Supplemental Materials

Masahito Hayashi[1,2,3*†], Longyang Cao[4,5*], Baichu Yu[4,6‡], Yuan-Yuan Zhao[4§]
[1]School of Data Science, The Chinese University of Hong Kong, Shenzhen, Longgang District, Shenzhen 518172, China
[2]International Quantum Academy, Futian District, Shenzhen 518048, China
[3]Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan
[4]Quantum Science Center of Guangdong-Hong Kong-Macao Greater Bay Area, Shenzhen 518045, China
[5]Shenzhen University, Shenzhen 518060, China
[6]Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Nanshan District, Shenzhen 518055, China

Operational assumptions and the meaning of "characterization-free"

Throughout this work, *characterization-free* means that the inference does not require an explicit calibrated operator-level description of the parties' devices (e.g., matrix representations, eigenbases, or an implementation model of the POVM and preparation operators). Instead, our guarantees rely on explicit *operational non-degeneracy* assumptions on the accessible settings, stated as (S1) and (S2) in the main text. We emphasize that "characterization-free" is not the same as device-independence in the usual sense, but only refers to not assuming calibrated operator descriptions.

More specifically, condition (S1) assumes that the chosen measurement and preparation operators span the relevant local Hermitian-operator spaces. Operationally, this presumes enough calibration/coordination to speak about spanning and linear independence of the implemented settings, while not requiring their explicit calibrated forms. Under (S1), Theorem 1 yields an if-and-only-if identification criterion, but requires $d^2_{(I,1)} d^2_{(I,2)} d^2_{(O,1)} d^2_{(O,2)}$ settings to generate tomographically complete correlations.

---

[*] These authors contributed equally to this work.
[†] E-mail: hmasahito@cuhk.edu.cn.
[‡] E-mail: yubc@sustech.edu.cn.
[§] E-mail: zhaoyuanyuan1@quantumsc.cn.

On the other hand, condition (S2) replaces the explicit spanning condition in (S1) by a *genericity* requirement: the implemented POVM elements and prepared states vary continuously in a sufficiently rich (full-dimensional) manner so that measure-zero "bad sets" are avoided almost surely. This requirement does *not* ask for calibrated operator forms, but it is not automatically easier to satisfy than (S1): achieving full-dimensional variability typically requires many independent degrees of freedom (scaling with the real dimension of the relevant operator manifolds, $\sim d^2$), and thus may necessitate knowledge of the local dimension (or at least an upper bound) when the variability must be engineered.

In the present work, we take (S2) in an engineered sense: the experimenter deliberately randomizes the local settings so as to ensure sufficiently rich (full-dimensional) exploration, which keeps the analysis characterization-free while still avoiding any calibrated operator-level characterization. A complementary *drift-driven* route may also realize (S2) effectively, where uncontrolled quasi-static calibration offsets are fixed within a run but vary across runs and explore sufficiently many independent directions; if this occurs, the dimension bound (or an upper bound) needed for engineering may no longer be necessary. Establishing practical diagnostic criteria to certify when such drift-driven richness holds is therefore an important direction for future work.

Proof of Theorem 1

To prove Theorem 1, we prepare the following lemma.

Lemma 1. *The distribution P_{J_1,J_2,K_1,K_2} satisfies the Markovian chains $J_1 - K_1 - J_2 - K_2$ and $K_1 - J_1 - K_2 - J_2$ if and only if it satisfies the chain $(K_1, J_1) \perp (J_2, K_2)$. Also, the distribution P_{J_1,J_2,K_1,K_2} satisfies the Markovian chains $(J_1, K_1) - J_2 - K_2$ and $K_1 - J_1 - (J_2, K_2)$ if and only if it satisfies the Markovian chain $K_1 - J_1 - J_2 - K_2$.*

Proof. For the first statement, the part of "if" is immediate. We show the part of "only if". Assume the Markovian chains $J_1 - K_1 - J_2 - K_2$ and $K_1 - J_1 - K_2 - J_2$, which imply the relations $P_{J_1|J_2,K_2}(j_1|j_2, k_2) = P_{J_1|J_2}(j_1|j_2)$ and $P_{J_1|J_2,K_2}(j_1|j_2, k_2) = P_{J_1|K_2}(j_1|k_2)$, respectively. Hence, we have $P_{J_1|J_2}(j_1|j_2) = \sum_{k_2} P_{K_2}(k_2)P_{J_1|K_2}(j_1|k_2) = P_{J_1}(j_1)$, and $P_{J_1|K_2}(j_1|k_2) = P_{J_1}(j_1)$, which implies the Markovian chain $K_1 - J_1 \perp J_2 - K_2$.

For the second statement, the part of "if" is immediate. We show the part of "only if". Assume the Markovian chains $(J_1, K_1) - J_2 - K_2$ and $K_1 - J_1 - (J_2, K_2)$ The Markov chain implies the relation $P_{J_1,K_1|J_2,K_2}(j_1, k_1|j_2, k_2) = P_{J_1,K_1|J_2}(j_1, k_1|j_2)$ which yields the relation $P_{K_1|J_2,K_2}(k_1|j_2, k_2) = P_{K_1|J_2}(k_1|j_2)$. Thus, $P_{K_1|J_1,J_2,K_2}(k_1|j_1, j_2, k_2) = \frac{P_{J_1,K_1|J_2,K_2}(j_1,k_1|j_2,k_2)}{P_{K_1|J_2,K_2}(k_1|j_2,k_2)} = \frac{P_{J_1,K_1|J_2}(j_1,k_1|j_2)}{P_{K_1|J_2}(k_1|j_2)} = P_{K_1|J_1,J_2}(k_1|j_1, j_2)$. Hence, we have $K_1 - (J_1, J_2) - K_2$. The Markovian chain $K_1 - J_1 - (J_2, K_2)$ implies the Markovian chain $K_1 - J_1 - J_2$. The combination of $K_1 - (J_1, J_2) - K_2$ and $K_1 - J_1 - J_2$ implies $K_1 - J_1 - J_2 - K_2$. □

For the relation between constraints of process matrices and the Markovian conditions for the distribution P_{J_1,J_2,K_1,K_2}, we have the following equivalence conditions.

