Quantum Position Verification with Remote Untrusted Devices

Gautam A. Kavuri,[1,2,*] Yanbao Zhang,[3,†] Abigail R. Gookin,[4,2] Soumyadip Patra,[5] Joshua C. Bienfang,[6] Honghao Fu,[7] Yusuf Alnawakhtha,[8] Dileep V. Reddy,[1,2] Michael D. Mazurek,[1,2] Carlos Abellán,[9] Waldimar Amaya,[9] Morgan W. Mitchell,[10,11] Carl A. Miller,[12,8] Richard P. Mirin,[13] Martin J. Stevens,[13] Scott Glancy,[14,1] Emanuel Knill,[14,1,15] and Lynden K. Shalm[1,13,16,‡]

[1]*Department of Physics, University of Colorado, Boulder, CO, 80309, USA*
[2]*Associate of the National Institute of Standards and Technology, Boulder, CO, 80305, USA*
[3]*Quantum Information Science Section, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA*
[4]*Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, 80309, USA*
[5]*Department of Mathematics, University of New Orleans, New Orleans, LA, 70148, USA*
[6]*Joint Quantum Institute, National Institute of Standards and Technology and University of Maryland, 100 Bureau Drive, Gaithersburg, MD, 20899, USA.*
[7]*Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, H3G 1M8, Canada*
[8]*Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD, 20742, USA*
[9]*Quside Technologies S.L., Castelldefels (Barcelona), Spain*
[10]*ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain.*
[11]*ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010, Barcelona, Spain*
[12]*Computer Security Division, National Institute of Standards and Technology, Gaithersburg, MD, 20899, USA*
[13]*Physical Measurement Laboratory, National Institute of Standards and Technology, Boulder, CO, 80305, USA*
[14]*Applied and Computational Mathematics Division, National Institute of Standards and Technology, Boulder, CO, 80305, USA*
[15]*Center for Theory of Quantum Matter, University of Colorado, Boulder, CO, 80305, USA*
[16]*Quantum Engineering Initiative, Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, 80309, USA*

(Dated: August 11, 2026)

arXiv:2601.16892v2 [quant-ph] 8 Aug 2026

Position information underpins many modern technologies, from navigation and timing to authentication and critical infrastructure [1–3]. However, classical methods of proving that information originates from a particular position are

* gautamk@duck.com
† zhangy2@ornl.gov
‡ krister.shalm@nist.gov

vulnerable to spoofing [4–6]. This limitation can be overcome with quantum technologies [7–16] but current protocols rely on trust in quantum hardware that can be undermined [17, 18], or require quantum computers and bounds on adversarial computation [19]. Nevertheless, there has been significant interest in experimental demonstrations, and aspects of these protocols have been implemented [20–22]. Here we introduce and experimentally demonstrate the Bell-test quantum position verification protocol for device-independent quantum position verification that guarantees security with only observed correlations from a loophole-free Bell test [23–25] across a quantum network. We experimentally implement a version of this device-independent protocol against adversaries who, before each trial, are weakly entangled. Our demonstration achieves a one-dimensional localization 2.47(2) times smaller than the best, necessarily non-remote, classical localization protocol. Compared to classical protocols with identical latencies, the localization volume is 4.53(5) times smaller, and represents a certifiable quantum advantage. The general Bell-test protocol is loss tolerant and secure against adversaries with significant quantum resources. This work allows digital security to be anchored to physically trusted locations, enabling new position-based authentication protocols for applications such as financial transactions, legal agreements, and securing critical infrastructure [26–29].

INTRODUCTION

Position is a fundamental aspect of security [27]. Sensitive systems and information are often stored in secure locations that can only be accessed in person, and permissions for certain actions are granted to entities based on where they are. Physical access controls, such as air-gapped systems [30] or analog locks, mitigate remote attacks [31, 32]. A central challenge is extending the security of a trusted location into the digital world—if a party can prove they are communicating from a particular place, they can be remotely granted comparable levels of access and trust. Such guarantees would underpin a wide range of applications, from authenticating that commands to a power grid originate from the utility's control center rather than a hostile actor abroad [2], to confirming that a GPS signal genuinely originates from a satellite at a known orbital position [33] rather than a nearby software-defined radio, to ensuring that a transaction request to a bank's servers comes from a customer at a verified branch [26, 28]. Securely verifying position is the first step to building many of these advanced position-based authentication applications [1, 10].

Accurately localizing a position can be achieved using ranging methods that measure the round-trip time of signals sent at the speed of light to a target region. Several trusted parties, known as verifiers, can then work together to triangulate position. However, the signals used in such ranging protocols can in principle always be intercepted and manipulated by adversaries to spoof a location they do not occupy [1, 4, 10]. Cryptographic authentication can mitigate this [5, 6, 34, 35], but only at the cost of trusted hardware and pre-shared keys or secrets. If those keys are leaked, hardware devices cloned, or if cryptographic hardness assumptions fail, a wide range of sensitive systems that depend on position, navigation, and timing become vulnerable to spoofing.

An alternative approach to avoiding spoofing in position verification is to use quantum states that cannot be cloned in combination with classical ranging [7, 8]. In the simplest

case of BB84 quantum position verification (QPV) [8, 10] two trusted verifiers work together to determine if a third party, the prover, is located within a target location. One verifier randomly prepares a qubit in one of several bases and sends it to the prover. The verifiers also send classical challenge bits to the prover who uses them to perform a predefined computation. The result of this computation sets the measurement basis for the prover. The prover performs this measurement and returns the result to the verifiers, who check that the result is correct and use the round-trip time of the signals to localize the position of the prover. While unconditional security against entangled attackers is impossible [10], it is possible to design protocols where the resources that the adversaries require to successfully attack the protocol increase with the number of classical challenge bits. Ideally, a protocol also would require an honest prover to manipulate only a single qubit while requiring more resources from successful attackers [12, 15]. It is also desirable to allow the quantum signals in the protocol to be arbitrarily slow and rely only on the timing of classical communication for ranging. Protocols with this feature are more practical for deployment over large distances [12, 14, 16].

Besides BB84 QPV a range of other QPV protocols have been proposed, including extensions of the original BB84 QPV [12, 14, 16], methods based on ideas from measurement-device-independent quantum key distribution (QKD) and entanglement swapping [11], qubit routing [8, 12] and SWAP protocols [13]. Importantly, the security of these protocols requires that the quantum devices used to perform specific state preparations and measurements are trusted and operating correctly. Such protocols are said to be device-dependent, and this dependence introduces opportunities for attackers to subvert the quantum devices or exploit side-channels. Such exploits have been demonstrated for the closely related QKD protocols [4–6, 17, 18]. Despite much theoretical development in device-dependent QPV protocols, their complete experimental demonstration has been lagging. Experiments include demonstrations of the basic principles [20], with distance simulated with fiber-spools [22], and of the main components of a high-complexity scheme based on advanced quantum computers [21].

Here we introduce the Bell-test QPV protocol, a device-independent protocol that is based on a loophole-free Bell test. Bell-test QPV uses the violation of a Bell inequality by a verifier and the prover to certify that the prover is operating a quantum device in a specified target region in spacetime, where the quantum device must be capable of non-trivial quantum actions conditional on classical input. No assumptions are made about the internal workings of the required quantum devices, including the quantum measurement devices and the entanglement source. Device-independent security is achieved based only on analysis of the inputs to and outputs from the experimental devices. Like all device-independent protocols such as those implemented in Refs. [36–42], Bell-test QPV requires trust in the classical devices used for recording results, measuring timing, and performing distance measurements. The protocol performance does not depend on the speed of the qubits provided by the entanglement source, which is treated as fully untrusted, but only on the round-trip time of the classical signals sent. As long as a verifier and the prover share an entangled state, be it entangled photons or matter-based qubits entangled via a quantum repeater, the protocol can proceed. The compatibility of the protocol with heralded sources of entanglement such as those reported in the recent work [41, 43] means that the protocol is scalable with distance over a quantum network. For precise localization, it suffices to use near-speed-of-light classical communication such as wireless communication or hollow-core fiber [22].

We implemented the Bell-test QPV protocol with two verifiers separated by 195.1(3) m, performing 335 experimental protocol instances, each lasting two or four minutes depending on the class of adversaries considered, over two days. Each protocol instance consists of a sequence of trials. Measurement outcomes and messages from the trials are recorded and determine whether the instance successfully verified position. While we do not limit the computation or communication power of adversaries during each experimental trial, our security analysis requires a limit on the amount of entanglement shared between adversaries prior to the trial. Throughout this work, we refer to such pre-shared entanglement as prior entanglement. A basic version of the protocol is secure against adversaries who share no prior entanglement. An extended protocol implemented in the experiment demonstrates security against adversaries possessing limited prior entanglement, bounded from above by an average robustness of entanglement [44] of 8×10^{-6} over the trials of a protocol instance. A coarse such bound was previously obtained for general QPV protocols in Ref. [45]. To protect against adversaries with significant quantum resources including prior entanglement, we prove that it suffices to increase the number of bits in the challenges. The security of our protocol is derived from the impossibility of faster-than-light signaling and constraints on the shareability of three-party non-signaling correlations [46]. These constraints can be effectively falsified by a loophole-free Bell violation in the configuration we experimentally implemented. Each trial requires the prover to compute a function of challenges received from the verifiers. We experimentally demonstrate the simplest nontrivial instance of such a function.

Conditional on successfully passing the Bell-test protocol, we can be confident that a prover must have performed an input-conditional quantum operation in a localized target spacetime region E. Ideally, the localization could approach arbitrary precision, independent of verifier separation. In practice, it is constrained by communication and processing latencies. Using low-latency processing and communication, we are able to certify the position of a prover to a volume 4.53(5) smaller than that achievable by a classical protocol that is constrained by the same latencies. Our results establish a new cryptographic primitive for quantum networks, allowing access to a secure physical location to be used to authenticate a network user. This bridges physical and network security and opens new directions for position-based cryptography and computation [1, 10].

POSITION VERIFICATION

Formally, position verification is a cryptographic protocol that involves two or more verifiers working together to ensure that only provers P with access to a specified target region E can pass the protocol. In this work, we consider the case of two verifiers A and B who use time-of-flight ranging to establish presence of a prover within the target region. We consider first a classical protocol where we assume that the verifiers are only interacting with a single prover, as shown in Fig. 1 a), the verifiers A and B send classical challenges u_{A} and u_{B} timed to arrive simultaneously at the prover. The prover performs a computation using the challenges from both verifiers and a predetermined, publicly known function $g(u_{\mathsf{A}}, u_{\mathsf{B}})$, and then returns the result. Conditioned on the prover computing the correct answer, each verifier can use the time difference between when the challenge was sent and when the answer was received to bound the maximum distance to the prover. The intersection of the regions within the maximum distances from each verifier localizes the position of the prover. If this intersection is contained in the region to be secured, the protocol passes. With light-speed

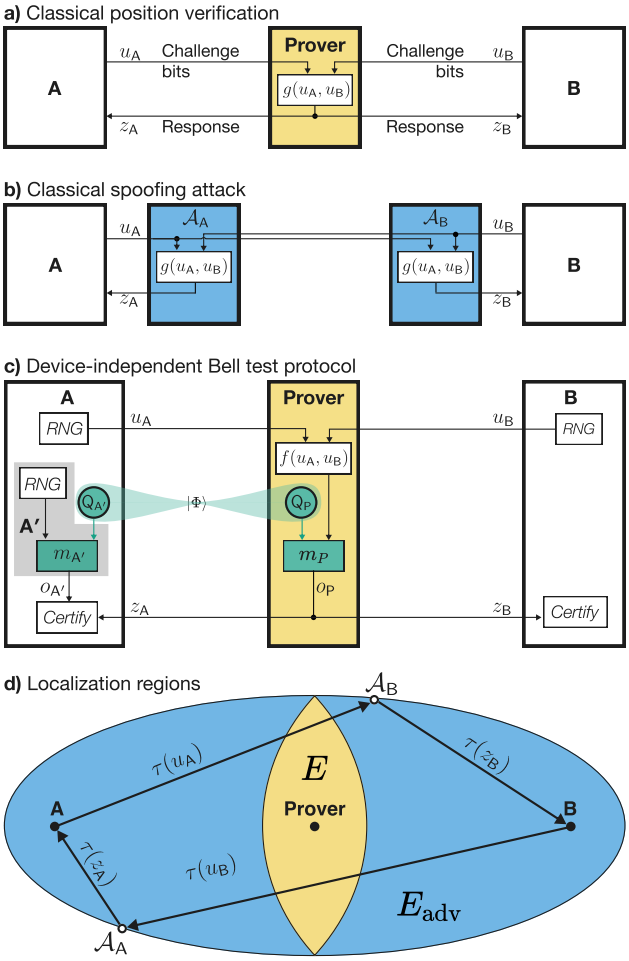

FIG. 1. Schematic representations of position verification protocols and possible attacks. Subfigure a) shows a single-prover classical protocol for position verification that enforces the target region E in yellow with classical challenges from the verifiers A and B. The verifiers desire that the computation $g(u_{\mathsf{A}}, u_{\mathsf{B}})$ is localized to the target region. Subfigure b) sketches a specific attack on the classical protocol, with colluding adversaries $\mathcal{A}_{\mathsf{A}}$ and $\mathcal{A}_{\mathsf{B}}$, situated in the blue regions outside the target regions, each performing the computation locally. By communicating the challenges to one another, they can spoof a successful pass of the classical protocol from anywhere inside E_{adv}. To ensure security against general attacks perpetrated by unentangled or weakly entangled adversaries, we introduce our quantum protocol, sketched in c). Here, the honest prover P and a quantum measurement station A′ located at A share an entangled state $|\Phi\rangle$. The state is measured using untrusted quantum hardware, with the measurement basis at A decided by a local random number generator and the basis at P decided by the output of $f(u_{\mathsf{A}}, u_{\mathsf{B}})$. Subfigure d) is a schematic of the spatial extents of the target regions enforceable by the classical (E_{adv}) and quantum (E) protocols, projected into two dimensions for a realistic two-verifier protocol.

signals and zero processing latency, the verifiers can localize the prover to a single point. However, any latencies in the classical signals or the processing increases the size of the smallest possible localization region, E. Projected to two dimensions, this forms a lenticular region, as shown by the yellow region in Fig. 1 d).

