We theoretically predict, and experimentally verify with entangled photons, that outcome communication is not enough for hidden-state models to reproduce quantum steering. Hidden-state models with outcome communication correspond, in turn, to the well-known instrumental processes of causal inference but in the 1-sided device-independent (1S DI) scenario of one black-box measurement device and one well-characterised quantum apparatus. We introduce 1S-DI instrumental inequalities to test against these models, with the appealing feature of detecting entanglement even when communication of the black box’s measurement outcome is allowed. We find that, remarkably, these inequalities can also be violated solely with steering, i.e. without outcome communication. In fact, an efficiently-computable formal quantifier – the robustness of non-instrumentality – naturally arises; and we prove that steering alone is enough to maximize it. Our findings imply that quantum theory admits a stronger form of steering than known until now, with fundamental as well as practical potential implications.
rally lead to a resource-theoretic measure efficiently computable via semi-definite programming: the robustness of non-instrumentality. Furthermore, we show that 1S-DI instrumental inequalities can be violated with little entanglement and purity at the common-cause node and, remarkably, without outcome communication, so that the violations are due solely to the states’ steering. We present an experimental demonstration in an entangled-photon platform. Finally, we prove an even stronger incompatibility between steering and 1SQI processes. Namely, that quantum steering alone is enough to attain any value of the robustness. Our findings imply that steering is a stronger quantum phenomenon than previously thought, beyond classical hidden-variable models even when equipped with output communication.

1-sided quantum and quantum-common-common instrumen-
tal processes. We start by introducing 1SQI causal networks, shown in Fig. 1 a). Nodes $X$, $A$, and $B$ are observable, while $\Lambda$ is hidden. Node $B$ is a quantum system with Hilbert space $H_B$, whereas all four other nodes encode classical random variables. Node $X$ takes $|X|$ possible values $x \in Z_X$, with the short-hand notation $Z_X \coloneqq \{0, \ldots, |X| - 1\}$ introduced, $A$ takes $|A|$ values $a \in Z_A$, and $\Lambda$ can w.l.o.g. be assumed to take $|\Lambda| = |A|^{|X|}$ possible values $\lambda \in Z_\Lambda$ (see App. A). A 1SQI causal model assigns a probability distribution $P_A \coloneqq \{P_a\}_\lambda$ to $A$, a conditional distribution $P_{A|X,\Lambda} \coloneqq \{P_{a|x,\lambda}\}_{a,x,\lambda}$ to $A$, and a quantum state $\varrho_{a,\lambda} \in \mathcal{B}(H_B)$ to $B$, i.e. such that $\varrho_{a,\lambda} \geq 0$ and $\text{Tr}[\varrho_{a,\lambda}] = 1$ for all $a$ and $\lambda$. We are interested in the statistics of $A$ and $B$ given $X$. Hence, the local statistics of $X$ is not explicitly considered here. We refer to the users at nodes $A$ and $B$ as Alice and Bob, respectively.

Since $A$ is classical and $B$ quantum, they are most-conveniently described jointly by subnormalised conditional states $\sigma_{a|x}$, which encapsulate both the probability $P_{a|x} \coloneqq \text{Tr}[\sigma_{a|x}]$ of $a$ given $x$ for Alice and the conditional state $\varrho_{a,x} \coloneqq P_{a|x}/P_{a|x}$ given $a$ and $x$ for Bob. 1SQI models produce ensembles $\Sigma_{A|X}^{(\text{inst})} \coloneqq \{\sigma_{a|x} \in \mathcal{B}(H_B)\}_{a,x}$, with

$$\sigma_{a|x}^{(\text{inst})} = \sum_{\lambda} P_\Lambda P_{a|x,\lambda} \varrho_{a,\lambda}.$$

We refer to $\Sigma_{A|X}^{(\text{inst})}$ as a 1SQI assemblage, and denote the set of all such ensembles by 1SQI. The term “assemblage” is native of the steering literature [52]. Its use here is not coincidental: there is a connection between 1SQI and steering. To see this, let us next introduce local hidden-state (LHS) models. These correspond to restricted 1SQI models without the causal influence from $A$ to $B$. An LHS as-

![Figure 1](image-url)

**Figure 1. a) and b) Hybrid (classical-quantum) instrumental processes. Causal structures are specified by directed acyclic graphs (DAGs). Each node encodes either a classical random variable or a quantum system, and each directed edge a causal influence. The term acyclic refers to the physical requirement of no causal loops. a) 1-sided quantum instrumental (1SQI) causal networks generalise the classical instrumental causal structure to the case where node $B$ is a quantum system. Removing the red dashed edge, in turn, leads to local hidden-state (LHS) models. b) Quantising also $\Lambda$ gives what we here refer to as quantum instrumental (QI) processes: $\Lambda$ is now a bipartite system in a possibly entangled quantum state, with one subsystem causally influencing $A$ and the other one $B$. Removing the red dashed edge here leads to quantum (Q) models. c) Pictorial representation of the inner geometry of the set QI. 1SQI is a strict subset of QI. The lower dimensional manifold NS of non-signalling assemblages coincides (for the bipartite case) with Q, in turn containing LHS. LHS and Q are the sets studied in steering theory. Surprisingly, Q is not contained in 1SQI. What is more, for any assemblage in QI and outside of 1SQI, there is an assemblage in Q as far away from 1SQI as the former.**

A 1SQI assemblage $\Sigma_{A|X}^{(\text{lhs})} \coloneqq \{\sigma_{a|x}^{(\text{lhs})} \in \mathcal{B}(H_B)\}_{a,x}$ has components $\sigma_{a|x}^{(\text{lhs})} = \sum_\lambda P_\Lambda P_{a|x,\lambda} \varrho_{a,\lambda}$, i.e. as Eq. (1) but with $\varrho_{a,\lambda}$ independent of $a$. We denote the set of all LHS assemblages by LHS, and call any assemblage $\Sigma_{A|X}$ steerable if $\Sigma_{A|X} \not\in$ LHS. Clearly, LHS $\subseteq$ 1SQI.

