Demonstration that Einstein-Podolsky-Rosen steering requires more than one bit of faster-than-light information transmission

Yu Xiang,[1,2,*] Michael D. Mazurek [*,3,4,†] Joshua C. Bienfang,[5] Michael A. Wayne,[6] Carlos Abellán,[7] Waldimar Amaya,[7] Morgan W. Mitchell,[7,8] Richard P. Mirin,[4] Sae Woo Nam,[4] Qiongyi He,[1,2,‡] Martin J. Stevens,[4] Lynden K. Shalm,[4,3] and Howard M. Wiseman[9,§]

[1]*State Key Laboratory for Mesoscopic Physics, School of Physics, Frontiers Science Center for Nano-optoelectronics, & Collaborative Innovation Center of Quantum Matter, Peking University, Beijing 100871, China*
[2]*Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China*
[3]*Department of Physics, 390 UCB, University of Colorado, Boulder, Colorado 80309*
[4]*National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA*
[5]*Joint Quantum Institute: University of Maryland and National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg MD 20899, USA*
[6]*National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg MD 20899, USA*
[7]*ICFO - Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*
[8]*ICREA-Institució Catalana de Recerca i Estudis Avançats, 08015 Barcelona, Spain*
[9]*Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Brisbane QLD 4111, Australia*

Schrödinger held that a local quantum system has some objectively real quantum state and no other (hidden) properties. He therefore took the Einstein-Podolsky-Rosen (EPR) phenomenon, which he generalized and called 'steering', to require nonlocal wavefunction collapse. Because this would entail faster-than-light (FTL) information transmission, he doubted that it would be seen experimentally. Here we report a demonstration of EPR steering with entangled photon pairs that puts—in Schrödinger's interpretation—a non-zero lower bound on the amount of FTL information transmission. We develop a family of *n*-setting loss-tolerant EPR-steering inequalities allowing for a size-*d* classical message sent from Alice's laboratory to Bob's. For the case $n = 3$ and $d = 2$ (one bit) we observe a statistically significant violation. Our experiment closes the efficiency and locality loopholes, and we address the freedom-of-choice loophole by using quantum random number generators to independently choose Alice's and Bob's measurement basis settings. To close the efficiency and locality loopholes simultaneously, we introduce methods for quickly switching between three mutually unbiased measurement bases and for accurately characterizing the efficiency of detectors. From the space-time arrangement of our experiment, we can conclude that if the mechanism for the observed bipartite correlations is that Alice's measurement induces wave-function collapse of Bob's particle, then more than one bit of information must travel from Alice to Bob at more than three times the speed of light.

I. INTRODUCTION

In 1935, Einstein, Podolsky and Rosen (EPR) put forward their famous paradox to declare the quantum mechanical description of physical reality incomplete [1]. Their argument relied on a remarkable feature of bipartite quantum entanglement, that the choice of measurement by one party (Alice) seems to affect, instantaneously, the type of state held by the second party (Bob). Einstein later called this phenomenon "spooky action-at-a-distance" [2]. Meanwhile, in the same year as EPR, Schrödinger called the phenomenon "steering" or "piloting" the remote state [3], and discussed the possibility of using arbitrarily many types of measurement. Reid [4] gave the first formulation of an experimental criterion for testing this effect based on EPR's example of position and momentum. More recently, Schrödinger's generalisation was formalised, as EPR steering, by one of us and co-workers [5, 6], as the experimental violation of the asymmetric model comprising a local hidden state (LHS) quantum model for Bob's

system and a local hidden variable (LHV) model to generate Alice's measurement results. Subsequently, broad classes of experimental criteria to demonstrate EPR steering, known as steering inequalities, were derived [7–10]. Moreover, EPR-steering was shown [5, 6] to be equivalent to the quantum information task of verifying entanglement in the absence of trust of Alice or her equipment. This is relevant for communication networks where reliability of devices and dishonest observers becomes an issue [11–14].

While the largest impacts of EPR steering have been in its myriad quantum information applications —including one-sided device-independent quantum cryptography [12–18], secure quantum teleportation with high fidelity [19–21] and subchannel discrimination [22–24]—there has also been considerable theoretical and experimental interest in its foundational implications [25–42], as Einstein and Schrödinger were concerned with. In particular, from the perspective of Schrödinger, who was convinced of the completeness and correctness of the quantum state as a description of a *local* system, EPR steering implied a genuinely superluminal effect (which is why he doubted that it would be seen experimentally [3]). This effect was first demonstrated in a way that closed the locality and efficiency loopholes for EPR steering in 2012 [27]. However, that experiment did not put any lower bound on *how much* classical information would have to be sent faster-than-light (FTL) for Alice to steer Bob's state, in

the absence of entanglement.

The question of how to establish a lower bound on the apparent FTL information transmission in EPR steering was theoretically investigated in Refs. [43, 44]. The authors showed that an infinite amount of classical communication (from an untrusted Alice to a trusted Bob), is necessary to simulate all EPR-steering correlations for any pure entangled state of two qubits. For any experimentally accessible correlations, the amount of classical communication required to violate the LHS model will be finite, and moreover is one measure of the strength of EPR steering.

In this paper, we provide the first demonstration of EPR-steering correlations that provably require more than one bit of FTL communication to simulate classically. We do this by introducing and subsequently violating an inequality that is derived assuming one bit of FTL communication is possible. Our experiment addresses the efficiency, locality, and freedom-of-choice loopholes. The requirements for experimentally testing our inequality while closing the efficiency and locality loopholes are many, but we highlight two which have interest outside this particular task. First, we achieve fast-switching between three mutually unbiased measurement bases with minimal losses by using a single Pockels cell. Second, we introduce a modification of Klyshko's method [45] to accurately characterize the efficiencies for Bob's two detectors.

This paper is structured as follows. In Sec. II we discuss how to close the loopholes in quantum nonlocality tests in general, in EPR-steering tests more particularly, and in our experiment allowing for one bit of FTL communication most particularly. In Sec. III we introduce a general family of inequalities for testing communication-assisted EPR steering, and in Sec. IV we derive a specific inequality for the case when Bob can implement a set of three measurement settings. Our derivation goes beyond the theory in Refs. [43, 44] in that it includes the possibility of null results by Alice (as required for closing the efficiency loophole) and is optimized for the experimental conditions (Alice's efficiency, and the three measurements Bob can switch between). In Sec. V we present an experimental test of our EPR-steering inequality using polarization-entangled photon pairs. Our experiment closes the locality, freedom-of-choice, and efficiency loopholes. We observe a statistically significant violation of our inequality, thus demonstrating correlations that—under the standard EPR-steering assumptions—require more than one bit of FTL information transmission.

II. CLOSING LOOPHOLES IN QUANTUM NONLOCALITY TESTS

There are four standard loopholes for tests of quantum nonlocality in a general sense: the locality loophole, the freedom-of-choice loophole, the efficiency (or detection) loophole, and the memory loophole [46]. These loopholes are best known for experimental tests of Bell-nonlocality with two parties, where they have the following meaning. The *locality loophole* is closed if and only if each party's choice of setting is space-like separated from the other party's observation of the measurement result. The *freedom-of-choice loophole* refers to the requirement that the setting choices by the two parties are 'free', i.e. not influenced by any hidden variables [47, 48]. In principle, the closure of this loophole can never be absolutely assured, but in practice each party can use a state-of-the-art random number generator for this purpose [49]. Closing the *efficiency loophole* requires that all runs of the experiment are included in the data analysis, not just a subensemble where both particles are detected, for example [50]. This can be achieved by having sufficiently high detection efficiency on both sides. Finally, the *memory loophole* applies to the analysis of finite experimental data. If all trials are assumed to be independent and identical it is possible for an adversary to use information (memory) about past trials to try and influence or bias the outcomes of future trials. This can lead to the assignment of greater confidence to a violation than is warranted [46]. All four loopholes have been simultaneously closed in a number of LHF Bell experiments [51–55].

EPR steering has the same loopholes, but asymmetrically. Because Bob's apparatus is trusted, *efficiency* only matters on Alice's side; it is permissible to consider the sub-ensemble where Bob detects a particle. Moreover, under some conditions it is only necessary to have one *free* measurement choice. For example, it can suffice for Bob to make only a single generalized measurement [28]. An alternative scheme, whereby Alice uses the same randomly generated setting as Bob, but can find it out only after Bob's photon is in his 'lab' (his trusted space), was implemented in Ref. [27]. In that experiment, Alice's and Bob's outcomes were space-like separated, as required by the *locality* loophole, and it also closed the efficiency loophole (three years before the first LHF Bell experiments [51–53]). Proving that an EPR-steering experiment is not affected by any potential *memory* loophole has never been done, as it would require the development of an entirely new and nontrivial analysis method. In the current work we close the efficiency, freedom-of-choice, and locality loopholes, but we analyze our data in a way that assumes that experimental trials are independent.

In our current experiment, we use measurement choice on both sides, and so the LHF nature is ensured in a similar way to the bipartite photonic LHF Bell experiments [52, 53], with two random number generators and fast electro-optic switching. The main difference is that, as in Ref. [27], Bob has a lab, as illustrated in Fig. 1, in which he trusts the operation of his devices and his quantum mechanical description of them. This lab is taken to be large enough that it contains the event E defined as the intersection of the future light-cone emanating from the event j of Bob's random number generation with the world-line of the photon which Bob will ultimately detect. Note this world-line only needs to be defined inside Bob's lab (where it is indeed defined because Bob trusts his apparatus) in order to define E. While E is a point in space-time, not a point in space, it is natural to define Bob's lab to be stationary in the rest frame of his apparatus, so that E effectively determines its extent. As in LHF Bell experiments, we must trust that Alice's and Bob's random number generators make their basis choices at the times and locations we set them to. Finally, event E must not be in the future light cone of Alice's random number generation k, while the event b of Bob's
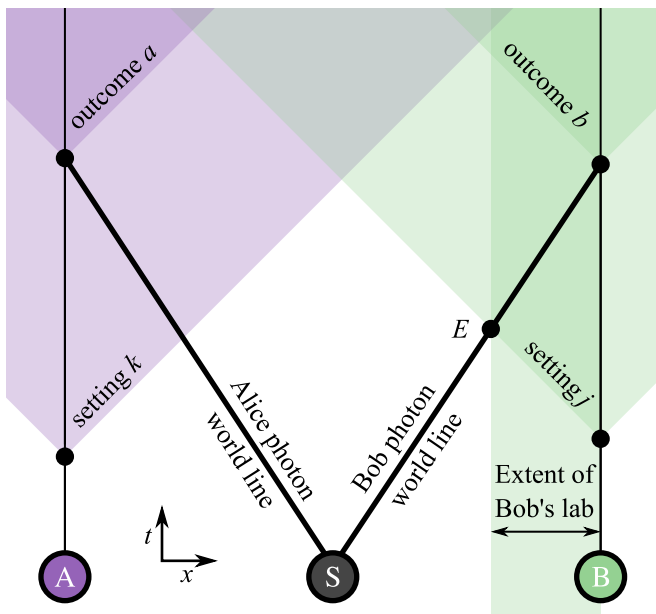
Figure 1. Minkowski diagram illustrating the relations between spacetime events (black dots) at Alice (A), Bob (B), and the source (S) that are needed to close the locality loophole, for a simple example in which all parties are arranged collinearly. A trial begins when the source emits an entangled pair of photons. If these photons travel in fiber, they will travel at roughly 2/3 light speed, as indicated by the slope of their world-lines. While the photons are in flight, measurement settings k and j are randomly chosen, defining their associated future light cones. The event E is the point in spacetime where the world line of Bob's photon intersects the light cone corresponding his setting choice j. Bob's 'lab' must be defined to include the event E, and the event E must lie outside of the future light cone of Alice's setting k. Finally, Alice must complete her measurement and record her outcome a outside of the future light cone of Bob's outcome b.

measurement result becoming macroscopic (and thus, we'll assume, potentially 'visible' to an adversary) must not be in the past light cone of the event a of Alice reporting her result.

The reason Bob needs a trusted space which includes event E satisfying the above conditions is as follows. In a steering scenario, Alice is not trusted, and she could have an accomplice, Fenella, who can act on her behalf anywhere except in Bob's trusted space. If Fenella were present at event E, or any subsequent point on the world-line of the photon going to Bob's detectors, then she could make the state of that photon depend on Bob's measurement choice j. This violates one of the fundamental assumptions behind EPR-steering inequalities, that j is uncorrelated with Bob's LHS. It would easily allow Alice to cheat as follows: Knowing the setting j, Fenella could prepare the state so that Bob always gets the result +1, and thus a steering inequality of the form we later consider could attain its logically maximum value (and therefore violate the EPR-steering bound) if Alice simply always says +1 too. This means that E has to be off-limits to Fenella. In other words, it has to be in Bob's trusted space. Note that this is in contrast to Bell inequalities, where there is no LHS assumption, and no trusted space.

The event E cannot be in the future light-cone of the gener-

ation of Alice's setting k because if it were then Fenella could make the state Bob is to detect depend on k, which again violates the fundamental assumption of EPR steering. In this case, it is analogous to the Bell condition that the event of Bob's outcome b cannot occur in the future light-cone of the generation of Alice's setting k. Similarly, an untrusted Alice cannot be allowed to delay the generation of her outcome a until it is in the future light-cone of Bob's result b appearing, which in this case is exactly the same in EPR steering as it is in the Bell case[1]. Finally, note that Bob's detectors being in his trusted space means that he is allowed to post-select on a successful photon-detection.