Proof of Theorem 1:

Step 1: We show the equivalence between Eq. (9) and $(J_1,K_1)-J_2-K_2$. The equivalence for the $2\to 1$ direction can be shown in the same way.

Since the condition (9) implies

$$\begin{aligned}
&P_{K_2,J_2,J_1,K_1}(k_2,j_2,j_1,k_1) = \mathrm{Tr}(W_S C[\Gamma_{1,(j_1,k_1)}]\otimes C[\Gamma_{2,(j_2,k_2)}]) \\
=&P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}(W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \rho_{k_2})) \\
=&P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)\otimes \rho_{mix,(O,2)})(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \rho_{k_2})] \\
=&\frac{P_{K_2|J_2}(k_2|j_2)}{d_{(O,2)}}\mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2})],
\end{aligned} \tag{E1}$$

then we have

$$P_{J_2,J_1,K_1}(j_2,j_1,k_1) = \frac{1}{d_{(O,2)}}\mathrm{Tr}(\mathrm{Tr}_{(O,2)}W_S)(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}), \tag{E2}$$

$$P_{K_2|J_2,J_1,K_1}(k_2|j_2,j_1,k_1) = P_{K_2|J_2}(k_2|j_2). \tag{E3}$$

Hence, the condition (9) implies the Markovian condition $(J_1,K_1)-J_2-K_2$, i.e., the condition that the random variable K_2 is decided only by the random variable J_2.

Next, we show that the Markovian condition $(J_1,K_1)-J_2-K_2$ implies the condition (9). We assume that the Markovian condition $(J_1,K_1)-J_2-K_2$ holds. The joint distribution P_{K_2,J_2,J_1,K_1} is written as

$$P_{K_2,J_2,J_1,K_1}(k_2,j_2,j_1,k_1) = P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \rho_{k_2})]. \tag{E4}$$

We have

$$P_{J_2,J_1,K_1}(j_2,j_1,k_1) = \mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \sum_{k_2} P_{K_2|J_2}(k_2|j_2)\rho_{k_2})]. \tag{E5}$$

Thus, the Markovian condition $(J_1,K_1)-J_2-K_2$ implies

$$P_{K_2|J_2}(k_2|j_2) = P_{K_2|J_2,J_1,K_1}(k_2|j_2,j_1,k_1) = \frac{P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \rho_{k_2})]}{\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \sum_{k_2'} P_{K_2|J_2}(k_2'|j_2)\rho_{k_2'})]} \tag{E6}$$

Thus, we have

$$\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \sum_{k_2'} P_{K_2|J_2}(k_2'|j_2)\rho_{k_2'})] = \mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}]\otimes M_{2,j_2}\otimes \rho_{k_2})]. \tag{E7}$$

That is, $\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}] \otimes M_{2,j_2} \otimes \rho_{k_2})]$ does not depend on k_2. Since ρ_{k_1} is tomographically complete,

$$\begin{aligned} &\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}] \otimes M_{2,j_2} \otimes \rho_{k_2})] \\ =&\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}] \otimes M_{2,j_2} \otimes \rho_{mix,(O,2)})] \\ =&\frac{1}{d_{(O,2)}}\mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(C[\Gamma_{1,(j_1,k_1)}] \otimes M_{2,j_2})], \end{aligned} \tag{E8}$$

which is equivalent to the condition (9).

Therefore, the condition (9) is equivalent to the Markovian condition $(J_1, K_1) - J_2 - K_2$.

Step 2: We show the equivalence between (7) and $J_1 - K_1 - J_2 - K_2$. The equivalence between the $2 \to 1$ direction can be shown in the same way.

Since the condition (7) implies

$$\begin{aligned} &\mathrm{Tr}[W_S(C[\Gamma_{1,(j_1,k_1)}] \otimes C[\Gamma_{2,(j_2,k_2)}])] = P_{K_1|J_1}(k_1|j_1)P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[W_S(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2} \otimes \rho_{k_2})] \\ =&P_{K_1|J_1}(k_1|j_1)P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[(\rho_1 \otimes C[\Lambda_C] \otimes I_{(O,2)})(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2} \otimes \rho_{k_2})] \\ =&P_{K_1|J_1}(k_1|j_1)P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}(\rho_1 M_{1,j_1})\mathrm{Tr}(C[\Lambda_C](\rho_{k_1} \otimes M_{2,j_2})), \end{aligned} \tag{E9}$$

which shows Markov conditions $J_1 - K_1 - J_2$ and $(J_1, K_1) - J_2 - K_2$, therefore condition (7) implies the Markovian condition $J_1 - K_1 - J_2 - K_2$.

Next, we show that the Markovian condition $J_1 - K_1 - J_2 - K_2$ implies the condition (7). We assume that the Markovian condition $J_1 - K_1 - J_2 - K_2$ holds.

