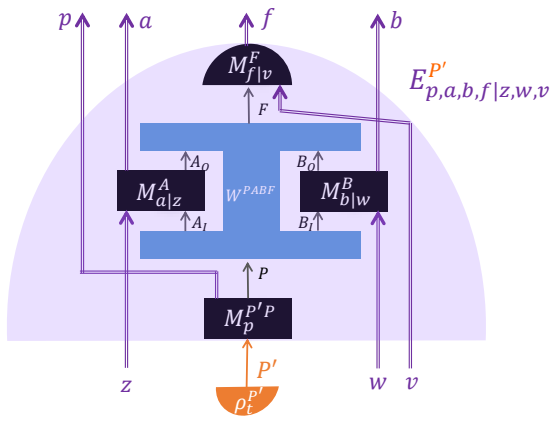

FIG. 4. $(P{+}2{+}F)$-partite SDI-QI scenario with quantum inputs solely for Phil. The trusted quantum input (orange) $\rho_t \in \mathcal{L}(\mathcal{H}^{P'})$ is prepared and effectively measured by the D-POVMs $(E^{P'}_{p,a,b,f|z,w,v})_{p,a,b,f}$ generated by Phil, Alice, Bob and Fiona's untrusted operations on the process matrix W^{PABF}.

$\mathbb{E}^{P'\prec B'} = (E^{P'\prec B'}_{p,b})_{p,b}$, $\mathbb{E}^{P'\,[A\prec B]} = (E^{P'\,[A\prec B]}_{p})_{p}$, $\mathbb{E}^{P'\,[B\prec A]} = (E^{P'\,[B\prec A]}_{p})_{p}$ are not D-POVMs as they do not sum to the identity in general. On the other hand, the set $\mathbb{E}^{P'\,[A\prec B]} \cup \mathbb{E}^{P'\,[B\prec A]}$ formed by all elements of $\mathbb{E}^{P'\,[A\prec B]}$ and $\mathbb{E}^{P'\,[B\prec A]}$ taken together forms a valid POVM.

Analogously to the bipartite and $(2{+}F)$-partite cases, this leads us to the following definition.

Definition 1. *A $(P{+}2{+}F)$-partite D-POVM $\mathbb{E}^{P'A'B'F'} = (E^{P'A'B'F'}_{p,a,b,f})_{p,a,b,f}$ that can be decomposed in the form*

$$\mathbb{E}^{P'A'B'F'} = \mathbb{E}^{P'\prec A'\prec B'\prec F'} + \mathbb{E}^{P'\prec B'\prec A'\prec F'}, \tag{E9}$$

where $\mathbb{E}^{P'\prec A'\prec B'\prec F'} = (E^{P'\prec A'\prec B'\prec F'}_{p,a,b,f})_{p,a,b,f}$ and $\mathbb{E}^{P'\prec B'\prec A'\prec F'} = (E^{P'\prec B'\prec A'\prec F'}_{p,a,b,f})_{p,a,b,f}$ satisfy

$$\begin{aligned}
&\sum_f E^{P'\prec A'\prec B'\prec F'}_{p,a,b,f} = E^{P'\prec A'\prec B'}_{p,a,b} \otimes \mathbb{1}^{F'}, \quad \sum_b E^{P'\prec A'\prec B'}_{p,a,b} = E^{P'\prec A'}_{p,a} \otimes \mathbb{1}^{B'}, \quad \sum_a E^{P'\prec A'}_{p,a} = E^{P'\,[A\prec B]}_{p} \otimes \mathbb{1}^{A'},\\
&\sum_f E^{P'\prec B'\prec A'\prec F'}_{p,a,b,f} = E^{P'\prec B'\prec A'}_{p,a,b} \otimes \mathbb{1}^{F'}, \quad \sum_a E^{P'\prec B'\prec A'}_{p,a,b} = E^{P'\prec B'}_{p,b} \otimes \mathbb{1}^{A'}, \quad \sum_b E^{P'\prec B'}_{p,b} = E^{P'\,[B\prec A]}_{p} \otimes \mathbb{1}^{B'},\\
&\sum_p (E^{P'\,[A\prec B]}_{p} + E^{P'\,[B\prec A]}_{p}) = \mathbb{1}^{P'}
\end{aligned} \tag{E10}$$

(with all operators $E^{\cdots}_{\cdots} \geq 0$) is said to be causally separable.