The extra feature of the experiment reported here is that we allow, in the theory, Alice to send one bit of information FTL to her accomplice Fenella. This means that Fenella can make, in any trial, Bob's photon have one of two possible polarization states after she knows Alice's setting k but before it enters Bob's lab. That the information would have to be sent FTL is guaranteed by the fact that event E (at which is present the photon Bob will detect) is space-like separated from Alice's random number generation k, so any event that can influence the state of the photon at E must also be space-like separated from k. Because it rules out one bit of FTL communication as an explanation for EPR steering, our experiment is stronger, and considerably more difficult to achieve, than that of Ref. [27]. Finally, we note that it would not be possible to do an experiment similar to the current one that rules out more than one bit of FTL communication in a Bell scenario, because—under the usual Bell assumptions in which both Alice and Bob are untrusted—one bit of communication suffices to simulate any correlations from projective measurements on a singlet state [56].

III. EPR-STEERING INEQUALITY WITH BOUNDED ONE-WAY COMMUNICATION

We consider an EPR-steering task with bounded one-way communication. That is, the model to disprove is one with a LHS for Bob, but where Alice can send a classical message, taking one of d possible values, to Bob's side to affect his state before his detection. Here, d is an integer between 1 (no message) and $n - 1$, where n is the number of settings Alice uses. There is no point considering $d \geq n$ because that would allow perfect FTL communication of Alice's setting k to Bob's side—technically, to event E—which would allow perfect correlations even with the LHS assumption, and so could never be disproven experimentally. It is convenient to define $d = 2^{H_0}$, where H_0 is the number of bits of communication in a max-entropy sense. Recently, Refs. [43, 44] studied the quantification of EPR steering according to the classical message size c (i.e. H_0) needed to simulate an assemblage

---

[1] Both of these conditions (that involving E and k, and that involving a and b) are automatically guaranteed by the conditions that are used in a LHF Bell scenario, namely that each party's choice of setting is space-like separated from the other party's observation. However, we note that these LHF-Bell conditions are not necessary for a LHF test of EPR-steering correlations.

without entanglement. Here an assemblage is a complete description of Alice's ability to steer Bob in a given experiment — the set of all Bob's states to which she can steer, and associated probabilities, indexed by her settings and outcomes. In Ref. [43] the set of assemblages which have an LHS model when c bits of communication are sent from Alice to Bob was shown to have an efficient semidefinite program (SDP) formulation. They proved that infinite communication is necessary to simulate the maximally entangled state. On the other hand, in Ref. [44] the LHS-robustness was shown to provide an upper bound on the amount of communication, and infinite communication cost was shown to be necessary even for some impure states that were not full-rank.

In this paper we adopt the approach of generalizing the linear steering inequalities of Refs. [26, 32, 57]. Bob is assumed to have some LHS ρ, and—under this assumption—a violation indicates that Alice's measurement choice must have an effect on this state. Our formulation establishes a correlation bound, $h^n_{H_0}$, beyond which one would require H_0 bits of superluminal communication to replicate the effect of Alice's choice, in a LHS model. The bound is determined by optimizing the following LHS (no-entanglement) protocol. (1) Bob sets his apparatus to measure the observable $\hat{B}_j$ with randomly chosen setting index $j \in \{1, \dots, n\}$. (2) At the same time (i.e. as a space-like separated event), Alice generates an outcome a_k (from a randomly chosen setting $k \in \{1, \dots, n\}$). (3) Alice sends an FTL message $l \in \{1, \dots, d\}$ to her accomplice, Fenella, located near Bob's laboratory. (4) Knowing l, Fenella generates a state $|l\rangle_F$, one of d possible states $\{|l\rangle_F\}^d_{l=1}$ she could prepare, and feeds it into Bob's laboratory. Note that the states $\{|l\rangle_F\}^d_{l=1}$ need not be mutually orthogonal, even for $d = 2$.

Two-outcome measurements yield a binary variable, +1 or −1. Physically, in our experiment, this corresponds to a click in one of two detectors, at each side[2]. Due to the limited transmission and detection efficiency of the photons, a third outcome 0 has to be considered to represent no detection. According to the LHS model, Bob trusts his apparatus, including his two detectors, so he can discard those experimental trials where he fails to detect a photon. That is, Bob can make the fair-sampling assumption without opening the efficiency loophole, and describe his output by the binary variable $b_j \in \{+1, -1\}$. However, because Alice is not trusted, she is not allowed to discard any results, and her output is a ternary variable $a_k \in \{+1, 0, -1\}$. After all measurement trials are complete, Alice sends her data to Bob. Bob discards all data for trials in which he did not detect an outcome, and then he computes an appropriate steering correlation function S_n and checks whether the inequality

$$S_n \leq h^n_{H_0} \tag{1}$$

is violated for the appropriately calculated bound $h^n_{H_0}$. If so, they have demonstrated EPR steering requiring more than H_0 bits of FTL communication. Note that the dependence on n

here is shorthand for dependence on the set of n observables $\{\hat{B}_j\}^n_{j=1}$ that Bob performs. Neither the correlation function nor the bound make any assumptions on the nature of the measurements that Alice performs, apart from their having outcomes $a_k \in \{+1, 0, -1\}$.

The task, then, is to choose a suitable steering correlation function, or family of correlation functions, and to compute the bound that can be achieved with an LHS model supplemented by allowing Alice H_0 bits of FTL communication. We design the steering correlation functions specifically for our experiment, in which the two-photon state shared by Alice and Bob is close to a maximally entangled singlet state $|\Psi^-\rangle$ that has been subjected to loss, and Alice's n measurement axes are intended to be the same as Bob's. For simplicity, we number the corresponding pairs of axes the same, and label the directions corresponding to the results (±1) in opposite directions. Thus, Alice's and Bob's results will be maximally correlated if they choose the same setting. Such correlations are best for demonstrating steering [32]. Also, for all j and k the mean of b_j and a_k will be close to zero, and $|b_j|$ is identically 1 (because Bob can post-select). Finally, we aim to implement a set of measurements $\{\hat{B}_j\}$ with a symmetry property that ensures that no measurement is special. Specifically, we choose a set of measurements for which, given any two measurements, there exists a rotation of the axes that can take the first to the second while leaving Bob's total set of measurement directions unchanged. That is, $\forall j, j' \exists \hat{U} : \hat{U}^\dagger \hat{B}_j \hat{U} = \hat{B}_{j'}$ and $\{\hat{U}^\dagger \hat{B}_i \hat{U}\}_i = \{\hat{B}_i\}_i$. (As a concrete example, the set of measurements that contains the three Pauli observables $\{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ satisfies this symmetry requirement.)

Based on these considerations we can impose, with almost no loss of effectiveness, simplifying symmetries on the form of our correlation function: invariance under interchanging the results ±1 for a and b jointly; and invariance under permutation of setting values j and k jointly. Imposing these symmetries means that the function weighs all correlations equally, and thus we obtain a family of steering correlation functions with only one real parameter, r_{H_0}:

$$S_n(r_{H_0}) = \frac{1}{n} \sum_{j,k=1}^n \delta_{j,k} \left(\langle a_k \hat{B}_j \rangle - r_{H_0} \langle |a_k| \rangle \right). \tag{2}$$

In Eq. 2, Alice's result a_k is to be considered a random variable, and so is Bob's result b_j, but we represent the latter by its associated Hermitian operator $\hat{B}_j$ to emphasize that (unlike in a Bell correlation) we trust this description of Bob's measurement process. That is, deriving the EPR-steering bound on this correlation (below) makes direct use of the operator properties of the $\hat{B}_j$. Note that the use of the above symmetries in motivating the above correlation function is for simplicity only; the validity of the inequality we will ultimately derive does not depend on assuming that these symmetries in the state and measurement settings are satisfied.

Once the parameter r_{H_0} is chosen, the bound $h^n_{H_0}$ can then be calculated via

$$h^n_{H_0} = \max_{\alpha, \ell, \{|l\rangle_F\}^d_{l=1}} \frac{1}{n} \sum_{j=1}^n \delta_{a,\alpha(j)} \left(-r_{H_0}|a| + \sum_{l=1}^d \delta_{l,\ell(j)} a \langle l|_F \hat{B}_j |l\rangle_F \right), \tag{3}$$

where ℓ indicates a deterministic function[3] from $\{1,\ldots,n\}$ to $\{1,\ldots,d\}$ (recall $d = 2^{H_0}$), which specifies the strategy for Fenella to prepare the states $\{|l\rangle_F\}_{l=1}^d$ related to Alice's measurement setting k, and α indicates a deterministic function from $\{1,\ldots,n\}$ to $\{+1, 0, -1\}$, which determines Alice's outcomes. Note that the linearity of the correlation function with respect to Bob's observables means that it can always be maximized by a deterministic strategy, allowing us to calculate the bound in this way.

We define Alice's heralding efficiency through the method outlined by Klyshko [45]. That is, it is the probability with which a detection by Bob heralds a detection by Alice. It is given by the expression $\frac{1}{n}\sum_{k=1}^{n}\langle|a_k|\rangle$, and in the below, we will simply call this Alice's efficiency, η_A. The purpose of r_{H_0} is to make the steering demonstration as loss-tolerant as possible. That is, to allow violation for a value of Alice's heralding efficiency which is as small as possible given the other parameters in the experiment (the degree of mixture in the lossy maximally entangled singlet state, the value of H_0, and the measurement axes).

IV. INEQUALITIES FOR OCTAHEDRAL SETTING ARRANGEMENT

In our experiment, we use $n = 3$, the minimum number required to rule out LHS models with up to 1 bit of FTL communication. We use an octahedral arrangement for the measurement axes. That is, the labels $\{1, 2, 3\}$ correspond to states $|0\rangle$,
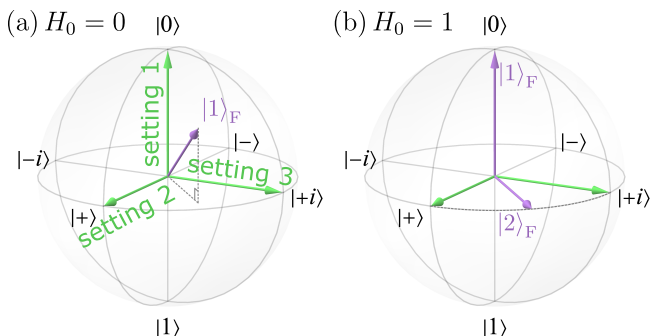

Figure 2. Sketch of Bob's three ideal measurement directions (green arrows represent the $+1$ eigenstate of each measurement setting) and Alice's optimized cheating ensembles generated by Fenella $\{|l\rangle_F\}$ (purple arrows) in the Bloch spheres. (a) Without any information transferred ($H_0 = 0$), there is only one optimized cheating state $|1\rangle_F$ for all three measurement directions; (b) When $H_0 = 1$ bit messages are transferred, there are many possible optimal cheating strategies. One such strategy is as follows: if Alice measures in setting $k = 1$, she sends Fenella the message $l = 1$ and Fenella prepares the state $|1\rangle_F$. Otherwise, Alice sends the message $l = 2$ and Fenella prepares the state $|2\rangle_F$.

---

[3] In principle, Alice and Fenella could update this function based on the results of past measurement trials. However, since we assume that all trials are independent and identical, we do not consider that here.

$|+\rangle$, and $|+i\rangle$, as illustrated in Fig. 2. This simplifies the derivation of the inequality, allowing us to obtain analytical bounds, which we present in this section. In a later section, where we allow for systematic and statistical errors in Bob's settings, we obtain the bounds numerically. In both cases we find the bounds by an exhaustive search over deterministic strategies. We begin by considering the no message ($H_0 = 0$) case, for which optimising r_{H_0} gives the same loss-tolerant inequalities previously derived for EPR-steering with no FTL message assistance [26, 57].

IV.1. Derivation of the inequality for the no-message case

If Alice, the untrusted party, had perfect detectors, she could declare non-null results for every trial and for all n settings, so that we would have $\langle|a_k|\rangle = 1$, and we do not need the parameter r_0. When there is no message transferred ($H_0 = 0$), Fenella generates the same state $\{|l\rangle_F\}$ ($l = 1$) for all three settings. The optimal orientation is towards the face-center of the octahedron used to define three measurement axes, as delineated in Fig. 2(a). This leads to the following EPR-steering inequality

$$\frac{1}{3}\sum_{j,\,k=1}^{3}\delta_{j,k}\langle a_k\hat{B}_j\rangle \le \frac{\sqrt{3}}{3},\qquad(4)$$

The bound $h_0^3 = \sqrt{3}/3$ coincides with the inequality previously given in Ref. [32].

If Alice sometimes reports null results, the parameter r_0 is now relevant, and a new search for the optimal solutions for $\alpha(j)$ and $\{|l\rangle_F\}$ must be undertaken for each value of r_0 in Eq. (3) (Note that in this no-message case, the function ℓ is trivial, so no search over this is needed). In addition, we wish to choose r_0 optimally for a given value of η_A, Alice's efficiency. Note that while we assume, for the purpose of optimizing r_0, that Alice's efficiency is independent of k ($\eta_A = \frac{1}{n}\sum_{k=1}^{n}\langle|a_k|\rangle$), and that Alice's two detectors have the same efficiency, these assumptions are not required for the inequality itself to be valid. An optimal r_0 for a given η_A is the one which allows the inequality to be violated most easily with our model of the state, a loss-depleted Werner state [58]

$$W_\mu = \mu|\Psi^-\rangle\langle\Psi^-| + (1 - \mu)\mathbf{I}/4\qquad(5)$$

with singlet-proportion $\mu \in [0, 1]$. For such a model, the residual between the left-hand-side and right-hand-side of Eq. (1) is given by

$$R_0^3 = \eta_A(\mu - r) - h_0^3(r).\qquad(6)$$

Maximizing this residual for any η_A (which is equivalent to minimizing the μ at which a positive residual is possible), we find that the optimal r_0, and the corresponding bound, is piecewise constant, changing at the points $\eta_A = 1/3$ and $2/3$. Examining these points analytically, and using the $\eta_A = 1$ analysis above, we are able to find the relevant values of r_0, and the bound, analytically. We show how this is done for one

example in Appendix A. We find

$$
r_0 = \begin{cases} \sqrt{2} - 1 & \text{if } \frac{1}{3} \leq \eta_A < \frac{2}{3}, \\ \sqrt{3} - \sqrt{2} & \text{if } \frac{2}{3} \leq \eta_A < 1, \end{cases}
$$

$$
h_0^3 = \begin{cases} \frac{2 - \sqrt{2}}{3} & \text{if } \frac{1}{3} \leq \eta_A < \frac{2}{3}, \\ \frac{3\sqrt{2} - 2\sqrt{3}}{3} & \text{if } \frac{2}{3} \leq \eta_A < 1. \end{cases} \quad (7)
$$

If Alice reports too many null results, so that $\eta_A < 1/3$, the optimal gain factor is $r_0 = 1$, and $h_0^3 = 0$. This means that the inequality is impossible to violate. This is because, below this detection-efficiency threshold, the steering correlation function S_3 cannot be positive even with the maximally entangled states [26].