In turn, allowing not only for a quantum $B$ but also for a quantum $\Lambda$ defines the set QI of quantum instrumental (QI) assemblages $\Sigma_{A|X}^{(\text{qinst})} \coloneqq \{\sigma_{a|x}^{(\text{qinst})} \in \mathcal{B}(H_B)\}_{a,x}$. More precisely, we allocate to $\Lambda$ a composite Hilbert space $H_\Lambda = H_{\Lambda_A} \otimes H_{\Lambda_B}$, such that subsystems $\Lambda_A$ and $\Lambda_B$ causally influence nodes $A$ and $B$, respectively [see Fig. 1 b)]. Hence, $\Lambda$ is now a quantum common cause [15, 16] for $A$ and $B$ [53]. Accordingly, for $\Lambda$ in a state $\varrho_\Lambda \in \mathcal{B}(H_\Lambda)$, the resulting QI conditional states are

$$\sigma_{a|x}^{(\text{qinst})} = \mathcal{E}_{B|A,\Lambda_B} \left(\text{Tr}_{\Lambda_A} \left[M_x^{(\alpha)} \otimes \mathbb{I}_{\Lambda_B} \varrho_\Lambda \right]\right).$$

Here, $\mathcal{E}_{B|A,\Lambda_B} : \mathcal{B}(H_{\Lambda_B}) \to \mathcal{B}(H_B)$ is an $\alpha$-dependent completely-positive trace-preserving map and $M_x^{(\alpha)}$ is the $\alpha$-th element of an $x$-dependent measurement $M_x :=$
\{ M^{(a)}_x \in B(\mathcal{H}_A) : M^{(a)}_x \geq 0, \sum_a M^{(a)}_x = \mathbb{1}_{\Lambda_A} \}\), with \(\mathbb{1}_{\Lambda_A}\) the identity on \(\mathcal{H}_A\). Clearly, 1SQI \(\subseteq\) QI, as Eq. (2) reduces to Eq. (1) for the specific case of \(q_A\) separable.

On the other hand, for the particular case of \(E_{B[a,A,B]}\) being the identity map for all \(a\) (no causal influence from \(A\) to \(B\)), QI reduces to the set \(Q\) of \textit{quantum assemblages}\(\Sigma^{(q)}_{A|X} := \{ \sigma^{(q)}_{a|x} \in B(\mathcal{H}_B) \}_{a,x}\) of components \(\sigma^{(q)}_{a|x} := \text{Tr}_{\Lambda_A} [M^{(a)}_x \otimes \mathbb{1}_{\Lambda_B} \otimes \rho_A]\). Hence, \(Q\) is to Q what LHS is to 1SQI. Clearly, \(Q \subseteq QI\). In addition, from steering theory, we know that LHS \(\subset Q\). On the contrary, it holds that 1SQI \(\not\subseteq Q\), as Q assemblages are non-signalling while 1SQI ones not. An assemblage \(\Sigma_{A|X}\) is said to be non-signalling if

\[
\rho_B := \sum_a \sigma^{(q)}_{a|x}
\]

(the reduced state of Bob) is independent of \(x\). Also due to non-signalling, it follows that, actually, LHS \(\subset 1\)SQI and \(Q \subseteq QI\). We call the set of non-signalling assemblages \(NS\). For the bipartite case under consideration, it is known that \(Q = NS\) [54]. In contrast, for instrumental causal models, Bob’s state can depend on \(x\) even after summing \(a\) out.

1S-DI \textit{instrumental inequalities, witnesses, and robustness.} Since 1SQI is convex, any \(\Sigma_{A|X} \not\in 1\)SQI is separated from 1SQI by a hyperplane, represented by an assemblage-like object \(W_{A|X} := \{ W_{a|x} \in B(\mathcal{H}_B) \}_{a,x}\), with \(W_{a|x}\) Hermitian, of fixed \textit{scale} \(s := \sum_{a,x} \text{Tr} [W_{a|x}]\), such that

\[
\langle W_{A|X}, \Sigma^{(\text{inst})}_{A|X} \rangle := \sum_{a,x} \text{Tr} [W_{a|x} \sigma^{(\text{inst})}_{a|x}] \leq \beta,
\]

for all \(\Sigma^{(\text{inst})}_{A|X} \in 1\)SQI, and \(\langle W_{A|X}, \Sigma_{A|X} \rangle > \beta \in \mathbb{R}\).

We refer to Eq. (4) as a 1S quantum instrumental inequality with \textit{1SQI bound} \(\beta\) (which depends solely on \(W_{A|X}\)). Thus, \(W_{A|X}\) plays a role analogous to the normal vector of a plane in Euclidean space. We refer to \(W_{A|X}\) as a \textit{non-instrumentality witness}. The separation is then quantified by the violation \(\langle W_{A|X}, \Sigma_{A|X} \rangle - \beta\). Finally, we say that \(W_{A|X}\) is an \textit{optimal non-instrumentality witness} for \(\Sigma_{A|X} \not\in 1\)SQI if \(\langle W_{A|X}, \Sigma_{A|X} \rangle - \beta \geq \langle W'_{A|X}, \Sigma_{A|X} \rangle - \beta' \) for all non-instrumentality witnesses \(W'_{A|X}\) with 1SQI bound \(\beta'\) and scale \(\sum_{a,x} \text{Tr} [W'_{a|x}] = s\). Remarkably, as shown in App. A, the optimal witness is obtained efficiently via semi-definite programming [52].