However, a pair of adversaries working together can spoof responses from inside E, as shown in Figure 1 b). In the attack, each adversary stores a copy of the challenge bits sent from the verifier closest to them. When the challenge bits from the other verifier arrive, they compute the response and return the answer to the verifier closest to them. In this way, the adversaries can convince the verifiers they are inside the target region E. In the ideal case where there is no excess latency, the spoofing adversaries must both be stationed on the line joining the verifiers, or the verifier axis. However, if there is processing or other excess latency in the system, then adversaries with negligible processing latency can exploit this to move away from the verifier axis in three dimensions and still respond in time. As described in the methods section, the region that successful adversaries are required to occupy is bounded by a set of geometric and timing constraints. In the symmetric case, where the target region is centered at the midpoint between the verifiers, and any timings are symmetric, the region that the adversaries are required to occupy is an ellipsoid with foci at the verifier positions, and whose major axis is twice the distance light travels in the time from when a challenge is emitted to when a response is received, at either verifier. The surface of the ellipsoid corresponds to the set of positions from which an adversary can just barely receive a challenge from one verifier, compute a response, and deliver it to the other verifier in time, as shown in Fig. 1 d). Any adversaries outside this ellipsoidal region E_{adv} are unable to spoof the protocol. This means the verifiers can achieve a degree of localization with a classical protocol, but this region contains the verifier positions and has diameter substantially larger than the verifier separation.

The device-independent Bell-test QPV protocol is based on the prover and verifiers jointly violating a loophole-free Bell inequality, as shown in Fig. 1 c). In a Bell test, two parties perform independent, space-like separated measurements on an entangled state. These measurements are repeated many times in a series of trials. If the joint measurement statistics violate a Bell inequality, it is possible to rule out the presence of local hidden variable models as an explanation for the results. In our protocol, in each trial, verifier A and prover P share an entangled Bell state. At the start of the trial, the verifiers send challenges u_{A} and u_{B} to the prover. When they arrive, the prover computes a measurement settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ whose value is used to choose the prover's measurement basis. Simultaneously, verifier A uses a random number generator to set their measurement basis. A and P then measure their respective entangled qubits in a space-like separated manner, and P responds to the challenges by sending their measurement outcome separately to A and B. For each trial, a Bell-like test statistic is evaluated on the trial record, which consists of the measurement basis choice and measurement outcome of A, and the prover measurement basis $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ and the two prover responses to A and B. The two responses may not be identical due to transmission error or other sources of noise. Intuitively, empirical violation of the inequality associated with the test statistic can be used to rule out the presence of spoofing adversaries. Security of the Bell-test QPV protocol places additional demands on the test statistic and its interpretation, see the next section. The round-trip time between the challenges being sent and the prover's response with the outcome of their measurement allows the verifiers to localize the prover. The target region E for the quantum protocol is determined in the Supplementary Information Sect. V B and is the same as the single-prover target region E in the classical protocol above, while being secure against colluding adversaries. The ratio between the achievable localization region E by the Bell-test QPV protocol and the localization region E_{adv} of the classical protocol represents a quantum advantage in localization. In principle, the quantum protocol's target region could be made arbitrar-

ily small with instantaneous processing and direct, speed-of-light classical communication. The only requirement is that the verifier and prover have established entanglement before the measurements need to be performed. This can be done, for instance, through direct transmission over a quantum channel or by using a quantum repeater.

While the preceding discussion has focused on the target region in the three spatial dimensions, it can be specified in all four dimensions including time, see Sect. V B in the Supplementary Information. Given that our experiment is performed as a series of trials, the localization has a periodic structure in time.

THEORY

We establish security of Bell-test QPV against several classes of adversaries, where colluding adversaries attempt to pass the Bell-test protocol without having any presence in the target region. The classes are distinguished by the resources available to the adversaries in a trial. The resources considered include prior entanglement in a trial, and quantum computation and communication. We do not restrict adversaries ability to move provided that they do not enter the target region's spacetime extent. Security is established under standard assumptions for device-independent protocols and allows for complete control over the source by the adversaries. The assumptions are summarized in the Methods section and a detailed list of assumptions is given in Supplementary Information Sect. II. Bell-test QPV depends on choosing measurement bases and challenge distributions, and the settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ of the challenges. Requirements on these distributions and their dependencies are included in the list of assumptions.

Security of QPV protocols requires soundness and completeness. Soundness with parameter δ requires that before a protocol starts, the maximum probability that adversaries in a given class can pass the protocol is at most δ. For high security, the soundness parameter must be extremely small. Completeness with parameter ϵ requires that the actual prover has probability at least ϵ of passing. It is desirable for completeness to be close to 1. For the experimental demonstration we choose $\delta = 2^{-64}$. For a summary of Bell-test QPV soundness and completeness results, see the Methods.

Completeness of Bell-test QPV is witnessed by any verifier-prover pair able to violate local realism in a Bell test with two measurement bases for the prover, such as we experimentally demonstrated. Scalability of the protocol requires that the number of trials grows sufficiently slowly as the soundness is reduced and the completeness is increased. Like most useful protocols, Bell-test protocols require only $O(\log(1/\delta))+O(\log(1/(1-\epsilon)))$ trials for soundness δ and completeness ϵ. The suppressed constants depend on the quality of the honest prover's performance.

Bell-test protocols require a departure from the conventional approach of classifying trials by whether the prover responses are correct or not. Instead we use test factors originally introduced for test of local realism [47], which are special test statistics W applied to each trial and calculated from the trial record. Test factors have recently been found to have broad applications in statistics for constructing so-called e-values, which are a chainable alternative to p-values [48, 49]. For position verification, these test statistics satisfy $W \geq 0$ and the trial expectation of W is at most 1 for all adversaries in a given class regardless of what happened in the past. To determine whether the protocol passed, the test factor W is applied to each trial's record and the resulting values are multiplied to obtain the test-factor product P. A key consequence of the properties of test factors is that the probability that

$P \geq 1/\delta$ for adversaries is at most δ. Soundness is therefore assured if the protocol is passed only when $P \geq 1/\delta$.

The use of Bell tests with test factors and the need to establish target regions in 3+1-dimensional spacetime require a detailed analysis of the spacetime capabilities of adversaries. We adapt the spacetime circuit method used for a parallel, one-round QPV protocol by Unruh [50] to determine the trial target region from the timing information for the verifiers sending out the challenges and receiving the responses. We then partition the adversary accessible regions to reduce adversary attacks to those that can be implemented by two adversaries who can communicate only once during a trial. See Supplementary Information Sect. V B. Such reductions to two adversaries are typical for QPV protocols but are usually only implemented in one spatial dimension, which omits several spacetime regions that could be exploited by adversaries and does not address the effects of realistic latencies and transmission delays. Because our treatment is fully relativistic, it is also necessary to determine the specific trial boundary at which we quantify the adversary entanglement resources. The trial boundary obtained is non-trivial and can be described as the boundary of a spacetime region constructed from lightcones, see Supplementary Information Fig. 9.

The basis for our soundness results against different classes of adversaries is an analysis of the power of adversaries with no prior entanglement, but unlimited classical and quantum computing and communication resources. Such an analysis forms the basis of the security of many other QPV protocols and was introduced in Ref. [10]. Subsequent improved analyses connected security of a class of existing QPV protocols to monogamy of entanglement games [51]. For Bell-test QPV, the analysis requires a different approach for simplifying adversary strategies that takes advantage of test-factor properties. See Supplementary Information Sect. V C. The simplification reduces the problem of test-factor design for Bell-test QPV to constructing test factors for three-party quantum non-signaling scenarios. For provers with two measurement basis choices, the problem can be further reduced to that of modifying test factors against local realism.

An advantage of test factors is their flexibility: they can be adapted during a protocol, enable heralding and early stopping, and can be rescaled against adversaries with prior entanglement. Heralded trials can be used to mitigate the problem of loss in photon transmission. In particular, previous proofs of loss tolerance by trial commitment by the prover [16] generalize to Bell-test QPV, see Supplementary Information Sect. V F. Rescaling can be used directly for the security against weakly entangled adversaries demonstrated in our experiment, see below. This is a quantified generalization of an observation in Ref. [45], see Supplementary Information Sect. V E. We quantify the entanglement by robustness of entanglement, which measures how far an entangled state is from the set of separable states. We show that the increase in the expectation of the test factor W due to the adversaries' prior entanglement can be bounded in terms of robustness of this prior entanglement, and this increase determines the required test-factor rescaling.

Rescaled test factors are also used against adversaries with significant entanglement or other quantum resources. This requires increasing the number l of bits in each of the challenges u_{A} and u_{B} and a suitable choice of settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$. That increasing the complexity of $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ is sufficient for security against adversaries with significant quantum resources has been established for some previous protocols [12, 15] by reduction to a communication-complexity problem. Test factors enable a tighter such reduction to a different communication-complexity problem, see Supplementary Information Sect. V G. We can protect against adversaries with prior entanglement in $(\log(l)/2 - O(1))$ qubits

and unlimited classical communication and local quantum computation, or adversaries with $O(l/(\log(q)+1))$ quantum operations on q qubits and unlimited classical communication. With our experimental parameters and an increase of the number of trials required by a factor of 6, if $l = 229$, then adversaries, in addition to having to occupy intercepting paths on two opposite sides of the target region and to intercept and control multiple signals, must use at least 1 qutrit for entanglement or quantum communication per trial to have a reasonable chance of passing the protocol. In the experiment, we used one weakly entangled qubit per trial. If $l = 999$, then adversaries with fewer than $q = 8$ qubits require more than 120 operations in a standard universal gate and measurement set to break security. There is a simple function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ that achieves this security and that is suitable for low-latency implementation. The function is described in the Methods.

EXPERIMENT

In our experiment, we demonstrate successful instances of the Bell-test protocol using two verifier stations, A and B, and a prover station, P. Each verifier is equipped with a random number generator to produce challenge bits and a time-tagging device to record detection events. Verifier A also houses an entangled photon source and a polarization measurement apparatus A'. As shown in Fig. 2, verifier stations A and B are separated by 195.1(3) m. The prover station is 92.8(1) m from verifier B. All three stations are arranged approximately collinearly, and much of our hardware is similar to past device-independent experiments [36, 37].

The protocol requires the accumulation of statistics over many trials. At the beginning of each trial, the quantum verifier A attempts to generate a pair of polarization-entangled photons in a state close to $0.383\,|HH\rangle + 0.924\,|VV\rangle$, where $|H\rangle$ and $|V\rangle$ represent horizontal and vertical polarizations respectively. The target state was chosen based on a simulation numerically maximizing Bell violation, taking measured background counts and photon efficiencies into account. One photon is delayed locally via a fiber loop, while the other is transmitted to the prover in fiber. Each verifier then sends a randomly generated challenge bit to the prover. The challenge bits are timed to arrive at the prover station within 3(2) ns of each other and are transmitted using specialized coaxial cables with signal transmission speeds about 86% the speed of light. We note that we allow adversaries to transmit signals at the speed of light. The implemented 86% transmission speed makes our localization region larger than it would be in an implementation with speed-of-light classical signal transmission.

When the challenge bits $u_{\mathsf{A}}, u_{\mathsf{B}}$ arrive at the prover station, an XOR circuit is used to compute a settings bit $m_{\mathsf{P}} = u_{\mathsf{A}} \oplus u_{\mathsf{B}}$. The prover selects their polarization measurement bases, defined as $\cos(\theta_P)\,|H\rangle + \sin(\theta_P)\,|V\rangle$, by using this bit to choose between two angles, $\theta_P \in \{6.7°, -29.5°\}$. The selection is implemented using a Pockels cell to rotate the polarization of the incoming photon, followed by a polarizer to perform a polarization measurement, see Fig. 2. The output of the polarizer is coupled into a single-mode fiber and incident on a superconducting nanowire detector. The result of the single-photon measurement (click or no-click) is sent back to the verifier stations via another set of coaxial cables, where the arrival times are recorded on time tagging devices. As depicted in Fig. 3 a), the incoming photon arrives just in time for the measurement. At nearly the same time that the prover performs its quantum measurement (see Fig. 3 a)), verifier A measures the polarization of the stored local photon $\mathsf{Q}_{\mathsf{A}'}$ using a similar setup. The verifier uses a local random number gener-

ator to select $m_{\mathsf{A}'}$, which sets the polarization measurement basis $\cos(\theta_{A'})\,|H\rangle + \sin(\theta_{A'})\,|V\rangle$ to the angle $\theta_{A'}$ to either $-6.7°$ or $29.26°$. The setting choice and outcome are recorded on a local time tagger. The quantum verifier and prover locations are sufficiently separated to ensure space-like separation of measurement settings and measurement outcomes at the other station, see Fig. 3.