IV.2. Derivation of the inequality when one FTL bit is transferred

Now allowing Alice the assistance of an $H_0 = 1$ bit message, we must include the function $\ell : \{1, 2, 3\} \to \{1, 2\}$ which allows Fenella to choose a state from $\{|l\rangle_F\}_{l=1}^2$ given Alice's measurement setting k. If Alice declares a non-null result in every round, then we find the inequality

$$
\frac{1}{3} \sum_{j,k=1}^3 \delta_{j,k} \langle a_k \hat{B}_j \rangle \leq \frac{1 + \sqrt{2}}{3}, \quad (8)
$$

with the optimized parameter $r_1 = 0$. Here, the bound $h_1^3 = (1 + \sqrt{2})/3$ can be achieved by appropriately choosing $\{|l\rangle_F\}_{l=1}^2$, $\alpha(j)$ and $\ell(j)$. For instance, the optimal strategy for Alice/Fenella with $\alpha(j) = +1$ for all j, and $\ell(1) = 1$, $\ell(2) = \ell(3) = 2$ is shown in Fig. 2(b).

Turning now to the case of where Alice has to deal with loss (or is trying to cheat by declaring null results), the maximization has to be done including the parameter r_1, and r_1 has to be chosen optimally. Again, we find the solution analytically:

$$
r_1 = \sqrt{2} - 1, \; h_1^3 = \frac{4 - 2\sqrt{2}}{3} \quad \text{if } \frac{2}{3} \leq \eta_A < 1. \quad (9)
$$

Now, when Alice's efficiency is less than $2/3$, the optimal $r_1 = 1$ and $h_1^3 = 0$, leading to a trivial (unable to be violated) inequality. This is as expected, as the assistance of transferred messages can make it easier for a lossy Alice to find a cheating strategy, which in turn means more stringent conditions for her efficiency to demonstrate EPR-steering. The residual from Eq. (1),

$$
R_1^3 = \eta_A(\mu - r) - h_1^3(r), \quad (10)
$$

shows that to have the $R_1^n > 0$, the required purity $\mu \to 1$ as η_A approaches $2/3$ from above. The details of the optimization procedure leading to Eqs. (8) and (9) can be found in Appendix A. When the measurements actually implemented by Bob are different from the above ideal case, we have to reoptimize the parameters r_{H_0} to obtain the corresponding bounds; this will be covered in more detail in Sec. V.2.
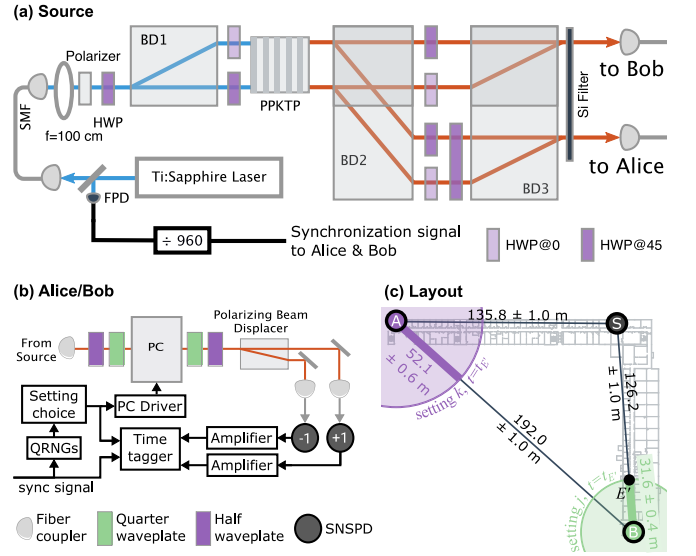

Figure 3. Experimental test of EPR steering with one bit of FTL communication. (a) The source produces polarization-encoded photon pairs in the maximally entangled singlet state via type-II spontaneous parametric downconversion. (b) Alice's and Bob's labs where the photons' polarizations are measured; a Pockels cell allows fast switching between three mutually unbiased measurement bases. (c) Layout of the experiment, showing the relative locations of Alice (A), Bob (B) and the source (S), overlaid on a floor plan of the section of the building where the experiment took place. The spacetime event E' (black dot) is our conservative estimate of the event E (see Sec. V.1 for details). At the time of E', $t_{E'}$, the region of space inside the future light cone of Bob's setting choice j is represented by the green circle, and the region of space inside the future light cone of Alice's setting choice k is represented by the purple circle. The faded rings outside the green and purple circles represent the uncertainties on their radii. We define Bob's lab (in which he trusts his characterization of his equipment) to contain his measurement apparatus plus the final 31.6 ± 0.4 m of fiber leading to his measurement apparatus. FPD: fast photodiode; SMF: single-mode fiber; BD: polarizing beam displacer; PPKTP: periodically poled potassium titanyl phosphate; QRNG: quantum random number generator; PC: Pockels cell; SNSPD: superconducting nanowire single-photon detector.

V. EXPERIMENTAL METHODS

We perform a test of our new EPR-steering inequality with polarization-entangled photon pairs that are distributed to two distant measurement stations we name Alice and Bob.

The entangled pair source [53] is displayed in Fig. 3(a). A Ti:Sapphire laser emits a beam of pump pulses with wavelength 775 nm at a repetition rate of 80 MHz. A small fraction of the beam is picked off and sent to a fast photodiode (FPD) which sends an electronic synchronization pulse to both Alice and Bob once every 960 laser pulses. The remainder of the pump beam is sent through a single-mode fiber to clean up its spatial mode. After exiting the fiber, the beam is gently focused and prepared in the diagonal polarization state before it is equally split into two paths by a polarizing beam displacer (BD1). In each path, photon pairs are produced via type-II spontaneous parametric downconversion in a 20 mm-long periodically poled potassium titanyl phosphate crystal placed at

the pump beam's focus. Polarizing beam displacers BD2 and BD3 and a series of half waveplates are used to recombine the beams into two paths (with one photon in each path). A silicon window filters out the pump beam, then each photon is coupled into single-mode fiber over which it is sent to Alice's or Bob's measurement station. By controlling the polarization rotation that occurs inside each fiber, we ensure that the pair arriving at Alice's and Bob's measurement stations is in the maximally entangled singlet state $\frac{1}{\sqrt{2}}(|H\rangle|V\rangle - |V\rangle|H\rangle)$, where $|H\rangle$ ($|V\rangle$) denotes horizontal (vertical) polarization.

The measurement stations at Alice and Bob are nominally identical, and one of them is represented in Fig. 3(b). We perform measurement trials at a rate of roughly 83 kHz – each electronic synchronization pulse from the source triggers a new measurement trial. At the beginning of each trial, a pair of quantum random number generators [49, 59] produces a pair of random bits to determine the setting for that trial. If these two output bits are 00, 01, or 10, we measure in setting 1, 2, or 3 respectively, and the setting choice determines which voltage is applied to a Pockels cell. (One fourth of the time the two output bits will be 11; we ignore these trials since we want each of the three settings to occur with equal probability.) When a photon arrives at a measurement station, it is coupled into free space and transmitted through four waveplates and a Pockels cell before its $|H\rangle$ and $|V\rangle$ components are separated with a polarizing beam displacer. Each output of the beam displacer is then coupled into single-mode fiber where it is sent to one of two superconducting nanowire single-photon detectors (SNSPDs), labelled[4] with the outcomes '+1' and '-1'. Output pulses from the SNSPDs are amplified before being recorded by a time tagger. We map the computational basis $\{|0\rangle, |1\rangle\}$ to polarization states with the assignments $|0\rangle \rightarrow |V\rangle$ and $|1\rangle \rightarrow |H\rangle$. The waveplate angles and Pockels cell voltages are set to (ideally) realize the measurements in Fig. 2. That is, ideally, setting 1 corresponds to a measurement in the $\{|0\rangle, |1\rangle\}$ ($\{|V\rangle, |H\rangle\}$) basis, setting 2 to the $\{|+\rangle, |-\rangle\}$ ($\{|D\rangle, |A\rangle\}$) basis, and setting 3 to $\{|+i\rangle, |-i\rangle\}$ ($\{|L\rangle, |R\rangle\}$). Here, $|D\rangle$, $|A\rangle$, $|L\rangle$, and $|R\rangle$ correspond to the diagonal, antidiagonal, left-circular, and right-circular polarization states. We provide full details of how we determine which waveplate angles and Pockels cell voltages to use in Appendix B.

V.1. Locality conditions

In our experiment, the pair source and Alice's and Bob's measurement stations are arranged in a triangular configuration (Fig. 3(c)). Before assessing if the locality conditions are met, we must first define the extent of Bob's lab, i.e. the size of the area in which Bob trusts his characterization of his equipment, which is defined by the spacetime event E. Recall that the spacetime event E is defined as the point in space-time at which the light cone emanating from Bob's random number

generators (and potentially carrying information about Bob's measurement setting choice) intersects with the world line of the photon travelling to Bob's lab (Fig. 1). One piece of information we need to find E is the exact path that Bob's photon takes while it is inside his lab. Because Bob's photon travels in fiber along the hallways and through the ceiling of the building we run the experiment in, we only approximately know where the fiber lies. We do, however, have an accurate characterization of the position of the Pockels cell in Bob's measurement setup, as well as the time that Bob's photon arrives at his Pockels cell. Therefore, we make the conservative assumption that, inside his lab, Bob's photon travels in a straight line in fiber between the source and his Pockels cell, which is shorter than the true path that Bob's photon actually takes. We then calculate the position of the spacetime point of the event E' (Fig. 3(c)), which would be equivalent to E if this assumption about the fiber inside Bob's lab were true, and we define Bob's lab to include the point E'. Defining Bob's lab in this way gives us a conservative *overestimate* of the size of Bob's lab (and thus ensures it contains the point E). We display E' and Bob's setting light cone at time $t_{E'}$ (i.e. the time of event E') in Fig. 3(c). The point E' is 31.6 ± 0.4 m from Bob's random number generators, so we define the trusted equipment in Bob's lab to include that final length of fiber.

To close the locality loophole, there are two conditions that need to be met. First, a light-speed signal carrying information about Alice's measurement setting k must not be able to reach Bob's lab before the time of the event E. Our definition of E' guarantees that E' occurs later in time than E. In addition, E' is closer to Alice's lab than any part of the fiber inside Bob's lab, and this guarantees that the distance between Alice's lab and E' is shorter than the distance between Alice's lab and E. As a result of these two facts, any signal carrying information about k needs to travel faster to influence E than it does to influence E', and this implies that if k and E' are space-like separated then k and E must be space-like separated as well. Alice's first random number generator fires 230 ± 2 ns before time $t_{E'}$, and we consider this the earliest moment that a light speed signal containing information about Alice's setting choice could begin propagating from her lab. Hence, at time $t_{E'}$, the light cone with information about Alice's setting k has traveled at most 52.1 ± 0.6 m from Alice's random number generators—and has not reached Bob's lab—ensuring that this first locality condition is satisfied (Fig. 3(c)).

The second condition that needs to be satisfied is that Alice's measurement outcome a must be recorded before the light cone carrying information about Bob's outcome b reaches Alice. We consider Alice's measurement complete (and her guess of Bob's outcome recorded) the moment the amplified electronic pulse from either of her detectors reaches her time tagger. We define the earliest possible time that Bob's measurement could be considered complete as the moment that his photon impinges on one of his SNSPDs. With these definitions, Alice records her measurement outcome at least 44.8 ± 1.3 ns before Bob completes his measurement, ensuring that the second locality condition is satisfied.

---

[4] In Fig. 3(b) we denote the labelling for Bob's measurement apparatus, where detections in the transmitted output port of the BD are labelled '+1' and detections in the displaced output port are labelled '-1'. In Alice's measurement apparatus, the labels are swapped.

V.2. Tomography on Bob's measurements and relevant bounds

To obtain rigorous EPR-steering inequalities we need to take into account that the measurements actually implemented by Bob are different from the ideal ones assumed for the analytical bounds obtained above. To account for this, we perform a parametric bootstrapping measurement tomography routine which provides us with estimates of the positive-operator valued measures (POVMs) representing each of the measurement settings implemented by Bob. To be conservative, we constrain our estimates of Bob's measurements to be both noiseless and projective (since this is the optimal scenario for a dishonest Alice[5].), meaning each POVM can be represented by the Bloch vector that represents the eigenstate corresponding to the +1 eigenvalue of the measurement operator. Our tomography method also provides estimates of the uncertainties on each of these Bloch vectors, in the form of an angular uncertainty on the direction that they point. We provide our best estimates of the Bloch vector representing each of Bob's measurement settings, as well as their uncertainties, in Table I. We illustrate these Bloch vectors and their uncertainties with cones in Fig. 4; the region enclosed by each cylinder represents the region that we expect—with five-sigma certainty—the actual Bloch vector representing that measurement setting to lie in. Full details of our measurement tomography technique are provided in Appendix C.

Table I. Results of tomography on Bob's measurements.

