Supplemental material: Experimental demonstration of measurement-device-independent measure of quantum steering

Yuan-Yuan Zhao,[1,2,3] Huan-Yu Ku,[4,5] Shin-Liang Chen,[5,6,*] Hong-Bin Chen,[7] Franco Nori,[4,8] Guo-Yong Xiang,[1,3,†] Chuan-Feng Li,[1,3] Guang-Can Guo,[1,3] and Yueh-Nan Chen[5,‡]

[1]*Key Laboratory of Quantum Information, University of Science and Technology of China, CAS, Hefei, 230026, China*
[2]*Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen 518055, China*
[3]*CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, 230026, P. R. China*
[4]*Theoretical Quantum Physics Laboratory, RIKEN Cluster for Pioneering Research, Wako-shi, Saitama 351-0198, Japan*
[5]*Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan*
[6]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[7]*Department of Engineering Science and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 70101, Taiwan*
[8]*Department of Physics, The University of Michigan, Ann Arbor, 48109-1040 Michigan, USA*

(Dated: July 13, 2020)

SUPPLEMENTARY NOTE 1: MEASUREMENT-DEVICE-INDEPENDENT WITNESSES FOR ALL STEERABLE ASSEMBLAGES

For textural completeness, we first recall the standard steering witness. The set of all local-hidden state LHS forms a convex set [1]. Therefore, given a steerable assemblage $\{\sigma^{\mathrm{S}}_{a|x}\}$, there always exists a set of positive semidefinite operators $\{F_{a|x}\succeq 0\}$, called a *steering witness* SW, such that [2–6]

$$\mathrm{Tr}\sum_{a,x}F_{a|x}\,\sigma^{\mathrm{S}}_{a|x}>\alpha:=\max_{\{\sigma^{\mathrm{US}}_{a|x}\}\in\mathsf{LHS}}\mathrm{Tr}\sum_{a,x}F_{a|x}\,\sigma^{\mathrm{US}}_{a|x},\quad(1)$$

while

$$\mathrm{Tr}\sum_{a,x}F_{a|x}\,\sigma^{\mathrm{US}}_{a|x}\leq\alpha\quad\forall\{\sigma^{\mathrm{US}}_{a|x}\}\in\mathsf{LHS}.\quad(2)$$

There two conditions above can be reformulated as follows:

$$\mathrm{Tr}\sum_{a,x}\left(F_{a|x}-\frac{\alpha}{|\mathcal{X}|}\mathbb{1}\right)\sigma^{\mathrm{S}}_{a|x}>0,\quad(3)$$

while

$$\mathrm{Tr}\sum_{a,x}\left(F_{a|x}-\frac{\alpha}{|\mathcal{X}|}\mathbb{1}\right)\sigma^{\mathrm{US}}_{a|x}\leq 0\quad\forall\{\sigma^{\mathrm{US}}_{a|x}\}\in\mathsf{LHS},\quad(4)$$

where $|\mathcal{X}|$ denotes the number of elements in $\mathcal{X}$, i.e., the number of measurement settings.

* shin.liang.chen@phys.ncku.edu.tw
† gyxiang@ustc.edu.cn
‡ yuehnan@mail.ncku.edu.tw

Motivated by the results from Ref. [7], here we show how to systematically construct a collection of steering witnesses (SW) in an measurement-device-independent (MDI) scheme, dubbed MDI-SWs. It is MDI since we certify steerability based only on the statistics $\{p(a,b|x,\tau_y)\}$ and on the fact that $\{\tau_y\}$ forms a tomographically complete set. In what follows, we would like to address the problem within the framework of the resource theory of steering [1], i.e., we will certify steerability of the underlying assemblage $\{\sigma_{a|x}\}$ instead of the quantum state ρ_{AB}.

Within the framework of the resource theory of steering [1], the correlation is obtained from Bob's joint measurement on the assemblage, i.e.,

$$p(a,1|x,\tau_y)=\mathrm{tr}(E_1\sigma_{a|x}\otimes\tau_y).\quad(5)$$

The average payoff of an assemblage can then be defined as

$$\mathcal{W}\Big(\mathbf{P},\beta\Big)=\sum_{a,x,y}\beta^{x,y}_{a,1}\,p(a,1|x,\tau_y).\quad(6)$$

where $\mathbf{P}:=\{p(a,1|x,\tau_y)\}$.

Now we show that for any given steerable assemblage, one can properly choose a set of coefficients $\beta:=\{\beta^{x,y}_{a,1}\}$, such that $\mathcal{W}\Big(\mathbf{P},\beta\Big)$ is a steering witness of the steerable assemblage. That is,

$$\begin{aligned}\text{given}\quad&\{\sigma_{a|x}\}\notin\mathsf{LHS},\quad\exists\beta:=\{\beta^{x,y}_{a,1}\}\\ \text{such that}\quad&\mathcal{W}\Big(\mathbf{P},\beta\Big)>0,\\ &\mathcal{W}\Big(\mathbf{P}(\{\sigma^{\mathrm{US}}_{a|x}\}),\beta\Big)\leq 0\quad\forall\{\sigma^{\mathrm{US}}_{a|x}\}\in\mathsf{LHS}.\end{aligned}\quad(7)$$

Proof. Since the set of Bob's input quantum states $\{\tau_y\}$ is a tomographically complete set, it can be used to span all Hermitian matrices of the same dimension:

$$\begin{aligned} &\text{given} \quad \{F_{a|x}\}\ \&\ \{\tau_y\}, \quad \exists\{\beta_{a,1}^{x,y}\} \\ &\text{such that} \quad F_{a|x} - \frac{\alpha}{|\mathcal{X}|}\mathbb{1} = \sum_y \beta_{a,1}^{x,y}\,\tau_y^{\intercal} \quad \forall a, x, \end{aligned} \tag{8}$$

where $\{F_{a|x}\}$ is a SW of the assemblage and $\{\beta_{a,1}^{x,y}\}$ is a set of some real numbers. The transposition $\intercal$ is for convenience, as will be shown later.