The system operates at a rate of approximately 250,000 trials per second, limited by the operating rate of the Pockels cells. Data was recorded in 60 s intervals over two days. At the start of each day, an automated calibration routine was performed to correct for polarization drifts in the optical fiber and to ensure high coupling efficiency for the entangled photons. The measured system detection efficiency was $> 81\%$ for both verifier A and the prover station P, sufficient for an analysis that does not rely on fair-sampling assumptions. Specifically, this means we do not post-select or perform background subtraction. Loss or background counts beyond the acceptable threshold would result in repeated protocol failure. The efficiency threshold for our protocol is 2/3, the same as a two-party Bell test [52, 53], see Supplementary Information Sect. V C. Loophole-free Bell tests have been successfully implemented on many platforms [23, 24, 54–57], making the Bell-test QPV protocol amenable to a wide range of quantum hardware.

ANALYSIS AND RESULTS

We first collected calibration data to set key parameters in our analysis. We then performed a series of QPV experiments on two separate days, with data collected in 60 s chunks with each chunk containing about 15 million experimental trials. Approximately eight hours of data was collected on September 20, 2024 (day 1) and seven hours of data on October 7, 2024 (day 2). The day 1 data was reserved for testing the basic protocol against adversaries without prior entanglement, while the day 2 data was used to test the extended protocol against adversaries with prior entanglement whose robustness averaged over trials is limited by a threshold $r_{\rm th}$. The soundness parameter for both protocols was chosen to be $\delta = 2^{-64}$, and the targeted completeness, or success rate, was set to $\epsilon = 0.97725$, corresponding to a Gaussian 2σ probability of a particular protocol run successfully passing. After evaluating the calibration data to determine the anticipated performance of the experiment, for the extended protocol the robustness threshold was fixed at $r_{\rm th} = 8 \times 10^{-6}$. We determined that for these parameters, 30 million trials (two minutes of data) were required per run of the basic protocol, and 60 million trials (four minutes of data) were needed per run of the extended protocol.

The robustness threshold of $r_{\rm th} = 8 \times 10^{-6}$ can be compared to the robustness per trial 2×10^{-3} of the entangled state designed to be emitted by the source. The latter robustness accounts for the probability of about 1/350 that a photon pair is emitted in a trial (see Supplementary Information Sect. III B). The robustness of entanglement of a Bell state is 1. With a version of the "garden-hose" attack [9] given in Supplementary Information Sect. V G, it is possible to break the security of the experimentally implemented version of Bell-test QPV with adversaries sharing two pairs of qubits, each with robustness of entanglement 2×10^{-3}. As explained above, for security against adversaries with more quantum resources than used in the experiment, it suffices to increase the number of classical challenge bits in each trial.

After running the two series of QPV experiments, the collected data were compressed and placed in blind storage prior to analysis. Occasional errors where our single-photon

nanowire detectors failed to properly reset, and therefore were not light sensitive for one or more of the trials during a chunk, were flagged. On both day 1 and day 2, the first 10 minutes of data free of detector errors were used to construct the initial test factor W. The trials from the subsequent two or four minutes, depending on whether running the basic or extended protocol, and irrespective of detector errors, were then un-blinded and analyzed to determine the product P of the values of W witnessed by each trial record. A product $P \geq 2^{64}$ results in a pass; otherwise, that instance is considered a fail. Subsequent instances were processed sequentially using the next two or four minutes of data if available. To track long-term drifts in the experiment, we updated the test factor W before each subsequent instance, using the preceding 10 minutes of processed data without detector errors. This yielded 232 runs of the basic protocol on day 1 and 103 runs of the extended protocol on day 2. Of the 232 evaluated for the basic analysis, 224 passed (96.5%), and of the 103 evaluated for the extended analysis, 102 succeeded (99.0%). Both of these agree well with the expected completeness of $\epsilon = 0.97725$, and are shown in Fig. 4.

In each of the successful instances, the target three-dimensional volume in which a prover must perform a quantum operation is 4.53(5) times smaller than the volume they would need to be in to pass a comparable classical protocol. This is a measure of our quantum position verification advantage ($\mathrm{QA} = E/E_{\mathrm{adv}}$, see above). This comparison can also be made in two-dimensions, $\mathrm{QA}_{\mathrm{area}} = 4.02(3)$, and in one-dimension, $\mathrm{QA}_{\mathrm{length}} = 4.48(2)$. For intuition on our metric, consider that in Fig. 1 for example, the (QA) is the ratio of the total volume of the blue and yellow regions to that of the yellow region, where the figure shows the two-dimensional projections of these regions. Finally, we can also compare with the ideal classical protocol in one dimension where the target region is the line segment between the two verifiers. In this case the advantage is $\mathrm{QA}_{\mathrm{length-ideal}} = 2.47(2)$. See Sect. IV B in the Supplementary Information for more details. Prior theory works [1, 12] often consider an ideal version of this protocol. In this case, the quantum target region reduces to a point, the best classical target region reduces to the line joining the verifiers, and the QA tends to infinity.

CONCLUSION

This work presents the first complete demonstration of a quantum position verification protocol—a major step forward in the development of secure quantum cryptographic primitives. Our implementation achieves strong localization guarantees and rules out powerful classes of adversaries without relying on assumptions about the internal functioning of the quantum devices involved. We have shown that Bell-test QPV is scalable and robust. Because the source is untrusted with no restriction on speed of transmission of qubits, transmission loss can be tolerated provided that trials are heralded via entanglement swapping or commitment. The effect of all noise sources including detector inefficiencies on completeness is the same as that on Bell tests: It suffices for the Bell test to be successful for the honest prover to be able to pass the protocol [53]. For future scalability, entangled photon pairs can be provided by quantum repeaters. No source of noise can affect soundness, and neither do adversary actions such as interference with communication and calibrations, including calibrations used to choose the test factors. Such actions at worst result in denial of service, an issue that is faced by all communication protocols.

While the protocol as demonstrated already offers robust security within minutes, there are several areas in which future experiments can be improved. At present, the Pockels cells

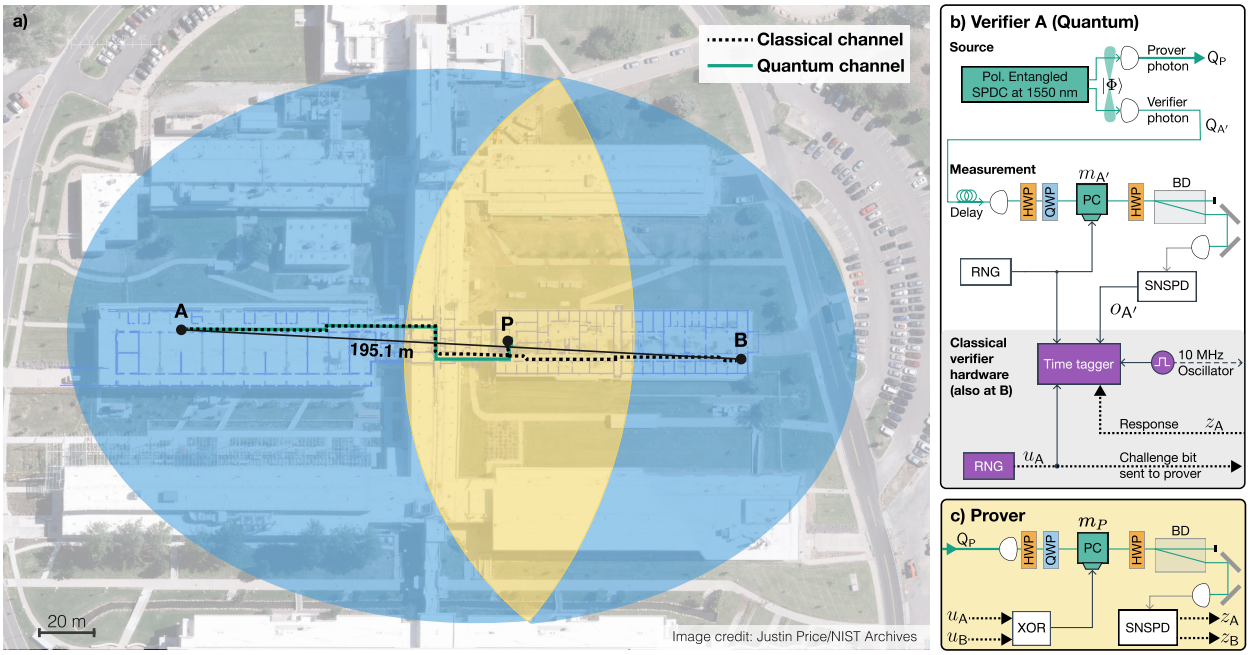

FIG. 2. Figure depicting the experimental setup and verifier and prover stations used for our demonstration. a) The quantum target region E in yellow, and the classical target region of a comparable classical protocol in blue overlaid on an aerial photograph of the NIST building, where the experiment is housed. b) Schematic of the hardware at verifier A. From an untrusted entangled photon pair source, one photon of the pair is sent to prover P, and the other to the measurement station A′ at verifier A. At measurement station A′, the incoming photon is delayed in a fiber loop and measured by a superconducting nanowire detector in a polarization basis determined by a fast Pockels cell and waveplates. The setting is chosen via a random number generator (RNG) [58], and the outcome is recorded on a time-tagger. c) Measurement station at the prover P. The choice of measurement basis here is set via an XOR of the bits u_{A} and u_{B} from the verifiers. The outcome of this measurement is sent back to both verifiers, who record it on their timetaggers. SNSPD: Superconducting nanowire single photon detector; XOR: Exclusive OR; PC: Pockels cell; HWP: Half-wave plate; QWP: Quarter-wave plate.

used to perform the measurements can operate only at 250 kHz due to the large switching voltages required, whereas all other components in the system can support trial rates of up to 100 MHz. Faster active measurement devices could reduce run times by two orders of magnitude, leading to position-verification times on the order of one second while maintaining the same security parameters.

Our localization accuracy is currently limited by transmission latencies as well as measurement processing time at the prover station. For larger distances on the metropolitan scale, low-latency line-of-sight classical communication links will be required, using either laser communications or wireless technologies. Both technologies are mature and should be readily adaptable to larger distances. The line-of-sight requirement for ranging means that transmitting signals through buried cables, whether high-velocity-factor cables or hollow-core fibers, will not be feasible in most situations. Buried cables rarely follow line-of-sight paths, leading to significant geometry-induced latencies. Using custom low-latency circuits and FPGAs, it should be possible with existing classical technology to reduce the com-

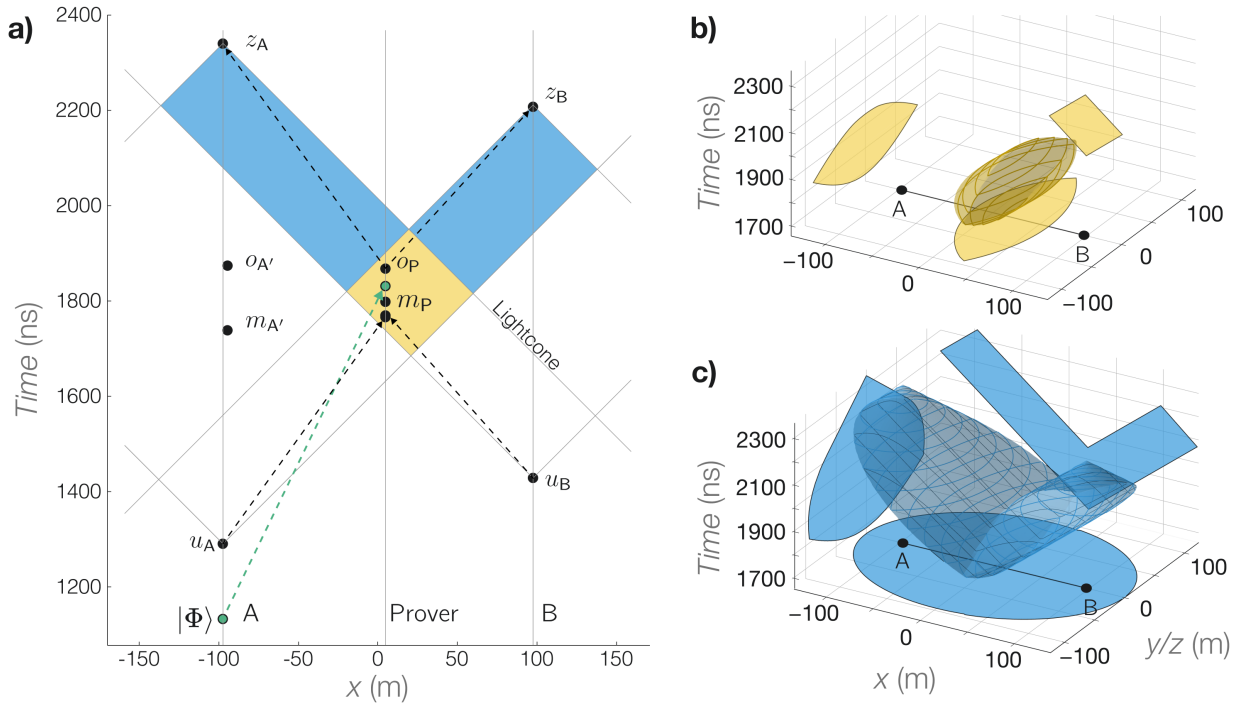

FIG. 3. Figure depicting the spacetime target regions in our demonstration. a) The two-dimensional spacetime slice, where the spatial dimension is along the line joining the provers, here assigned the x-axis. Both the quantum (yellow), and comparable classical (blue) regions are completely defined by forward and backward lightcones from the spacetime events corresponding to the release of u_{A} and u_{B} and the last moments when z_{A} and z_{B} are accepted. Events $o_{\mathsf{A}'}$ and $m_{\mathsf{A}'}$ occur at A', which is slightly displaced from A. The region in blue is also the region from which adversaries with sufficient quantum resources can successfully attack the quantum protocol. Subfigures b) and c) show three-dimensional projections of the quantum and classical target regions, with the time dimension vertically up on the page, and the spatial dimensions parallel and perpendicular to the verifier axis depicted in the other two directions. The regions are all rotationally symmetric about the verifier axis, and so are identical in the spatial dimensions (y and z) perpendicular to the verifier axis. At the displayed scale, error bars indicating uncertainty on the positions and timings are too small to be discernible. A full accounting of uncertainties and their effect on our results is detailed in the Supplementary Information.

bined transmission and measurement latency to below 30 ns, even over distances of many kilometers. This would lead to prover regions that are well contained within a large room or a single building along the verifier axis. Such low-latency classical communications will also be valuable for future quantum repeater networks, as they can significantly reduce the feed-forward time between repeater nodes, thereby reducing the storage times required by quantum memories.