	Setting 1	Setting 2	Setting 3
X	−0.0502	0.9984	0.1019
Y	0.0419	0.0559	0.9944
Z	0.9978	−0.0089	−0.0276
Angular uncertainty, σ_j (rad)	0.0114	0.0114	0.0114

Using the above tomographic results, we recalculate Alice's optimal cheating ensembles based on Bob's real measurements, as delineated in Fig. 4 (a) and (b). For a completely rigorous EPR-steering test we must also include uncertainties in our characterization of Bob's actual settings, at the five-sigma level; see Appendix C.3 for more details. By using high-efficiency detectors, Alice's average heralding efficiency over all three settings — as defined at the end of Section III to be $\eta_A = \frac{1}{3} \sum_{k=1}^{3} \langle |a_k| \rangle$ — is measured to be $(74.8 \pm 0.1)\%$. For this efficiency level and for Bob's measurements (including errors), the gain factor r_{H_0} and the bound $h_{H_0}^n$ are reoptimized, yielding

$$r_0 = 0.4046, \quad h_0^3 = 0.2548,$$
$$r_1 = 0.5930, \quad h_1^3 = 0.2713. \tag{11}$$

---

[5] Alice and Fenella's strategy of controlling the state that is sent to Bob in order to influence his measurement outcomes will work best if Bob's measurements are both noiseless and projective. Adding noise to Bob's measurements only decreases how accurately his outcomes can be predicted, which in turn decreases Alice's and Fenella's probability of success [25]

We can now ask how good must our effective state preparation (i.e. purity) and propagation (i.e. loss avoidance) be if we wish to violate a communication-assisted-EPR-steering inequality. In our experiment, we created, as closely as possible, a singlet state at the source. After distribution to Alice and Bob, this is well approximated by a loss-depleted Werner state. In Fig. 5, we display, as a function of Alice's efficiency, the minimum required purity $\mu_{\min} = h_{H_0}^n/\eta_A + r_{H_0}$ for a loss-depleted Werner state to demonstrate EPR steering when Bob's measurement axes are ideal (black curves), and with the actual settings implemented in our experiment, allowing for uncertainties (red curves) without and with $H_0 = 1$ message assistance. The horizontal dashed lines give the predicted minimum required singlet-proportion parameters $\mu_{\min}$ for Alice's average heralding efficiency $\eta_A = (74.8 \pm 0.1)\%$ for each case. We predict that we require a minimum pure-fraction $\mu_{\min}$ of 0.745 and 0.956, to demonstrate EPR-steering with no message and $H_0 = 1$ bit of messages transferred, respectively. Achieving a state preparation purity well above 0.956 might not seem too onerous, since visibilities > 99.6% in the H/V and D/A bases have been demonstrated for Bell experiments [53]. However it must be emphasized that the figure of 0.956 relates to the *effective* state purity; imperfections in the polarization rotations by Alice's and Bob's Pockels cells are the dominant source of effective impurity in our experiment.

V.3. Results

Here we establish the correspondence between the EPR-steering inequality (1) and the event counts in the experiment. During each trial, Alice and Bob randomly choose among one of three measurement settings then write down their own re-
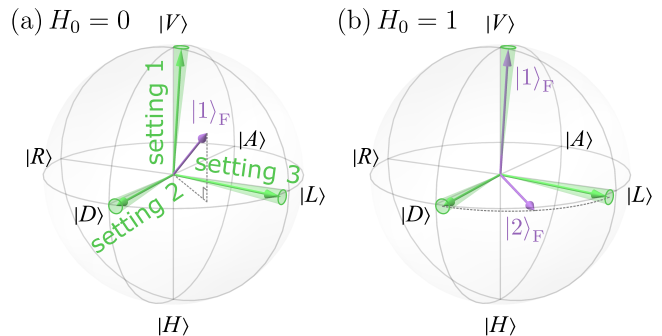

Figure 4. Characterization of Bob's three actual measurement directions (green arrows) and Alice's optimized cheating ensembles generated by Fenella {$|l\rangle_F$} (purple arrows) in the Bloch sphere. The green arrows represent our best estimate of Bob's actual measurement settings and with five-sigma certainty, the actual Bloch vectors representing each measurement setting lie inside the corresponding cones of uncertainty. (a) Alice's optimal cheating state $|1\rangle_F$ if no information is transferred ($H_0 = 0$); (b) Alice's optimal cheating states $|1\rangle_F$ and $|2\rangle_F$ if an $H_0 = 1$ bit message is transferred. Details of how we calculate Alice's optimal cheating ensembles are summarized in Appendix C.3.
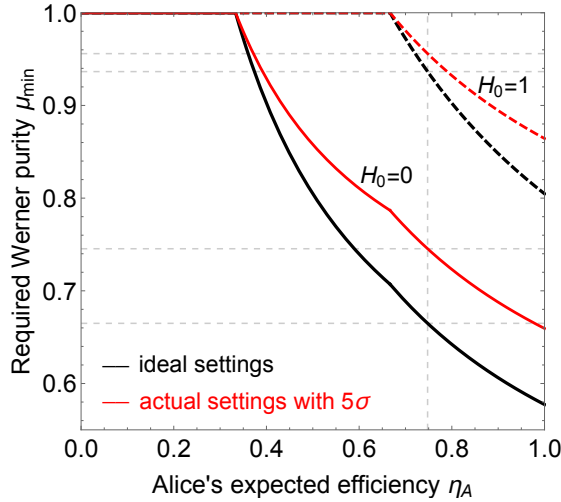
Figure 5. Loss-dependent EPR-steering bounds with one-way communication assistance. The solid curves are the theoretical bounds for required singlet-proportion to demonstrate EPR-steering without message assistance under different measurement conditions. The dashed curves are the required purity with $H_0 = 1$ bit of messages assistance. The vertical dashed line represents Alice's measured efficiency $\eta_A = (74.8 \pm 0.1)\%$.

sults. Alice (Bob) can record either a '+1' if she (he) observes a detection event on her (his) single H (V) detector or a '-1' if she (he) observes a detection event on her (his) single V (H) detector. In addition, Alice can record a '0' to represent no detection. In some experimental trials both of Alice's and/or Bob's detectors click[6], and Alice and Bob treat these trials differently. If both of Alice's detectors click, Alice records the outcome '0' for that trial. If both of Bob's detectors click, Bob randomly chooses to record either '+1' or '-1' as the outcome for that trial[7], a procedure introduced and justified in Ref. [25]. We use '$ab|kj$' to express the trial outcome, which indicates that Alice records outcome a in setting configuration k and Bob records outcome b in setting configuration j, and we use $N_{ab|kj}$ to denote the total number of times we obtain the outcome $ab|kj$.

Bob's task is now to use his and Alice's data to calculate $\tilde{S}_3(r_{H_0})$, which is his experimental estimate of the steering correlation function $S_3(r_{H_0})$, Eq (2). First, Bob discards data from all trials in which he and Alice measure in different bases (i.e. all trials for which $k \neq j$), as well as from all trials in which he does not detect a photon. Because Bob trusts his characterization of his measurement apparatus, and therefore his characterization of his measurement efficiencies, he can normalize the data based on these efficiencies to estimate what

data would have been collected if his measurement apparatus worked with unit efficiency. Bob estimates the normalized data by computing $\mathcal{N}_{ab|jj} = N_{ab|jj}/\eta_{B(b|j)}$. Here $\eta_{B(b|j)}$ is the efficiency with which Bob's outcome b detector will click when he measures in setting j — it includes the transmission efficiency of the path photons take from the edge of Bob's lab to his polarizing beam displacer, the transmission efficiency from the output of the beam displacer to the outcome-b detector, and that detector's internal efficiency. We find that these efficiencies depend on both the outcome and the measurement setting.

To characterize these efficiencies we modify Klyshko's method [45] (which applies to systems with two photons and two detectors) to apply to our two-photon, four-detector setup. We apply this method to data that is independent from, but acquired in parallel with, the data used to test the steering inequality (1). A full description of our method is provided in Appendix C.4. After normalizing the data in this way, Bob calculates the normalized total number of trials in which both he and Alice measure in setting j, given by $\mathcal{N}_{jj} = \mathcal{N}_{+1+1|jj} + \mathcal{N}_{-1+1|jj} + \mathcal{N}_{0+1|jj} + \mathcal{N}_{+1-1|jj} + \mathcal{N}_{-1-1|jj} + \mathcal{N}_{0-1|jj}$. Now, Bob is ready to calculate $\tilde{E}_j^c$, which is his estimate of $\langle a_j \hat{B}_j \rangle$:

$$\tilde{E}_j^c = \frac{\mathcal{N}_{+1+1|jj} + \mathcal{N}_{-1-1|jj} - \mathcal{N}_{-1+1|jj} - \mathcal{N}_{+1-1|jj}}{\mathcal{N}_{jj}}, \quad (12)$$

and he can also calculate $\tilde{E}_j^a$, which is his estimate of $|a_j|$:

$$\tilde{E}_j^a = \frac{\mathcal{N}_{+1+1|jj} + \mathcal{N}_{+1-1|jj} + \mathcal{N}_{-1+1|jj} + \mathcal{N}_{-1-1|jj}}{\mathcal{N}_{jj}}. \quad (13)$$

Finally, Bob's estimate $\tilde{S}_3(r_{H_0})$, and the experimental EPR-steering inequality, are given by:

$$\tilde{S}_3(r_{H_0}) = \frac{1}{3} \sum_{j=1}^{3} \left(\tilde{E}_j^c - r_{H_0} \tilde{E}_j^a \right) \leq h_{H_0}^3. \quad (14)$$

Our data were taken in 90 one-minute chunks, and over the total dataset there are 936,848 experimental trials in which at least one of Bob's detectors clicks. By using the parameters in Eq. (11), the experimental residuals between the left-hand-side and right-hand-side of the above inequality (14) are $R_0^3 = 0.1646 \pm 0.0003$ and $R_1^3 = 0.0071 \pm 0.0003$. The uncertainties here represent the statistical uncertainty caused by Poissonian counting statistics. These two violations are 478 and 25 standard deviations above zero, respectively. While we violate both inequalities by a very large number of standard deviations, an analysis method that closes the memory loophole could reduce the apparent statistical significance of this result. However, it is important to remember that the derivation of these inequalities includes our estimates of the uncertainty on our characterization of Bob's measurement settings up to the five-sigma level. Thus, our confidence in these two inequality violations is also at the five-sigma level (and not 478 or 25 sigma). This significant violation of the $H_0 = 1$ inequality conclusively demonstrates EPR-steering correlations that require more than one bit of FTL communication to simulate classically.

---

[6] Of all trials in which Bob records at least one click, Alice measures two clicks 0.26% of the time, and Bob measures two clicks 0.24% of the time.

[7] Intuitively, this can be understood as Bob randomly choosing one of the two photons he received to be 'the' photon for that trial, and this is consistent with the outcome Bob reports when two photons are received by a single detector (which may also happen), since in that case both photons have the same outcome.

VI. DISCUSSION

Our experiment implies that if the mechanism of accounting for the quantum correlations is steering (that is, if Alice's measurement induces wave-function collapse in Bob's lab) then that mechanism requires more than one bit of information to be transmitted, faster than light, from Alice's to Bob's lab. If one is willing to accept that information is in fact being transmitted faster than light, it is interesting to ask exactly how fast this information needs to travel. We calculate the speed at which a signal would need to travel if Alice sent it from her lab the moment her setting choice k was chosen and it was received by Fenella at the spacetime point E' before Bob's photon enters his lab — this is a lower bound on the speed a signal would need to travel between Alice's QRNGs and the spacetime point E. (See Sec. V.1 for the distinction between these two events.) The point E' is 161.3 ± 1.5 m from Alice's random number generators, and Alice's first random number generator fires 230 ± 2 ns before time $t_{E'}$. Thus, this hypothetical FTL signal would have to travel at least $(9.84 \pm 0.14) \times 10^8$ m/s, or 3.28 ± 0.05 times the speed of light[8].

In order to fairly evaluate the strength of the nonlocality between Alice and Bob in our experiment, we compare R_0^3 and R_1^3 with their corresponding Tsirelson bounds T_0^3 and T_1^3. Like the maximum quantum violation of the Clauser-Horne-Shimony-Holt Bell inequality $2\sqrt{2}$ [60], the Tsirelson bounds represent the maximum allowable values according to quantum theory, which can be achieved with a Bell state ($\mu = 1$) and ideal projective measurements with perfect detection efficiency at Alice ($\eta_A = 1$). Recall that if $r_{H_0} = 0$, the bounds for the $H_0 = 0$ and $H_0 = 1$ inequalities are $h_0^3 = \sqrt{3}/3$ and $h_1^3 = (1+\sqrt{2})/3$ respectively. Thus, the maximum possible violations allowed by quantum mechanics are $T_0^3 = (3 - \sqrt{3})/3$ and $T_1^3 = (2 - \sqrt{2})/3$. Comparing our results to the maximum violations, we see that $R_0^3/T_0^3 \approx 0.389$ and $R_1^3/T_1^3 \approx 0.037$, respectively. The smallness of the latter number reflects the difficulty of demonstrating the violation of communication-assisted EPR-steering, limited mainly by Alice's total photon detection efficiency being only ~75%.

The conceptualization of EPR-steering as a quantum information task—to verify entanglement in the partial absence of trust—was key to its modern formulation [5] and its generalization to multi-party networks [61]. Pioneering works showed theoretically [57] and experimentally [26, 31, 62] that, through the use of multiple measurement settings, EPR-steering could be robust to loss even in the presence of some noise. Here, by working to minimize both noise and loss, we have been able to show the robustness of EPR steering

to (hypothetical) FTL communication. Our communication-assisted and loss-tolerant EPR-steering inequalities may have applications in secure entanglement distribution with untrusted parties. In particular, in the situation where the parties are not space-like separated, adversarial communication need not be FTL and so need not be hypothetical. To achieve the greatest security in such schemes, it will also be important to analyze the data in a way that does not assume experimental trials are independent. Otherwise, a malicious actor could exploit the memory loophole and compromise the security of the protocol.