As noted above, a causally separable $(P{+}2{+}F)$-partite process matrix can only generate a causally separable $(P{+}2{+}F)$-partite D-POVM. *A contrario*, if one finds that, for some choice of operations for each party, the induced D-POVM is causally nonseparable, this certifies in a SDI-QI manner that the process matrix that generated it is itself causally nonseparable. Similarly to the bipartite and the $(2{+}F)$-partite cases, one can construct a causal witness—and associated witness inequality—that can be measured by providing tomographically complete quantum inputs to the parties, thereby allowing one to certify the causal non-separability of $(P{+}2{+}F)$-partite D-POVMs [10, 15, 24].

3. SDI-QI certification with quantum inputs only for Phil

SDI-QI scenarios can also be considered, where not all parties receive quantum inputs: some may receive classical inputs (or no input at all). For example, in [15] we considered the $(2{+}F)$-partite case where only Alice and Bob (but not Fiona) receive quantum inputs—a situation that could readily be extended to the $(P{+}2{+}F)$-partite case.

Let us here consider instead the case where only Phil receives quantum input states $\rho^{P'}_t \in \mathcal{L}(\mathcal{H}^{P'})$, while Alice, Bob and Fiona are given classical inputs z, w and v respectively; see Fig. 4. Conditioning their instruments explicitly on

the classical inputs,[10] the correlations established by the four parties are given here by

$$P(p,a,b,f|\rho_t^{P'},z,w,v) = \left(M_p^{P'P}\otimes M_{a|z}^A\otimes M_{b|w}^B\otimes M_{f|v}^F\right)*\left(\rho_t^{P'}\otimes W^{PABF}\right)$$
$$= E_{p,a,b,f|z,w,v}^{P'} * \rho_t^{P'}, \tag{E11}$$

with the D-POVM elements[11]

$$E_{p,a,b,f|z,w,v}^{P'} = \left(M_p^{P'P}\otimes M_{a|z}^A\otimes M_{b|w}^B\otimes M_{f|v}^F\right)*W^{PABF}. \tag{E12}$$

Assuming that W^{PABF} is causally separable, one obtains a similar decomposition to Eqs. (E5)–(E8), for the whole family of D-POVMs $\{\mathbb{E}_{|z,w,v}^{P'} = (E_{p,a,b,f|z,w,v}^{P'})_{p,a,b,f}\}_{z,w,v}$ (i.e., involving the D-POVMs for all classical inputs z,w,v). Similarly to the previous case, we introduce here the following definition.

Definition 2. *A family of (P+2+F)-partite D-POVMs* $\{\mathbb{E}_{|z,w,v}^{P'} = (E_{p,a,b,f|z,w,v}^{P'})_{p,a,b,f}\}_{z,w,v}$ *that can be decomposed in the form*

$$\mathbb{E}_{|z,w,v}^{P'} = \mathbb{E}_{|z,w,v}^{P'\,[A\prec B]} + \mathbb{E}_{|z,w,v}^{P'\,[B\prec A]}, \tag{E13}$$

for all z,w,v, *where* $\mathbb{E}_{|z,w,v}^{P'\,[A\prec B]} = (E_{p,a,b,f|z,w,v}^{P'\,[A\prec B]})_{p,a,b,f}$ *and* $\mathbb{E}_{|z,w,v}^{P'\,[B\prec A]} = (E_{p,a,b,f|z,w,v}^{P'\,[B\prec A]})_{p,a,b,f}$ *satisfy*