(i) First we prove the second requirement of Eq. (7). Each component in the correlation $\{p(a,1|x,\tau_y)\}$ admitting a LHS model can be expressed as

$$\begin{aligned} p(a,1|x,\tau_y) &= \mathrm{Tr}\left[E_1(\sigma_{a|x}\otimes\tau_y)\right] \\ &= \sum_\lambda p(\lambda)\,p(a|x,\lambda)\,\mathrm{Tr}\left[(\tilde{E}_{1,\lambda}^{B_0}\tau_y)\right], \end{aligned} \tag{9}$$

where

$$\tilde{E}_{1,\lambda}^{B_0} := \mathrm{Tr}_B[E_1(\sigma_\lambda\otimes\mathbb{1})] \tag{10}$$

is an effective POVM element. The payoff of the assemblage is then written as

$$\begin{aligned} &\mathcal{W}\Big(\mathbf{P},\beta\Big) := \sum_{a,x,y}\beta_{a,1}^{x,y}p(a,1|x,\tau_y) \\ &= \sum_{a,x,\lambda}p(\lambda)p(a|x,\lambda)\,\mathrm{Tr}\left[\tilde{E}_{1,\lambda}^{B_0}\left(\sum_y\beta_{a,1}^{x,y}\tau_y\right)\right] \\ &= \mathrm{Tr}\left[\sum_{a,x}\left(F_{a|x}-\frac{\alpha}{|\mathcal{X}|}\mathbb{1}\right)\sum_\lambda p(\lambda)p(a|x,\lambda)(\tilde{E}_{1,\lambda}^{B_0})^{\intercal}\right] \\ &\leq 0, \end{aligned} \tag{11}$$

where the inequality holds due to Eq. (4).

(ii) Now we prove the first requirement of Eq. (7). We choose the joint measurement performed by Bob to be the projection onto the maximally entangled state

$$|\Phi_1^{BB_0}\rangle = 1/\sqrt{d_B}\sum_{i=1}^{d_B}|i\rangle\otimes|i\rangle. \tag{12}$$

Therefore, each component of the correlation can be expressed as

$$\begin{aligned} p(a,1|x,\tau_y) &= \mathrm{Tr}\left[E_1(\sigma_{a|x}^{\mathrm{S}}\otimes\tau_y)\right] \\ &= \mathrm{Tr}\left[(|\Phi_1^{BB_0}\rangle\langle\Phi_1^{BB_0}|)(\sigma_{a|x}^{\mathrm{S}}\otimes\tau_y)\right] \\ &= \mathrm{Tr}\left[\tau_y^{\intercal}\ \sigma_{a|x}^{\mathrm{S}}\right]/d_B. \end{aligned} \tag{13}$$

The average payoff is reformulated as

$$\begin{aligned} \mathcal{W}\Big(\mathbf{P},\beta\Big) &:= \sum_{a,x,y}\beta_{a,1}^{x,y}\,p(a,1|x,\tau_y) \\ &= \sum_{a,x}\mathrm{Tr}\left[\left(\sum_y\beta_{a,1}^{x,y}\tau_y^{\intercal}\right)\sigma_{a|x}^{\mathrm{S}}\right]/d_B \\ &= \sum_{a,x}\mathrm{Tr}\left[\left(F_{a|x}-\frac{\alpha}{|\mathcal{X}|}\mathbb{1}\right)\sigma_{a|x}^{\mathrm{S}}\right]/d_{\mathrm{B}} > 0, \end{aligned} \tag{14}$$

where the inequality holds according to Eq. (3). ■

SUPPLEMENTARY NOTE 2: EQUIVALENCE BETWEEN THE MDI-SM AND THE STEERING FRACTION

Let us now rewrite the definition of the MDI steering measure (MDI-SM), i.e., Eqs. 3 and 4 in the main text

$$\mathcal{S}_1 := \max\{\mathcal{W}_1 - 1, 0\}, \tag{15}$$

where

$$\begin{aligned} \mathcal{W}_1 &:= \sup_{\beta,\mathbf{P}}\frac{\mathcal{W}(\mathbf{P},\beta)}{\mathcal{W}_{\mathrm{LHS}}(\beta)} \\ &= \sup_{\beta,\mathbf{P}}\frac{\sum_{a,x,y}\beta_{a,1}^{x,y}\,p(a,1|x,\tau_y)}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}}\ \sum_{a,x,y}\beta_{a,1}^{x,y}\,\bar{p}(a,1|x,\tau_y)}. \end{aligned} \tag{16}$$

To show that $\mathcal{S}_1$ is a steering monotone and find out the optimal $\mathbf{P}$ in Eq. (16), we rewrite Eq. (16). First, if the trusted quantum inputs $\{\tau_y\}$ form a tomographically complete set, one can choose a set of coefficients β satisfying the spanned relation

$$F_{a|x} = \sum_y\beta_{a,1}^{x,y}\tau_y^{\intercal} \quad \forall a, x, \tag{17}$$

for any positive semidefinite operator $\{F_{a|x}\succeq 0\}$. On the other hand, the optimization over $\mathbf{P}$ is carried out by Bob choosing a proper measurement, described by a POVM $\{E_1, \mathbb{1}-E_1\}$. With these in mind, Eq. (16) can be reformulated as

$$\mathcal{W}_1 = \sup_{\mathbf{F}\succeq 0, E_1\succeq 0}\frac{\sum_{a,x}\mathrm{Tr}\left[E_1(\sigma_{a|x}\otimes F_{a|x}^{\intercal})\right]}{\sup_{\omega\in\mathsf{LHS}}\ \sum_{a,x}\mathrm{Tr}\left[E_1(\omega_{a|x}\otimes F_{a|x}^{\intercal})\right]}, \tag{18}$$

where $\mathbf{F}$ and ω, respectively, denote $\{F_{a|x}\}$ and $\{\omega_{a|x}\}$ for brevity.