It is natural to ask if one can demonstrate EPR-steering correlations requiring an even larger amount of FTL communication to simulate classically. This is certainly possible in principle. In general, if one is performing measurements on an entangled state of two qubits, one needs to perform an experiment with at least $2^{H_0} + 1$ choices of measurement setting in order to exclude FTL messages of up to H_0 bits. As the number of measurement directions increases, some measurements necessarily become closer to each other and therefore the minimum visibility required to demonstrate an inequality violation increases. In addition, for a large enough number of settings, each measurement station will require multiple Pockels cells (or some other method of quickly switching between settings). For the specific case of ruling out messages of size $H_0 = 2$, five measurement settings are needed. One could choose these settings to have Bloch vectors corresponding to five of the vertices of a dodecahedron (similar to how in our three setting experiment we chose settings with Bloch vectors corresponding to three of the vertices of an octahedron). With the proper driving electronics, one could in principle use a single Pockels cell to implement fast switching between these five settings. A more significant challenge for ruling out messages of size $H_0 = 2$ bits is that the minimum efficiency required for such an experiment is 4/5, much higher than the 2/3 threshold needed for $H_0 = 1$ bits, and at the limit of what has currently been demonstrated in loophole-free experiments with entangled photons [64]. Finally, there are many open questions about the role of communication in multipartite or high-dimensional EPR-steering scenarios, where many different communication patterns can be considered.

---

[8] This figure is a lower bound on the speed of the signal in the rest frame of the building containing the two laboratories. In a different reference frame, a different value would be inferred and, because the speed is superluminal, it could be anything from just above the speed of light to beyond infinite (backwards-in-time information transmission). This can be contrasted with the experiment of Salart *et al.* [63], which gave, under certain assumptions, a lower bound on the speed v of superluminal information transmission in a Bell test, regardless of the preferred frame.

Appendix A: The derivations of Eqs. (7), (8) and (9)

To derive the expression of Eq. (7) in the main text, we first rewrite the bound h_0^3 in this scenario. When there is no message transferred, only one state $|1\rangle_F$ can be generated by Fenella for all three settings, and the deterministic function ℓ in Eq. (3) is also fixed from $\{1, 2, 3\}$ to $\{1\}$, thus

$$h_0^3 = \max_\alpha \frac{1}{3} \sum_{j=1}^{3} \delta_{a,\alpha(j)} \left(-r_0|a| + a\langle 1|_F \hat{B}_j |1\rangle_F \right). \quad (A1)$$

where α still indicates a deterministic function from $\{1, 2, 3\}$ to $\{+1, 0, -1\}$. It has been proven that the optimal deterministic bounds related to the second term in above equation is [57]

$$g_0^3 = \max_\alpha \left[\lambda_{\max} \left(\frac{1}{3} \sum_{j=1}^{3} \alpha(j) \hat{B}_j \right) \right], \quad (A2)$$

where $\lambda_{\max}$ denotes the maximum eigenvalue of the operator within the following brackets. Then, for a lossy maximally entangled singlet state, the optimal r_0 in the first term should be chosen by

$$\arg\min_{r_0} \left[\max_\alpha \left(g_0^3 - r_0 \frac{m}{3} \right) - \eta_A(\mu - r_0) \right]. \quad (A3)$$

where m is the total number of events when $\alpha(j) \neq 0$. For example, if Alice reports null for two of the three settings, *e.g.*, $\alpha(1) = \alpha(2) = 0$, it means that $m = 1$. Thus, for a given η_A, the parameter r_0 and the bound h_0^3 can be optimized by searching all possible cases numerically. It is reasonable that several forms of $\alpha(j)$ can give the optimized bound. For example, when $2/3 \leq \eta_A < 1$, both $\alpha(1) = 0, \alpha(2) = \alpha(3) = +1$ and $\alpha(1) = -1, \alpha(2) = +1, \alpha(3) = 0$ can get the expressions in Eq. (7).

When we introduce the classical messages into the above scenario, the derivation will become more complicated. When $H_0 = 1$, there are two states $|1\rangle_F$ and $|2\rangle_F$ for three settings to choose. Now the bound is written as

$$h_1^3 = \max_{\alpha, \ell, \{|l\rangle_F\}_{l=1}^2} \frac{1}{3} \sum_{j=1}^{3} \delta_{a,\alpha(j)} \left(-r_1|a| + \sum_{l=1}^{2} \delta_{l,\ell(j)} a\langle l|_F \hat{B}_j |l\rangle_F \right), \quad (A4)$$

where both ℓ and α are no longer fixed. Then the optimal deterministic bounds related to the second term in the above equation is

$$g_1^3 = \max_{\alpha, \ell} \left[\frac{1}{3} \sum_{l=1}^{2} \lambda_{\max} \left(\sum_{j=1}^{3} \delta_{l,\ell(j)} \alpha(j) \hat{B}_j \right) \right], \quad (A5)$$

The value of g_1^3 depends on both Alice's result $\alpha(j)$ and the strategy of Fenella $\ell(j)$.

First, let us consider an idealized scenario where Alice always declares non-null results, i.e., $\alpha(j) = \pm 1$. Similarly, we also need to search all possible forms for $\ell(j)$ to find the optimized r^1 by minimizing Eq. (A3). Suppose that $\ell(1) = 1, \ell(2) = \ell(3) = 2$, i.e., for $j = 1, 2, 3$ measurements, the states prepared by Fenella are $|1\rangle_F, |2\rangle_F, |2\rangle_F$, respectively, then we have

$$g_1^3 = \max_\alpha \left\{ \frac{1}{3} \left[\lambda_{\max}(\alpha(1) \hat{B}_1) + \lambda_{\max}(\alpha(2) \hat{B}_2 + \alpha(3) \hat{B}_3) \right] \right\}. \quad (A6)$$

By substituting $\alpha(j) = +1$ for all j into this case, the optimal $r_1 = 0$ and the maximum value of h_1^3 reaches $(1 + \sqrt{2})/3$. There are likely other combinations with different choices of $\ell(j)$ and $\alpha(j)$ that can reach the same bound in Eq. (8).

Then, taking loss into account, m in Eq. (A3) will change related to the function of $\alpha(j)$. We can still search all possible cases numerically to find the bound. Here we give one solution for the $2/3 \leq \eta_A < 1$ case. By keeping $\ell(1) = 1, \ell(2) = \ell(3) = 2$ and substituting $\alpha(1) = 0, \alpha(2) = \alpha(3) = +1$, we can get the optimal expressions in Eq. (9).

Appendix B: Operation of Pockels cell

We can obtain the largest violation of the steering inequality if Alice and Bob each choose three measurement settings that correspond to projective measurements in three mutually unbiased bases. Closing the locality loophole requires switching between these settings quickly. This appendix describes how we use a combination of waveplates, a single Pockels cell (PC), and a polarizing beam displacer (BD) to accomplish this.

First we define the target measurement operators we aim to measure in the experiment (at both Alice and Bob) as $\hat{\sigma}_z$, $\hat{\sigma}_x$, and $\hat{\sigma}_y$ for settings 1, 2, and 3, respectively. Here, $\hat{\sigma}_z$, $\hat{\sigma}_x$, and $\hat{\sigma}_y$ are the Pauli matrices, which are defined in the z-basis, $\{|0\rangle, |1\rangle\}$, as:

$$\hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad (B1)$$
$$\hat{\sigma}_x = |+\rangle\langle +| - |-\rangle\langle -|, \quad (B2)$$
$$\hat{\sigma}_y = |+i\rangle\langle +i| - |-i\rangle\langle -i|, \quad (B3)$$

where we use the convention $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and $|\pm i\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$. Using the Bloch representation, we display the '+' eigenstates of the three measurement settings in Fig. B1.

The states in our experiment are encoded in polarization, and thus we must map the Pauli operators to measurements on the polarization degree of freedom. We map the z-basis
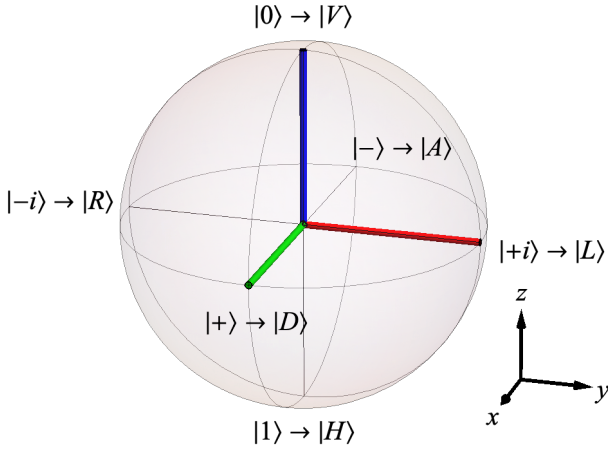
Figure B1. Bloch sphere representation of target measurements in our experiment. The blue, green, and red Bloch vectors represent the '+' eigenstates of the target measurements for settings 1, 2, and 3 respectively. The figure also denotes how we map the Bloch sphere to polarization states. As a result of our choice of coordinate system, vertically polarized light $|V\rangle$ is represented by the Bloch vector $(0, 0, 1)$, diagonally polarized light has Bloch vector $(1, 0, 0)$, and left-circularly polarized light has Bloch vector $(0, 1, 0)$.

to horizontal ($|H\rangle$) and vertical ($|V\rangle$) polarization with the assignments $|0\rangle \to |V\rangle$ and $|1\rangle \to |H\rangle$. As a result, the x-basis is mapped to the diagonal and anti-diagonal polarization basis via $|+\rangle \to |D\rangle$ and $|-\rangle \to |A\rangle$, and the y-basis is mapped to the left- and right-circular polarizations via $|+i\rangle \to |L\rangle$ and $|-i\rangle \to |R\rangle$.

To realize the target measurements, we need a scheme for which the overall effect of Bob's measurement is that, for setting $j = 1$, polarization state $|V\rangle$ remains $|V\rangle$, for $j = 2$, $|D\rangle$ is rotated to $|V\rangle$, and for $j = 3$, $|L\rangle$ becomes $|V\rangle$. It is easiest to describe our scheme if we consider how a vertically polarized beam is affected as it travels *backwards* through the optics in Bob's measurement setup, beginning from the '+' detector. For the j-th measurement setting, vertically polarized light will be transformed to some other polarization state as it propagates backwards, and the polarization state that the beam ends up in is the '+' eigenstate of this j-th measurement.

We represent the polarization of the light beam at various stages of backwards-propagation in Fig. B2. Initially, the light transmitted by the beam displacer is vertically polarized (Fig. B2(a)). Next the beam propagates backwards through two waveplates, and it is rotated to an elliptical polarization state (Fig. B2(c)). Then the beam encounters the PC. Ideally, when no voltage is applied to the PC, it has no birefringence and thus has no effect on polarization. The PC becomes birefringent when a voltage is applied, and it is aligned such that it induces a relative phase θ between the $|D\rangle$ and $|A\rangle$ polarization states. In our setup, setting $j = 1$ corresponds to no voltage applied to the PC, $j = 2$ corresponds to a positive voltage which applies a phase of $+\theta$, and $j = 3$ corresponds to a negative voltage that applies a phase $-\theta$. We choose $\theta = 120°$, which ensures the three different polarization states exiting the PC are eigenstates of three different mutually unbiased measurement operators. Finally, the light backwards-propagates

through the final QWP and HWP and we can see that the three resulting Bloch vectors (Fig. B2(f)) are equal to the Bloch vectors of the '+' eigenstates of the three target measurement operators (Fig. B2).

The above scheme works for an idealized PC. In reality, the PC's we used in Alice and Bob's measurement stations exhibited a small amount of birefringence even when no voltage was applied to them. To correct for this, we performed quantum process tomography to characterize the Jones matrix of each PC when no voltage was applied. Instead of using the idealized waveplate angles given in the caption of Fig. B2, we used a numerical optimization to find the optimal angles for each measurement station.

Appendix C: Quantum characterization of Bob's measurements

In order to perform the steering test, we need an accurate quantum description of the measurements performed in Bob's laboratory for each of the three settings, and we detail how we find this description in this appendix. We begin by defining the family of quantum models we use to represent Bob's measurement apparatus in subsection C.1 of this appendix. We perform two types of measurements in order to learn a model that accurately represents Bob's apparatus. First, we characterize the polarization-rotation optics in Bob's measurement setup by performing a bootstrapped version of standard quantum measurement tomography with parametric resampling. We explain our tomography procedure in subsection C.2, and we explain how we modify the steering inequalities to deal with imperfections in Bob's measurement settings in subsection C.3. Second, we measure the path efficiencies for each of Bob's detection outcomes (and track these efficiencies in real time) using a modified version of Klyshko's efficiency measurement, which we detail in subsection C.4.

C.1. Quantum model of Bob's measurement apparatus

A simplified diagram of our experiment is given in Fig. C1(a). Bob's measurement setup consists of two waveplates, followed by a PC, followed by two more waveplates, and a polarizing beam displacer (BD) with a detector in each output port. We model the combined effect of the waveplates and PC as a polarization rotation that depends on Bob's choice of measurement setting, j. The BD splits the beam into two paths; vertically polarized light is directly transmitted into the '+' path, and horizontally polarized light is deflected into the '-' path. Bob records a '+' outcome if the detector in the '+' path clicks, a '-' outcome if the other detector clicks, and a null outcome if neither detector clicks. To account for non-unit detection efficiency we model the detectors as ideal detectors preceded by non-polarizing beamsplitters; the transmission probability of the beamsplitter in the '±' path is given by $\beta_j^{(\pm)}$.