Witnesses are, in turn, connected with robustness measures [56, 57]. Here, we consider the \textit{robustness of non-instrumentality} \(R_{ni}\), defined, for any \(\Sigma_{A|X} \in QI\), as

\[
R_{ni}(\Sigma_{A|X}) := \min \left\{ t \in \mathbb{R}_{\geq 0} : \frac{\Sigma_{A|X} + t \Pi_{A|X}}{1 + t} \in 1\text{SQI} \right\}
\]

It measures the minimal mixing \(t/(1 + t)\) with any \(\Pi_{A|X} \in 1\)SQI that \(\Sigma_{A|X}\) tolerates before the mixture enters 1SQI. Interestingly, \(R_{ni}\) is a measure of non-instrumentality in the formal, resource-theoretic sense [58–60], as we explicitly show in App. D. Moreover, we note that other choices of “noise” types are possible, giving rise to different variants of \(R_{ni}\). However, the choice \(\Pi_{A|X} \in 1\)SQI is particularly convenient as it yields the robustness efficiently computable through a semi-definite programming optimisation. In fact, such optimisation shows that \(R_{ni}(\Sigma_{A|X}) = \langle V_{A|X}, \Sigma_{A|X} \rangle - 1\), where \(V_{A|X}\) is the optimal for \(\Sigma_{A|X}\) over a simple subclass of non-instrumentality witnesses (see App. E for details). We call \(V_{A|X}\) the \textit{optimal robustness witness} for \(\Sigma_{A|X}\).

Fig. 2 shows \(R_{ni}\), together with the usual steering robustness [52], for Q assemblages obtained from local measure-
ments (by Alice) on states with different degrees of entanglement and purity. Some assemblages in the figure have positive steering robustness and $R_{ni} = 0$. This confirms that $\text{LHS} \subset 1\text{SQI}$: Alice’s outcome signalling indeed provides the models with more descriptive power. However, the figure also shows assemblages with $R_{ni} > 0$. This implies the following theorem, proven also analytically in App. B.

**Theorem 1** ($Q \not\subseteq 1\text{SQI}$). Outcome signalling is not enough for LHS models to reproduce quantum steering.

It is instructive to compare with the fully DI case, where Bob’s instrument is also a black box. There, usual Bell-like correlations (where Alice and Bob have independent inputs) obtained without outcome communication are known to be stronger than local hidden-variable models augmented with outcome signalling from Alice [19, 48, 61]. However, for DI instrumental processes, in contrast, if Bob does not actively exploit Alice’s output his measurement setting is fixed [62]. Hence, no matter how entangled $g_A$, or what measurements Alice makes, the resulting correlations will be automatically compatible with local hidden-variable models with no input for Bob, a subclass of classical instrumental models. In other words, if Bob applies a measurement that does not depend on $a$, the correlations trivially fulfill any DI instrumental inequality, including the recent one of Ref. [36]. Thus, incompatibility with the instrumental DAG without output signalling is a distinctive feature of the 1-sided DI case.

Finally, $R_{ni}$ could in principle attain higher values over QI than over Q. After all, the former allows for signalling while the latter does not. Surprisingly, this is false. The following theorem, proven in App. F, holds instead.

**Theorem 2** ($R_{ni}(Q) = R_{ni}(\text{QI})$). For every quantum instrumental assemblage, there exists a $Q$ assemblage with the same non-instrumentality robustness.

From a practical viewpoint, the theorem provides a significant computational shortcut in the task of, given a fixed value of $R_{ni}$, finding an assemblage with that robustness. Because the theorem allows one to restrict the search to quantum assemblages, instead of searching over all QI ones [63]. From a fundamental perspective, in turn, it has implications in the inner geometry of QI. Namely, it tells us that, for any point $\Sigma_{A|X}^{(\text{qinst})} \in \text{QI} \setminus 1\text{SQI}$, there is always a point $\Sigma_{A|X}^{(q)} \in Q$ as far away from 1SQI as $\Sigma_{A|X}^{(\text{qinst})}$. This does not contradict the fact that $\text{Q} \subset \text{QI}$, because $Q$ is a lower-dimensional manifold (the NS one) of QI. Theorem 2 thus suggests that $\Sigma_{A|X}^{(q)}$ is a kind of projection of $\Sigma_{A|X}^{(\text{qinst})}$ onto NS [see Fig. 1 c].

**Experimental demonstration.** We verified the gap between quantum steering and non-instrumentality using entangled-photon pairs produced by spontaneous parametric down conversions in a BBO crystal [64–66] (see Fig. 3). Alice’s device is taken as the black box, so her measurement settings (wave-plate angles) and outcomes encode the bits $x$ and $a$, respectively. Each measurement by her accounts for the preparation of the assemblage element $\sigma_{a|x}^{(q)}$. In contrast, Bob’s measurements allow for state tomography of each $\sigma_{a|x}^{(q)}$. We produce assemblages of different purities and steering by applying local dephasing of different strengths $V$ on Bob’s qubit and having Alice measure the observables $-\sigma_x + \sigma_z) / \sqrt{2}$, $\sigma_x$, or $\sigma_z$, where $\sigma_x$ and $\sigma_z$ are respectively the first and third Pauli matrices. These three observables correspond to the input choices $x = 0$, 1, and 2, respectively. We tomographically reconstruct each produced assemblage and evaluate $R_{ni}$, whose values are those displayed in Fig. 2.

**Final discussion.** 1SQI processes are a generalisation of both classical instrumental processes (native of causal inference [1, 2]) to the case where the final node is quantum and of local hidden-state models (native of steering theory [39, 52]).
thus unifying two previously disconnected topics. We introduced inequalities and witnesses to test against such 1SQI causal models. These can, in particular, detect entanglement with a single trusted device even in the presence of 1-way outcome signalling. Strikingly, they can also be violated with entanglement alone, i.e. without using outcome communication. Hence, outcome communication is not enough to explain quantum steering. Interestingly, this is a distinctive feature of the 1S DI setting: For DI instrumental processes, if Bob’s setting is independent of Alice’s outcome, any state produces local hidden-variable correlations, automatically compatible with classical instrumental models.