The joint distribution P_{K_2,J_2,J_1,K_1} is written as

$$P_{K_2,J_2,J_1,K_1}(k_2, j_2, j_1, k_1) = P_{K_1|J_1}(k_1|j_1)P_{K_2|J_2}(k_2|j_2)\mathrm{Tr}[W_S(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2} \otimes \rho_{k_2})]. \tag{E10}$$

Since the Markovian condition $J_1 - K_1 - J_2 - K_2$ is a special case of the Markovian condition $(J_1, K_1) - J_2 - K_2$, the conclusion of Step 1 implies that the condition (9) holds. Thus, we have

$$\begin{aligned} &\mathrm{Tr}W_S(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2} \otimes \rho_{k_2}) \overset{(a)}{=} \mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2})](\mathrm{Tr}\rho_{mix,(O,2)}\rho_{k_2}) \\ =&\frac{1}{P_{K_1|J_1}(k_1|j_1)d_{(O,2)}}\mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(M_{1,j_1} \otimes P_{K_1|J_1}(k_1|j_1)\rho_{k_1} \otimes M_{2,j_2})] \overset{(b)}{=} \frac{P_{J_1,K_1,J_2}(j_1,k_1,j_2)}{P_{K_1|J_1}(k_1|j_1)} \end{aligned} \tag{E11}$$

$$=\frac{P_{J_1}(j_1)P_{J_1,K_1,J_2}(j_1,k_1,j_2)}{P_{K_1,J_1}(k_1,j_1)} = P_{J_1}(j_1)P_{J_2|K_1,J_1}(j_2|k_1,j_1) \overset{(c)}{=} P_{J_1}(j_1)P_{J_2|K_1}(j_2|k_1), \tag{E12}$$

where (a) follows from (9), (b) is derived by substituting $P_{K_1|J_1}(k_1|j_1)M_{1,j_1} \otimes \rho_{k_1}$ into $C[\Gamma_{1,(j_1,k_1)}]$ in (E2), (c) follows from the Markovian condition $J_1 - K_1 - J_2 - K_2$. Thus, (E12) implies

$$\begin{aligned} &\mathrm{Tr}[(\mathrm{Tr}_{(I,1)(O,2)}W_S)(\rho_{k_1} \otimes M_{2,j_2})] = \mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(I_{(I,1)} \otimes \rho_{k_1} \otimes M_{2,j_2})] \\ =&\sum_{j_1}\mathrm{Tr}[(\mathrm{Tr}_{(O,2)}W_S)(M_{1,j_1} \otimes \rho_{k_1} \otimes M_{2,j_2})] = \sum_{j_1}P_{J_1}(j_1)P_{J_2|K_1}(j_2|k_1) = P_{J_2|K_1}(j_2|k_1). \end{aligned} \tag{E13}$$

The combination of (E12) and (E13) yields

$$\frac{\mathrm{Tr}[W_S(M_{1,j_1}\otimes\rho_{k_1}\otimes M_{2,j_2}\otimes\rho_{k_2})]}{\mathrm{Tr}[(\mathrm{Tr}_{(I,1)(O,2)}W_S)(\rho_{k_1}\otimes M_{2,j_2})]}=\frac{P_{J_1}(j_1)P_{J_2|K_1}(j_2|k_1)}{P_{J_2|K_1}(j_2|k_1)}=P_{J_1}(j_1). \tag{E14}$$

Hence,

$$\begin{aligned}&\mathrm{Tr}[W_S(M_{1,j_1}\otimes\rho_{k_1}\otimes M_{2,j_2}\otimes\rho_{k_2})]=P_{J_1}(j_1)\mathrm{Tr}[(\mathrm{Tr}_{(I,1)(O,2)}W_S)(\rho_{k_1}\otimes M_{2,j_2})]\\=&P_{J_1}(j_1)\mathrm{Tr}[((\mathrm{Tr}_{(I,1)(O,2)}W_S)\otimes I_{(O,2)})(\rho_{k_1}\otimes M_{2,j_2}\otimes\rho_{k_2})].\end{aligned} \tag{E15}$$

Since $M_{2,j_2}\otimes\rho_{k_2}$ is tomographically complete, there exist coefficients $t(j_2,k_2)$ such that $\sum_{j_2,k_2}t(j_2,k_2)M_{2,j_2}\otimes\rho_{k_2}=I_{(I,2)(O,2)}$. Thus, we have

$$\begin{aligned}&\mathrm{Tr}(\mathrm{Tr}_{(I,2)(O,2)}W_S)(M_{1,j_1}\otimes\rho_{k_1})=\mathrm{Tr}W_S(M_{1,j_1}\otimes\rho_{k_1}\otimes I_{(I,2)(O,2)})\\=&\mathrm{Tr}W_S(M_{1,j_1}\otimes\rho_{k_1}\otimes\sum_{j_2,k_2}t(j_2,k_2)M_{2,j_2}\otimes\rho_{k_2})\\\overset{(a)}{=}&P_{J_1}(j_1)\mathrm{Tr}((\mathrm{Tr}_{(I,1)(O,2)}W_S)\otimes I_{(O,2)})(\rho_{k_1}\otimes\sum_{j_2,k_2}t(j_2,k_2)M_{2,j_2}\otimes\rho_{k_2})=P_{J_1}(j_1)\mathrm{Tr}((\mathrm{Tr}_{(I,1)(O,2)}W_S)\otimes I_{(O,2)})(\\=&P_{J_1}(j_1)\mathrm{Tr}[((\mathrm{Tr}_{(I,1)(I,2)(O,2)}W_S)\rho_{k_1}]\mathrm{Tr}I_{(O,2)}=P_{J_1}(j_1)d_{(O,2)}\mathrm{Tr}((\mathrm{Tr}_{(I,1)(I,2)(O,2)}W_S)\rho_{k_1},\end{aligned} \tag{E16}$$