Finally, to exclude adversaries possessing more entanglement than that prepared by the source, we have shown that it suffices to increase the number of bits used for the challenges $u_{\mathsf{A}}, u_{\mathsf{B}}$. Recent work in [22] has demonstrated the feasibility of 20-bit challenges for each of u_{A} and u_{B}, together with random functions $f(u_{\mathsf{A}}, u_{\mathsf{B}})$. Adapting these classical processing techniques will enable systems based on our protocol to safeguard against more powerful

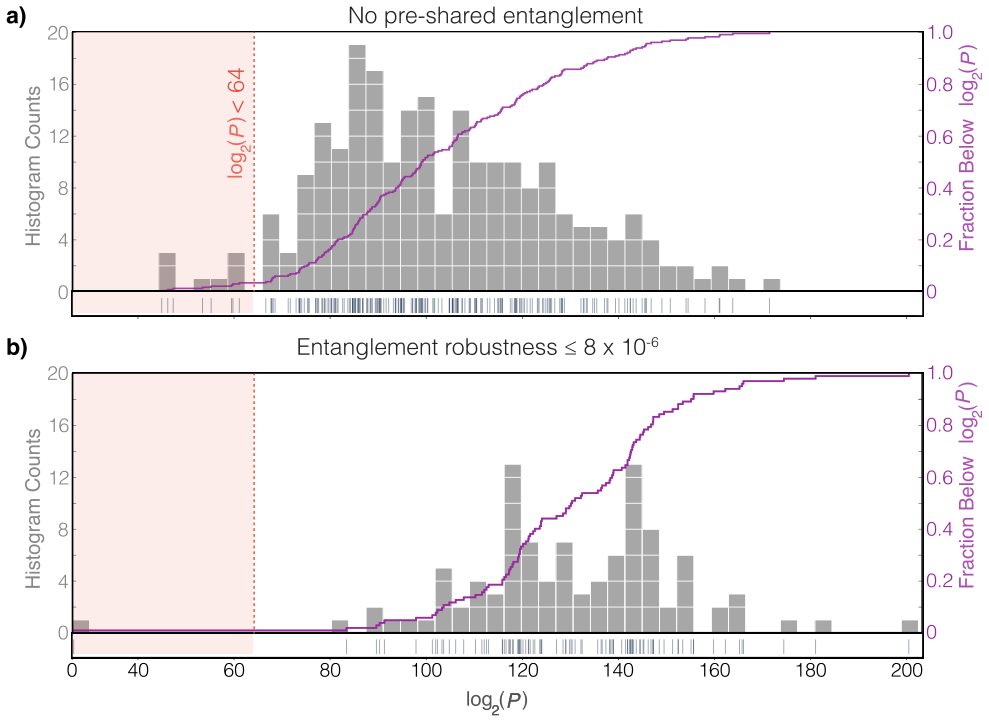

FIG. 4. Plots of results from our experimental runs. Two sets of protocol instances were run as part of our demonstration: a) is day 1, analyzed for security against adversaries pre-sharing no entanglement; and b) is day 2, analyzed against adversaries pre-sharing entanglement such that the average robustness over an instance's trials is upper bounded by $r_{\text{th}} = 8 \times 10^{-6}$. The soundness target for all instances was $\delta = 2^{-64}$. Plotted are histograms of the data binned based on the calculated $\log_2(P)$, where P is the product of the test-factor values over the trials in a protocol instance. The $\log_2(P)$ value for each instance is indicated with a tick directly above the x-axis and below the histograms on both plots. Any instances with a $\log_2(P)$ less than $-\log_2(\delta) = 64$ are declared failures, and correspond to ticks inside the red regions. Also plotted in purple is the experimental fraction of instances below $\log_2(P)$ as a function of $\log_2(P)$. At a higher $\log_2(P)$, which corresponds to a stricter choice for the soundness, a larger fraction of the data would be considered a failure, reducing success probability. The plots both share the same x-axis.

quantum adversaries.

While there is still much work to be done before QPV systems can be widely deployed in real-world settings, our work makes significant progress toward demonstrating the requirements for a complete protocol capable of achieving a quantum advantage in a nontrivial distributed environment.

METHODS

Assumptions

Completeness and soundness of Bell-test QPV are established under the assumptions detailed in Supplementary Information Sect. II. The assumptions are standard for device-

independent protocols [36–38] and include secure classical computing, known verifier laboratory locations, and sufficiently accurate event timing. The verifier laboratories must be physically secured against unauthorized access. According to device-independence, the quantum measurement apparatus and the entangled photon source are untrusted: no assumptions are made about their internal operation. Like other device-independent protocols [40], we require shielding of the measurement station's external interface. In the no-prior-entanglement setting, it is desirable to add shielding to the source to prevent emissions outside defined time windows. For significant prior entanglement, source contributions are accounted for by the adversary entanglement bounds, which exceed the entanglement required by the verifiers and honest prover.

Protocol specific assumptions include conditions on measurement settings and challenge distributions, defined target regions, trial boundaries, and specified classes of adversaries. Verifier settings and challenges are drawn from fixed distributions satisfying two independence conditions: (1) the prover setting M_{P} and verifier setting $M_{\mathsf{A}'}$ are independent of A's challenge U_{A}; and (2) $M_{\mathsf{A}'}$ is conditionally independent of both challenges given M_{P}. These hold, for example, when the challenges are uniform and independent of $M_{\mathsf{A}'}$, and the settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ is chosen appropriately.

Depending on the settings function and test-factor construction, Bell-test QPV is sound against different classes of adversaries. For all classes, adversaries can perform instantaneous computation and speed-of-light communication. Each class is defined by a set of assumptions on adversaries required for soundness. All classes of adversaries allow for unlimited classical communication. For the simplest class, we assume no prior entanglement but allow unlimited quantum communication. For classes of adversaries with significant prior entanglement quantum communication is replaced by quantum teleportation and contributes to the prior entanglement. The maximum amount of prior entanglement allowed while ensuring protocol soundness is determined by the number of challenge bits and the settings function used in each trial. See the completeness and soundness results below for details.

Completeness and Soundness Results

Protocol completeness: *For all $\epsilon \in (0, 1)$, Bell-test QPV is ϵ-complete. To witness completeness it suffices for the verifiers and honest prover to violate local realism for some Bell test configuration with two measurement settings for the prover. For every class of adversaries under consideration and a given violation of local realism, the number of trials required for completeness ϵ and soundness δ is $O(\log(1/\delta)) + O(\log(1/(1-\epsilon)))$.*

The proof of this result is in Supplementary Information Sect. V C. Because there is no restriction on how the entangled state for the Bell-test is obtained by the verifiers and prover before each trial, completeness is preserved if a slow quantum network is used to prepare these states. The quantum link may be heralded and can support prover commitment strategies [16]. Therefore loss in photonic links is well tolerated. See Supplementary Information Sect. V F for how the protocol supports heralded sources and commitment.

The soundness results depend on the class of adversaries under consideration. In each case, soundness holds for all spatial dimensions, and the source is untrusted and may be controlled by one of the adversaries. Classical communication by the adversaries is always unconstrained. The first soundness result involves the class of adversaries with no prior entanglement.

Protocol soundness against adversaries with no prior entanglement: *For all $\delta > 0$,*

Bell-test QPV is δ-sound against adversaries who have no prior entanglement for each trial. The adversaries can freely use classical and quantum computation and communication during each trial.

The proof is contained in Supplementary Information Sect. V B, Sect. V C and Sect. V D. The first step of the proof follows the traditional path of reducing the adversary attacks to those involving two adversaries on opposite sides of the target region [10]. Bell-test QPV requires an analysis that is independent of spatial dimension, takes into account the time required by the honest prover to perform its measurement task, and supports device-independence with test factors. For this we adapt the spacetime circuit analysis introduced in Ref. [50] and fully formalize the reduction to two adversaries. The subsequent analysis of adversary attacks deviates from the traditional path in order to reduce the statistics of adversarial trial records to a three-party non-signaling scenario for which we can construct test factors.

The absence of prior entanglement for the above soundness result implies a constraint on the location of the source entanglement at the trial boundary: The source entanglement must be located on the side closer to verifier A', which requires that we prevent the quantum system destined for the prover from leaking to the other side. This constraint on the location of the source entanglement is mitigated by the next soundness result and absent in the last two soundness results.

The second soundness result allows for small amounts of prior entanglement.

Protocol soundness against adversaries with small amounts of prior entanglement: *For all $\delta > 0$, the Bell-test protocol of Sect. Sect. II is δ-sound against adversaries with small amounts of prior entanglement for each trial or on average over the trials. The adversaries can freely use classical and quantum computation and communication during each trial.*

This is proven in Supplementary Information Sect. V E. The amount of entanglement allowed depends on the quality of the Bell test used by the verifiers and prover. The amount of prior entanglement is quantified either as maximum robustness of entanglement over trials, or as average robustness of entanglement over trials. For the experiment, the upper bounds on allowed prior entanglement are smaller than the entanglement created by the source, limiting the strength of the result. The technique used here can be viewed as a quantified strengthening of Lemma 5.3 of Ref. [45]. It is also related to the entanglement forging method of Ref. [59]. Neither our above bounds nor those based on Ref. [45] have so far led to bounds on prior entanglement of adversaries that are higher than the resources used by the verifiers and honest prover when quantum computation and communication during a trial is unrestricted.

To prove soundness against adversaries with significant amounts of prior entanglement requires increasing the complexity of the settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$. For the next two results, we set $f(u_{\mathsf{A}}, u_{\mathsf{B}}) = g_0(u_{\mathsf{A}} \oplus u_{\mathsf{B}})$, where g_0 is the semi-bent function $g_0(z) = z_0 + \sum_{i=1}^{l-1} z_i\, z_{i+l-1} \mod 2$ and the bit-length l of u_{A} and u_{B} is odd. The semi-bent function is explicit, efficiently computable, and therefore scalable.

Protocol soundness against adversaries with logarithmic prior entanglement: *There are efficiently implementable settings functions $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ such that for all $\delta > 0$, Bell-test QPV is δ-sound against adversaries for whom the sum of the number of qubits quantum communicated and the number of qubits used for prior entanglement is bounded by $\log(l)/2 - O(1)$. The adversaries may use arbitrary local quantum computation and classical communication during each trial.*

The proof is in Supplementary Information Sect. V G. Since quantum communication can be replaced by prior entanglement and classical communication, the soundness can be interpreted as being against adversaries with $\log(l)/2 - O(1)$ qubits of prior entanglement and no quantum communication. To support practical applications, we fully quantify the allowed prior-entanglement bound. To maintain a given soundness and completeness, it suffices to increase the number of trials by a constant factor. We choose a factor of 6 in the examples mentioned in the main text. The suppressed constant in the bound depends on the test factor used and its gain. See the paragraphs starting at Eq. (54) in Supplementary Information Sect. V G. The proof of this soundness result is inspired by the proof of Prop. 6.1 in Ref. [12]. A specific analysis of the relationship between adversarial test-factor expectations and guessing probabilities for the settings function is required to efficiently preserve completeness under all circumstances. A benefit of our result is that there is no constraint on classical communication, and there are no added constraints on honest-prover performance. Refs. [12, 14] show non-constructively that there are device-dependent protocols where the verifiers and prover require one qubit, but adversaries need at least a linear total amount of quantum memory, classical memory and classical communication to pass the protocol. However, it is an open problem to have a better than logarithmic gap between the qubit resource required by the verifiers and honest prover, and the adversary prior entanglement in qubits, when the classical memory and communication accessible by adversaries are unrestricted. For Bell-test QPV, there are similar open questions.

Instead of bounding prior entanglement, we can bound qubits and gates used by the adversary nearer B.

Protocol soundness against adversaries with linear quantum operations: *There are efficiently implementable settings functions $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ such that for all $\delta > 0$, Bell-test QPV is δ-sound against adversaries with arbitrary classical communication but no quantum communication, and where the adversary nearer A can use arbitrary local quantum resources, but the adversary nearer B is restricted to $O(l/(\log(q) + 1))$ quantum operations in a basic universal set, where q is the number of qubits used by this adversary.*

The proof of this result is based on the proof of the previous soundness result and given at the end of Supplementary Information Sect. V G. Again, to support practical applications, we fully quantify the bound on quantum resources allowed by adversaries. The proof is motivated by the argument given in Ref. [15] to prove a similar result for BB84 and routing QPV protocols. The suppressed constant in the bound $O(l/(\log(q) + 1))$ depends on the size of the basic universal set of quantum operations, which includes qubit measurements.