In our model of Bob's measurements, we will make two main assumptions. First, we assume that in Bob's measurement apparatus there is no polarization-dependent loss before the BD; this allows us to model all the loss in the experiment

(a) **(b)** **(c)**

(d) **(e)** **(f)**

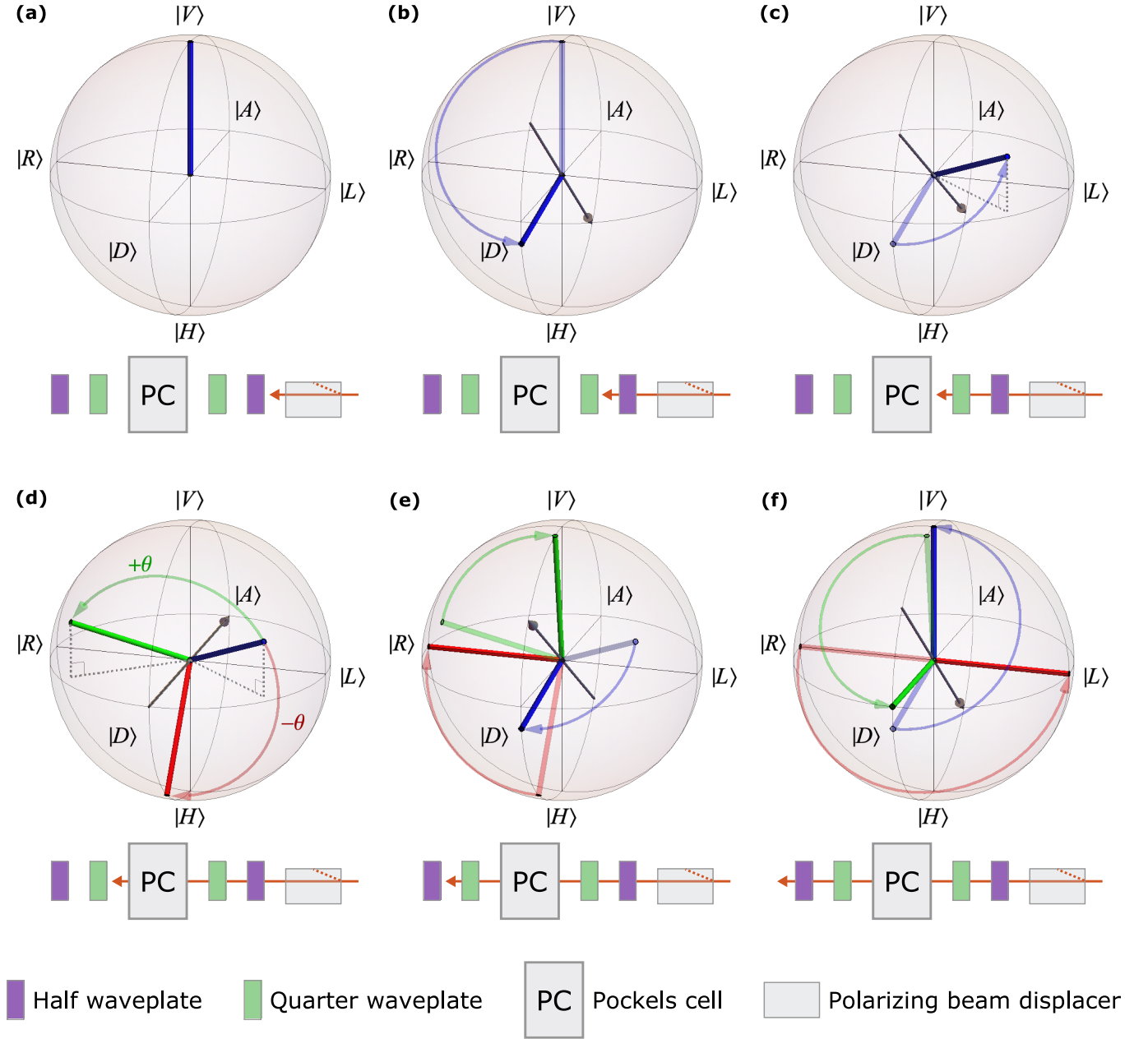Half waveplate   Quarter waveplate   PC   Pockels cell   Polarizing beam displacer

Figure B2. Visualization of polarization rotation light experiences as it propagates backwards through Bob's measurement setup. In each subfigure we plot the beam's polarization state on the Bloch sphere, after it has propagated backwards through part of Bob's setup. (a) After propagating through the beam displacer, the beam is vertically polarized, as represented by the blue Bloch vector. Next the polarization state is rotated by (b) a HWP with fast-axis −24.94° from horizontal and (c) a QWP at an angle of −27.38°. The grey vector in the linear-polarization plane of the Bloch sphere indicates the angle of the waveplate; the action of a half (quarter) waveplate is to rotate the Bloch vector by an angle of 180° (90°) about this grey vector. Since we are considering *backwards* propagation, the angle of rotation is determined by the *left*-hand rule (the Bloch vector rotates in the direction of the curl of the fingers of the left hand when the left thumb points in a direction parallel to the grey vector). (d) When the beam travels through the PC, one of three rotations are applied to the beam, depending on the measurement setting. We either apply no rotation to the Bloch vector ($j = 0$, blue Bloch vector), apply a rotation of $+\theta$ ($j = 1$, green Bloch vector), or apply a $-\theta$ rotation ($j = 2$, red vector). Each of the three resulting Bloch vectors is orthogonal to the other two, indicating that they each correspond to the '+' eigenstate of one of a set of three mutually unbiased bases. Finally the beam travels through (e) a QWP at 62.62° and (f) a HWP at −24.94°. The overall effect of backwards propagation through Bob's measurement is that, if $j = 0$, the polarization state $|V\rangle$ remains $|V\rangle$, if $j = 1$, $|V\rangle$ is rotated to $|D\rangle$, and when $j = 2$, $|V\rangle$ becomes $|L\rangle$.
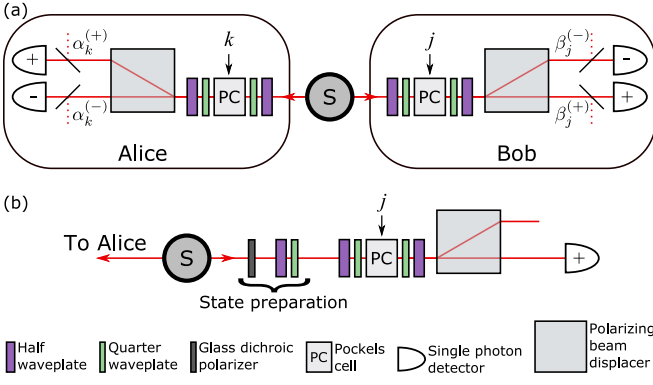
Figure C1. Simplified diagrams of the experimental setups used for (a) testing the steering inequality and (b) performing tomography on Bob's measurement apparatus. (a) A source emits a pair of photons; one travels to Alice and one to Bob. At each measurement station, photons propagate through polarization-rotating optics, including a PC with an applied voltage that depends on the measurement settings. In this simplified model, we assume there is no polarization-dependent loss before the light encounters a polarizing beam displacer (BD). We model loss in the experiment by first modeling all optics and detectors with perfect efficiency, and introducing non-polarizing beamsplitters with setting-dependent transmission probabilities $\alpha_k^{(+)}$, $\alpha_k^{(-)}$, $\beta_j^{(+)}$, and $\beta_j^{(-)}$. (b) We add a polarizer and two waveplates at Bob's station to perform the measurements required for quantum measurement tomography.

as occurring after the BD. When characterizing Bob's measurement apparatus, we find no significant evidence of polarization dependent loss in the waveplates and Pockels cell. Second, we assume that Bob's measurements are lossy versions of projective measurements. We measure an excellent extinction ratio greater than 10^5:1 with the beam displacer, providing evidence in favour of this assumption. We note that this second assumption is conservative, because if it were the case that Alice and Bob *do not* share an entangled state, then the bound of the steering inequality can only be achieved if $E_j^{(\pm)}$ are projectors.

In quantum theory, the positive-operator valued measure (POVM) representing Bob's three-outcome measurement for setting j, is:

$$M_j^B = \left\{ \beta_j^{(+)} \hat{E}_j^{(+)}, \beta_j^{(-)} \hat{E}_j^{(-)}, \hat{E}_j^{(\text{null})} \right\}. \quad \text{(C1)}$$

The measurement effects satisfy $\beta_j^{(+)} \hat{E}_j^{(+)} + \beta_j^{(-)} \hat{E}_j^{(-)} + \hat{E}_j^{(\text{null})} = \mathbb{I}$, where $\mathbb{I}$ is the 2×2 identity matrix. For a measurement of a single photon with polarization state ρ, the probability that detector $\pm$ clicks is $P(\pm|\rho, j) = \beta_j^{(\pm)} \text{Tr}\left(\rho \hat{E}_j^{(\pm)} \right)$, and the probability that neither detector clicks is $P(\text{null}|\rho, j) = \text{Tr}\left(\rho \hat{E}_j^{(\text{null})} \right)$. We represent the measurement effect operators $\hat{E}_j^{(\pm)}$ as:

$$E_j^{(\pm)} = \frac{1}{2} \left(e_{j,0}^{(\pm)} \mathbb{I} + e_j^{(\pm)} \cdot \boldsymbol{\sigma} \right), \quad \text{(C2)}$$

where $\boldsymbol{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ is a vector of Pauli matrices, and $e_j^{(\pm)} = \left(e_{j,1}^{(\pm)}, e_{j,2}^{(\pm)}, e_{j,3}^{(\pm)} \right)$ is the Bloch vector consisting of three real numbers that specify the basis of the measurement. Because of our assumption that Bob's measurements are lossy

projective measurements, we set $e_{j,0} = 1$, and the Bloch vectors $e_j^{(\pm)}$ must satisfy $\left| e_j^{(\pm)} \right| = 1$. Finally, since in our model we assume there is no polarization-dependent loss before the BD, the measurement effects must satisfy $\hat{E}_j^{(+)} + \hat{E}_j^{(-)} = \mathbb{I}$, implying $e_j^{(+)} = -e_j^{(-)}$. Thus, in order to fully specify measurement M_j^B it is sufficient to know the Bloch vector $e_j^{(+)}$, and the two efficiencies $\beta_j^{(+)}$ and $\beta_j^{(-)}$.

C.2. Bootstrapped measurement tomography with parametric resampling

We slightly modify our setup to perform tomography on Bob's measurement device (see Fig. C1(b)). We tune the source to produce the separable state $|VH\rangle$, so that a vertically polarized photon is sent to Alice's measurement station, and a horizontally polarized one to Bob's. We place a horizontal polarizer in Bob's measurement bridge, followed by a half and a quarter waveplate which we can rotate to prepare one of the six polarization states $|H\rangle$, $|V\rangle$, $|D\rangle$, $|A\rangle$, $|R\rangle$, or $|L\rangle$. We turn Alice's PC off and rotate her waveplates to maximize the rate of detections in her '+' detector. We set Bob's PC to randomly switch between his three measurement settings, and acquire data for one minute. The PC runs at a rate of 100 kHz, such that Bob performs roughly 3.3×10^4 measurement trials per second, per measurement setting. We acquire data over a total of 120 one-minute intervals, and each minute we set the input state to one of the above six polarization states. For each minute of data we record the input polarization state, the number of coincident detections recorded between Alice's and Bob's '+' detectors while Bob's PC was set to each of the three measurement settings, as well as the number of trials the PC was set to each setting. For each setting, the final dataset is a list of 120 measured pairs of input states and coincidence count rates.

We analyze the data for each setting independently, using a bootstrap method with parametric resampling. For each measurement setting j we perform 10,000 analysis trials. Each analysis trial begins by sampling, with replacement, 120 pairs of input states and coincidence rates; this is the bootstrapping. Next, to account for statistical errors in the count rates we measure, as well as systematic errors caused by imperfections in the waveplates used to prepare the input states, we perform a parametric resampling step. We resample each measured count rate from a Poisson distribution with mean equal to the raw count rate that we measured. Based on the manufacturer's specified tolerance, we assume that the phase of each state-preparation waveplate is accurate to within ±0.005 wavelengths, and we choose an error value for each waveplate from a uniform distribution within this range. We also assume that we can find the zero angle for each waveplate to within an accuracy of 0.1 degrees, and we choose an error value from a normal distribution with standard deviation 0.1 for each waveplate. Then we recalculate the six input states based on the randomly assigned waveplate imperfections. After both the bootstrapping and parametric resampling steps we are left with a new list of data consisting of 120 (imperfect) input states and resampled coincidence count rates.

Finally, using this new list of data we find $\tilde{e}_j^{(+)}$, the maximum-likelihood estimate of $e_j^{(+)}$. Because we are making the conservative assumption that the $E_j^{(\pm)}$ are projectors, we enforce the constraint $\left|\tilde{e}_j^{(+)}\right| = 1$.

After performing all analysis trials, we are left with a list of 10,000 normalized maximum likelihood estimates of $e_j^{(+)}$. The mean of this list, $\langle \tilde{e}_j^{(+)} \rangle$, is our final estimate of $e_j^{(+)}$. When the 10,000 max-likelihood estimates are plotted, their endpoints lie on the surface of the Bloch sphere, and they are distributed around $\langle \tilde{e}_j^{(+)} \rangle$ in the shape of a 2-D Gaussian ellipse. We calculate σ_j, the standard deviation of angular displacements from the mean vector in the direction along the semi-major axis of this ellipse. The results of the measurement tomography analysis are summarized in Table I in the main text. When testing the steering inequality, we assume that the angle between the true Bloch vector describing Bob's measurement, $e_j^{(+)}$, and our best-guess vector $\langle \tilde{e}_j^{(+)} \rangle$ is no larger than $5\sigma_j$.