where (a) follows from (E15). Since ρ_{k_1} is tomographically complete, there exist coefficients $t'(k_1)$ such that $\sum_{k_1}t'(k_1)\rho_{k_1}=I_{(O,1)}$,

$$\begin{aligned}&\mathrm{Tr}(\mathrm{Tr}_{(O,1)(I,2)(O,2)}W_S)M_{1,j_1}=\mathrm{Tr}(\mathrm{Tr}_{(I,2)(O,2)}W_S)(M_{1,j_1}\otimes I_{(O,1)})=\mathrm{Tr}(\mathrm{Tr}_{(I,2)(O,2)}W_S)(M_{1,j_1}\otimes\sum_{k_1}t'(k_1)\rho_{k_1}\\\overset{(a)}{=}&d_{(O,2)}P_{J_1}(j_1)\mathrm{Tr}((\mathrm{Tr}_{(I,1)(I,2)(O,2)}W_S)\sum_{k_1}t'(k_1)\rho_{k_1}=d_{(O,2)}P_{J_1}(j_1)\mathrm{Tr}((\mathrm{Tr}_{(I,1)(I,2)(O,2)}W_S)I_{(O,1)}\\=&d_{(O,2)}P_{J_1}(j_1)\mathrm{Tr}W_S=d_{(O,2)}^2d_{(O,1)}P_{J_1}(j_1),\end{aligned} \tag{E17}$$

where (a) follows from (E16).

The combination of (E14) and (E17) implies

$$\frac{\mathrm{Tr}W_S(M_{1,j_1}\otimes\rho_{k_1}\otimes M_{2,j_2}\otimes\rho_{k_2})}{\mathrm{Tr}(\mathrm{Tr}_{(I,1)(O,2)}W_S)(\rho_{k_1}\otimes M_{2,j_2})}=d_{(O,1)}^{-1}d_{(O,2)}^{-2}\mathrm{Tr}(\mathrm{Tr}_{(O,1)(I,2)(O,2)}W_S)M_{1,j_1}. \tag{E18}$$

Under the tomographically complete condition, Eq. (E18) implies that we can always find an input state ρ_1 and a channel Λ_C such that $W_S=\rho_1\otimes C[\Lambda_C]\otimes I_{(O,2)}$, which is exactly the condition (7).

Therefore, the condition (7) is equivalent to the Markovian condition $J_1-K_1-J_2-K_2$.

Step 3: We show the equivalence between (6) and $K_1-J_1-J_2-K_2$. The condition (6) is equivalent to the combination of the condition (9) for $W_{Q,1\to 2}$ and the analogous condition for

$W_{Q,2\to 1}$. These two conditions are equivalent to $(J_1, K_1) - J_2 - K_2$ and $K_1 - J_1 - (J_2, K_2)$, respectively. Hence, The condition (6) is equivalent to the combination of $(J_1, K_1) - J_2 - K_2$ and $K_1 - J_1 - (J_2, K_2)$, which is equivalent to $K_1 - J_1 - J_2 - K_2$, as shown in Lemma 1.

We show the equivalence between (4) and $K_1 - J_1 \perp J_2 - K_2$. The condition (4) is equivalent to the combination of the condition (7) for $W_{N,1\to 2}$ and the analogous condition for $W_{N,2\to 1}$. These two conditions are equivalent to $J_1 - K_1 - J_2 - K_2$ and $K_1 - J_1 - K_2 - J_2$, respectively. Hence, The condition (4) is equivalent to the combination of $J_1 - K_1 - J_2 - K_2$ and $K_1 - J_1 - K_2 - J_2$, which is equivalent to $K_1 - J_1 \perp J_2 - K_2$, as shown in Lemma 1.

□

Combining the above methods, we are able to identify all strategy classes in (14).

Proof of theorem 2

We prove a useful lemma before the proof of the theorem.

Lemma 2. *Let ρ_{AB} be a density operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ such that $\rho_{AB} \neq \rho_A \otimes \rho_B$, where $\rho_A := \mathrm{Tr}_B[\rho_{AB}]$ and $\rho_B := \mathrm{Tr}_A[\rho_{AB}]$. Let $\mathcal{M}_A = \{A_0, A_1, \ldots, A_{n-1}\}$ and $\mathcal{M}_B = \{B_0, B_1, \ldots, B_{m-1}\}$ be POVMs on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Assume that the random choice of $(\mathcal{M}_A, \mathcal{M}_B)$ is continuous in the following sense: there exist finite-dimensional local coordinates for the POVM sets such that the joint law of the coordinates is absolutely continuous with respect to Lebesgue measure, and the choices of $\mathcal{M}_A$ and $\mathcal{M}_B$ are independent.*

Define $p_{AB}(a,b) := \mathrm{Tr}[\rho_{AB}(A_a \otimes B_b)]$ and marginals $p_A(a) := \sum_b p_{AB}(a,b)$, $p_B(b) := \sum_a p_{AB}(a,b)$. Then, with probability one, the outcomes are not independent, i.e., there exists at least one pair (a,b) such that $p_{AB}(a,b) \neq p_A(a)p_B(b)$. Equivalently, the event $\{p_{AB}(a,b) = p_A(a)p_B(b)\ \forall a,b\}$ has probability 0.