Target-region comparison

To benchmark our demonstration's performance, we compare the size of the target region in our demonstration to the smallest achievable target region with a classical protocol having the same latencies. In the classical protocol illustrated in Fig. 1 a), verifiers A and B start by broadcasting challenge bits u_{A} and u_{B} at times s_{A} and s_{B}, timed to arrive simultaneously at some position between them. Appropriately positioned provers are assumed to have access to these bits as soon as they are released. An honest prover P, or adversaries, can receive the challenge bits and use them as inputs to a predetermined function $g(u_{\mathsf{A}}, u_{\mathsf{B}})$ that is publicly known. The value of this function determines the specific measurement to be performed. The measurement outcome is then sent back to A and B, who receive z_{A} and z_{B}, respectively, and jointly validate the results.

The verifiers choose the target region by deciding when they send challenges and accept responses. There are four relevant times: s_{A} and s_{B} are the times when A and B send their challenge bits, and r_{A} and r_{B} are the latest times when the verifiers accept the responses. This procedure is repeated multiple times to build confidence in the honest running of the protocol. Repeated late responses, or trials where $z_{\mathsf{A}} \neq z_{\mathsf{B}}$ will result in failure with high likelihood. We show in the Supplementary Information Sect. V B that target regions for the classical protocol and our quantum protocol are determined by the speed of light c and the four time intervals

$$\Delta\tau_1 = (r_{\mathsf{A}} - s_{\mathsf{A}})/2, \tag{1}$$
$$\Delta\tau_2 = (r_{\mathsf{B}} - s_{\mathsf{B}})/2, \tag{2}$$
$$\Delta\tau_3 = r_{\mathsf{A}} - s_{\mathsf{B}}, \tag{3}$$
$$\Delta\tau_4 = r_{\mathsf{B}} - s_{\mathsf{A}}. \tag{4}$$

The impossibility of superluminal signaling constrains the locations from which a stationary prover can receive u_{A} and reply to verifier A to points in a sphere $\mathsf{S}(\mathsf{A}, c\Delta\tau_1)$ of radius $c \times \Delta\tau_1$ centered on A. Similarly, $c \times \Delta\tau_2$ is the radius of an analogous sphere $\mathsf{S}(\mathsf{B}, c\Delta\tau_2)$ centered on B. The set of all locations from which a stationary prover can receive u_{B} (released at s_{B}) and respond to verifier A is a prolate ellipsoid $\mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_3)$ with the verifier A and B locations as foci, and major axis $c \times \Delta\tau_3$. Prolate ellipsoids are the set of all points such that the sum of the distances from the verifiers is less than or equal to the major axis. This means any prover outside this shape could not receive u_{B} and reply to A within r_{A}. An analogous ellipsoid $\mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_4)$ constrains receiving u_{A} and replying to B.

If the verifiers can ensure that they only ever interact with a single prover, the target region is the intersection of all the regions listed above, $E = \mathsf{S}(\mathsf{A}, c\Delta\tau_1) \cap \mathsf{S}(\mathsf{B}, c\Delta\tau_2) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_3) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_4)$. A projection of this region to a 2-D plane for a particular choice of release and receive times is indicated in yellow in Fig. 1 d). This single-prover target region could be made arbitrarily small if one could approach zero-latency processing and direct speed-of-light transmission. However, the verifiers have no way to enforce the single-prover assumption in practice, and must allow for multiple interacting provers. As we show in the Supplementary Information Sect. V B, multiple adversaries can use the "classical spoofing" protocol, shown in Fig. 1 b), to pass the protocol, provided they have access to the larger region $E_{\text{adv}} = \Big(\mathsf{S}(\mathsf{A}, c\Delta\tau_1) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_3)\Big) \cup \Big(\mathsf{S}(\mathsf{B}, c\Delta\tau_2) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_4)\Big) \supseteq E$. This spoofing is indistinguishable from an honest single-prover strategy from the verifiers' point of view. A 2-D projection of E_{adv} is indicated with blue in Fig. 1 c). The target region for the classical protocol E_{adv} precludes remote position verification because the target region includes the verifier positions themselves in addition to the space between the verifiers. While arbitrarily small target regions that include the verifier positions are always possible with straightforward ranging [60], no classical protocols admit remote position verification [1]

Note on Uncertainties

All of the uncertainties reported in the main and supplementary texts are standard uncertainties (1-sigma). When we report a quantity with uncertainty as (for example) 774.3(2) nm, the number in parentheses is the numerical value of the combined standard uncertainty referred to the corresponding last digits of the quoted result.

USE OF AI TOOLS

AI tools were used in editing the text to help improve readability, to assist in producing schematics, and as coding assistants for implementing analysis code. The authors thoroughly reviewed text, schematics and code produced by the AI tools and take full responsibility for the final results.

DATA AVAILABILITY

All experimental data from the quantum position verification protocol shown in this manuscript are available upon request from the corresponding authors.

ACKNOWLEDGMENT

The authors would like to thank Peter Bierhorst and Michael Grayson for useful discussions, Yi-Kai Liu, Adam McCaughan and Aneesh Ramaswamy for close reading of the manuscript, and Daniel Sorensen for help with cable routing. The authors acknowledge support from the National Science Foundation QLCI OMA-2016244, National Science Foundation NQVL:QSTD:Pilot 2435378, the National Science Foundation ECCS 2330228, the National Science Foundation OSI 2328800, the University of Colorado Quantum Engineering Initiative, the National Institute of Standards and Technology, and the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research (Field Work Proposal ERKJ381). Yusuf Alnawakhtha acknowledges support from the Crown Prince International Scholarship Program and the Arabian Gulf University.

DISCLAIMER

Any mention of commercial products within the manuscript is for information only; it does not imply recommendation or endorsement by NIST. This work includes contributions of the National Institute of Standards and Technology, which are not subject to U.S. copyright. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation thereon.

Additionally, this manuscript has been co-authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

DISCLOSURES

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: G.A.K., L.K.S., E.K., Y.Z., S.G., M.J.S., H.F., C.A.M., Y.A., A.R.G., S.P., J.C.B.; Investigation: G.A.K., A.R.G., L.K.S., M.J.S., M.D.M; Methodology: Y.Z., G.A.K., L.K.S, E.K., S.G.; Software: G.A.K., Y.Z., L.K.S.; Formal analysis: Y.Z., S.P., E.K., G.A.K., S.G.C., L.K.S.; Security analysis: E.K., Y.Z., C.A.M., Y.A.; Funding acquisition: L.K.S., R.P.M., M.J.S.; Project administration: L.K.S., E.K., M.J.S.; Resources: D.V.R., J.C.B., C.A., W.A., M.W.M.; Supervision: L.K.S., E.K., S.G., M.J.S., R.P.M.; Visualization: G.A.K., L.K.S., E.K., Y.Z.; Writing – original draft: G.A.K, E.K., L.K.S., Y.Z., S.G.; Writing – review & editing: G.A.K, E.K., Y.Z., S.G, L.K.S., S.P., A.R.G., H.F., Y.A., M.W.M., C.A.M.

[1] N. Chandran, V. Goyal, R. Moriarty, and R. Ostrovsky, Position-Based Cryptography, SIAM Journal on Computing **43**, 1291 (2014).

[2] S. Bhamidipati, T. Y. Mina, and G. X. Gao, GPS time authentication against spoofing via a network of receivers for power systems, in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)* (IEEE, Monterey, CA, 2018) pp. 1485–1491.

[3] F. Ardizzon, L. Crosara, N. Laurenti, S. Tomasin, and N. Montini, Authenticated Timing Protocol Based on Galileo ACAS, Sensors **22**, 6298 (2022).

[4] N. O. Tippenhauer, K. B. Rasmussen, C. Pöpper, and S. Čapkun, Attacks on public WLAN-based positioning systems, in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services* (ACM, Kraków Poland, 2009) pp. 29–40.

[5] S. Brands and D. Chaum, Distance-Bounding Protocols, in *Advances in Cryptology — EUROCRYPT '93*, edited by T. Helleseth (Springer, Berlin, Heidelberg, 1994) pp. 344–359.

[6] S. Drimer and S. J. Murdoch, Keep your enemies close: Distance bounding against smartcard relay attacks, in *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, SS'07 (USENIX Association, USA, 2007) pp. 1–16.

[7] R. A. Malaney, Location-dependent communications using quantum entanglement, Phys. Rev. A **81**, 042319 (2010).

[8] A. Kent, W. J. Munro, and T. P. Spiller, Quantum tagging: Authenticating location via quantum information and relativistic signaling constraints, Phys. Rev. A **84**, 012326 (2011).

[9] H. Buhrman, S. Fehr, C. Schaffner, and F. Speelman, The garden-hose model, in *Proceedings of the 4th conference on Innovations in Theoretical Computer Science* (2013) pp. 145–158.

[10] H. Buhrman, N. Chandran, S. Fehr, R. Gelles, V. Goyal, R. Ostrovsky, and C. Schaffner, Position-Based Quantum Cryptography: Impossibility and Constructions, SIAM Journal on Computing **43**, 150 (2014).

[11] C. C. W. Lim, F. Xu, G. Siopsis, E. Chitambar, P. G. Evans, and B. Qi, Loss-tolerant quantum secure positioning with weak laser sources, Physical Review A **94**, 032315 (2016).

[12] A. Bluhm, M. Christandl, and F. Speelman, A single-qubit position verification protocol that is secure against multi-qubit attacks, Nature Physics **18**, 623 (2022).

[13] R. Allerstorfer, H. Buhrman, F. Speelman, and P. V. Lunel, Towards practical and error-robust quantum position verification (2022), arXiv:2106.12911 [quant-ph].

[14] L. Escolà-Farràs and F. Speelman, Single-Qubit Loss-Tolerant Quantum Position Verification Protocol Secure against Entangled Attackers, Physical Review Letters **131**, 140802 (2023).

[15] V. Asadi, R. Cleve, E. Culf, and A. May, Linear gate bounds against natural functions for position-verification, Quantum **9**, 1604 (2025).

[16] R. Allerstorfer, A. Bluhm, H. Buhrman, M. Christandl, L. Escolà-Farràs, F. Speelman, and P. Verduyn Lunel, Making existing quantum position verification protocols secure against arbitrary transmission loss, Physical Review Letters **135**, 260801 (2025).

[17] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nature Photonics **4**, 686 (2010).

[18] H. Weier, H. Krauss, M. Rau, M. Fürst, S. Nauerth, and H. Weinfurter, Quantum eavesdropping without interception: An attack exploiting the dead time of single-photon detectors, New Journal of Physics **13**, 073024 (2011).

[19] F. Kaleoglu, M. Liu, K. Chakraborty, D. Cui, O. Amer, M. Pistoia, and C. Lim, On the Equivalence between Classical Position Verification and Certified Randomness (2025), arXiv:2410.03982 [quant-ph].

[20] K. Kanneworff, M. Poortvliet, D. Bouwmeester, R. Allerstorfer, P. V. Lunel, F. Speelman, H. Buhrman, P. Steindl, and W. Löffler, Towards experimental demonstration of quantum position verification using true single photons (2025), arXiv:2502.04125 [quant-ph].

[21] M. Liu, P. Niroula, M. DeCross, C. Foreman, W. Y. Kon, I. W. Primaatmaja, M. S. Allman, J. P. Campora, A. Isanaka, K. Singhal, O. Amer, S. Chakrabarti, K. Chakraborty, S. F. Cooper, R. D. Delaney, J. M. Dreiling, B. Estey, C. Figgatt, C. Foltz, J. P. Gaebler, A. Hall, Z. He, C. A. Holliman, T. S. Humble, S.-H. Hung, A. A. Husain, Y. Jin, F. Kaleoglu, C. J. Kennedy, N. Kotibhaskar, N. K. Lysne, I. S. Madjarov, M. Mills, A. R. Milne, K. Milner, L. Narmour, S. Omanakuttan, A. J. Park, M. A. Perlin, A. P. Reed, C. N. Self, M. Steinberg, D. T. Stephen, J. Sullivan, A. Chernoguzov, F. J. Curchod, A. Ransford, J. G. Bohnet, B. Neyenhuis, M. Foss-Feig, R. Otter, and R. Shaydulin, Certified randomness amplification by dynamically probing remote random quantum states (2025), arXiv:2511.03686 [quant-ph].

[22] G.-J. Fan-Yuan, Y.-G. Shan, C. Zhang, Y.-L. Wang, Y.-X. Fan, W.-X. Xie, D.-Y. He, S. Wang, Z.-Q. Yin, W. Chen, *et al.*, Relativistic position verification with coherent states (2026), arXiv preprint arXiv:2602.01787.

[23] 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. Abellán, 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, Physical Review Letters **115**, 250402 (2015).

[24] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. 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, Physical Review Letters **115**, 250401 (2015).

[25] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[26] A. Alabdulatif, R. Samarasinghe, and N. N. Thilakarathne, A Novel Robust Geolocation-Based Multi-Factor Authentication Method for Securing ATM Payment Transactions, Applied Sciences **13**, 10743 (2023).

[27] D. E. Denning and P. F. MacDoran, Location-based authentication: Grounding cyberspace for better security, Computer Fraud & Security **1996**, 12 (1996).

[28] F. Zhang, A. Kondoro, and S. Muftic, Location-Based Authentication and Authorization Using Smart Phones, in *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (IEEE, Liverpool, United Kingdom, 2012) pp. 1285–1292.

[29] M. A. Alawami and H. Kim, LocAuth: A fine-grained indoor location-based authentication system using wireless networks characteristics, Computers & Security **89**, 101683 (2020).