Since E_1 is a POVM element, it is diagonalizable and can be taken as a linear combination of rank-1 projectors with coefficients lying between 0 and 1. Since any rank-k projector can be produced by acting a separable operation on the maximally entangled state, E_1 can then be written as

$$\begin{aligned} E_1 &= \sum_{k,i}u(k)\tilde{A}_i^k\otimes\tilde{B}_i^k|\Phi\rangle\langle\Phi|\tilde{A}_i^{k\dagger}\otimes\tilde{B}_i^{k\dagger}, \\ &= \sum_{k,i}A_i^k\otimes B_i^k|\Phi\rangle\langle\Phi|A_i^{k\dagger}\otimes B_i^{k\dagger}, \end{aligned} \tag{19}$$

where $u(k)$ denotes the coefficients between 0 and 1,

$$A_i^k\otimes B_i^k = \sqrt{u(k)}\tilde{A}_i^k\otimes\tilde{B}_i^k \tag{20}$$

is the redefined Kraus operators for each i, and (for brevity) $|\Phi\rangle$ denotes $|\Phi_1^{BB_0}\rangle$. Then, we can proceed to write $\mathcal{W}_1$ as

$$
\begin{aligned}
&\sup_{\mathbf{F}\succeq 0, A_i^k, B_i^k} \frac{\sum_{a,x,k,i} \mathrm{Tr}\left[A_i^k \otimes B_i^k |\Phi\rangle\langle\Phi| A_i^{k\dagger} \otimes B_i^{k\dagger}(\sigma_{a|x} \otimes F_{a|x}^{\intercal})\right]}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x,k,i} \mathrm{Tr}\left[A_i^k \otimes B_i^k |\Phi\rangle\langle\Phi| A_i^{k\dagger} \otimes B_i^{k\dagger}(\omega_{a|x} \otimes F_{a|x}^{\intercal})\right]} \\
&= \sup_{\mathbf{F}\succeq 0, A_i^k, B_i^k} \frac{\sum_{a,x,k,i} \langle\Phi| A_i^{k\dagger} \otimes B_i^{k\dagger}(\sigma_{a|x} \otimes F_{a|x}^{\intercal}) A_i^k \otimes B_i^k |\Phi\rangle}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x,k,i} \langle\Phi| A_i^{k\dagger} \otimes B_i^{k\dagger}(\omega_{a|x} \otimes F_{a|x}^{\intercal}) A_i^k \otimes B_i^k |\Phi\rangle} \\
&= \sup_{\mathbf{F}\succeq 0, A_i^k, B_i^k} \frac{\sum_{a,x,k,i} \langle\Phi| (A_i^{k\dagger}\sigma_{a|x} A_i^k) \otimes (B_i^{k\dagger} F_{a|x}^{\intercal} B_i^k) |\Phi\rangle}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x,k,i} \langle\Phi| (A_i^{k\dagger}\omega_{a|x} A_i^k) \otimes (B_i^{k\dagger} F_{a|x}^{\intercal} B_i^k) |\Phi\rangle} \\
&= \sup_{\mathbf{F}\succeq 0, A_i^k, B_i^k} \frac{\sum_{a,x,k,i} \mathrm{Tr}\left[A_i^{k\dagger}\sigma_{a|x} A_i^k \cdot B_i^{k\intercal} F_{a|x} B_i^{k\dagger^{\intercal}}\right]}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x,k,i} \mathrm{Tr}\left[A_i^{k\dagger}\omega_{a|x} A_i^k \cdot B_i^{k\intercal} F_{a|x} B_i^{k\dagger^{\intercal}}\right]} \\
&= \sup_{\mathbf{F}\succeq 0, A_i^k, B_i^k} \frac{\sum_{a,x} \mathrm{Tr}\left[\sigma_{a|x} \sum_{k,i} A_i^k B_i^{k\intercal} F_{a|x} B_i^{k\dagger^{\intercal}} A_i^{k\dagger}\right]}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x} \mathrm{Tr}\left[\omega_{a|x} \sum_{k,i} A_i^k B_i^{k\intercal} F_{a|x} B_i^{k\dagger^{\intercal}} A_i^{k\dagger}\right]} \\
&\leq \sup_{\mathbf{F}\succeq 0} \frac{\sum_{a,x} \mathrm{Tr}\left[\sigma_{a|x} F_{a|x}\right]}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x} \mathrm{Tr}\left[\omega_{a|x} F_{a|x}\right]}.
\end{aligned}
\tag{21}
$$

The inequality right above is due to the fact that the convex set $\mathbf{F}$ is a superset of the set after performing the completely positive map, i.e.,

$$
\mathbf{F}' := \{\sum_{k,i} A_i^k B_i^{k\intercal} F_{a|x} B_i^{k\dagger^{\intercal}} A_i^{k\dagger}\}_{a,x}. \tag{22}
$$

The last formula in Eq. (21) is exactly the steering fraction shifting with a coefficient equal to one one [8] which we will explicitly define in the next section. Now, if Bob's measurement is chosen as the projection onto the maximally entangled state, then

$$
\begin{aligned}
\mathcal{W}_1 &:= \frac{\mathcal{W}(\mathbf{P},\beta)}{\mathcal{W}_{\mathrm{LHS}}(\beta)}\Big|_{E_1=|\Phi\rangle\langle\Phi|} \\
&= \sup_{\mathbf{F}\succeq 0} \frac{\sum_{a,x} \langle\Phi|\sigma_{a|x} \otimes F_{a|x}^{\intercal}|\Phi\rangle}{\sup_{\omega\in\mathsf{LHS}} \ \sum_{a,x} \langle\Phi|\omega_{a|x} \otimes F_{a|x}^{\intercal}|\Phi\rangle} \\
&= \sup_{\mathbf{F}\succeq 0} \frac{\sum_{a,x} \mathrm{Tr}\left[\sigma_{a|x} F_{a|x}\right]}{\sup_{\omega\in\mathsf{LHS}} \sum_{a,x} \mathrm{Tr}\left[\omega_{a|x} F_{a|x}\right]},
\end{aligned}
\tag{23}
$$

which achieves the upper bound. Therefore, the projection onto the maximally entangled state is the optimal measurement for Bob to achieve the value of $\mathcal{S}_1$, i.e., the supremum of Eq. (16), and $\mathcal{S}_1$ is equivalent to the steering fraction. We show below the proof of the equivalence between the steering fraction and the steering robustness.

SUPPLEMENTARY NOTE 3: EQUIVALENCE BETWEEN THE STEERING FRACTION AND THE STEERING ROBUSTNESS

In this section, we explicitly prove the equivalence between the steering fraction and the steering robustness, although their equivalence is implicitly mentioned in some references (see Ref. [4]).