Proof. Set $\Delta := \rho_{AB} - \rho_A \otimes \rho_B$. By assumption, $\Delta \neq 0$.
Step 1 (Fixing one pair (a,b) is sufficient). Let $\mathcal{E}$ be the event that $p_{AB}(a,b) = p_A(a)p_B(b)$ holds for all (a,b). For any fixed pair $(a^\star, b^\star)$, define $\mathcal{E}_{a^\star b^\star} := \{p_{AB}(a^\star,b^\star) = p_A(a^\star)p_B(b^\star)\}$. Then $\mathcal{E} \subseteq \mathcal{E}_{a^\star b^\star}$, hence it suffices to show $\Pr(\mathcal{E}_{a^\star b^\star}) = 0$ for one fixed pair. We fix such a pair and write it as (a,b).
Step 2 (Independence implies a bilinear constraint). We have $p_{AB}(a,b) = \mathrm{Tr}[\rho_{AB}(A_a \otimes B_b)]$ and $p_A(a)p_B(b) = \mathrm{Tr}[(\rho_A \otimes \rho_B)(A_a \otimes B_b)]$. Thus $p_{AB}(a,b) = p_A(a)p_B(b)$ is equivalent to

$$\mathrm{Tr}\big[\Delta\,(A_a \otimes B_b)\big] = 0. \tag{E19}$$

Step 3 (The solution set is Lebesgue-null). Given a fixed labels (a,b), we have randomized operators A_a, B_b; they are random Hermitian matrices. Choose real bases $\{E_i\}_{i=1}^{d_A^2}$ of $\mathsf{Herm}(\mathcal{H}_A)$ and $\{F_j\}_{j=1}^{d_B^2}$ of $\mathsf{Herm}(\mathcal{H}_B)$, and write $A_a = \sum_{i=1}^{d_A^2} x_i E_i, \qquad B_b = \sum_{j=1}^{d_B^2} y_j F_j$ by using $x \in \mathbb{R}^{d_A^2}$ and $y \in \mathbb{R}^{d_B^2}$. Then, by setting $c_{ij} := \mathrm{Tr}\big[\Delta(E_i \otimes F_j)\big]$, the left-hand side of (E19) becomes a real polynomial (indeed bilinear) $g(x,y) := \mathrm{Tr}\big[\Delta(A_a \otimes B_b)\big] = \sum_{i,j} c_{ij}\, x_i y_j$ of $(x,y) \in \mathbb{R}^{d_A^2} \times \mathbb{R}^{d_B^2}$. Because $\Delta \neq 0$ and $\{E_i \otimes F_j\}_{i,j}$ is a basis of $\mathsf{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$, at least one coefficient c_{ij} is nonzero, and consequently g is not the zero polynomial.

By a standard measure-theoretic fact, the set that a nonzero polynomial in $\mathbb{R}^N$ takes the value zero has Lebesgue measure zero (*1*). Therefore, the set $\{(x,y) : g(x,y) = 0\}$ is a Lebesgue-null subset of $\mathbb{R}^{d_A^2 + d_B^2}$.

Step 4 (Lebesgue-null ⇒ probability zero under continuous sampling). By the continuity assumption on the sampling (absolute continuity w.r.t. Lebesgue measure in local coordinates), any Lebesgue-null set is hit with probability 0. Hence, for the fixed pair (a,b), the probability that (E19) holds is 0, i.e., $\Pr(\mathcal{E}_{ab}) = 0$. Combining with Step 1 yields $\Pr(\mathcal{E}) = 0$. □

Now we prove the theorem.

Proof of Theorem 2. Throughout this proof, Charlie's strategy is represented by a fixed process matrix W_S (independent of Alice's and Bob's random choices), as assumed in item (i) of the theorem. Alice and Bob choose their local MP operations according to condition (S2), i.e., by independent continuous random variables $T = (T_1, T_2, T_3, T_4)$. We write the corresponding collection of local operators as

$$\Theta = \big(\{M_{1,j_1}\}, \{M_{2,j_2}\}, \{\rho_{k_1}\}, \{\rho_{k_2}\}\big).$$

Step 1. Fix a target class S_X. We prove the contrapositive statement: if $S \notin S_X$, then the induced distribution of (J_1, K_1, J_2, K_2) violates the Markovian condition $Mar(S_X)$ with probability 1 with respect to the randomness of T. Equivalently, the set of random choices Θ for which $Mar(S_X)$ still holds is a Lebesgue-null (measure-zero) subset of the parameter space of Θ.

Step 2. The observed statistics are given by the generalized Born rule for process matrices (*2*):

$$P(j_1, j_2, k_1, k_2) = \mathrm{Tr}\Big[W_S\big(C[\Gamma_{1,(j_1,k_1)}] \otimes C[\Gamma_{2,(j_2,k_2)}]\big)\Big],$$

where $C[\cdot]$ denotes the Choi representation (*3*). Markovian conditions in Eq. (15) of the main text are conditional-independence relations among (J_1, K_1, J_2, K_2). Each such relation can be tested by checking whether certain *induced bipartite operators* (defined below) have a tensor-product (product-matrix) form across a specified bipartition.