C.3. Dealing with imperfections in Bob's measurements

In order for Bob to make the most conservative characterization of his measurements, we make Bob's measurement Bloch vectors as close as possible after the rotations by corresponding angles with 5σ uncertainties. The idea is that if Bob's settings were all closer together (i.e. closer to a "common" direction than he thought they were), then Alice could convince him that she was steering him even if she wasn't. At the extreme, if Bob's Pockels cell stopped working so that he was always measuring the same direction, then Alice could, in each run, simply dial up which result ($+1$ or -1) she wanted Bob to get by having Fenella send a photon aligned or anti-aligned with this direction. She could get any correlation she desired with no entanglement at all. This is why we consider the worst case for Bob's directions, when they are all closer to a common direction.

For the no message case, Alice/Fenella's best strategy is to prepare a state aligned/anti-aligned with the mean (normalized) direction $\vec{m}$ of Bob's three measurement settings, then send it to Bob. We can flip the measurement directions $\vec{B}_j$ (e.g., from $\vec{B}_1$ to its opposite $-\vec{B}_1$) and the results (from $+1$ to -1) without changing anything physically. Depending on the sign of each $\vec{B}_j$ we choose, there are 8 different combinations and their respective normalized mean vectors. We need to find, out of these 8 groups, the one that will have the largest inner product between their own means and each of the $\vec{B}_j$ that go into it–that is $\{-\vec{B}_1, \vec{B}_2, \vec{B}_3\}$. Having identified this, we can rotate each $\{-\vec{B}_1, \vec{B}_2, \vec{B}_3\}$ towards $\vec{m} = (-\vec{B}_1 + \vec{B}_2 + \vec{B}_3)/|-\vec{B}_1 + \vec{B}_2 + \vec{B}_3|$ by their respective uncertainties ($5\sigma_j$) through the Rodrigues formula.

When an $H_0 = 1$ bit message is transferred, Alice has $d = 2$ level of information that can encode something about the measurement setting Bob will choose. For instance, we assume that Alice's first level of information indicates that Bob will measure along $\vec{B}_3$, and the second level message means that Bob will measure either $\vec{B}_1$ or $\vec{B}_2$ (each with 50% probability). In this situation, Alice's optimal cheating strategy would

Table CI. Conservative estimate of Bob's measurements after rotations by 5σ uncertainties.

		Setting 1	Setting 2	Setting 3
$H_0 = 0$	X	0.0913	0.9942	0.1424
	Y	-0.0024	0.0943	0.9875
	Z	-0.9958	-0.0508	-0.0671
$H_0 = 1$	X		0.9936	0.1584
	Y		0.1127	0.9869
	Z		-0.0104	-0.0278

be to ask Fenella to prepare the state $|1\rangle_F$ aligned with $\vec{B}_3$ if her message is on the first level, and to prepare the state $|2\rangle_F$ along $\vec{m} = (\vec{B}_1 + \vec{B}_2)/|\vec{B}_1 + \vec{B}_2|$ if her message is on the second level. Now, to avoid Alice's optimal cheating strategy, Bob should find the closest two of the three settings among 12 different combinations of $\{\pm\vec{B}_u, \pm\vec{B}_v\}$ (u, $v \in \{1, 2, 3\}$). After checking their respective inner products, we find $\vec{B}_2$ and $\vec{B}_3$ are closest to each other, then rotate them towards their mean vector $(\vec{B}_2 + \vec{B}_3)/|\vec{B}_2 + \vec{B}_3|$ (each by $5\sigma_j$). Note that in this case it doesn't matter how much Bob rotates his $\vec{B}_1$ by, because Alice/Fenella can always prepare a state that perfectly aligns with $\vec{B}_1$. The results of Bob's measurement settings after rotation are summarized in Table CI. The optimal parameter r and bounds $h_n^{H_0}$ are obtained based on these rotated settings.

C.4. Measurement of Bob's detection efficiencies

David Klyshko invented a widely used method to estimate overall detection efficiency in experiments with two correlated photons that each travel to one detector [45]. In our experiment each photon can travel to one of *two* detectors, and we modify Klyshko's method to allow us to estimate the overall detection for efficiency for each of the four detectors in this type of setup.

C.4.a. Modified Klyshko equations

To explain our method, we refer again to the simplified model of our experiment in Fig. C1(a). Our goal is to characterize the efficiencies $\beta_j^{(+)}$ and $\beta_j^{(-)}$. We do this by first deriving equations for the count rates Alice and Bob expect to observe. For now, we will assume that any loss in the setup affects both polarizations of light equally. We will revisit this assumption later. Each photon pair emitted by the source may take one of four paths towards the detectors at Alice's and Bob's measurement stations; Alice's and Bob's photons might each travel towards a '+' detector, Alice's can travel to her '+' detector and Bob's to his '-' detector, Alice's could travel to her '-' detector and Bob's to his '+' detector, or finally both photons can travel towards the '-' detectors. Each of these four paths has an associated joint probability represented by $p_{k.j}^{a,b}$, with $a, b \in \{+, -\}$, representing Alice's and Bob's outcomes,

and $k, j \in \{1, 2, 3\}$ representing their settings.[9]

Using these joint probabilities, we can write down the expected detection rates if Alice and Bob measure light from a source that produces N pairs each second. We assume that the source emits at most one pair per coincidence window, and we also assume that the detectors have negligibly small rates of background counts. We will also revisit these two assumptions later. With this, the expected rates of coincident detections $C_{k,j}^{(a,b)}$ between detector a at Alice and detector b at Bob are:

$$C_{k,j}^{(a,b)} = N\alpha_k^{(a)}\beta_j^{(b)} p_{k,j}^{(a,b)}. \tag{C3}$$

The expected rates of single detections at detector a at Alice, $A_k^{(a)}$, are:

$$A_k^{(a)} = N\alpha_k^{(a)} \left(p_{k,j}^{(a,+)} + p_{k,j}^{(a,-)} \right). \tag{C4}$$

Finally, the expected rates of single detections at detector b at Bob, $B_j^{(b)}$, are:

$$B_j^{(b)} = N\beta_j^{(b)} \left(p_{k,j}^{(+,b)} + p_{k,j}^{(-,b)} \right). \tag{C5}$$

Rearranging Eqs. (C3)-(C5) we find that Alice's efficiency $\alpha_k^{(a)}$ is given by:

$$\alpha_k^{(a)} = \frac{C_{k,j}^{(a,\bar{a})} C_{k,j}^{(\bar{a},a)} - C_{k,j}^{(a,a)} C_{k,j}^{(\bar{a},\bar{a})}}{C_{k,j}^{(\bar{a},a)} B_j^{(\bar{a})} - C_{k,j}^{(\bar{a},\bar{a})} B_j^{(a)}}, \tag{C6}$$

and Bob's efficiency $\beta_j^{(b)}$ by:

$$\beta_j^{(b)} = \frac{C_{k,j}^{(b,\bar{b})} C_{k,j}^{(\bar{b},b)} - C_{k,j}^{(b,b)} C_{k,j}^{(\bar{b},\bar{b})}}{C_{k,j}^{(b,\bar{b})} A_k^{(\bar{b})} - C_{k,j}^{(\bar{b},\bar{b})} A_k^{(b)}}, \tag{C7}$$

where the overbars on the a's and b's in the superscripts in the above two equations denote "the opposite outcome" (e.g. if $b = $ '$-$', then $\bar{b} = $ '$+$'). Equations C6 and C7 provide a method for computing the efficiency of each of Alice's and Bob's detection outcomes from the pair source to each detector, using only the measured singles and coincidence rates in the experiment. Importantly, the probabilities $p_{k,j}^{(a,b)}$ all factor out, and thus we can characterize the efficiencies of all four detectors without having details of the setup such as the quantum state

---

[9] Another way to understand the meaning of the joint probabilities $p_{k,j}^{a,b}$ is by considering the case of an experiment with no loss (i.e. consider $\alpha_k^{(\pm)} = \beta_j^{(\pm)} = 1$); in this case, $p_{k,j}^{(a,b)}$ is the probability that, in an experiment in which Alice and Bob measure a single pair of photons with settings k and j, respectively, Alice obtains outcome a and Bob obtains outcome b. In a real experiment with non-zero loss, the probability of Alice and Bob obtaining the joint outcome (a, b) is actually $\alpha_k^{(a)}\beta_j^{(b)} p_{k,j}^{(a,b)}$. In principle the probabilities $p_{k,j}^{(a,b)}$ could be computed given a complete quantum description of Alice and Bob's measurement apparatuses as well as the state shared between them. However while the $p_{k,j}^{(a,b)}$'s are a useful concept for deriving the modified Klyshko equations, they do not need to be known in order to find the efficiencies $\alpha_k^{(a)}$ and $\beta_j^{(b)}$.

of the photon pairs or the measurement operators describing Alice's and Bob's measurement stations.

While the accuracy of Eqs. (C6)-(C7), does not depend on the state of the photon pair or on the specific measurements performed, the precision does—this is due to the subtraction of two terms in the numerators and denominators. Specifically, these equations are most precise when Alice's and Bob's measurement outcomes are strongly correlated (or anticorrelated) and least precise when their outcomes are uncorrelated. For example, if the source emits the maximally entangled singlet state and Alice and Bob perform measurements in the same basis, then the positive term in each of the numerators and denominators will be much larger than the negative one, and Eqs. (C6)-(C7) will be a ratio of two large numbers with error bars determined by Poissonian counting statistics. On the other hand, if Alice and Bob measure the singlet state in two mutually unbiased bases, then both terms in the numerators and denominators will be of similar orders of magnitude, and the numerators and denominators will both approach 0, leading to a larger overall uncertainty in the final estimate of the efficiencies.

In order to test Eq. (14) it suffices to know the *ratio* of efficiencies of Bob's two detectors. To get this ratio, we divide the $b = $ '$+$' form of Eq. (C7) by its $b = $ '$-$' form. The detector outcomes between Alice and Bob are maximally anticorrelated—and the modified Klysho equations are most precise—when Alice measures in the same setting as Bob ($k = j$). Thus, to find the ratio of the detection efficiencies for each of Bob's outcomes, we use the formula:

$$\beta_j^{\text{ratio}} = \frac{\beta_j^{(+)}}{\beta_j^{(-)}} = \frac{C_{j,j}^{(-,+)} A_j^{(+)} - C_{j,j}^{(+,+)} A_j^{(-)}}{C_{j,j}^{(+,-)} A_j^{(-)} - C_{j,j}^{(-,-)} A_j^{(+)}}. \tag{C8}$$

We will now revisit the assumption that any lossy components of our experimental setup affect both polarizations of light equally. When characterizing our experiment, we do not find any evidence of polarization-dependent loss through any of the individual optics in our experiment. However, there is still some polarization-dependent loss at the source. Each photon produced in the source is split into two paths (the $|H\rangle$ component of the photon travels down one path, and the $|V\rangle$ component down another other), and the efficiency with which these two paths are coupled into the single-mode fibers leading to Alice's and Bob's measurements stations are not equal. To account for this, we simulate the count rates we would expect to measure in the presence of some polarization-dependent loss, and see what size of systematic error these simulated rates introduce into the equation for β_j^{ratio}. We find that if one polarization experiences $x\%$ more loss than the orthogonal polarization, then this can introduce an error of at most about $\pm x\%$ in Eq. C8. The exact amount of systematic error introduced depends on the bases that Alice and Bob measure in. If we can bound the polarization-dependent loss in the setup to be below a specific level, we can also bound the systematic error in our estimate of β_j^{ratio} to that same level.
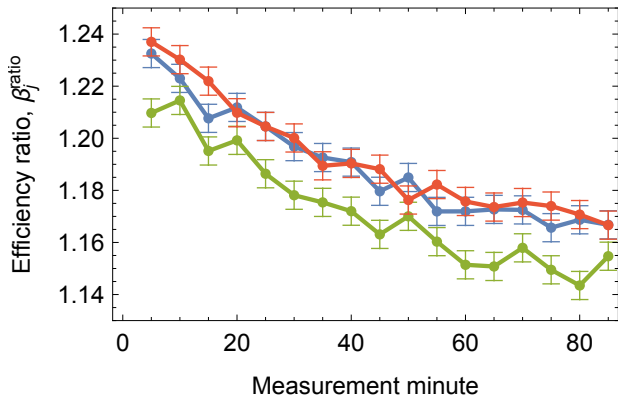
Figure C2. Ratio of system detection efficiency between Bob's +1 and −1 measurement outcomes throughout the running time of the experiment, for measurement settings $j=1$ (blue), $j=2$ (green), and $j=3$ (orange-red). Each point is calculated from five minutes of data, and error bars represent a 1-sigma statistical uncertainty estimated from the Allan deviation of the data.

Table CII. Measurement and 5-sigma conservative estimate of β_j^{ratio} for each measurement setting.

Setting	Measured value closest to 1	Conservative estimate
1	1.1657 ± 0.0054	1.115
2	1.1435 ± 0.0054	1.094
3	1.1666 ± 0.0054	1.116

C.4.b. Efficiency measurements

During our experiment, each time a voltage is applied to the PC it is applied for approximately 150 ns, which is enough time for 12 laser pulses to pass through the PC. We record data for all of these pulses, but only use the data from one pulse to test our steering inequality. We call this pulse the *inequality pulse*. We average the data from the other 11 pulses (which we call *calibration pulses*) and use these averaged data to estimate the efficiency of Bob's measurements. Using the calibration pulses allows us to keep track of β_j^{ratio} in real time while testing the steering inequality. We use Eq. (C8) to find β_j^{ratio} for each of Bob's three measurement settings.

Calibration pulses were collected in 85 separate one-minute intervals[10]. Calculating the Allan deviation [65] of the data implies that the smallest uncertainty on β_j^{ratio} is obtained if the data is averaged over intervals of approximately five minutes. The total dataset can be divided into 17 five-minute windows. To visualize how β_j^{ratio} varies over the running time of the experiment, we calculate it for each of the five-minute windows, and plot this is in Fig. C2. We notice a clear difference in β_j^{ratio} between setting 2 and settings 1 and 3. We also notice a significant drift in β_j^{ratio} throughout the experiment. The β_j^{ratio}'s we measure are all greater than one. For the specific $H_0 = 0$ and $H_0 = 1$ inequalities we test, it becomes easier to violate the inequality as β_j^{ratio} increases. Thus, the conservative choice is to underestimate the value of each β_j^{ratio}. We are confident that the amount of polarization-dependent loss in the source is no larger than 2%. For each setting we conservatively estimate β_j^{ratio} as the lowest measured value from among the five-minute windows, minus the maximum 2% systematic error we expect from polarization-dependent loss, minus five times the estimated standard error on that measurement. The conservative estimates of β_j^{ratio} we use to test the steering inequality are listed in Table CII.