[30] R. Langner, Stuxnet: Dissecting a Cyberwarfare Weapon, IEEE Security & Privacy Magazine **9**, 49 (2011).

[31] M. E. Whitman and H. J. Mattord, *Principles of Information Security*, 3rd ed. (Thomson Course Technology, Boston, Mass, 2009).

[32] R. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems*, third edition ed. (John Wiley & Sons, Indianapolis, 2020).

[33] I. Fernández-Hernández, V. Rijmen, G. Seco-Granados, J. Simon, I. Rodríguez, and J. D. Calle, A Navigation Message Authentication Proposal for the Galileo Open Service: NMA proposal for Galileo OS, Navigation **63**, 85 (2016).

[34] G. Hancke and M. Kuhn, An RFID Distance Bounding Protocol, in *First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05)* (IEEE, Athens, Greece, 2005) pp. 67–73.

[35] G. Cowperthwaite, A. Kent, and D. Pitalua-Garcia, Towards practical quantum position verification (2023), arXiv:2309.10070 [quant-ph].

[36] L. K. Shalm, Y. Zhang, J. C. Bienfang, C. Schlager, M. J. Stevens, M. D. Mazurek, C. Abellán, W. Amaya, M. W. Mitchell, M. A. Alhejji, H. Fu, J. Ornstein, R. P. Mirin, S. W. Nam, and E. Knill, Device-independent randomness expansion with entangled photons, Nature Physics **17**, 452 (2021).

[37] G. A. Kavuri, J. Palfree, D. V. Reddy, Y. Zhang, J. C. Bienfang, M. D. Mazurek, M. A. Alhejji, A. U. Siddiqui, J. M. Cavanagh, A. Dalal, C. Abellán, W. Amaya, M. W. Mitchell, K. E. Stange, P. D. Beale, L. T. A. N. Brandão, H. Booth, R. Peralta, S. W. Nam, R. P. Mirin, M. J. Stevens, E. Knill, and L. K. Shalm, Traceable random numbers from a non-local quantum advantage, Nature **642**, 916 (2025).

[38] W.-Z. Liu, M.-H. Li, S. Ragy, S.-R. Zhao, B. Bai, Y. Liu, P. J. Brown, J. Zhang, R. Colbeck, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent randomness expansion against quantum side information, Nature Physics **17**, 448 (2021).

[39] M.-H. Li, X. Zhang, W.-Z. Liu, S.-R. Zhao, B. Bai, Y. Liu, Q. Zhao, Y. Peng, J. Zhang, Y. Zhang, W. J. Munro, X. Ma, Q. Zhang, J. Fan, and J.-W. Pan, Experimental Realization of Device-Independent Quantum Randomness Expansion, Physical Review Letters **126**, 050503 (2021).

[40] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, Experimental quantum key distribution certified by Bell's theorem, Nature **607**, 682 (2022).

[41] B.-W. Lu, C.-W. Yang, R.-Q. Wang, B.-F. Gao, Y.-Z. Zhen, Z.-G. Wang, J.-K. Shi, Z.-Q. Ren, T. A. Hahn, E. Y.-Z. Tan, X.-P. Xie, M.-Y. Zheng, X. Jiang, J. Zhang, F. Xu, Q. Zhang, X.-H. Bao, and J.-W. Pan, Device-independent quantum key distribution over 100 km with single atoms, Science **391**, 592 (2026).

[42] A. Kulikov, S. Storz, J. D. Schär, M. Sandfuchs, R. Wolf, F. Berterottière, C. Hellings, A. Wallraff, and R. Renner, Experimental randomness amplification, Nature **653**, 1033 (2026).

[43] W.-Z. Liu, Y.-B. Zhou, J.-P. Chen, B. Wang, A. Teng, X.-W. Han, G.-C. Liu, Z.-J. Zhang, Y. Yang, F.-G. Liu, C.-H. Xue, B.-W. Yang, J. Yang, C. Zeng, D.-R. Pan, M.-Y. Zheng, X.-J. Zhang, C. Shen, Y.-Z. Zhen, Y. Xiao, H. Li, L.-X. You, X.-F. Ma, Q. Zhao, F. Xu, Y. Wang, Y. Wan, Q. Zhang, and J.-W. Pan, Long-lived remote ion-ion entanglement for scalable quantum repeaters, Nature **652**, 51 (2026).

[44] G. Vidal and R. Tarrach, Robustness of entanglement, Physical Review A **59**, 141 (1999).

[45] S. Beigi and R. König, Simplified instantaneous non-local quantum computation with applications to position-based cryptography, New Journal of Physics **13**, 093036 (2011).

[46] Ll. Masanes, A. Acin, and N. Gisin, General properties of nonsignaling theories, Physical Review A **73**, 012112 (2006).

[47] Y. Zhang, S. Glancy, and E. Knill, Asymptotically optimal data analysis for rejecting local realism, Phys. Rev. A **84**, 062118/1 (2011), for supporting code, see arXiv:1108.2468.

[48] V. Vovk and R. Wang, E-values: Calibration, combination and applications, The Annals of Statistics **49**, 1736 (2021).

[49] P. Grünwald, R. de Heide, and W. Koolen, Safe testing, Journal of the Royal Statistical Society Series B: Statistical Methodology **86**, 1091 (2024).

[50] D. Unruh, Quantum position verification in the random oracle model, in *Advances in Cryptology–CRYPTO 2014: 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II 34* (Springer, 2014) pp. 1–18.

[51] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, A monogamy-of-entanglement game with applications to device-independent quantum cryptography, New Journal of Physics **15**, 103002 (2013).

[52] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, Experimental violation of a Bell's inequality with efficient detection, Nature **409**, 791 (2001).

[53] P. H. Eberhard, Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment, Physical Review A **47**, R747 (1993).

[54] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015).

[55] 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, Physical Review Letters **119**, 010402 (2017).

[56] 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).

[57] S. Storz, J. Schär, A. Kulikov, P. Magnard, P. Kurpiers, J. Lütolf, T. Walter, A. Copetudo, K. Reuer, A. Akin, J.-C. Besse, M. Gabureac, G. J. Norris, A. Rosario, F. Martin, J. Martinez, W. Amaya, M. W. Mitchell, C. Abellan, J.-D. Bancal, N. Sangouard, B. Royer, A. Blais, and A. Wallraff, Loophole-free Bell inequality violation with superconducting circuits, Nature **617**, 265 (2023).

[58] C. Abellán, W. Amaya, D. Mitrani, V. Pruneri, and M. W. Mitchell, Generation of fresh and pure random numbers for loophole-free Bell tests, Physical Review Letters **115**, 250403 (2015).

[59] A. Eddins, M. Motta, T. P. Gujarati, S. Bravyi, A. Mezzacapo, C. Hadfield, and S. Sheldon, Doubling the size of quantum simulators by entanglement forging, PRX Quantum **3**, 010309 (2022).

[60] N. Sastry, U. Shankar, and D. Wagner, Secure verification of location claims, in *Proceedings of the 2nd ACM Workshop on Wireless Security* (ACM, San Diego CA USA, 2003) pp. 1–10.

[61] G. Shafer, A. Shen, N. Vereshchagin, and V. Vovk, Test martingales, Bayes factors and p-values, Statistical Science **26**, 84 (2011).

[62] Y. Zhang, S. Glancy, and E. Knill, Efficient quantification of experimental evidence against local realism, Phys. Rev. A **88**, 052119/1 (2013).

[63] A. Bluhm, M. Christandl, and F. Speelman, A single-qubit position verification protocol that is secure against multi-qubit attacks, Nature Physics **18**, 623 (2022).

[64] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C. C.-W. Lim, and H. Weinfurter, A device-independent quantum key distribution system for distant users, Nature **607**, 687 (2022).

[65] F. Xu, Y.-Z. Zhang, Q. Zhang, and J.-W. Pan, Device-Independent Quantum Key Distribution with Random Postselection, Physical Review Letters **128**, 110506 (2022).

[66] S. Das and G. Siopsis, Practically secure quantum position verification, New Journal of Physics **23**, 063069 (2021).

[67] Y. Zhang, L. K. Shalm, J. C. Bienfang, M. J. Stevens, M. D. Mazurek, S. W. Nam, C. Abellán, W. Amaya, M. W. Mitchell, H. Fu, *et al.*, Experimental low-latency device-independent quantum randomness, Physical review letters **124**, 010505 (2020), arXiv:1812.07786.

[68] Y. Zhang, E. Knill, and P. Bierhorst, Certifying Quantum Randomness by Probability Estimation, Physical Review A **98**, 040304 (2018), arXiv:1811.11928 [quant-ph].

[69] T. Cooke, Coaxial Cables, in *The Global Cable Industry*, edited by G. Beyer (Wiley, 2021) 1st ed., pp. 331–349.

[70] H. Buhrman, N. Chandran, S. Fehr, R. Gelles, V. Goyal, R. Ostrovsky, and C. Schaffner, Position-based quantum cryptography: Impossibility and constructions, SIAM Journal on Computing **43**, 150 (2014), arXiv:1009.2490.

[71] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[72] N. Chandran, V. Goyal, R. Moriarty, and R. Ostrovsky, Position based cryptography, in *Annual International Cryptology Conference* (Springer, 2009) pp. 391–407.

[73] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Reviews of Modern Physics **86**, 419 (2014).

[74] L. Masanes, A. Acin, and N. Gisin, General properties of nonsignaling theories, Phys. Rev. A **73**, 012112 (2006).

[75] P. H. Eberhard, Background level and counter efficiencies required for a loophole-free einstein-podolsky-rosen experiment, Physical Review A **47**, R747 (1993).

[76] N. Brunner, N. Gisin, V. Scarani, and C. Simon, Detection loophole in asymmetric Bell experiments, Physical Review Letters **98**, 220403 (2007).

[77] S. Pironio, J.-D. Bancal, and V. Scarani, Extremal correlations of the tripartite no-signaling polytope, Journal of Physics A: Mathematical and Theoretical **44**, 065303 (2011).

[78] G. Vidal and R. Tarrach, Robustness of entanglement, Physical Review A **59**, 141 (1999).

[79] O. Rudolph, Further results on the cross norm criterion for separability, Quantum Information Processing **4**, 219 (2005).
[80] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics, Nature **414**, 413 (2001).
[81] P. Verduyn Lunel, *Quantum Position Verification: Loss-tolerant Protocols and Fundamental Limits*, Ph.D. thesis, Universiteit van Amsterdam (2024).
[82] L. Escolà-Farràs and F. Speelman, Single-qubit loss-tolerant quantum position verification protocol secure against entangled attackers, Physical Review Letters **131**, 140802 (2023).
[83] C. A. Miller and Y. Alnawakhtha, Perfect cheating is impossible for single-qubit position verification (2024), arXiv preprint arXiv:2406.20022.
[84] R. Vershynin, *High-dimensional probability*, 2nd ed. (Cambridge University Press, 2026).
[85] O. S. Rothaus, On "bent" functions, Journal of Combinatorial Theory, Series A **20**, 300 (1976).
[86] C. Carlet, Boolean functions for cryptography and error-correcting codes, in *Boolean models and methods in mathematics, computer science, and engineering*, Encyclopedia of Mathematics and its Applications (Cambridge University Press, 2010) pp. 257–397.
[87] S. Chee, S. Lee, and K. Kim, Semi-bent functions, in *Proceedings of the 4th International Conference on the Theory and Applications of Cryptology: Advances in Cryptology* (1994) pp. 107–118.
[88] Z. Zhang and Y. Shi, Communication complexities of symmetric xor functions, Quantum information and computation **9**, 255 (2009).
[89] R. Cleve, W. Van Dam, M. Nielsen, and A. Tapp, Quantum entanglement and the communication complexity of the inner product function, Theoretical Computer Science **486**, 11 (2013).
[90] A. Montanaro and T. Osborne, On the communication complexity of xor functions (2009), arXiv preprint arXiv:0909.3392.
[91] N. Linial and A. Shraibman, Lower bounds in communication complexity based on factorization norms, in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing* (2007) pp. 699–708.
[92] R. Bhatia, *Matrix analysis* (Springer Science & Business Media, 2013).
[93] E. Knill, Y. Zhang, and P. Bierhorst, Quantum randomness generation by probability estimation with classical side information (2017), arXiv:1709.06159.

SUPPLEMENTARY INFORMATION

CONTENTS

I. Preliminaries 26

II. Bell-test protocol and assumptions 30

III. Experiment 35
A. Overview 35
B. Protocol Implementation 36
C. Hardware and Timing 39

IV. Classical Spacetime Measurements 40

A. Physical Distances 40
B. Target Regions in the Spatial Dimensions 42
1. Experimental Parameters for Determining Target Regions 43
2. Computing Target Regions 44
3. Comparing Classical and Quantum Target Regions 44

V. Theory 47
A. Statements of results 47
B. Spacetime analysis 49
C. Reduction to three-party non-signaling and connection to two-party local realism 57
D. Test-factor construction 66
E. Test factors against small amounts of prior entanglement 68
F. Heralded sources and commitment to overcome transmission loss 69
G. Security against adversaries with significant quantum resources 71

VI. Analysis of experimental results 87
A. Parameter Determination 87
B. Calibration and Analysis 91

I. PRELIMINARIES

The Bell-test quantum position verification (QPV) protocol involves three verifiers A, A' and B, and a source of entangled pairs of arbitrary quantum systems. One member of each entangled pair is sent to A', who performs measurements on this member. The other member of the pair is sent to a target region E. Verifiers A and B send challenges and receive prover responses. In our experimental implementation, A and A' are co-located in the same lab, and we sometimes call A' a "measurement station". The entangled quantum systems are optical wavepacket modes designed to contain a polarization-entangled pair of photons. We employ the label $\mathsf{A}_{\text{source}}$ for the source of entangled photons. In the experiment, the source is co-located with A' and A. The goal of the Bell-test protocol is to verify that a prover is in the target region E. The protocol consists of a sequence of n trials, and success of the verification and the protocol is determined by analyzing the statistics of the trial records. For provers whose movements are unrestricted, success of the protocol certifies that they are in the target region during at least one trial of the protocol. If the protocol succeeds for a prover, we say that the prover passes the protocol. Our analysis shows that to succeed, the prover must perform some non-trivial quantum operation in the target region. The honest prover P is in the target region during every trial of the protocol. The protocol is designed so that such an honest prover can ensure that the protocol succeeds with a reasonably high probability, while for any cooperating collection of provers who are never in the target region the protocol fails with high probability. Provers attempting to pass the Bell-test protocol without being in the target region are called adversaries. We demonstrate the honest prover strategy experimentally by implementing a measurement station P in a lab physically separated from the verifier A and B labs.