Define the following "slice" operators obtained by partially contracting W_S with a local element and tracing out a subsystem:

$$W^S_{M_{2,j_2}} := \mathrm{Tr}_{I_2}\Big[W_S\big(I_{I_1O_1} \otimes M_{2,j_2} \otimes I_{O_2}\big)\Big], \tag{E20}$$

$$W^S_{(\rho_{k_1}, I_{O_2})} := \mathrm{Tr}_{O_1O_2}\Big[W_S\big(I_{I_1} \otimes \rho_{k_1} \otimes I_{I_2O_2}\big)\Big], \tag{E21}$$

$$W^S_{(M_{1,j_1}, I_{O_2})} := \mathrm{Tr}_{I_1O_2}\Big[W_S\big(M_{1,j_1} \otimes I_{I_1O_1O_2}\big)\Big]. \tag{E22}$$

Step 3. Let A and B be two (finite-dimensional) Hilbert spaces. An operator X on $A \otimes B$ is said to be a *product matrix across* $\mathcal{B}(A) \otimes \mathcal{B}(B)$ if it can be written as

$$X = X_A \otimes X_B$$

for some Hermitian operators X_A on A and X_B on B. This is the operator-level analogue of statistical independence for joint distributions generated by local measurements. Lemma 2 states that if a bipartite density operator is not a product, then random independent local POVMs produce a non-independent joint distribution with probability 1. Because the random choice in (S2) is independent and continuous, Lemma 2 applies whenever one of the slice operators in (E20)–(E22) fails to be a product matrix across the relevant bipartition.

Concretely: if for some fixed conditioning index (e.g., a particular j_2 or k_1) the corresponding slice operator is non-product across a specified bipartition, then the set of random local choices Θ for which the induced outcomes would *fake* the required independence is measure zero; hence the Markovian condition fails with probability 1.

Step 4. We now list, for each strategy class S_X in Eq. (14) of the main text, a product-structure property of the slices that is necessary for $S \in S_X$. If $S \notin S_X$, at least one such property is violated, and then Step 4 implies that $Mar(S_X)$ fails almost surely.

(i) The classes $S_{Q,1\to2}$ and $S_{Q,2\to1}$. For $S_{Q,1\to2}$, the defining one-way causal/memory structure implies that for every j_2, the slice $W^S_{M_{2,j_2}}$ must be a product matrix across

$$\mathcal{B}(\mathcal{H}_{I_1} \otimes \mathcal{H}_{O_1}) \;\otimes\; \mathcal{B}(\mathcal{H}_{O_2}).$$

If $S \notin S_{Q,1\to2}$, then there exists some j_2 for which this product property fails. By Step 4 (Lemma 2), the corresponding independence relation in $Mar(S_{Q,1\to2})$ is violated with probability 1. The case $S_{Q,2\to1}$ is analogous.

(ii) The classes $S_{N,1\to2}$ and $S_{N,2\to1}$. For $S_{N,1\to2}$ (sequential without memory), the structure implies two product requirements: for all k_1, the slice $W^S_{(\rho_{k_1}, I_{O_2})}$ must be a product matrix across

$$\mathcal{B}(\mathcal{H}_{I_1}) \;\otimes\; \mathcal{B}(\mathcal{H}_{I_2}),$$

and for all j_2, the slice $W^S_{M_{2,j_2}}$ must be a product matrix across

$$\mathcal{B}(\mathcal{H}_{I_1} \otimes \mathcal{H}_{O_1}) \;\otimes\; \mathcal{B}(\mathcal{H}_{O_2}).$$

If $S \notin S_{N,1\to2}$, then at least one of these properties fails (for some k_1 or j_2), and Step 4 yields that $Mar(S_{N,1\to2})$ is violated with probability 1. The direction $2 \to 1$ is analogous.

(iii) The class S_Q (quantum parallel). For S_Q, the parallel structure implies that at least one of the following product conditions must hold for all relevant local indices (depending on the slicing convention): for all j_1, the slice $W^S_{(M_{1,j_1}, I_{O_2})}$ is product across

$$\mathcal{B}(\mathcal{H}_{O_1}) \otimes \mathcal{B}(\mathcal{H}_{I_2}),$$

and for all j_2, the slice $W^S_{M_{2,j_2}}$ is product across

$$\mathcal{B}(\mathcal{H}_{I_1} \otimes \mathcal{H}_{O_1}) \otimes \mathcal{B}(\mathcal{H}_{O_2}).$$

If $S \notin S_Q$, then at least one of these product properties fails for some index, and Step 4 implies that $Mar(S_Q)$ fails with probability 1.

(iv) The class S_I (individual). Finally, S_I requires that the process matrix itself factorizes across the two parties:

$$W_S = W_{I_1O_1} \otimes W_{I_2O_2} \quad \text{(up to an overall scalar)},$$

i.e., W_S is a product matrix across $\mathcal{B}(\mathcal{H}_{I_1} \otimes \mathcal{H}_{O_1}) \otimes \mathcal{B}(\mathcal{H}_{I_2} \otimes \mathcal{H}_{O_2})$. If $S \notin S_I$, then W_S is non-product across this bipartition, and Step 4 yields that $Mar(S_I)$ fails with probability 1.
Step 6 (Conclusion). In all cases above, if $S \notin S_X$, then at least one relevant slice operator is non-product across the partition that corresponds to $Mar(S_X)$. By Lemma 2 and the independent continuous random choice of local operations in (S2), the probability that the induced distribution satisfies $Mar(S_X)$ is 0. Equivalently, it violates $Mar(S_X)$ with probability 1. This completes the proof. □

Information of the Settings

In this part, we give the specific information of Alice and Bob's 7 different MP channel, and the parameters of realizing Charlies' strategies as well. Table E1 gives the 7 settings of measurement operators and preparation states for testing Markovian conditions in Eq. (15) of the main text (step 1). In this case, each party makes a two-outcome POVM, and prepares a output state out of two candidates. Except for setting 1 which is specially chosen as a reference, the operators in the remaining 6 settings are randomly chosen according to the Haar measure.