$$\sum_f E_{p,a,b,f|z,w,v}^{P'\,[A\prec B]} = E_{p,a,b|z,w}^{P'\,[A\prec B]},\quad \sum_b E_{p,a,b|z,w}^{P'\,[A\prec B]} = E_{p,a|z}^{P'\,[A\prec B]},\quad \sum_a E_{p,a|z}^{P'\,[A\prec B]} = E_p^{P'\,[A\prec B]},$$
$$\sum_f E_{p,a,b,f|z,w,v}^{P'\,[B\prec A]} = E_{p,a,b|z,w}^{P'\,[B\prec A]},\quad \sum_a E_{p,a,b|z,w}^{P'\,[B\prec A]} = E_{p,b|w}^{P'\,[B\prec A]},\quad \sum_b E_{p,b|w}^{P'\,[B\prec A]} = E_p^{P'\,[B\prec A]},$$
$$\sum_p (E_p^{P'\,[A\prec B]} + E_p^{P'\,[B\prec A]}) = \mathbb{1}^{P'} \tag{E14}$$

(with all operators $E_{\dots}^{\dots} \geq 0$*) is said to be* causally separable.

As always, a causally separable $(P{+}2{+}F)$-partite process matrix W^{PABF} can only generate a causally separable family of $(P{+}2{+}F)$-partite D-POVMs, so that obtaining a causally nonseparable family of $(P{+}2{+}F)$-partite D-POVMs (which can be verified using some form of causal witnesses) certifies in a SDI-QI manner the causal nonseparability of W^{PABF}.

4. SDI-QI certification of the quantum switch in the "$(P{+}2{+}F)$-partite" scenario

The quantum switch [3] is the canonical example of a quantum process with indefinite causal order, which can be understood as a quantum circuit with quantum control of causal order [13]. Despite being causally nonseparable, however, the quantum switch can only generate so-called causal correlations [10, 11], so that its causal nonseparability cannot be certified in the standard DI way.

As recalled previously, in [15] we showed that the causal nonseparability of the quantum switch could nevertheless be certified in a $(2{+}F)$-partite SDI-QI scenario where Alice and Bob receive quantum inputs. This scenario can be seen as a particular case of a $(P{+}2{+}F)$-partite one, with a trivial role for the past party Phil. Introducing a nontrivial P allows one to also provide the latter with quantum inputs. We will show here that the causal nonseparability of the quantum switch can still be certified if only Phil receives quantum inputs.

In the most general form of the quantum switch, as described in the main text, Phil sends and Fiona receives both "target" and "control" systems, in spaces $\mathcal{H}^{P_\mathrm{t}}$, $\mathcal{H}^{P_\mathrm{c}}$, $\mathcal{H}^{F_\mathrm{t}}$ and $\mathcal{H}^{F_\mathrm{c}}$, respectively. Here, let us take the target system to be initialised in the state $|0\rangle\langle 0|^{P_\mathrm{t}} \in \mathcal{L}(\mathcal{H}^{P_\mathrm{t}})$ and traced out in the future space $\mathcal{H}^{F_\mathrm{t}}$, and the qubit control system

[10] Classical inputs can also be fitted into the framework with quantum inputs by using orthogonal quantum input states. Indeed, one can encode (e.g., for Alice) the classical inputs into some orthonormal basis $\{|z\rangle^{A'}\}_z$, and introduce the instruments $(M_a^{A'A})_a = \left(\sum_z |z\rangle\langle z|^{A'}\otimes M_{a|z}^A\right)_a$ [15]. One can use this, for instance, to obtain the analogous constraints to Eqs. (E5)–(E8) for classical inputs z,w,v, and to thereby justify the decomposition in Definition 2—and in particular, the "no-signalling constraints" (e.g., that $\sum_f E_{p,a,b,f|z,w,v}^{P'\,[A\prec B]} = E_{p,a,b|z,w}^{P'\,[A\prec B]}$ does not depend on v, that $\sum_b E_{p,a,b|z,w}^{P'\,[A\prec B]} = E_{p,a|z}^{P'\,[A\prec B]}$ does not depend on w, etc.).