The proposed quantifier (robustness) of non-instrumentality is efficiently computable via semi-definite programming. Moreover, we prove in the appendix that it is a formal, resource-theoretic monotone. With it, we showed an even stronger incompatibility between steering and 1SQI processes: that any value of the robustness can be attained with steering alone. We experimentally verified our predictions in an entangled-photon platform. The experiment is simple, but proves that quantum states can be steerable in a stronger way than previously reported.

Finally, the fact that quantum mechanics allows one to falsify – with quantum control only at a single lab – classical explanations even when these exploit output signalling is not only relevant from the perspectives of quantum foundations and causal inference but also promising from an applied one. More precisely, quantum-nonlocality applications are experimentally less demanding in the 1S DI regime than the fully DI one, as already mentioned. To this, our findings now add the possibility of steering-based protocols with the additional experimentally-appealing feature of no need for space-like separation. We note that, even if $B$ is in the future light-cone of $A$ (and, therefore, also of $X$), direct causal influences from $X$ to $B$ can be ruled out (and so an underlying instrumental causal structure guaranteed) with interventions on $A$, e.g. There, cryptographic or randomness-generation protocols based on 1S quantum instrumental inequality violations are conceivable. It would thus be interesting to explore steering beyond outcome signalling as a potential resource for information processing in comparison with conventional steering-based schemes [38] requiring space-like separation. Our results open new venues for research in that direction.

**ACKNOWLEDGMENTS**

We are indebted to Stephen P. Walborn for the experimental infrastructure and thank Ana B. Sainz for noticing an error in a previous version. LA and RC acknowledge support from the Brazilian ministries MCTIC and MEC. LA, MMT, and RVN acknowledge support from the Brazilian agencies CNPq, FAPERJ, and INCT-IQ; and LA also from FAPESP.

[1] P. Spirtes, N. Glymour, and R. Scheines, *Causation, Prediction, and Search, 2nd ed.* (The MIT Press, 2001).
[2] J. Pearl, *Causality* (Cambridge University Press, 2009).
[3] P. G. Wright et al., *Tariff on animal and vegetable oils* (The Macmillan Co., 1928).
[4] A. Balke and J. Pearl, “Bounds on Treatment Effects From Studies With Imperfect Compliance,” *Journal of the American statistical Association* 92, 1171 (1997).
[5] S. Greenland, “An introduction to instrumental variables for epidemiologists,” *International Journal of Epidemiology* 29, 722 (2000).
[6] J. D. Angrist, G. W. Imbens, and D. B. Rubin, “Identification of causal effects using instrumental variables,” *Journal of the American statistical Association* 91, 444–455 (1996).
[7] S. Creel and M. Creel, “Density dependence and climate effects in rocky mountain elk: an application of regression with instrumental variables for population time series with sampling error,” *Journal of Animal Ecology* 78, 1291–1297 (2009).
[8] J. Pearl, “On the testability of causal models with latent and instrumental variables,” in *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence* (Morgan Kaufmann Publishers Inc., 1995) pp. 435–443.
[9] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Theoretical framework for quantum networks,” *Phys. Rev. A* 80, 022339 (2009).
[10] O. Oreshkov, F. Costa, and C. Brukner, “Quantum correlations with no causal order,” *Nature Communications* 3, 1092 (2012).
[11] M. S. Leifer and R. W. Spekkens, “Towards a formulation of quantum theory as a causally neutral theory of bayesian inference,” *Phys. Rev. A* 88, 052130 (2013).
[12] J. Henson, R. Lal, and M. F. Pusey, “Theory-independent limits on correlations from generalized bayesian networks,” *New J. Phys.* 16, 113043 (2014).
[13] R. Chaves, C. Majenz, and D. Gross, “Information–theoretic implications of quantum causal structures,” *Nat. Commun.* 6, 5766 (2015).
[14] J. Pienaar and C. Brukner, “A graph-separation theorem for quantum causal models,” *New J. Phys.* 17, 073020 (2015).
steering,” Phys. Rev. A 94, 012308 (2016).
51] S. Nagy and T. Vértesi, “EPR Steering inequalities with Communication Assistance,” Sci. Rep. 6, 21634 (2016).
52] D. Cavalcanti and P. Skrzypczyk, “Quantum steering: a review with focus on semidefinite programming,” Reports on Progress in Physics 80, 024001 (2016).
53] Importantly, node A is kept classical throughout. The term “quantum” in QI is used here as a short form to refer to “quantum common-cause”. Thus, QI should not be mistaken with the fully quantum case where all nodes, including A, are quantised.
54] A. B. Sainz, N. Brunner, D. Cavalcanti, P. Skrzypczyk, and T. Vértesi, “Postquantum steering,” Phys. Rev. Lett. 115, 190403 (2015).
55] M. M. Taddei, R. V. Nery, and L. Aolita, “Necessary and sufficient conditions for multipartite Bell violations with only one trusted device,” Phys. Rev. A 94, 032106 (2016).
56] Guifré Vidal and Rolf Tarrach, “Robustness of entanglement,” Phys. Rev. A 59, 141–155 (1999).
57] M. Piani and J. Watrous, “Necessary and sufficient quantum information characterization of Einstein-Podolsky-Rosen steering,” Phys. Rev. Lett. 114, 060404 (2015).
58] R. Gallego and L. Aolita, “Resource theory of steering,” Phys. Rev. X 5, 041008 (2015).
59] R. Gallego and L. Aolita, “Nonlocality free wirings and the distinguishability between Bell boxes,” Phys. Rev. A 95, 032118 (2017).
60] B. Amaral, A. Cabello, M. T. Cunha, and L. Aolita, “Noncontextual wirings,” arXiv: 1705.07911 (2017).
61] M. Ringbauer and R. Chaves, “Probing the non-classicality of temporal correlations,” arXiv: 1704.05469 (2017).
62] Note that, interestingly, instrumental causal models display the same DAG for both the DI and the 1-sided DI cases. This is not the case for LHV and LHS models. The former need an extra (input) node with respect to the latter.
63] Interestingly, optimising the violation of a fixed NI witness over all quantum assemblages can – thanks to the no-signalling constraint – be directly recast as an SDP problem. In contrast, this is not true for QI assemblages.
64] O. J. Farías, G. H. Aguilar, A. Valdés-Hernández, P. H. Souto Ribeiro, L. Davidovich, and S. P. Walborn, “Observation of the emergence of multipartite entanglement between a bipartite system and its environment,” Phys. Rev. Lett. 109, 150403 (2012).
65] G. H. Aguilar, O. J. Farías, A. Valdés-Hernández, P. H. Souto Ribeiro, L. Davidovich, and S. P. Walborn, “Flow of quantum correlations from a two-qubit system to its environment,” Phys. Rev. A 89, 022339 (2014).
66] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. Souto Ribeiro, and S. P. Walborn, “Detection of entanglement in asymmetric quantum networks and multipartite quantum steering.” Nature communications 6, 7941 (2015).
67] Programs used to perform the analyses of the assemblages can be found at https://git.io/vbHqN.
Appendix A: 1-sided quantum instrumentality as a semi-definite programming membership problem