The conditional probability distributions $p(K_i|J_i)$ follows the condition

$$p(K_i = J_i) = 0.65 \tag{E23}$$

for $i = 1, 2$. That is, when Alice obtains a measurement result $J_1 = 1$, she prepares the state corresponding to $K_1 = 1$ with probability $p = 0.65$, and prepares the state corresponding to $K_1 = 2$ with $1 - p = 0.35$. When Alice obtains $J_1 = 2$, she prepares the state $K_1 = 2$ with probability $p = 0.65$, and prepares the state corresponding to $K_1 = 1$ with $1 - p = 0.35$. Bob's scheme is analogous.

Setting	$\mathbf{M_{1,j_1}}$	$\mathbf{M_{2,j_2}}$	$\rho_{\mathbf{1,k_1}}$	$\rho_{\mathbf{2,k_2}}$
1	$(0, 0)$	$(45, 0)$	$(0, 0)$	$(0, 0)$
	$(180, 0)$	$(-135, 0)$	$(45, 0)$	$(45, 0)$
2	$(36, 103)$	$(15, 340)$	$(0, 0)$	$(0, 0)$
	$(144, 283)$	$(165, 160)$	$(152, 158)$	$(45, 0)$
3	$(18, 76)$	$(87, 50)$	$(0, 0)$	$(0, 0)$
	$(162, 258)$	$(93, 230)$	$(149, 17)$	$(45, 0)$
4	$(2, 177)$	$(62, 143)$	$(0, 0)$	$(0, 0)$
	$(178, 357)$	$(118, 323)$	$(38, 85)$	$(45, 0)$
5	$(12, 212)$	$(1, 271)$	$(0, 0)$	$(0, 0)$
	$(168, 32)$	$(179, 91)$	$(172, 236)$	$(45, 0)$
6	$(25, 163)$	$(8, 270)$	$(0, 0)$	$(0, 0)$
	$(155, 343)$	$(172, 90)$	$(37, 248)$	$(45, 0)$
7	$(12, 292)$	$(134, 32)$	$(0, 0)$	$(0, 0)$
	$(168, 112)$	$(46, 212)$	$(75, 110)$	$(45, 0)$

Table E1: Choices of Alice's and Bob's measurement operators and preparation states. Each operator (pure) is represented by the angle of its Bloch vector in (θ, ϕ) format. In each setting, the operators in the first row correspond to $J_i, K_i = 1$ (the first choices of the measurement operator and prepared state), the ones in the second row correspond to $J_i, K_i = 2$ (the second choices of the measurement operator and prepared state). Note that since Bob's state preparation does not influence the result, we always set his two state as $(0, 0)$ and $(45, 0)$.

To test the nonlocality across J_1 and J_2, each of Alice and Bob makes two two-element POVMs. Since the state preparation of Alice and Bob does not influence this test, we do not consider the information of preparation states. The way we choose the 7 different measurement settings is as follow. We randomly generated many different measurement settings, and post-select 7 settings that lead to violation of CHSH inequality with maximally entangled states. We perform post-selection in step 2 to increase the probability of detecting nonlocality, because with randomly chosen measurement settings, the chance of a CHSH violation is only about 40%. In Table E2, we give the information of the measurement operators. Each measurement operator is a pure operator represented by by the angle of its Bloch vector in (θ, ϕ) format.

Setting	$\mathbf{M_{1,j_1}}$	$\mathbf{M_{2,j_2}}$
1	$\{(0,0), (180,0)\}$ $\{(90,0), (90,180)\}$	$\{(45,0), (135,180)\}$ $\{(135,0), (45,180)\}$
2	$\{(36,103), (144,283)\}$ $\{(49,46), (131,226)\}$	$\{(15,340), (165,160)\}$ $\{(79,211), (101,31)\}$
3	$\{(18,76), (162,256)\}$ $\{(79,95), (101,275)\}$	$\{(87,50), (93,230)\}$ $\{(48,291), (132,111)\}$
4	$\{(2,177), (178,357)\}$ $\{(61,19), (119,199)\}$	$\{(62,143), (118,323)\}$ $\{(28,59), (152,239)\}$
5	$\{(12,212), (168,32)\}$ $\{(18,252), (162,72)\}$	$\{(1,271), (179,91)\}$ $\{(68,262), (112,82)\}$
6	$\{(25,163), (155,343)\}$ $\{(88,300), (92,120)\}$	$\{(8,270), (172,90)\}$ $\{(109,24), (71,204)\}$
7	$\{(12,292), (168,112)\}$ $\{(91,330), (89,150)\}$	$\{(134,32), (46,212)\}$ $\{(33,58), (147,238)\}$

Table E2: Choices of Alice's and Bob's measurement operators to detect nonlocality. Each operator (pure) is represented by the angle of its Bloch vector in (θ, ϕ) format.

Now we introduce the realization of Charlie's strategies. Firstly we consider the parallel ones. For independent strategy $\mathcal{S}_I$, we let $\rho_1 = \rho_2 = |+\rangle\langle+|$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. For $\mathcal{S}_C$ we let $\rho_{12} = \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$. For $\mathcal{S}_Q$ we let $\rho_{12} = |\Phi\rangle\langle\Phi|$, where $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Note that the channel Λ_C on the output system does not influence the strategy so we just ignore it (or take it as an identity channel).