[11] While the D-POVM acts on a single Hilbert space $\mathcal{H}^{P'}$, accessed by Phil only, it is still distributed in the sense that the classical inputs and outputs are associated with distributed parties, even if no quantum systems are associated with them.

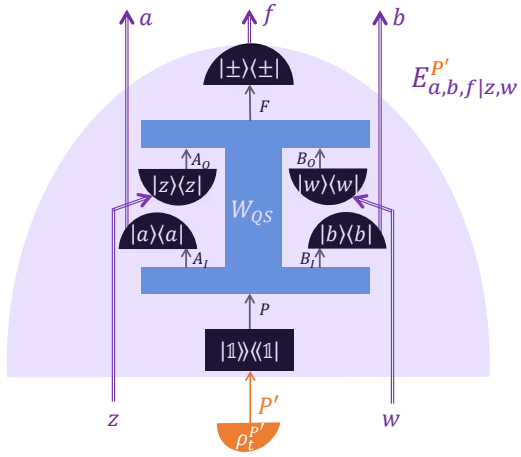

FIG. 5. D-POVM generated by the quantum instruments of Eq. (E17), which can be used to certify the causal nonseparability of the quantum switch in a SDI-QI manner.

to be prepared by Phil in the global past and given to Fiona in the global future. The quantum switch can then be written as a $(P{+}2{+}F)$-partite process matrix as

$$W_{\mathrm{QS}} = |0\rangle\langle 0|^{P_{\mathrm{t}}} * \mathrm{Tr}_{F_{\mathrm{t}}}(|w_{\mathrm{QS}}\rangle\langle w_{\mathrm{QS}}|^{PP_{\mathrm{t}}ABF_{\mathrm{t}}F}), \tag{E15}$$

with

$$|w_{\mathrm{QS}}\rangle^{PP_{\mathrm{t}}ABF_{\mathrm{t}}F} = |0\rangle^{P}\,|\mathbb{1}\rangle\rangle^{P_{\mathrm{t}}A_I}\,|\mathbb{1}\rangle\rangle^{A_OB_I}\,|\mathbb{1}\rangle\rangle^{B_OF_{\mathrm{t}}}\,|0\rangle^{F} + |1\rangle^{P}\,|\mathbb{1}\rangle\rangle^{P_{\mathrm{t}}B_I}\,|\mathbb{1}\rangle\rangle^{B_OA_I}\,|\mathbb{1}\rangle\rangle^{A_OF_{\mathrm{t}}}\,|1\rangle^{F} \tag{E16}$$

where $|\mathbb{1}\rangle\rangle^{XX'} = \sum_i |i\rangle^X \otimes |i\rangle^{X'}$ (and with implicit tensor products) [13].[12]

Let us then consider the operations

$$M^{P'P} = |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|^{P'P}\,, \quad M^A_{a|z} = |a\rangle\langle a|^{A_I} \otimes |z\rangle\langle z|^{A_O}\,, \quad M^B_{b|w} = |b\rangle\langle b|^{B_I} \otimes |w\rangle\langle w|^{B_O}\,, \quad M^F_{f=\pm} = |\pm\rangle\langle\pm|^{F}\,, \tag{E17}$$

which can be interpreted as follows (see Fig. 5): Phil simply implements an identity channel that directly forwards the quantum input into the quantum switch, with no (or just a trivial) output p. Alice and Bob measure the target system (taken here to be a qubit) in the computational basis with classical outcomes $a, b = 0, 1$, and re-prepare it in the computational basis state that corresponds to their input $z, w = 0, 1$. Fiona has no (or just a trivial) input v and she simply measures the control qubit in the $\{|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)\}_\pm$ basis.