The steering robustness (SR) is that the noisy assemblage mixing with a given assemblage $\{\sigma_{a|x}\}$ such that the steerability of the total assemblage is destroyed i.e.,

$$
\begin{aligned}
&\mathrm{SR}(\{\sigma_{a|x}\}) = \min \quad \mu \\
&\text{s.t.} \quad \frac{\sigma_{a|x} + \mu\pi_{a|x}}{1+\mu} = \sum_\lambda p(a|x,\lambda)p(\lambda)\sigma_\lambda, \\
&\{\pi_{a|x}\} \quad \text{is an assemblage,}
\end{aligned}
\tag{24}
$$

which can be formulated as the following semidefinite program (SDP):

$$
\mathrm{SR}(\{\sigma_{a|x}\}) + 1 = \min_{\{\tilde{\sigma}_\lambda\}} \sum_\lambda \mathrm{Tr}\left(\tilde{\sigma}_\lambda\right) \tag{25a}
$$

$$
\text{s.t.} \sum_\lambda D(a|x,\lambda)\tilde{\sigma}_\lambda \succeq \sigma_{a|x} \quad \forall\ a,x, \tag{25b}
$$

$$
\tilde{\sigma}_\lambda \succeq 0 \quad \forall\ \lambda, \tag{25c}
$$

where $D(a|x,\lambda) := \delta_{a,\lambda(x)}$ is the deterministic probability distribution [5, 6]. We note that one can further define the steering robustness of a given "quantum state" ρ_{AB}, which is obtained by optimizing over all possible assemblages $\{\sigma_{a|x}\}$ Bob can obtain. It is equivalent with the optimization over all Alice's possible measurements $\{E_{a|x}\}$ due to the relation $\sigma_{a|x} = \mathrm{Tr}_{\mathrm{A}}(E_{a|x} \otimes \mathbb{1}\ \rho_{AB})$ for all a, x. Apparently, $\mathrm{SR}(\{\sigma_{a|x}\})$ is a lower bound on $\mathrm{SR}(\rho_{AB})$.

On the other hand, the steering fraction (SF) is defined as [8]

$$\mathrm{SF}(\{\sigma_{a|x}\}) + 1 = \max_{\mathbf{F}\succeq 0} \quad \frac{\mathrm{Tr}\sum_{a,x} F_{a|x}\,\sigma_{a|x}}{\max_{\omega\in\mathsf{LHS}} \mathrm{Tr}\sum_{a,x} F_{a|x}\,\omega_{a|x}}, \tag{26}$$

which can rewritten as

$$\mathrm{SF}(\{\sigma_{a|x}\}) + 1 = \max_{\tilde{\mathbf{F}}\succeq 0} \quad \mathrm{Tr}\sum_{a,x} \tilde{F}_{a|x}\,\sigma_{a|x}, \tag{27}$$

where

$$\tilde{F}_{a|x} := \frac{F_{a|x}}{\max_{\omega\in\mathsf{LHS}} \mathrm{Tr}\sum_{a,x} F_{a|x}\,\omega_{a|x}} \succeq 0. \tag{28}$$

Therefore, to prove the equivalence between Eqs. (24) and (26), it is equivalent to prove

$$\sum_{a,x} D(a|x,\lambda)\tilde{F}_{a|x} \preceq \mathbb{1} \quad \forall\lambda. \tag{29}$$

Proof. For each λ, the quantity $\mathbb{1} - \sum_{a,x} D(a|x,\lambda)\tilde{F}_{a|x}$ is multiplied by a subnormalized quantum state $\rho_\lambda \geq 0$. We take the trace, and sum over all λ:

$$\begin{aligned} &\mathrm{Tr}\sum_{\lambda}\left(\mathbb{1} - \frac{\sum_{a,x} D(a|x,\lambda)F_{a|x}}{\max_{\omega\in\mathsf{LHS}} \mathrm{Tr}\sum_{a,x} F_{a|x}\omega_{a|x}}\right)\rho_\lambda \\ &= 1 - \frac{\mathrm{Tr}\sum_{a,x} F_{a|x}\sigma^{\mathrm{US}}_{a|x}}{\max_{\omega\in\mathsf{LHS}} \mathrm{Tr}\sum_{a,x} F_{a|x}\omega_{a|x}}, \end{aligned} \tag{30}$$

which is non-negative for *all* $\rho_\lambda \succeq 0$ and λ. Since the only constraint between the free parameters ρ_λ is $\mathrm{Tr}\sum_\lambda \rho_\lambda = 1$, we derive the condition

$$\mathbb{1} - \frac{\sum_{a,x} D(a|x,\lambda)F_{a|x}}{\max_{\omega\in\mathsf{LHS}} \mathrm{Tr}\sum_{a,x} F_{a|x}\omega_{a|x}} \succeq 0 \quad \forall\lambda, \tag{31}$$

which is exactly the same as Eq. (29). With above, we complete the proof. ■

SUPPLEMENTARY NOTE 4: SEMIDEFINITE PROGRAMMING FOR MEASUREMENT-DEVICE-INDEPENDENT STEERING MEASURE

In this section, we show how to arrive the formulation of the SDP described by Eq. 6 from the definition of Eq. 5 in the main text:

$$\mathcal{S}_1^{\mathrm{LB}}(\mathbf{P}) := \max\left\{\mathcal{W}_1^{\mathrm{LB}}(\mathbf{P}) - 1, 0\right\} \tag{32}$$

The first step is to consider the quantity

$$\begin{aligned} \mathcal{W}_1^{\mathrm{LB}}(\mathbf{P}) &:= \sup_\beta \frac{\mathcal{W}(\mathbf{P},\beta)}{\mathcal{W}_{\mathrm{LHS}}(\beta)} \\ &= \frac{\sum_{a,x,y} \beta^{x,y}_{a,1} p(a,1|x,\tau_y)}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}} \sum_{a,x,y} \beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)} \end{aligned} \tag{33}$$

and redefine the set of coefficients $\{\beta^{x,y}_{a,b}\}$ as

$$\tilde{\beta}^{x,y}_{a,1} := \frac{\beta^{x,y}_{a,1}}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}} \sum_{a,x,y} \beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)}. \tag{34}$$