In this section, we consider the problems of how to determine if a given arbitrary assemblage \( \Sigma_{A|X} \) is or not in 1SQI and how to determine its optimal non-instrumentality witness, together with its corresponding violation. As in the membership problem for LHS and the determination of optimal steering witnesses of standard steering theory [52], these problems turn out to admit a formulation as a semi-definite program (SDP). SDPs deal with optimisations of a linear objective function over a matrix space defined by linear and positive-semidefinite constraints. Because of this, SDPs are exact in the sense that the solutions they return are guaranteed not to get stuck at local maxima or minima [68].

To this end, we express the conditional states in Eq. (1) as

\[
\tilde{\sigma}_{a,\lambda} \equiv \sum_{\lambda} D_{\lambda}(a|x) \tilde{\sigma}_{a,\lambda},
\]

where \( D_{\lambda} \) is the \( \lambda \)-th deterministic response function and \( \tilde{\sigma}_{a,\lambda} \equiv \sum_{\lambda} \tilde{P}_{\lambda} P_{a,\lambda} \), with \( \tilde{P}_{\lambda} \) defined such that \( P_{a,\lambda} \equiv \sum_{\lambda} \tilde{P}_{\lambda} D_{\lambda}(a|x) \). There are as many such functions as hidden-variable values, i.e. \( |A| = |\Lambda| = |A|^{[X]} \). In addition, the conditional states \( \tilde{\sigma}_{a,\lambda} \) are subnormalised such that

\[
\text{Tr} \left[ \tilde{\sigma}_{a,\lambda} \right] = \sum_{\lambda} \tilde{P}_{\lambda} P_{a,\lambda} =: \tilde{P}_{a},
\]

for all \( a \in \mathbb{Z}_A \) and all \( \lambda \in \mathbb{Z}_\Lambda \), i.e., their trace is independent of \( a \). Note that the distribution \( \tilde{P}_{\lambda} \equiv \{ \tilde{P}_{\lambda} \}_\lambda \) is automatically normalized if so is \( \Sigma_{A|X} \).

We can then recast the membership problem of \( \Sigma_{A|X} \) for 1SQI, i.e., whether \( \Sigma_{A|X} \) admits or not a decomposition as in Eq. (1), directly as an SDP feasibility test. This can be conveniently expressed by the following optimisation.

\[
\text{Given } \Sigma_{A|X} = \{ \sigma_a | x \}_a, \quad \text{min} \quad \mu, \quad \text{s. t.} \quad \sigma_a | x = \sum_{\lambda} D_{\lambda}(a|x) \tilde{\sigma}_{a,\lambda}, \quad \text{Tr} [\tilde{\sigma}_{a,\lambda}] = \text{Tr} [\tilde{\sigma}_{0,\lambda}], \quad \tilde{\sigma}_{a,\lambda} \geq -\mu I_B.
\]

Eqs. (A3b) and (A3c) encode the constraints in Eqs. (A1) and (A2), respectively. Hence, the minimisation in Eq. (A3a) amounts to finding an 1SQI decomposition in terms of conditional states \( \widetilde{\sigma}_{a,\lambda} \) as positive as possible, in the sense of satisfying the constraint of Eq. (A3d) with \( \mu \) as negative as possible. When the objective function \( \mu \) reaches a non-positive value, a decomposition as in Eq. (A1) is feasible with some \( \tilde{\sigma}_{a,\lambda} \geq 0 \), and vice versa. That is, any value \( \mu > 0 \) returned by the optimisation is equivalent to an 1SQI-decomposition being infeasible for \( \Sigma_{A|X} \), i.e. to \( \Sigma_{A|X} \notin \text{1SQI} \).