As it turns out, when implemented on a quantum switch, these instruments generate a causally nonseparable family of D-POVMs $\{\mathbb{E}^{P'}_{|z,w} = (E^{P'}_{a,b,f|z,w})_{a,b,f}\}_{z,w}$, according to Definition 2 (with the trivial p and v omitted in this special case). Indeed, using a semidefinite programming approach, we can show that these D-POVMs cannot be decomposed as in Eqs. (E13)–(E14), and a causal witness can be constructed to test this [10, 15, 24]. More specifically, we find that a "depolarised" version of the quantum switch $W_{\mathrm{QS}}(r) = \frac{1}{1+r}(W_{\mathrm{QS}} + r\,\mathbb{1}^\circ)$, obtained by mixing it with some fully white noise described by the process matrix $\mathbb{1}^\circ = \mathbb{1}^{PABF}/8$ (by some amount $r \geq 0$), induces with the above instruments a causally nonseparable family of D-POVMs for all $r < 2 - \sqrt{2}$ (up to numerical precision).

Hence, as we claimed, the quantum switch can be certified in a SDI-QI manner by simply letting Phil receive quantum inputs. Of course, providing more parties with quantum inputs (or here, also providing Fiona with some input, or letting Phil also have outputs) could only make the certification more efficient.

Finally, as for any SDI-QI scenario considered in this work, the SDI-QI certification of a causally nonseparable process in the $(2{+}F)$- and $(P{+}2{+}F)$-partite scenarios can readily be turned into a NDI certification by introducing new parties that share maximally entangled states with whichever parties receive quantum inputs, and by self-testing these states and the new parties' measurements. For the quantum switch, in the case where only Phil receives a quantum input that he directly feeds into the quantum switch, as considered above (see Eq. (E16)), this leads to a very similar scenario to that considered in Ref. [30], which we shall discuss in the next appendix.

---

[12] When considering isomorphic Hilbert spaces $\mathcal{H}^X$ and $\mathcal{H}^{X'}$, we take their computational bases $\{|i\rangle^{X^{(\prime)}}\}_i$ to be in one-to-one correspondence.

Appendix F: Comparing the NDI and DRF certifications of the quantum switch

In this appendix we provide some further comparison between our NDI certification of the causal nonseparability of the quantum switch and the DRF certification of Ref. [30]. This complements the discussion in the main text of the underlying assumptions behind these two approaches.

As mentioned at the end of Appendix E 4, the SDI-QI certification of the quantum switch described therein can be transformed into a NDI certification by following the same procedure as for the bipartite scenario discussed in detail in the main text. In particular, the SDI-QI certification is transformed first into a NSDI-QI certification by adding an extra separated party, Emily, who shares a maximally entangled state with Phil and remotely prepares his quantum inputs $\rho_t^{P'}$, and then into a NDI certification by self-testing both the state Phil and Emily share, and Emily's measurements. The certification is then obtained by violating a witness inequality using the observed correlations $P_*(p,a,b,f|t,z,w) := P(p,a,b,f|u{=}*,t,z,w,e{=}0) = P'(p,a,b,f|\rho_t^{P'},z,w)$, analogously to the bipartite setting. This NDI scenario is shown in Fig. 6.