For the sequential strategies, we need to determine the input state ρ_1 and the intermediate channel Λ_C. For strategy $\mathcal{S}_{N,1\to 2}$, we let $\rho_1 = |0\rangle\langle 0|$, and Λ_C be an identity channel. For strategy $\mathcal{S}_{C,1\to 2}$, we introduce a common uniformly distributed binary random variable $\lambda = \{0, 1\}$ between ρ_1 and Λ_C. Let

$$\rho_{1,\lambda} = \begin{cases} |0\rangle\langle 0|, \ \lambda = 0 \\ |1\rangle\langle 1|, \ \lambda = 1 \end{cases}, \tag{E24}$$

and let Λ_C be identity channel when $\lambda = 0$ and be the bit flip channel when $\lambda = 1$. To realize the strategy $\mathcal{S}_{Q,1\to 2}$, we perform a quantum parallel strategy S_Q with probability $p = 0.75$. And with probability $1 - p = 0.25$, we prepare a maximally entangled state $\rho_{1X} = |\Phi\rangle\langle\Phi|$, where subscript X represents an auxiliary system, and the channel Λ_C in this case is chosen to be a controlled-NOT gate, for which the control is the auxiliary system X and the target is the output system $\mathcal{H}_{O,1}$ of Alice.

Table E3 gives the parameter of each optical device (the angles of half-wave plates and quarter-wave plates) to realize the strategies of Charlie's.

	H1	Q1	H2	Q2	H3	H4	H5	H6	H7
S_I	22.5	45	22.5	45	45	0	0	45	0
$S_C[1]$	0	0	0	0	45	0	0	45	0
$S_C[2]$	22.5	45	22.5	45	45	0	0	45	0
S_Q	22.5	45	0	0	0	0	0	45	0
$S_{N,1\to 2}$	0	0	22.5	45	0	0	0	0	0
$S_{C,1\to 2}[1]$	0	0	0	0	0	0	0	0	0
$S_{C,1\to 2}[2]$	0	0	22.5	45	0	0	0	0	0
$S_{Q,1\to 2}$	22.5	45	0	0	0	0	45	0	0

Table E3: Parameters to realize Charlie's different strategies on the optical platform. The Hi $(i = 1, 2, 3, 4, 5, 6, 7)$ and Qj $(j = 1, 2)$ are the half-wave plates and quarter-wave plates in Fig. 3 of the main text.

Detailed Results of The χ^2 Tests

This part reports the χ^2 values derived from the experimental statistics (in step 1) for the strategies S_I, S_C, S_Q, $S_{N,1\to 2}$, $S_{C,1\to 2}$, and $S_{Q,1\to 2}$.

In step 1, we need to check the Markovian condition for each strategy from lower levels to higher levels according to Fig. 2 in the main text. Note that condition $J_1 - K_1 - J_2 - K_2$ is checked by $J_1 - K_1 - J_2$ and $(J_1 - K_1) - J_2 - K_2$, condition $K_1 - J_1 - J_2 - K_2$ is checked by $K_1 - J_1 - J_2$ and $(J_1 - K_1) - J_2 - K_2$, condition $K_1 - J_1 - K_2 - J_2$ is checked by $J_1 - K_2 - J_2$ and $K_1 - J_1 - (J_2 - K_2)$. A strategy class is identified until it is accepted as a null-hypothesis, that is, the χ^2 of its corresponding Markovian condition is smaller than the critical value χ^2_{crit} under significance level $\alpha = 0.05$. The values of χ^2_{crit} are determined by the χ^2-table. The χ^2 values derived from the experimental statistics are listed in Table. E4.

	Markovian conditions	Set.1	Set.2	Set.3	Set.4	Set.5	Set.6	Set.7
S_I	①	0.00	0.09	0.03	0.00	0.00	0.00	0.00
S_C	①	472.15	51.23	51.78	17.15	70.90	44.83	53.82
	②	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	⑤	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S_Q	①	3031.15	1536.00	42.67	938.71	3442.62	2996.07	1127.69
	②	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	⑤	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{N,1\to 2}$	①	225.49	3307.40	39.90	146.60	3497.90	223.16	371.42
	③	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	⑤	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{C,1\to 2}$	②	55.46	700.06	22.62	176.12	1418.99	44.38	102.25
	③	121.23	633.89	48.42	188.03	1262.27	1164.19	278.38
	④	7.90	15.65	9.76	82.81	38.38	712.16	404.27
	⑤	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{Q,1\to 2}$	②	55.68	360.85	8.06	26.47	905.55	24.04	23.52
	③	3196.84	2770.93	25.00	1588.82	4014.75	5777.58	2087.61
	④	1919.97	2223.12	283.98	776.05	3921.35	4429.29	2664.50
	⑤	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table E4: The χ^2 values for each strategy under 7 different settings of MP channel. The Markovian conditions are labeled as: ① $(K_1, J_1) \perp (J_2, K_2)$, ② $K_1 - J_1 - J_2$, ③ $J_1 - K_1 - J_2$, ④ $J_1 - K_2 - J_2$, ⑤ $(J_1 - K_1) - J_2 - K_2$, ⑥ $K_1 - J_1 - (J_2 - K_2)$. The critical value $\chi^2_{crit} = 16.919$ for testing condition ①, $\chi^2_{crit} = 5.991$ for testing conditions ②-④, and $\chi^2_{crit} = 12.592$ for testing conditions ⑤-⑥.

References

1. Boris S. Mityagin. The zero set of a real analytic function. *Mathematical Notes*, 107(3):529–530, 2020.

2. Ognyan Oreshkov, Fabio Costa, and Caslav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092, 2012.

3. Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.