In each trial of the Bell-test protocol, the source prepares two quantum systems $\mathsf{Q}_{\mathsf{A}'}$ and Q_{P} and sends $\mathsf{Q}_{\mathsf{A}'}$ to A' and Q_{P} to the target region. In spacetime, the target region E consists of a disjoint union of target regions E_k, where E_k is contained inside the time interval for the k'th trial. For the k'th trial, A and B choose challenges $u_{\mathsf{A}k}$ and $u_{\mathsf{B}k}$ from

finite sets I_{u_A} and I_{u_B} according to a probability distribution specified by the protocol. In the experiment, the two challenges are bits and the distribution is uniform. When it is necessary to distinguish random variables (RVs) from their values, we use standard capitalization conventions. For example, $U_{\mathsf{A}k}$ and $U_{\mathsf{B}k}$ are the random variables whose instances have values $u_{\mathsf{A}k}$ and $u_{\mathsf{B}k}$, respectively. We use $\mathbb{P}(\ldots)$ to denote probabilities of events and $\mathrm{Exp}(\ldots)$ to denote expectations of real-valued RVs. We assume that the values of random variables belong to finite sets of at least two elements. In the experiment, $U_{\mathsf{A}k}$ and $U_{\mathsf{B}k}$ are binary, independent and uniformly random. Unless explicitly stated, we do not assume that the trials are independent or identically distributed. Random variables without trial-index subscripts refer to the random variables of a generic trial and are used when the specific trial does not matter. In particular, for the remainder of this section, we consider a generic trial of the Bell-test protocol and drop the trial-index subscripts.

The honest prover receives u_A and u_B and computes an agreed-upon and publicly known prover settings function $f(u_\mathsf{A}, u_\mathsf{B})$ with range $\{1, \ldots, k_\mathsf{P}\}$. In parallel, A' chooses a measurement setting $m_{\mathsf{A}'} \in \{1, \ldots, k_{\mathsf{A}'}\}$. A' applies the measurement according to the setting $m_{\mathsf{A}'}$ and obtains measurement outcome $o_{\mathsf{A}'} \in \{1, \ldots, c_{\mathsf{A}'}\}$. The honest prover chooses their measurement setting m_P according to $m_\mathsf{P} = f(u_\mathsf{A}, u_\mathsf{B})$, applies the corresponding measurement and obtains measurement outcome $o_\mathsf{P} \in \{1, \ldots, c_\mathsf{P}\}$. The honest prover then sends o_P to each of A and B. In general, A and B may receive responses from imperfect provers or adversaries, in which case the two responses may not agree. The responses actually received by A and B are denoted by z_A and z_B, where these values may belong to a larger set than o_P, for example, to account for absence of responses and invalid responses.

In the experiment $f(u_\mathsf{A}, u_\mathsf{B}) = (u_\mathsf{A} - 1) \oplus (u_\mathsf{B} - 1) + 1$, where u_A and u_B have values $u_\mathsf{A}, u_\mathsf{B} \in \{1, 2\}$ and $\oplus$ denotes the exclusive OR operation. This is just the regular exclusive OR with inputs and outputs in $\{1, 2\}$. The measurement setting and outcomes of A' also have values $m_{\mathsf{A}'} \in \{1, 2\}$ and $o_{\mathsf{A}'} \in \{1, 2\}$. The prover outcome o_P is binary, and z_A and z_B are binned to be binary as well.

For each trial, the trial record R_{trial} consists of the verifier-visible random variables, which are $U_\mathsf{A}, U_\mathsf{B}, M_{\mathsf{A}'}, O_{\mathsf{A}'}, M_\mathsf{P}, Z_\mathsf{A}, Z_\mathsf{B}$. These are the challenges from A and B, the measurement setting and outcome at A', the computed measurement setting at the prover P, and the responses to the verifiers A and B, respectively. All random variables in play, including those involving multiple trials, are jointly distributed. To determine whether the prover passed the protocol successfully ("pass") or failed ("fail"), the verifiers perform a statistical test on the sequence of trial records. The security of the protocol for a given class of adversaries is quantified by its soundness and completeness parameters. For $\delta \in [0, 1)$, the protocol is δ-sound if for adversaries in the given class, before the protocol starts, the probability that the protocol will succeed is at most δ. To ensure good security, the soundness parameter δ should be extremely small. For $\epsilon \in (0, 1]$, the protocol is ϵ-complete for the actual, honest prover if this prover succeeds with probability at least ϵ. The soundness parameter is established by proof. The completeness parameter for the actual, honest prover can only be estimated, but usefulness of the protocol requires that there exist honest provers with completeness much larger than the soundness. Since we cannot be sure of the value of the completeness parameter for the actual, honest prover, to interpret the security of the protocol, it is desirable to be conservative and assume that ϵ itself is very small. On the other hand, to ensure that the honest prover is securely distinguishable from adversaries, it is necessary that ϵ/δ is very large, requiring δ to be much smaller than the conservative value of ϵ. In our experimental implementation, there are non-trivial tradeoffs between the

soundness, completeness, and the time required to run one instance of the protocol or the number of trials per instance. See Sect. VI.

The target region is determined by the positions of the verifiers and the timing of verifier events, see Sect. V B for details. In the experiment, the verifiers are stationary in a common inertial frame of a flat (Minkowski) spacetime $\mathcal{M}$. The space positions of verifiers A', A and B are $\boldsymbol{x}_{\mathsf{A}'}$, $\boldsymbol{x}_{\mathsf{A}}$ and $\boldsymbol{x}_{\mathsf{B}}$. Our spacetime and security analysis does not depend on stationary verifiers, but it simplifies the visualization of target regions and represents the situation implemented in the experiment. In the experiment, the honest prover P is at rest at position $\boldsymbol{x} = 0$. For simplicity, the theoretical analysis assumes that the relative timing of events in each trial is the same. The timing measurements are described in Sect. IV and summarized in Table I. Let t be the time at which the trial starts. We let $t + s_{\mathsf{A}}$ and $t + s_{\mathsf{B}}$ be the times at which A and B begin to emit their challenges u_{A} and u_{B} toward the prover. The final times at which the verifiers accept the prover's responses are $t + r_{\mathsf{A}}$ and $t + r_{\mathsf{B}}$. The verifiers determine a response value after these times based on the responses or their absence. The trial, including recording of the trial record, is complete at time $t + t_{\text{trial}}$. We require $0 \leq s_{\mathsf{A}} \leq r_{\mathsf{A}} \leq t_{\text{trial}}$ and $0 \leq s_{\mathsf{B}} \leq r_{\mathsf{B}} \leq t_{\text{trial}}$. In principle, we can also specify the position of the source of the entangled pair of quantum systems and the time at which it sends the two systems. For our analysis this is subsumed by the constraints on prior entanglement of adversaries as explained below.

We use statistical tests based on the test-factor strategy [47, 61, 62] for testing the hypothesis that the responses are from adversaries in a given class. Given a general composite hypothesis, a test factor for a trial is a non-negative function of the trial record whose expectation is at most 1 under any prediction of that hypothesis. The use of test factors simplifies the security proofs and is required to enable use of the Bell-test protocol with experimentally available loophole-free Bell tests. Test factors can be productively used in most other QPV protocols to simplify security proofs and protocol operation while supporting otherwise difficult-to-implement features such as early stopping and heralding. See Sect. V F. For our purposes, a trial's test factor W for detecting a class of adversaries is a non-negative function of the trial record whose expectation is guaranteed to be at most 1 for all adversaries in the given class. To have completeness much larger than soundness, it is necessary that implementations of an honest prover strategy can achieve an expectation value $\text{Exp}(W) > 1$, as we experimentally demonstrate. In this work, we require that W is a function of $M_{\mathsf{A}'}, O_{\mathsf{A}'}, M_{\mathsf{P}}, Z_{\mathsf{A}}, Z_{\mathsf{B}}$ only, so that it depends on $U_{\mathsf{A}}, U_{\mathsf{B}}$ only through the settings function $M_{\mathsf{P}} = f(U_{\mathsf{A}}, U_{\mathsf{B}})$. The adversarial bound of 1 on the expectation of W must hold conditional on the past of the trial. The past includes all classical and quantum information available before the trial begins, as specified in the spacetime analysis below. In a generic trial, adversaries may either have no prior entanglement or pre-share bounded entanglement. We construct test factors for both scenarios. Test factors can be chained over multiple trials by multiplication. According to the theory of test factors, the product $\mathcal{T}$ of test factors over the trials of the protocol has expectation at most 1 for all adversary strategies in the given class. A key property of such test-factor products is that the expectation of $\mathcal{T}$ remains upper-bounded by 1 even when adversaries adapt their actions based on past information. By Markov's inequality, $\mathcal{T}^{-1}$ is a p-value for adversaries of the given class. For an experimental implementation of the protocol, an honest prover strategy that ensures $\text{Exp}(\mathcal{T}) \gg 1$ is desirable. We use the convention that a non-negative random variable R is a p-value against a composite null hypothesis if the probability that $R \leq p$ given a model in the null hypothesis is at most p. We do not require the null hypothesis

to contain models that achieve this bound, nor do we restrict non-negative p-values to be in the interval $[0, 1]$. We define the $\log(p)$-value for the protocol as $\log(\mathcal{T})$. Note that the "$\log(p)$-value" is the negative of the logarithm of the p-value. The higher the $\log(p)$-value, the stronger the evidence against the allowed adversaries. Given the desired soundness parameter δ, the protocol passes if the $\log(p)$-value is at least $\log(1/\delta)$ and fails otherwise. Thus, the protocol passes only if the null hypothesis that the responses are from adversaries in the given class can be rejected according to the p-value $\mathcal{T}^{-1}$ at significance level δ. Unless specified otherwise, the base of the logarithm is e.

Test factors can be chosen to maximize the trial gain for a specific honest prover P. The trial gain g is defined as the expectation $g = \mathrm{Exp}(\log(W))$ for P of the logarithm of the test factor. The gain is the expected contribution per trial to the $\log(p)$-value. The test factor is chosen by considering the honest prover that is implemented in the experiment. Since the implemented prover is designed to produce independent and identically distributed (i.i.d.) trial records, the value of $\log(\mathcal{T})$ after n trials has expectation ng with variance $O(n)$. Consequently, the number of trials required for P to succeed with reasonable probability can be estimated as $\log(1/\delta)/g$. Specific estimates based on asymptotic normality are used to determine protocol parameters, see Sect. VI.

II. BELL-TEST PROTOCOL AND ASSUMPTIONS

The Bell-test QPV protocol is detailed in Protocol 1 below. A version of this protocol was implemented in the experiment. The protocol is general and can be applied to arbitrary configurations involving pairs of entangled quantum systems for which choices of measurement settings exist, and for which the verifiers and honest prover yield trial-record distributions that are outside a convex set induced by the convex set of three-party quantum non-signaling distributions as described in Sect. V C. It is proven secure against unentangled or weakly entangled adversaries with no restrictions on quantum communication during a trial, provided that the settings function satisfies the assumptions given below. To address the issue of adversaries with significant amounts of prior entanglement, we apply an enhanced version of the ϵ-net construction pioneered in Ref. [63] to the Bell-test protocol. Given an an initial test factor against unentangled adversaries, we show in Sect. V G how to modify this test factor for scenarios where the adversaries pre-share significant bounded entanglement and utilize arbitrary classical communication but no quantum communication during trials. The completeness and soundness results are summarized in Sect. V A.

Protocol 1: Bell-Test Quantum Position Verification (QPV) Protocol

`// See Sects. V D, V E and V I A for the required preparatory work for running this protocol.`

Given:

δ — desired soundness parameter.
n — number of trials to be executed.
ν_U — joint challenge distribution.
$f(u_{\mathsf{A}}, u_{\mathsf{B}})$ — settings function.
$\mu_{\mathsf{A}'}$ — measurement settings distribution of verifier A'.
W — trial-wise test factor to be used. W depends on the class of adversaries and the Bell test used.
t_0 — initiation time.
t_{trial} — trial duration.
$s_{\mathsf{A}}, s_{\mathsf{B}}, r_{\mathsf{A}}, r_{\mathsf{B}}$ — verifier send and receive times.