C.4.c. Systematic uncertainties

There are two other types of systematic error that Eq. (C8) does not take into account, namely background counts from the detectors and multi-pair emissions from the source. We measure a background rate of approximately 200 counts/second, and the probability that any single laser pulse creates a photon pair is about 0.0072, which means that the probability of creating two pairs is $0.0072^2 \approx 5 \times 10^{-5}$. We simulate more realistic detection rates that take the background counts and multi-pair emission probability into account, and after evaluating Eq. (C8) with these simulated rates we see that these systematic errors lead to an *underestimation* of β_j^{ratio} by approximately 0.0006. Because these two sources of error combine to introduce a systematic uncertainty that is roughly 10% the magnitude of the statistical uncertainty on our measurement of β_j^{ratio}, and also because they lead to an even more conservative estimate of β_j^{ratio}, we can safely neglect them.

---

[1] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[2] A. Einstein, M. Born, and H. Born, *The Born-Einstein Letters: The Correspondence between Max & Hedwig Born and Albert Einstein 1916/1955*, 1st ed. (The MacMillan Press, London and Basingstoke, 1971).

[3] E. Schrödinger, Discussion of probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**, 555 (1935); *ibid.* **32**, 446 (1936).

[4] M. D. Reid, Demonstration of the Einstein-Podolsky-Rosen paradox using nondegenerate parametric amplification, Phys. Rev. A **40**, 913–923 (1989).

[5] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox, Phys. Rev. Lett. **98**, 140402 (2007).

---

[10] We collected 90 one-minute sets of inequality-pulse data, however due to an error in recording the data we only collected calibration data for 85 of those minutes.

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

[7] 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).

[8] J. Schneeloch, C. J. Broadbent, S. P. Walborn, E. G. Cavalcanti, and J. C. Howell, Einstein-Podolsky-Rosen steering inequalities from entropic uncertainty relations, Phys. Rev. A **87**, 062103 (2013).

[9] I. Kogias, P. Skrzypczyk, D. Cavalcanti, A. Acín, and G. Adesso, Hierarchy of steering criteria based on moments for all bipartite quantum systems, Phys. Rev. Lett. **115**, 210401 (2015).

[10] H. C. Nguyen, H.-V. Nguyen, and O. Gühne, Geometry of Einstein-Podolsky-Rosen correlations, Phys. Rev. Lett. **122**, 240401 (2019).

[11] 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, Nat. Commun. **6**, 7941 (2015).

[12] S. Armstrong, M.Wang, R. Y. Teh, Q. H. Gong, Q. Y. He, J. Janousek, H. A. Bachor, M. D. Reid, and P. K. Lam, Multipartite Einstein-Podolsky-Rosen steering and genuine tripartite entanglement with optical networks, Nat. Phys. **11**, 167–172 (2015).

[13] T. Gehring, V. Händchen, J. Duhme, F. Furrer, T. Franz, C. Pacher, R. F. Werner, and R. Schnabel, Implementation of continuous-variable quantum key distribution with composable and one-sided-device independent security against coherent attacks, Nat. Commun. **6**, 8795 (2015).

[14] N. Walk, S. Hosseini, J. Geng, O. Thearle, J. Y. Haw, S. Armstrong, S. M. Assad, J. Janoušek, T. C. Ralph, T. Symul, H. M. Wiseman, and P. K. Lam, Experimental demonstration of Gaussian protocols for one-sided device-independent quantum key distribution, Optica **3**, 634–642 (2016).

[15] I. Kogias, Y. Xiang, Q. Y. He, and G. Adesso, Unconditional security of entanglement-based continuous-variable quantum secret sharing, Phys. Rev. A **95**, 012315 (2017).

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

[17] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, One-sided device-independent quantum key distribution: security, feasibility, and the connection with steering, Phys. Rev. A **85**, 010301(R) (2012).

[18] Y. Xiang, I. Kogias, G. Adesso, and Q. Y. He, Multipartite Gaussian steering: Monogamy constraints and quantum cryptography applications, Phys. Rev. A **95**, 010101(R) (2017).

[19] M. D. Reid. Signifying quantum benchmarks for qubit teleportation and secure quantum communication using Einstein-Podolsky-Rosen steering inequalities, Phys. Rev. A, **88**, 062338 (2013).

[20] Q. He, L. Rosales-Zárate, G. Adesso, and M. D. Reid, Secure continuous variable teleportation and Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **115**, 180502 (2015).

[21] C.-Y. Chiu, N. Lambert, Teh-Lu Liao, F. Nori, and C.-M. Li, No-cloning of quantum steering, NPJ Quantum Information **2**, 16020 (2016).

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

[23] 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).

[24] K. Sun, X.-J. Ye, Y. Xiao, X.-Y. Xu, Y.-C. Wu, J.-S. Xu, J.-L. Chen, C.-F. Li, and G.-C. Guo, Demonstration of Einstein-Podolsky-Rosen steering with enhanced subchannel discrimination, npj Quantum Inf. **4**, 12 (2018).

[25] D. H. Smith, G. Gillett, M. P. de Almeida, C. Branciard, A. Fedrizzi, T. J. Weinhold, A. Lita, B. Calkins, T. Gerrits, H. M. Wiseman, S. Woo Nam, and A. G. White, Conclusive quantum steering with superconducting transition-edge sensors, Nat. Commun. **3**, 625 (2012).

[26] A. J. Bennet, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, Arbitrarily loss-tolerant Einstein-Podolsky-Rosen steering allowing a demonstration over 1 km of optical fiber with no detection loophole, Phys. Rev. X **2**, 031003 (2012); Phys. Rev. X **6**, 019902(E) (2016).

[27] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. Wiseman, R. Ursin, and A. Zeilinger, Loophole-free Einstein-Podolsky-Rosen experiment via quantum steering, New J. Phys. **14**, 053030 (2012).

[28] D. J. Saunders, M. S. Palsson, G. J. Pryde, A. J. Scott, S. M. Barnett, and H. M. Wiseman, The simplest demonstrations of quantum nonlocality, New J. Phys. **14**, 113020 (2012).

[29] E. G. Cavalcanti, C. J. Foster, M. Fuwa, and H. M. Wiseman, Analog of the Clauser-Horne-Shimony-Holt inequality for steering, J. Opt. Soc. Am. B **32**, A74 (2015).

[30] M. Fuwa, S. Takeda, M. Zwierz, H. M. Wiseman, and A. Furusawa, Experimental proof of nonlocal wavefunction collapse for a single particle using homodyne measurements, Nat. Commun. **6**, 6665 (2015).

[31] N. Tischler, F. Ghafari, T. J. Baker, S. Slussarenko, R. B. Patel, M. M. Weston, S. Wollmann, L. K. Shalm, V. B. Verma, S. W. Nam, H. C. Nguyen, H. M. Wiseman, and G. J. Pryde, Conclusive experimental demonstration of one-way Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **121**, 100401 (2018).

[32] D. J. Saunders, S. J. Jones, H. M. Wiseman, and G. J. Pryde, Experimental EPR-steering using Bell-local states, Nat. Phys. **6**, 845 (2010).

[33] H. M. Wiseman, and J. M. Gambetta, Are Dynamical quantum jumps detector dependent?, Phys. Rev. Lett. **108**, 220402 (2012).

[34] Q. Y. He and M. D. Reid, Genuine multipartite Einstein-Podolsky-Rosen steering. Phys. Rev. Lett. **111**, 250403 (2013).

[35] V. Händchen, T. Eberle, S. Steinlechner, A. Samblowski, T. Franz, R. F. Werner, and R. Schnabel, Observation of one-way Einstein-Podolsky-Rosen steering, Nat. Photonics **6**, 596–599 (2012).

[36] S. Wollmann, N. Walk, A. J. Bennet, H. M. Wiseman, and G. J. Pryde, Observation of genuine one-way Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **116**, 160403 (2016).

[37] X. W. Deng, Y. Xiang, C. Tian, G. Adesso, Q. Y. He, Q. H. Gong, X. L. Su, C. D. Xie, and K. C. Peng, Demonstration of monogamy relations for Einstein-Podolsky-Rosen steering in Gaussian cluster states, Phys. Rev. Lett. **118**, 230501 (2017).

[38] A. Cavaillès, H. Le Jeannic, J. Raskop, G. Guccione, D. Markham, E. Diamanti, M. D. Shaw, V. B. Verma, S. W. Nam, and J. Laurat, Demonstration of Einstein-Podolsky-Rosen steering using hybrid continuous- and discrete-variable entanglement of light, Phys. Rev. Lett. **121**, 170403 (2018).

[39] M. Fadel, T. Zibold, B. Décamps, P. Treutlein, Spatial entanglement patterns and Einstein-Podolsky-Rosen steering in Bose-Einstein condensate. Science **360**, 409-413 (2018).

[40] P. Kunkel, M. Prüfer, H. Strobel, D. Linnemann, A. Frölian, T. Gasenzer, M. Gärttner, M. K. Oberthaler, Spatially distributed multipartite entanglement enables EPR steering of atomic clouds. Science **360**, 413-416 (2018).

[41] Q. Y. He, Q. H. Gong, and M. D. Reid, Classifying directional Gaussian entanglement, Einstein-Podolsky-Rosen steering, and discord, Phys. Rev. Lett. **114**, 060402 (2015).

[42] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).

[43] S. Nagy, and T. Vértesi, EPR-steering inequalities with communication assistance, Scientific Reports **6**, 21634 (2016).

[44] A. B. Sainz, L. Aolita, N. Brunner, R. Gallego, and P. Skrzypczyk, Classical communication cost of quantum steering, Phys. Rev. A **94**, 012308 (2016).

[45] D. N. Klyshko, Use of two-photon light for absolute calibration of photoelectric detectors, Sov. J. Quantum Electron. **10**, 1112 (1980).

[46] J.-Å. Larsson, Loopholes in Bell inequality tests of local realism, J. Phys. A: Math. Theor. **47**, 424003 (2014).

[47] J. S. Bell, in *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 2004), p. 232.

[48] The BIG Bell Test Collaboration., C. Abellán, A. Acín, et al., Challenging local realism with human choices, Nature **557**, 212–216 (2018).

[49] C. Abellán, W. Amaya, D. Mitrani, P. Valerio, and M. W. Mitchell, Generation of Fresh and Pure Random Numbers for Loophole-Free Bell Tests, Phys. Rev. Lett., **115**, 250403 (2015).

[50] P. M. Pearle, Hidden-variable example based upon data rejection, Phys. Rev. D **2**, 1418 (1970).

[51] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, C. Abellán *et al*., Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometers, Nature (London) **526**, 682 (2015).

[52] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, Significant Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. **115**, 250401 (2015).

[53] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellan, W. Amaya, V. Pruneri, T. Jennewein, M.

W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. **115**, 250402 (2015).

[54] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes, Phys. Rev. Lett. **119**, 010402 (2017).

[55] M.-H. Li, C. Wu, Y. Zhang, W.-Z. Liu, B. Bai, Y. Liu, W. Zhang, Q. Zhao, H. Li, Z. Wang, L. You, W. J. Munro, J. Yin, J. Zhang, C.-Z. Peng, X. Ma, Q. Zhang, J. Fan, and J.-W. Pan, Test of Local Realism into the Past without Detection and Locality Loopholes, Phys. Rev. Lett. **121**, 080404 (2018).

[56] B. F. Toner and D. Bacon, Communication Cost of Simulating Bell Correlations, Phys. Rev. Lett. **91**, 187904 (2003).

[57] D. A. Evans, E. G. Cavalcanti, and H. M. Wiseman, Loss-tolerant Tests of Einstein-Podolsky-Rosen Steering, Phys. Rev. A **88**, 022106 (2013).

[58] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. Phys. Rev. A **40**, 4277-4281 (1989).

[59] M. A. Wayne, A. L. Migdall, Z. H. Levine, and J. C. Bienfang, A post-processing-free single-photon random number generator with ultra-low latency, Opt. Express **26**, 32788-32801 (2018).

[60] B. S. Cirel'son [Tsirelson], Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[61] Q. Y. He and M. D. Reid, Genuine multipartite Einstein-Podolsky-Rosen steering, Phys. Rev. Lett. **111**, 250403 (2013).

[62] M. M. Weston, S. Slussarenko, H. M. Chrzanowski, S. Wollmann, L. K. Shalm, V. B. Verma, M. S. Allman, S. W. Nam, and G. J. Pryde, Heralded quantum steering over a high-loss channel, Sci. Adv. **4**, e1701230 (2018).

[63] D. Salart, A. Baas, C. Branciard, N. Gisin, and H. Zbinden, Testing the speed of 'spooky action at a distance', Nature **454**, 861-864 (2008).

[64] Y. Liu, X. Yuan, M.-H. Li, W. J. Zhang, Q. Zhao, J. Q. Zhong, Y. Cao, Y.-H. Li, L.-K. Chen, H. Li, T. Y. Peng, Y.-A. Chen, C.-Z. Peng, S.-C. Shi, Z. Wang, L. X. You, X. F. Ma, J. Y. Fan, Q. Zhang, and J.-W. Pan, High-Speed Device-Independent Quantum Random Number Generation without a Detection Loophole, Phys. Rev. Lett. **120**, 010503 (2018).

[65] D. W. Allan, Statistics of atomic frequency standards, in *Proceedings of the IEEE* **54**, pp. 221-230 (1966).