The quantify $\sup_\beta \mathcal{W}_1^{\mathrm{LB}}(\mathbf{P}) - 1$ in Eq. (32) can then be written as

$$\max_{\tilde{\beta}} \sum_{a,x,y} \tilde{\beta}^{x,y}_{a,1} p(a,1|x,\tau_y) - 1, \tag{35}$$

This optimization problem can be solved by the following semidefinite program:

$$\begin{aligned} &\text{given} \quad \{p(a,1|x,\tau_y)\} \text{ and } \{\tau_y\} \\ &\max_{\tilde{\beta}} \sum_{a,x,y} \tilde{\beta}^{x,y}_{a,1} p(a,1|x,\tau_y) - 1 \\ &\text{s.t. } d\mathbb{1} - \sum_{a,x,y} D(a|x,\lambda)\tilde{\beta}^{x,y}_{a,1}\tau_y \succeq 0 \quad \forall\lambda \\ &\quad \sum_y \tilde{\beta}^{x,y}_{a,1}\tau_y \succeq 0 \quad \forall a,x, \end{aligned} \tag{36}$$

Proof. For each λ, the quantity $d\mathbb{1} - \sum_{a,x,y} D(a|x,\lambda)\tilde{\beta}^{x,y}_{a,1}$ is multiplied by a positive semidefinite operator $\mathrm{Tr}_B\left[|\Phi\rangle\langle\Phi|(\sigma_\lambda \otimes \mathbb{1})\right]$, and $\mathrm{Tr}\sum_\lambda \sigma_\lambda = 1$ with $\sigma_\lambda \succeq 0\ \forall\lambda$. After taking the trace and summing over all λ, we obtain

$$\begin{aligned}
&\operatorname{Tr}\sum_{\lambda}\left\{\left(d\mathbb{1}-\sum_{a,x,y}D(a|x,\lambda)\tilde{\beta}^{x,y}_{a,1}\tau_y\right)\cdot\operatorname{Tr}_B\left[|\Phi\rangle\langle\Phi|(\sigma_\lambda\otimes\mathbb{1})\right]\right\}\\
&=\operatorname{Tr}\sum_{\lambda}\left[|\Phi\rangle\langle\Phi|(\sigma_\lambda\otimes\mathbb{1})\cdot d\right]-\frac{\sum_{a,x,y}\beta^{x,y}_{a,1}\sum_\lambda D(a|x,\lambda)\operatorname{Tr}\left[|\Phi\rangle\langle\Phi|(\sigma_\lambda\otimes\tau_y)\right]}{\max_{\bar{\mathsf{P}}\in\mathsf{LHS}}\sum_{a,x,y}\beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)}\\
&=\operatorname{Tr}\sum_{\lambda}\sigma_\lambda-\frac{\sum_{a,x,y}\beta^{x,y}_{a,1}p^{\mathrm{LHS}}(a,1|x,\tau_y)}{\max_{\bar{\mathsf{p}}\in\mathsf{LHS}}\sum_{a,x,y}\beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)}\geq 0.
\end{aligned} \tag{37}$$

Since the inequality holds for *any* positive semidefinite operator $\operatorname{Tr}_B\left[|\Phi\rangle\langle\Phi|(\sigma_\lambda\otimes\mathbb{1})\right]$, we obtain the first constraint in Eq. (36). The second equality in the above equation comes from the fact that

$$\operatorname{Tr}(|\Phi\rangle\langle\Phi|(A\otimes B))=\operatorname{Tr}(A\cdot B^{\mathrm{T}})/d, \tag{38}$$

and that the numerator of the second term in the second line can be treated as a correlation obtained by Bob applying his measurement (corresponding to $|\Phi\rangle\langle\Phi|$) on an unsteerable assemblage, i.e.,

$$\sum_{a,x,y}\beta^{x,y}_{a,1}\operatorname{Tr}\left[|\Phi\rangle\langle\Phi|\left(\sum_\lambda D(a|x,\lambda)\sigma_\lambda\otimes\tau_y\right)\right], \tag{39}$$

leading to an unsteerable correlation $\{p^{\mathrm{LHS}}(a,1|x,\tau_y)\}$. The last inequality holds because $\operatorname{Tr}\sum_\lambda\sigma_\lambda=1$ and

$$\frac{\sum_{a,x,y}\beta^{x,y}_{a,1}p^{\mathrm{LHS}}(a,1|x,\tau_y)}{\max_{\bar{\mathsf{p}}\in\mathsf{LHS}}\sum_{a,x,y}\beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)}\leq 1. \tag{40}$$

The second constraint in Eq. (36) is due to the relation

$$F^{\mathsf{T}}_{a|x}=\sum_y\tilde{\beta}^{x,y}_{a,1}\tau_y\succeq 0 \tag{41}$$

between the coefficient $\{\beta^{x,y}_{a,1}\}$ and the standard steering witness $\{F_{a|x}\}$, which is chosen to be positive semidefinite when constructing the MDI-SM. ■

With the above semidefinite program, Eq. (32) can be computed, and it provides a lower bound on the MDI-SM $\mathcal{S}_1$. Note that if one obtains an optimal set of coefficients $\tilde{\beta}^*$ for a correlation $\mathbf{P}_1$, this set $\tilde{\beta}^*$ is still a valid set, although may not be optimal, for any other correlation $\mathbf{P}_2$. That is, $\tilde{\beta}^*$ satisfies the constraints in Eq. (36) for either $\mathbf{P_1}$ or $\mathbf{P_2}$. This means that when one obtains an optimal set $\{\beta^{*,x,y}_{a,1}\}$ for a given correlation, this set is also a steering witness for some other steerable assemblages. Therefore, we can define MDI steering witnesses with the following general formulation:

$$\sum_{a,x,y}\beta^{*,x,y}_{a,1}p(a,1|x,\tau_y)\leq 1\quad\forall\,\mathbf{P}\in\mathsf{LHS}. \tag{42}$$