By virtue of the duality theory of semi-definite programming [52, 68], every such SDP admits a dual, equivalent formulation as follows.

\[
\text{Given } \Sigma_{A|X} = \{ \sigma_a | x \}_a, x, \quad \max \quad \mu, \quad \text{s. t.} \quad \sum_{a,x} \text{Tr} \left[ W_{a|x} \sigma_a | x \right], \quad \sum_{a,\lambda} C_{a,\lambda} \text{Tr} [\tilde{\sigma}_{a,\lambda}] = 0, \quad \sum_{a,\lambda,x} \text{Tr} \left[ W_{a|x} D_{\lambda}(a|x) \right] = -1.
\]

where \( C_{a,\lambda} \in \mathbb{R} \) for all \( a \in \mathbb{Z}_A \) and all \( \lambda \in \mathbb{Z}_\Lambda \). Eqs. (A4b) and (A4c) together imply that \( \Sigma_{a,x} \text{Tr} \left[ W_{a|x} \sum_{\lambda} D_{\lambda}(a|x) \tilde{\sigma}_{a,\lambda} \right] \leq 0 \), for any conditional states \( \tilde{\sigma}_{a,\lambda} \) satisfying Eq. (A2). So, these two equations encode the constraints that the assemblage-like object \( W_{A|X} = \{ W_{a|x} \in B(H_B) \}_a \) of Hermitian operators \( W_{a|x} \) returned by the optimisation is a non-instrumentality witness for some \( \beta \geq 0 \), Eq. (A4d) fixes the scale of \( W_{A|X} \), which prevents the maximisation in Eq. (A4a) from diverging to \( \infty \). Indeed, using the fact that \( \sum_{\lambda} D_{\lambda}(a|x) = |A|^{[X]-1} \), Eq. (A4d) yields

\[
s = -\frac{1}{|A|^{[X]-1}}.
\]

Other choices of scaling are valid, but they must be accompanied by a corresponding rescaling factor for the primal objective function in Eq. (A3a). Finally, the maximisation in Eq. (A4a) guarantees that \( W_{A|X} \) is the optimal non-instrumentality witness for \( \Sigma_{A|X} \) and that it is therefore (due to the convexity of 1SQI) tight – i.e. that \( \beta = 0 \) –, as we wanted to show. In other words, the maximisation returns a positive value if, and only if, \( \Sigma_{A|X} \notin \text{1SQI} \).

Finally, it is important to mention that the primal and dual SDPs, given respectively by Eqs. (A3) and (A4), satisfy a
convenient property called *strong duality* [52, 68]. By virtue of this, the primal and dual objective functions, i.e. \( \mu \) and \( \sum_{a, x} \text{Tr} [W_{a|x} \sigma_{a|x}] \), respectively, converge to the same optimal values. That is, the minimum of Eq. (A3a) and the maximum of Eq. (A4a) are guaranteed to coincide.

### Appendix B: Proof of theorem 1

Here, we analytically prove theorem 1. We give a non-instrumentality witness \( W_{A|X} \) and a quantum assemblage \( \Sigma_{A|X}^{(q)} \) such that \( \Sigma_{A|X}^{(q)} \) violates the 1S quantum instrumental inequality defined by \( W_{A|X} \). The assemblage we use is the same as that of Fig. 2 in the main text for \( V = 1 \) (no dephasing), i.e. the one obtained from \( \Phi^+ = ((00) + (11))/\sqrt{2} \) through projective local measurements by Alice in the bases \(- (\sigma_x + \sigma_z)/\sqrt{2}, \sigma_x, \sigma_z\). As our witness, we use the assemblage itself multiplied by a factor 2 for normalization purposes, i.e. \( W_{A|X} = 2 \Sigma_{A|X}^{(q)} \). Below, we show that the 1SQI bound corresponding to this witness is \( \beta = 2 + 1/\sqrt{2} \). On the other hand, it is immediate to see that

\[
\langle W_{A|X}, \Sigma_{A|X}^{(q)} \rangle = 3 > 2 + 1/\sqrt{2}. \tag{B1}
\]

This implies that \( \Sigma_{A|X}^{(q)} \notin 1SQI \) and, therefore, that \( Q \notin 1SQI \).

To prove that \( \beta = 2 + 1/\sqrt{2} \), we analytically maximise \( \langle W_{A|X}, \Sigma_{A|X}^{(inst)} \rangle \) over all \( \Sigma_{A|X}^{(inst)} \in 1SQI \) and obtain the claimed maximum \( 2 + 1/\sqrt{2} \). To this end, note first that the components of the quantum assemblage \( \Sigma_{A|X}^{(q)} \) in question are rank 1, so that \( \text{Tr} [W_{a|x} W_{a'|x}] = \delta_{a, a'} \). Then, using that \( \sigma_{a|x}^{(inst)} = \sum_{\lambda \in \Lambda} P(\lambda) D_{\lambda}(a|x) \rho_{a, \lambda} \), one gets

\[
\langle W_{A|X}, \Sigma_{A|X}^{(inst)} \rangle = \frac{3}{2} \frac{1}{2} \sum_{\lambda \in \Lambda} \text{Tr} [D_{\lambda}(a|x) R_{a, \lambda}],
\]

where \( R_{a, \lambda} := \sum_{j=1}^3 \text{Tr} [\rho_{a, \lambda} \sigma_j] e_j \) and

\[
v_{a, \lambda} := \sum_{x=0}^2 D_{\lambda}(a|x) u_x,
\]

with \( u_x := \sum_{j=1}^3 \text{Tr} [W_{0,x} \sigma_j] e_j \) and \( e_j \) the \( j \)-th element of the canonical orthogonal basis of \( \mathbb{R}^3 \), for \( j = 1, 2, \) or 3. Finally, optimising over \( P(\lambda) \) and \( R_{a, \lambda} \) yields

\[
\langle W_{A|X}, \Sigma_{A|X}^{(inst)} \rangle = \frac{3}{2} + \frac{1}{2} \sum_{\lambda \in \Lambda} \| v_{0, \lambda} \| + \| v_{1, \lambda} \|. \tag{B4}
\]

For the vectors \( u_x \) used here (namely \( u_0 = -(e_1 + e_3)/\sqrt{2}, u_1 = e_1, \) and \( u_2 = e_3 \)) we obtain \( \max_{x \in \Lambda} \| v_{0, \lambda} \| + \| v_{1, \lambda} \| = 1 + 1/\sqrt{2} \), from which the value \( \beta = 2 + 1/\sqrt{2} \) follows and the proof is completed.