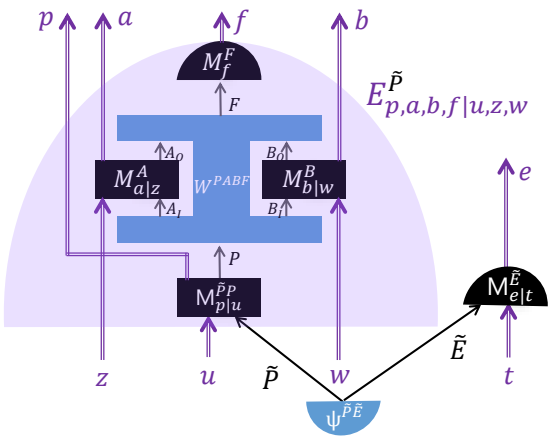

FIG. 6. Scenario for the NDI certification of the quantum switch with a global past. The remote preparation of a quantum input $\rho_t^{P'}$ is self-tested by first self-testing the (untrusted) shared state $\psi^{\tilde{P}\tilde{E}}$ (to contain a maximally entangled state) and Emily's measurements $\mathsf{M}^{\tilde{E}}_{e|t}$ (to effectively be tomographically complete projectors) from the correlations $P(p,e|u,t)$ shared by Phil and Emily. One can then certify the causal nonseparability of the family of D-POVMs $\{(E^{\tilde{P}}_{p,a,b,f|u,z,w})_{p,a,b,f}\}_{u,z,w}$ and thus the causal nonseparability of the process matrix W^{PABF}.

As we already noted, Ref. [30] showed how to device-independently certify the causal indefiniteness of the quantum switch by the violation of an inequality derived from the assumptions of Definite causal order, Relativistic causality, and Free interventions. The structure of the "DRF scenario" they considered is very similar to that of our NDI certification, but there are some important differences. In the DRF scenario, a (space-like) separated Emily is also added to a standard quantum switch scenario. However, the scenario contains no Phil and, in contrast to the NDI certification, Fiona must be provided with a classical input (which we could also have done, but did not need to). By having Alice share an entangled state with the control system in the global past of a quantum switch (directly incorporated in the process, instead of being provided by Phil), the parties can generate correlations violating the DRF inequality.

The DRF inequality itself, rephrased here in notation consistent with Fig. 6 and with v denoting Fiona's classical input, has the form [30]

$$\mathcal{I}_{\mathrm{DRF}} = P(e=0,b=z|t=0) + P(e=1,a=w|t=0) + P(f\oplus e = vt|z=w=0) \leq \frac{7}{4}. \tag{F1}$$

One can interpret this inequality as certifying (i) a space-like separated "control of causal order" (the first two terms of $\mathcal{I}_{\mathrm{DRF}}$), and (ii) the non-classical, indeterministic behavior of the control (the last term of $\mathcal{I}_{\mathrm{DRF}}$). The latter of these can be interpreted as a CHSH inequality tested between Fiona (acting after the quantum switch, and who is thus provided with inputs) and Emily. It can be violated if the control system passes transparently through the process, which can be ensured to be the case if Alice and Bob perform appropriate operations (on inputs $z=w=0$). On the other hand, while the NDI certification also relies on the violation of a Bell inequality, it is tested between Emily and Phil—who is in the causal past of the quantum switch and has classical inputs and outputs—to self-test the quantum

state they share. Fiona, in this case, needs no classical input and performs a fixed measurement. Curiously, however, we were able to obtain NDI certification using exactly the same measure-and-prepare instruments for Alice and Bob (see Eq. (E17)) as were used to violate the DRF inequality of Eq. (F1) in [30].

Finally, although we saw that our approach can, in the SDI-QI setting, certify the causal nonseparability of the quantum switch even in the presence of substantial noise on the quantum switch, the NDI certification is less tolerant to noise on the entangled state shared between Phil and Emily. Indeed, although we saw in Appendix C that the NDI certification is "robust", the actual reduction in Bell-inequality violation that can be tolerated in the self-testing procedure is typical small in absolute terms [17].[13] The DRF inequality, on the other hand, appears to be much more robust to noise on the entangled state shared between Emily and Fiona after passing through the quantum switch, as any violation of the classical bound of 3/4 of the last term in $\mathcal{I}_{\mathrm{DRF}}$ (corresponding to a CHSH inequality between them) can potentially lead to a violation of the DRF inequality. A more precise quantitive comparison of the robustness of the NDI and DRF certifications under different noise models would be an interesting direction of future study.

[13] We leave it as an open question to calculate the exact robustness, e.g. for the optimal witness inequality for the certification of the D-POVM obtained from the quantum switch from the operations of Eq. (E17). Note, however, that with a noiseless maximally entangled state shared between Phil and Emily, the NDI certification can tolerate the same amount of noise on the quantum switch as the SDI-QI certification.