SUPPLEMENTARY NOTE 5: THE OPTIMAL TWO-QUBIT JOINT MEASUREMENT FOR BOB

Recall again that the original proposed MDI-SM is written as (see Eqs. (15) and (16) in)

$$\mathcal{S}_1:=\max\{\mathcal{W}_1-1,0\}, \tag{43}$$

with

$$\mathcal{W}_1:=\sup_\beta\frac{\sum_{axy}\beta^{x,y}_{a,1}p^*(a,1|x,\tau_y)}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}}\ \sum_{axy}\beta^{x,y}_{a,1}\bar{p}(a,1|x,\tau_y)}, \tag{44}$$

where the set of probability distributions

$$p^*(a,1|x,\tau_y)=\operatorname{Tr}\left[E^*_1(\sigma_{a|x}\otimes\tau_y)\right]\quad\forall a,x,y \tag{45}$$

is the optimal correlation obtained by performing the optimal projection E^*_1 of Bob's joint measurement on the assemblage $\{\sigma_{a|x}\}$ and the quantum inputs $\{\tau_y\}$. We have proved that the projection onto the maximally entangled state $\frac{1}{\sqrt{d}}\sum_i|i\rangle\otimes|i\rangle$ is the optimal one for Bob. In what follows, we show that for Bob's assemblage $\{\sigma_{a|x}\}$ being a qubit, the four projections of the Bell-state measurement, i.e.,

$$\begin{aligned}
|\Phi_1\rangle&=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle),\ |\Phi_2\rangle=\frac{1}{\sqrt{2}}(|00\rangle-|11\rangle),\\
|\Phi_3\rangle&=\frac{1}{\sqrt{2}}(|01\rangle+|10\rangle),\ |\Phi_4\rangle=\frac{1}{\sqrt{2}}(|01\rangle-|10\rangle).
\end{aligned} \tag{46}$$

are all the optimal ones providing the optimal correlation $\{p^*(a,b|x,\tau_y)\}$ if the set of tomographically complete quantum inputs is composed of the eigenstates of the three Pauli matrices. That is,

$$\{\tau_y\}=\{|0\rangle,|1\rangle,|V\rangle,|H\rangle,|L\rangle,|R\rangle\}, \tag{47}$$

where $\{|0\rangle,|1\rangle\}$, $\{|V\rangle,|H\rangle\}$, $\{|L\rangle,|R\rangle\}$, are, respectively, the eigenstates of the Pauli matrices Z, X, and Y. Indeed, the four Bell states in Eq. (46) can be transformed into each other by applying some Pauli gates on them, i.e.,

$$|\Phi_b\rangle\langle\Phi_b|=(\mathbb{1}\otimes U_b)|\Phi_1\rangle\langle\Phi_1|(\mathbb{1}\otimes U_b^\dagger)\quad\forall b, \tag{48}$$

where $U_b\in\{\mathbb{1},X,Y,Z\}$. Therefore, when Bob's measurement outcomes correspond to the other three projections (i.e., $b \neq 1$), the obtained correlation becomes

$$\begin{aligned} p^{**}(a,b|x,\tau_y) &:= \mathrm{Tr}\left[E_b(\sigma_{a|x}\otimes\tau_y)\right] \\ &= \mathrm{Tr}\left[\mathbb{1}\otimes U_b|\Phi_1\rangle\langle\Phi_1|\mathbb{1}\otimes U_b^\dagger(\sigma_{a|x}\otimes\tau_y)\right] \\ &= \mathrm{Tr}\left[|\Phi_1\rangle\langle\Phi_1|\mathbb{1}\otimes U_b^\dagger(\sigma_{a|x}\otimes\tau_y)\mathbb{1}\otimes U_b\right] \\ &= \mathrm{Tr}\left[|\Phi_1\rangle\langle\Phi_1|(\sigma_{a|x}\otimes U_b^\dagger\tau_y U_b)\right] \\ &= \mathrm{Tr}\left[|\Phi_1\rangle\langle\Phi_1|(\sigma_{a|x}\otimes\tau_{y'})\right] \\ &= p^*(a,1|x,\tau_{y'}) \quad \forall a,b,x,y. \end{aligned} \tag{49}$$

It is easy to see that the elements of the set $\{\tau_{y'}\}$ remain the same as that of the set $\{\tau_y\}$. Therefore, the components of the correlation $\{p^{**}(a,b\neq 1|x,\tau_y)\}$ are just a permutation of the components of $\{p^*(a,1|x,\tau_y)\}$, which means that these correlations can all achieve the value of $\mathcal{S}_1$, i.e.,

$$\begin{aligned} &\mathcal{W}_b := \\ &\sup_{\mathbf{P}^{**},\{\beta_{a,b}^{x,y}\}_{a,x,y}} \frac{\sum_{axy}\beta_{a,b}^{x,y}p^{**}(a,b|x,\tau_y)}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}}\ \sum_{axy'}\beta_{a,b}^{x,y}\bar{p}(a,b|x,\tau_y)} \\ &= \sup_{\mathbf{P}^{*},\{\beta_{a,1}^{x,y'}\}_{a,x,y'}} \frac{\sum_{axy'}\beta_{a,1}^{x,y'}p^{*}(a,1|x,\tau_{y'})}{\sup_{\bar{\mathbf{P}}\in\mathsf{LHS}}\ \sum_{axy'}\beta_{a,1}^{x,y'}\bar{p}(a,1|x,\tau_{y'})} \\ &=: \mathcal{W}_1 \quad \forall b=2,3,4. \end{aligned} \tag{50}$$

SUPPLEMENTARY NOTE 6: BOUND RELATIONS BETWEEN THE STEERING ROBUSTNESS, THE ENTANGLEMENT ROBUSTNESS, AND THE INCOMPATIBILITY ROBUSTNESS

For readers' reference, in this section we briefly review the detailed formulation of the quantities mentioned in the main text, including the steering robustness [6], the entanglement robustness [9, 10], and the incompatibility robustness [11]. We also briefly review their bound relations proposed in Ref. [6, 12–14].