### Appendix C: Device-independent instrumental inequality

We use the linear inequality derived in [35] and recently revisited in [36] to test for violations of the classical instrumental model by our assemblages, under an \( a \)-dependent measurement \( M_a \) by Bob (his input is equal to the output of Alice’s black box). The inequality can be expressed as

\[
\sum_{a, b \in \{0, 1\}} \left[ (-1)^a \rho_0 - (-1)^b \rho_1 + (-1)^{a+b} \right] P(a, b|0) + 2(-1)^b P(a, b|1) + 2(-1)^{a+b} P(a, b|2) \leq 3,
\]

where \( P(a, b|x) := \text{Tr} [M_a^{(b)} \sigma_{a|x}] \), with \( M_a^{(b)} \) the \( b \)-th element of the \( a \)-th measurement of Bob’s. The optimal measurements by Bob for the maximal violation are obtained through the analytical technique of Ref. [55].

### Appendix D: Monotonicity of \( R_{ni} \) under free operations of non-instrumentality

In this section we prove that \( R_{ni} \) is a *non-instrumentality monotone* for any linear class of free operations of non-instrumentality. We leave the details of the resource theory of non-instrumentality (in particular the explicit form of the corresponding free operations) for future work and prove monotonicity solely from the abstract generic properties of free operations. That is, we prove that \( R_{ni} \) is monotonomous (non-increasing) under any linear map satisfying the essential free-operation requirement that \( \mathcal{M}(\Sigma_{A|X}) \in 1SQI \) for all \( \Sigma_{A|X} \in 1SQI \). The proof is similar to that [58] of steering monotonicity for the steering robustness [52].

By definition, \( R_{ni}(\Sigma_{A|X}) \) is the minimal value \( t^* \in \mathbb{R}_{\geq 0} \) such that

\[
\Sigma_{A|X} + t^* \Pi_{A|X} = (1 + t^*) \Sigma_{A|X}^{(inst)}, \tag{D1}
\]

for some \( \Pi_{A|X} \in 1SQI \) and \( \Sigma_{A|X}^{(inst)} \in 1SQI \). Applying \( \mathcal{M} \) to both sides of this equation gives

\[
\mathcal{M}(\Sigma_{A|X}) + t^* \mathcal{M}(\Pi_{A|X}) = (1 + t^*) \mathcal{M}(\Sigma_{A|X}^{(inst)}), \tag{D2}
\]
where the linearity of $\mathcal{M}$ has been used. Now, since $\mathcal{M}(\Pi_{A|X})$, $\mathcal{M}(\Sigma_{A|X}^{(\text{inst})}) \in 1\text{SQI}$ (because $\mathcal{M}$ is a free operation of QI), Eq. (D2) realises a particular decomposition for $\mathcal{M}(\Sigma_{A|X})$ of the form of that of Eq. (D1) for $\Sigma_{A|X}$. Thus, $t^*$ must necessarily be larger or equal than the corresponding minimum for $\mathcal{M}(\Sigma_{A|X})$. That is,

$$R_{ni}(\mathcal{M}(\Sigma_{A|X})) \leq R_{ni}(\Sigma_{A|X}), \quad (D3)$$

which proves that $R_{ni}$ is a non-instrumentality monotone.

**Appendix E: Robustness of non-instrumentality as an SDP optimisation**

Eq. (5) can be re-expressed as $R_{ni}(\Sigma_{A|X}) = t^*$, with $t^*$ defined by Eq. (D1). This implies that

$$\sigma_{a|x} = (1 + t^*) \sum_{\lambda} D_{\lambda}(a|x) \tilde{\sigma}_{a,\lambda} - t^* \pi_{a|x}, \quad (E1)$$

with $\Pi_{A|X} = \{\pi_{a|x} \geq 0\}_{a,x} \in 1\text{SQI}$. Hence,

$$\pi_{a|x} = \sum_{\lambda} D_{\lambda}(a|x) \sigma'_{a,\lambda}, \quad (E2)$$

for $\sigma'_{a,\lambda} \geq 0$, with $\text{Tr}[\sigma'_{a,\lambda}] = \text{Tr}[\sigma_{a',\lambda}]$, $\forall a, a' \in A$. Both $D_{\lambda}(a|x)$ and $\tilde{\sigma}_{a,\lambda}$ are also defined as in Eq. (A1). The problem of finding $t^*$ is then expressed explicitly as the following SDP:

$$\begin{array}{ll}
\text{min} & \sum_{\lambda} D_{\lambda}(a|x) (\tilde{\chi}_{a,\lambda} - \chi_{a,\lambda}), \\
\text{s. t.} & \sigma_{a|x} = \sum_{\lambda} D_{\lambda}(a|x) \sigma'_{a,\lambda}, \\
& 1 + t^* = \sum_{\lambda} \text{Tr}[\tilde{\chi}_{a,\lambda}], \\
& \text{Tr}[\tilde{\chi}_{a,\lambda}] = \text{Tr}[\tilde{\chi}_{0,\lambda}], \\
& \tilde{\chi}_{a,\lambda} \geq 0, \\
& \chi_{a,\lambda} \geq 0.
\end{array} \quad (E3a-e)$$

Like Eq. (A3), the test (E3) admits a dual formulation, which takes the following form.