The entanglement robustness [9, 10] of a given quantum state $\mathrm{ER}(\rho_{AB})$ is the minimum amount the noisy state one has to mix with, such that the mixture becomes a separable state. That is,

$$\begin{aligned} \mathrm{ER}(\rho_{AB}) = \min \quad & t \\ \text{s.t.} \quad & \frac{\rho_{AB}+t\omega_{AB}}{1+t} \quad \text{is separable}, \\ & \omega_{AB} \quad \text{is a quantum state.} \end{aligned} \tag{51}$$

In general, it is hard to characterize the set of separable states. However, one can still relax this set to the positive-partial-transposition states. Through this way, a lower bound on the above solution can be obtained by solving the following semidefinite program [15]:

$$\begin{aligned} \min_{\tilde{\omega}_{AB}} \quad & \mathrm{Tr}(\tilde{\omega}_{AB})-1 \\ \text{s.t.} \quad & \tilde{\omega}_{AB}^{\mathrm{T_A}} \succeq 0, \quad \tilde{\omega}_{AB}\succeq\rho_{AB}, \end{aligned} \tag{52}$$

where $\succeq$ denotes a matrix being positive semidefinite and $\mathrm{T_A}$ the partial transposition of the operator with respect to the Hilbert space of A. In particular, if the given state ρ_{AB} is a qubit-qubit or a qubit-qutrit state, which is also the case we consider in this work, it has been shown that this lower bound is tight [16].

In quantum theory, not all observables can be measured simultaneously. Such a property can be formulated as that there is no single POVM describing a non-jointly measurable measurement [17], i.e.,

$$E_{a|x} \neq \sum_\lambda p(a|x,\lambda)G_\lambda, \tag{53}$$

for some a,x, where $\{E_{a|x}\}_a$ is the POVM representing the measurement input x and a is a measurement outcome. Note that $G_\lambda \succeq 0\ \forall\lambda$ and $\sum_\lambda G_\lambda = \mathbb{1}$. Here, $p(a|x,\lambda)$ is a probability distribution, and can be chosen, without loss of generality, to be $p(a|x,\lambda)=D(a|x,\lambda):=\delta_{a,\lambda(x)}$. A way to quantify the incompatibility of given measurements is to minimize the ratio of noisy measurements one has to mix with, such that the mixture becomes jointly measurable. This *incompatibility robustness* is formulated as [11],

$$\begin{aligned} \mathrm{IR}(\{E_{a|x}\}) = \min \quad & r \\ \text{s.t.} \quad & \frac{E_{a|x}+rN_{a|x}}{1+r} = \sum_\lambda p(a|x,\lambda)G_\lambda \quad \forall a,x, \\ & \{N_{a|x}\}_a \quad \text{is a POVM} \quad \forall x, \end{aligned} \tag{54}$$

which can be solved by the following semidefinite program:

$$\begin{aligned} \mathrm{IR}(\{E_{a|x}\}) = \min_{\{\tilde{G}_\lambda\}} \quad & \frac{1}{d}\sum_\lambda \mathrm{Tr}[\tilde{G}_\lambda]-1 \\ \text{s.t.} \quad & \sum_\lambda D(a|x,\lambda)\tilde{G}_\lambda \succeq E_{a|x} \quad \forall\ a,x, \\ & \tilde{G}_\lambda \succeq 0 \quad \forall\ \lambda, \\ & \sum_\lambda \tilde{G}_\lambda = \mathbb{1}\frac{1}{d}\sum_\lambda \mathrm{Tr}[\tilde{G}_\lambda], \end{aligned} \tag{55}$$

where d is the dimension of $E_{a|x}$.

Finally, let us review the bound relations used in our work. In Ref. [6], it has been shown that the steering robustness of the underlying quantum state is a lower bound on the entanglement robustness, i.e.,

$$\mathrm{ER}(\rho_{AB}) \geq \mathrm{SR}(\rho_{AB}) \geq \mathrm{SR}(\{\sigma_{a|x}\}). \tag{56}$$

On the other hand, it has been shown that the steering robustness of the assemblage is a lower bound on the incompatibility robustness of the involved measurements [12–14], i.e.,

$$\mathrm{IR}(\{E_{a|x}\}) \geq \mathrm{SR}(\{\sigma_{a|x}\}). \quad (57)$$

With the above bound relations, we use the quantity $\mathcal{S}_1^{\mathrm{LB}}(\mathbf{P})$ in the main text, with the fact $\mathcal{S}_1^{\mathrm{LB}}(\mathbf{P}) \leq \mathcal{S}_1 = \mathrm{SR}(\{\sigma_{a|x}\})$, to estimate the degree of entanglement of the underlying state and the incompatibility of the involved measurements.

SUPPLEMENTARY NOTE 7: EXPERIMENTAL DETAILS

In this section, we give a detailed description of our experimental setup, including the preparation of the system state, the auxiliary input states, and the implementation of the Bell-state measurement. In addition, a simple circuit diagram for our experimental MDI-SM scenario is presented in Fig. 1.

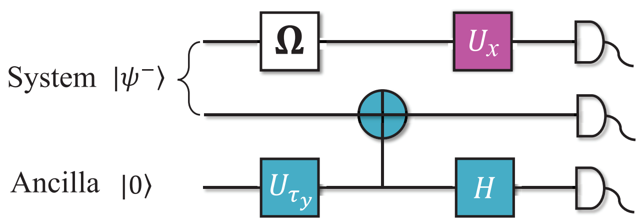

FIG. 1. **Quantum circuit diagram for our experimental MDI-SM scenario.** The symbols marked with pink and blue color represent Alice's and Bob's operations respectively. The Werner state ρ_{AB} is prepared by the operation Ω which generates the white noise on the system $|\psi^-\rangle = 1/\sqrt{2}(|HV\rangle - |VH\rangle)$. On Alice's side, the qubit-measurement with classical input $\{x\}$ is realized by the operator $\{U_x\}$ and the measurement on the computational basis. On Bob's side, the ancilla state $\{\tau_y\}$ is prepared by $\{U_{\tau_y}\}$, then a Bell-state measurement is implemented by the CNOT gate and the joint measurement of the control qubit and the target qubit. More details can be referred to the setup diagram of Fig. 2 in the main text.