Given $\Sigma_{A|X} = \{\sigma_{a|x}\}_{a,x}$,

$$\begin{array}{ll}
\max & \sum_{a,x} \text{Tr}[V_{a|x} \sigma_{a|x}] - 1, \\
\text{s. t.} & B'_{a,\lambda} \mathbb{1}_B \leq \sum_{x} V_{a|x} D_{\lambda}(a|x) \leq B_{a,\lambda} \mathbb{1}_B, \\
& \sum_{a} B_{a,\lambda} = 1, \\
& \sum_{a} B'_{a,\lambda} = 0.
\end{array} \quad (E4a-d)$$

where $B_{a,\lambda}, B'_{a,\lambda} \in \mathbb{R}$ for all $a \in Z_A$ and all $\lambda \in Z_A$. With the same arguments as in the discussion right after Eqs. (A4), one sees that Eq. (E4a) returns a positive maximum (the one defining $t^*$) if, and only if, $\Sigma_{A|X} \notin 1\text{SQI}$. In fact, using that $\Sigma_{A|X}$ is well-normalised, the term $-1$ in the objective function can be absorbed into the witnesses’ definition with the variable change $V_{A|X} \rightarrow W_{A|X} := \{V_{a|x} - \mathbb{1}_B / |X|\}_{a,x}$. The resulting SDP (for the redefined witness $W_{A|X}$) is similar to the one in Eqs. (A4), but with an extra constraint coming from the left-hand side inequality of Eq. (E4b), and with the witness scale no longer fixed. Thus, the robustness is given by the violation of the optimal $V_{A|X}$ over all non-instrumentality witnesses with $1\text{SQI}$ bound $\beta = 1$ and subject to the specific constraints given by Eqs. (E4).

**Appendix F: Proof of theorem 2**

Consider an arbitrary $\Sigma^{(\text{qinst})}_{A|X} \in \text{QI}$. By definition, it admits a decomposition as in Eq. (2). Here, we use the short-hand notation $\Sigma^{(\text{qinst})}_{A|X} = \mathcal{E}_{B|A,A_B}^{(q)}(\Sigma^{(\text{q})}_{A|X})$, where $\mathcal{E}_{B|A,A_B}^{(q)} := \{\mathcal{E}_{B|A,A_B}^{(q)}\}_{a}$, to represent Eq. (2). In addition, we denote by $\mathcal{E}_{B|A,A_B}^{(\text{adj})} := \{\mathcal{E}_{B|A,A_B}^{(\text{adj})}\}_{a}$ the collection of dual (adjoint) maps $\mathcal{E}_{B|A,A_B}^{(\text{adj})}$ of each completely-positive trace-preserving (CPTP) map $\mathcal{E}_{B|A,A_B}$. This has the property that

$$\langle V_{A|X}, \mathcal{E}_{B|A,A_B}(\Sigma^{(\text{q})}_{A|X}) \rangle = \langle \mathcal{E}_{B|A,A_B}^{(\text{adj})}(V_{A|X}), \Sigma^{(\text{q})}_{A|X} \rangle,$$

for any $V_{A|X}$ and $\Sigma^{(\text{q})}_{A|X}$. Then, if $V_{A|X}$ is the optimal robustness witness of $\Sigma^{(\text{q})}_{A|X}$, defined by Eqs. (E4), it holds...
that

\[ R_{ni} \left( \Sigma^{(q\text{inst})}_{A|X} \right) = \left\langle V_{A|X}, \Sigma^{(q\text{inst})}_{A|X} \right\rangle - 1 \]
\[ = \left\langle V_{A|X}, E^*_{B|A,A_B} \left( \Sigma^{(q)}_{A|X} \right) \right\rangle - 1 \]
\[ = \left\langle \mathcal{E}^*_{B|A,A_B} \left( V_{A|X} \right), \Sigma^{(q)}_{A|X} \right\rangle - 1. \quad \text{(F1)} \]

Now, assume, for a moment, that \( E^*_{B|A,A_B} \left( V_{A|X} \right) \) is also a valid robustness witness. Then, denoting by \( U_{A|X} \) the optimal robustness witness for \( \Sigma^{(q)}_{A|X} \), it must hold that

\[ \left\langle \mathcal{E}^*_{B|A,A_B} \left( V_{A|X} \right), \Sigma^{(q)}_{A|X} \right\rangle - 1 \leq \left\langle U_{A|X}, \Sigma^{(q)}_{A|X} \right\rangle - 1. \quad \text{(F2)} \]

The left-hand side of this equation equals \( R_{ni} \left( \Sigma^{(q\text{inst})}_{A|X} \right) \), whereas the right-hand side equals \( R_{ni} \left( \Sigma^{(q)}_{A|X} \right) \), thus giving

\[ R_{ni} \left( \Sigma^{(q)}_{A|X} \right) \geq R_{ni} \left( \Sigma^{(q\text{inst})}_{A|X} \right). \]

So, the only missing ingredient is to show that \( E^*_{B|A,A_B} \left( V_{A|X} \right) \) is, in fact, a valid robustness witness.

To prove this we note that, for any \( \Sigma^{(\text{inst})}_{A|X} \in 1\text{SQI} \), \( E^*_{B|A,A_B} \left( \Sigma^{(\text{inst})}_{A|X} \right) \) is also in 1SQI. Hence,

\[ \left\langle \mathcal{E}^*_{B|A,A_B} \left( V_{A|X} \right), \Sigma^{(\text{inst})}_{A|X} \right\rangle = \left\langle V_{A|X}, E^*_{B|A,A_B} \left( \Sigma^{(\text{inst})}_{A|X} \right) \right\rangle \leq 1, \text{ for all } \Sigma^{(\text{inst})}_{A|X} \in 1\text{SQI}. \]

This implies that \( E^*_{B|A,A_B} \left( V_{A|X} \right) \) is a non-instrumentality witness with \( \beta E^*_{B|A,A_B} \left( V_{A|X} \right) = 1 \). Also, given that each \( E^*_{B|a,A,B} \) is completely-positive (CP) and unital, since it is the dual of a CPTP map, applying these dual maps to any robustness witness does not invalidate its defining constraints, in Eqs. (E4b)-(E4d). This means that the resulting object after the application of the dual map is also a valid 1SQI-robustness witness in the SDP formulation. \( \square \)