A. Preparation of the system state

A 100 mW continuous laser beam passes through a HWP@404 nm to make the horizontally polarized (H) component and vertically polarized (V) component balanced. The beam is focused on two type-I phase-matched β-barium borate crystals (BBO) (0.5 mm $\times$6 mm$\times$6 mm), whose optical axes are normal to each other, to produce a pair of entangled photons with 808 nm. The photons are sent to Alice and Bob through the single-mode fibers. The set of components marked as Ω is where the photon is reflected by or transmitted through a 50:50 BS. In our experiment, the photons are filtered by 3 nm bandwidth interference filters (IF), creating a coherence length of about 269λ. When the photon is reflected and takes the long path, the wave packet of the photon will be split into four incoherent parts via the three 386λ quartz plates (QP) and a 22.5° rotated HWP, and the system state will dephase to a completely mixed state [18, 19]. At last, the reflected part, combined with the transmission part, incoherently prepare the Werner state, and the visibility v can be tuned by the attenuators. Note that, here, the path difference between the two BSs is 0.15 m, which is much longer than the coherence length. Therefore, the prepared Werner state is an incoherent mixture, instead of a coherent superposition.

In our experiment, the detailed forms of the prepared states are obtained by standard tomography, and the local measurements are realized by properly adjusting the configuration of the experimental setup shown in Fig. 2 of the main text. To be specific, Bob adjusts H2 to the angle of 0° to make the photon pass through Path-1 entirely, and then uses the QWP, HWP combined with the following polarizing beam splitter (PBS), to complete the standard polarization analysis; while Q1, H1 and the PBS are used on Alice's side. In our experiment, we prepare the Werner states with the visibilities $v = 0.9934(11), 0.8575(56), 0.7250(72), 0.5870(77)$ and $0.4689(72)$, and the corresponding fidelities are $0.996(1), 0.980(7), 0.958(6), 0.959(12)$ and $0.977(2)$ respectively. By projecting onto $|HH\rangle\langle HH|$ and $|VV\rangle\langle VV|$, the visibilities of the Werner states can be obtained through

$$v = 1 - 2(\mathrm{Tr}[\rho_{AB}(|HH\rangle\langle HH| + |VV\rangle\langle VV|)]. \quad (58)$$

B. Preparation of the auxiliary input state

On Bob's side, the quantum input τ_y is encoded on the path degree of freedom of Bob's particle. The blue box labelled τ_y in Fig. 2 (the detailed structure is shown below) behaves like a non-polarization beam splitter. Here, the main component is the designed beam displacer (BD), which can make the V light pass through it directly and make the H light pass through it with a 4 mm displacer at behaves 808 nm parallel with V. First, the photons are separated into two beams with the first BD; then a cut HWP is used to unify the polarization of the photons. The second BD splits the $H(V)$ component of the input light once again into $0H$ and $1H$ ($0V$ and $1V$) with the ratio $\cos^2\theta/\sin^2\theta$, where θ is the rotation angle of the half-wave plate H2. At last, the third BD combines the $0H$ and $0V$ components into the output light 0, and $1H$ and $1V$ components into the output light 1. By slightly tilting the third BD, we can compensate the phase of the two-photon state $|HV\rangle - |VH\rangle$. At the same time, the phase between $1H$ and $1V$ is controlled by tilting the HWP.

C. Bell state measurement

Now let us illustrate the way to implement Bob's optimal joint measurement, i.e., Bell-state measurement. The photons in path 1 undergo a bit-flip operation while the photons in path 0 undergo an identity operation. The two operations together are equivalent to a controlled-NOT (CNOT) gate. The following steps are joint measurements of the control qubit and the target qubit of the CNOT gate. The ports D1, D2, D3 and D4 correspond to the measurements in the basis $H\otimes(0+1), H\otimes(0-1), V\otimes(0+1)$, and $V\otimes(0-1)$, respectively, implementing a complete Bell-state measurement.

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

[2] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, "Experimental criteria for steering and the Einstein-Podolsky-Rosen paradox," Phys. Rev. A **80**, 032112 (2009).

[3] M. F. Pusey, "Negativity and steering: A stronger Peres conjecture," Phys. Rev. A **88**, 032313 (2013).

[4] D. Cavalcanti and P. Skrzypczyk, "Quantum steering: a review with focus on semidefinite programming," Rep. Prog. Phys. **80**, 024001 (2017).

[5] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, "Quantifying Einstein-Podolsky-Rosen steering," Phys. Rev. Lett. **112**, 180404 (2014).

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

[7] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, "Measurement-device-independent entanglement witnesses for all entangled quantum states," Phys. Rev. Lett. **110**, 060405 (2013).

[8] C.-Y. Hsieh, Y.-C. Liang, and R.-K. Lee, "Quantum steerability: Characterization, quantification, superactivation, and unbounded amplification," Phys. Rev. A **94**, 062120 (2016).

[9] G. Vidal and R. Tarrach, "Robustness of entanglement," Phys. Rev. A **59**, 141–155 (1999).

[10] M. Steiner, "Generalized robustness of entanglement," Phys. Rev. A **67**, 054305 (2003).

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

[12] S.-L. Chen, C. Budroni, Y.-C. Liang, and Y.-N. Chen, "Natural framework for device-independent quantification of quantum steerability, measurement incompatibility, and self-testing," Phys. Rev. Lett. **116**, 240401 (2016).

[13] D. Cavalcanti and P. Skrzypczyk, "Quantitative relations between measurement incompatibility, quantum steering, and nonlocality," Phys. Rev. A **93**, 052112 (2016).

[14] S.-L. Chen, C. Budroni, Y.-C. Liang, and Y.-N. Chen, "Exploring the framework of assemblage moment matrices and its applications in device-independent characterizations," Phys. Rev. A **98**, 042127 (2018).

[15] L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM Review **38**, 49–95 (1996).

[16] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," Phys. Lett. A **223**, 1 (1996).

[17] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, "Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity," Physics Reports **506**, 1 – 39 (2011).

[18] E.-M. Laine, H.-P. Breuer, J. Piilo, C.-F. Li, and G.-C. Guo, "Nonlocal memory effects in the dynamics of open quantum systems," Phys. Rev. Lett **108**, 210402 (2012).

[19] B. Qi, Z. Hou, Y. Wang, D. Dong, H.-S. Zhong, L. Li, G.-Y. Xiang, H. M. Wiseman, C.-F. Li, and G.-C. Guo, "Adaptive quantum state tomography via linear regression estimation: Theory and two-qubit experiment," npj Quantum Information **3**, 19 (2017).