Let $t_1 \geq t_0$ be the first trial's start time;
for $k = 1$ **to** n **do**
  Perform a QPV trial starting at time t_k:
  1. The source prepares a joint state of $\mathsf{Q}_{\mathsf{A}'}$ and Q_{P} (event #4 in table I), and sends $\mathsf{Q}_{\mathsf{A}'}$ to A' and Q_{P} to the target region. This state preparation can be done before t_k, provided that Assumption 7 is satisfied.
  2. A and B generate challenges u_{A} and u_{B} according to the joint distribution ν_U and A' chooses measurement setting $m_{\mathsf{A}'}$ (event #5 in table I) according to probability distribution $\mu_{\mathsf{A}'}$. The challenges and settings for this trial may be generated before t_k provided that Assumption 10 is satisfied.
  3. A and B send the challenges u_{A} and u_{B} at times $t_k + s_{\mathsf{A}}$ and $t_k + s_{\mathsf{B}}$ from specified locations. See Assumption 4. These are events #1 & #2 in table I.
  4. A' measures $\mathsf{Q}_{\mathsf{A}'}$ according to setting $m_{\mathsf{A}'}$, producing outcome $o_{\mathsf{A}'}$ (event #6 in table I). For device-independence based on causality, the period between applying the setting to recording the measurement outcome is space-like separated from the trial's target region.
  5. A and B receive response values z_{A} and z_{B} from the provers at times $t_k + r_{\mathsf{A}}$ and $t_k + r_{\mathsf{B}}$ and at specified locations (events #13 & #14 in table I). See Assumption 4.
  6. The verifiers save the reduced trial record $r_{\text{trial}k} = (m_{\mathsf{A}'}, o_{\mathsf{A}'}, f(u_{\mathsf{A}}, u_{\mathsf{B}}), z_{\mathsf{A}}, z_{\mathsf{B}})$ and compute $w_k = W(r_{\text{trial}k})$. In the experiment w_k was computed later, after unblinding the data.

  Let $t_{k+1} \geq t_k + t_{\text{trial}}$ be the next trial's start time;
end
Let $p = (\prod_{k=1}^{n} w_k)^{-1}$; `// This is the` p`-value against adversaries`
if $p \leq \delta$ **then**
  return *success*; `// Protocol succeeded`
else
  return *fail*; `// Protocol failed`
end

The assumptions required for the security of the protocol and to ensure that soundness is satisfied for the given soundness parameter δ are listed next.

Assumptions (see Sect. V B for definitions and terms):

1 **Classical computation and communication:** All classical computation is error-free and satisfies the protocol specifications. This is a standard assumption in device-independent quantum protocols [40, 64]. Classical communication between the verifiers is also error-free. Most classical communication between the verifiers can be deferred until after the protocol completed and can be implemented by authenticated classical channels.

2 **Classical spacetime and worldlines:** The background spacetime is Minkowski. Worldlines are classical. This means that the spacetime circuit of the relevant entities is determined in each trial given the past. In principle, the computations at the vertices could make choices of where to send the outputs. We assume that there is a finite, bounded number of such choices, in which case we can arrange for the fixed spacetime circuit to have edges and vertices for every possible combination of such choices. For this purpose, the systems on the edges may need to be enlarged by adding a "vacuum" state that indicates that the edge was not used. Vertex operations are extended accordingly.

3 **Classical timing and location:** For the theoretical analysis of the protocol we assume that trial timing and location measurements are error-free. The timing and location measurements determine the target region. In the experiment, timing and location measurements have uncertainties as described in Sect. III, which affect the boundaries of the target region. The effect of these uncertainties on localization is illustrated in Fig. IV.

4 **Verifier lab security:** The labs of the three verifiers are secured to disallow interaction with non-verifier entities except for those that are part of the protocol. In particular, the only outgoing edges from the labs of verifiers A and B to non-verifiers are those that jointly carry the classical challenges and are causally consistent with the challenge timing. In practice, this means that the classical challenges do not "leak" out of the verifier labs apart from the established exit points at the appropriate challenge exit time, see Fig. 6. In addition, there are no outgoing edges from A' to non-verifiers, which prevents information about the settings and outcomes at A' from leaking to adversaries during a trial. To ensure this condition, we assume that outgoing communications from A' other than to the verifiers are blocked during a trial. We further require that within a trial, there are no incoming edges from a non-verifier that causally precede edges from the verifiers to a non-verifier. For A and B, this can be enforced by blocking incoming communications before the challenges are sent. A' receives a quantum system from the source, which is treated as untrusted. We require that the verifiers' processing of challenges and responses is completely classical. In particular, challenges are classical and the responses are immediately decohered in the computational basis at the moment they enter the lab. Such assumptions of lab security are standard in device-independent and other cryptographic protocols [40, 64, 65].

5 **Challenge and response events:** We identify specific events in spacetime where the challenges leave the labs and the responses enter the labs. The only requirement on these events is that the target region as defined in Sect. V B is non-empty. In the protocol and the experiment, these events have the specified space positions of A and

B (see Fig. 6), which are stationary with respect to the earth's surface. The challenges leave the lab at or after times $t_k + s_{\mathsf{A}}$ and $t_k + s_{\mathsf{B}}$. The responses enter the labs at or before times $t_k + r_{\mathsf{A}}$ and $t_k + r_{\mathsf{B}}$. All relevant experimentally measured timings are listed in table I.

6 **Adversary communication and entanglement:** We consider three complementary sets of assumptions. The sets have different requirements for the bit-length of the challenges u_{A} and u_{B}. For the first, length-1 bit strings suffice. For the second and third, u_{A} and u_{B} are bit strings of sufficient length l and the settings function has sufficient complexity, see Sect. V G. In all cases, the settings function and challenge distributions must satisfy Assumptions 9 and 10 below, and prior entanglement of adversaries for each trial is defined on a specific time slice determined by the trial timing as described in the spacetime analysis of Sect. V B. This time slice defines the trial boundary for the purpose of quantifying the prior adversary entanglement. The protocol assumes that this prior entanglement is zero or satisfies bounds described in Sect. V E or in Sect. V G. The test factors and the pass/fail decision depend on which assumption is used. The constraint on prior entanglement includes the source and the emitted or to be emitted entangled pair of quantum systems on the trial boundary.
First set of assumptions: Except for spacetime causality, there are no restrictions on adversary communication capabilities during a trial. The adversaries can communicate both classically and quantumly. Quantum communication and memory between trials are indirectly restricted by the constraints on adversary entanglement. Here we assume that prior adversary entanglement is zero or satisfies bounds described in Sect. V E.
Second set of assumptions: There is no constraint on classical communication between the adversaries, but quantum communication and prior entanglement is constrained. By shifting quantum communication to prior entanglement and classical communication, the constraint is expressed in terms of a bound on the rank of the reduced density matrix of the prior adversary entanglement at the adversary nearer B. As shown in Sec. V G, the bound on this dimension is proportional to $\sqrt{l}$.
Third set of assumptions: As in the second set of assumptions, classical communication is not constrained and quantum communication is shifted to prior entanglement. The constraint is on the number of qubits, q, and the number of quantum gates, G, used by the adversary nearer B. In Sec. V G, the bound is of the form $\log(q)G \leq cl$ for some constant c.

7 **Quantum source:** We treat the source as an untrusted component that may be under the control of adversaries. This implies that the source cannot be in the target region. In addition, for the first set of assumptions in Assumption 6, the restriction on prior entanglement of the adversaries implicitly constrains the source and the entangled system prepared and communicated by the source. To prevent the adversaries from exploiting the source for shared entanglement it is necessary to secure the location of source and the timing of the source emissions. Because the quantum system $\mathsf{Q}_{\mathsf{A}'}$ is treated as part of the adversary closer to A, the source needs to be located in the region closer to A. To prevent the source from contributing entanglement to the adversaries, it is desirable for the source emission of Q_{P} to the prover location to be constrained so that the spacetime worldline of Q_{P} does not cross the trial boundary adjacent to the part of the adversary nearer B. If the protocol is secure against adversaries with prior entanglement greater than that available from the source, then there is no need to secure the source location or emission. We show in Sect. V G that such security is

satisfied with sufficiently complex challenges and settings functions.

8 **Quantum measurement:** A' uses a quantum apparatus in their lab to measure $\mathsf{Q}_{\mathsf{A}'}$. The apparatus is untrusted. Security of the protocol requires that the setting and the challenges are not known to the apparatus beforehand. For the analysis here, we assume that the only incoming message during a trial is the quantum system $\mathsf{Q}_{\mathsf{A}'}$ from the source. We assume that outgoing messages are blocked, except for the eventual message of the measurement outcomes to the other verifiers. We assume and ensure in the experiment (see Fig. 7) that the interval between applying a measurement setting and recording the measurement outcome is in the region where the challenge from B is not visible and from which no signal can reach B at or before the time point $t_k + r_{\mathsf{B}}$. In particular this interval is space-like separated from the target region of a trial. In Fig. 7 and using the conventions introduced before Eq. 1, the verifier measurement region consists of the subregions labeled $*0*1$, where $*$ denotes either 0 or 1. An illustration of these subregions is provided in Fig. 9.

9 **Settings function:** The main constraints on the settings function are for those implied by the assumptions on the settings and challenges in Assumption 10 below. For adversaries with significant quantum resources, the bit length of the challenges must be large enough and the settings function must have sufficient complexity.
In QPV, increasing the complexity of the challenges and the settings function leads to higher bounds on the minimum prior entanglement required for adversaries to successfully pass the protocol [63, 66]. It can also be used to protect against computationally bounded adversaries without prior entanglement constraints [50]. Our inclusion in the experiment of a challenge from A and use of the exclusive OR function anticipates the implementation of such strategies.

10 **Verifier settings and challenge distributions:** For our analysis, we assume that verifier settings and challenge distributions are fixed and independent of the state coming into the trial. In the spacetime circuit analysis given in Sect. V B below, this means that the probability distributions are independent of the state of the adversary including the source and the quantum systems that the source emitted on the time slice on which prior adversary entanglement is determined. For the reduction from general adversaries to two adversaries whose quantum messages to each other do not depend on the challenges, we require that 1) $M_{\mathsf{A}'}$ and M_{P} are independent of U_{A} so that $\mathbb{P}(m_{\mathsf{A}'}, m_{\mathsf{P}}, u_{\mathsf{A}}) = \mathbb{P}(m_{\mathsf{A}'}, m_{\mathsf{P}})\mathbb{P}(u_{\mathsf{A}})$, and 2) $M_{\mathsf{A}'}$ is conditionally independent of U_{A} and U_{B} given M_{P} so that $\mathbb{P}(m_{\mathsf{A}'}, u_{\mathsf{A}}, u_{\mathsf{B}}|m_{\mathsf{P}}) = \mathbb{P}(m_{\mathsf{A}'}|m_{\mathsf{P}})\mathbb{P}(u_{\mathsf{A}}, u_{\mathsf{B}}|m_{\mathsf{P}})$. See Sect. V C for how these independence assumptions are used. The assumptions are satisfied if U_{A} and U_{B} are uniformly random, independent of $M_{\mathsf{A}'}$, and the cardinality $|\{u_{\mathsf{B}}|f(u_{\mathsf{A}}, u_{\mathsf{B}}) = m_{\mathsf{P}}\}|$ does not depend on u_{A} or m_{P}. For simplicity of test-factor construction we choose uniform verifier and prover settings distributions for the experiment.
It is possible to relax the assumption that the joint distribution of the verifier settings $M_{\mathsf{A}'}$ and the prover settings M_{P} is fixed and independent of the past. This can be done by allowing mixtures of products of incoming states and joint distributions of $M_{\mathsf{A}'}$ and M_{P}, where the latter come from a fixed convex set of distributions near a target distribution. The chosen test factors then need to be designed to have worst-case expectation of 1 for adversaries over the convex set of distributions. See Ref. [67] for realizations of this strategy in the context of device-independent quantum randomness generation.
It is noteworthy that for rejecting unentangled adversaries, the assumptions on the

settings and challenges can be satisfied with a settings function that does not depend on u_{A}, in which case the challenge u_{A} can be omitted. In this case, the role of s_{A} is to define the trial boundary and the target region. Because $\mathsf{Q}_{\mathsf{A}'}$ is attributed to the adversary nearer A, the restriction on entanglement requires that Q_{P} cannot reach the adversary nearer B before the trial boundary as defined by s_{A}. In the experiment this is satisfied by transmitting Q_{P} in fiber and accounting for the speed of light in the fiber.

III. EXPERIMENT

A. Overview

Fig. 5 shows a schematic of the components of the experiment. In the experiment, verifier A that sends challenge bits and receives responses, verifier A' that measures one of the entangled photons, and the entangled photon source $\mathsf{A}_{\text{source}}$ are in the same room. Verifier B is in another room of the National Institute of Standards and Technology (NIST) building that is about 200 m away.

The entangled photons $\mathsf{Q}_{\mathsf{A}'}$ and Q_{P} generated at $\mathsf{A}_{\text{source}}$ are distributed through single-mode optical fiber to measurement station A' and the honest prover, P. For the successful protocol runs presented here, the honest prover hardware is stationary and located in a room about 90 m from verifier A. Subsequently, verifier A' and the honest prover P measure the photons in one of two measurement bases, performing a Bell test. We label the measurement settings of A' and P determined by their measurement bases by $m_{\mathsf{A}'}$ and m_{P}. Much of the hardware for this experiment builds on loophole-free Bell tests and device-independent random number generation (DIRNG) setups [36, 68]. For example, the measurements performed at A' are the same as at "Bob" in Ref. [37], with the measurement basis $m_{\mathsf{A}'}$ chosen using the same source of local randomness as in that work. The resulting measurement outcomes are recorded to a timetagger. The measurement basis at the honest prover P, however is decided by the output of an exclusive OR (XOR) of random challenge bits $m_{\mathsf{P}} = (u_{\mathsf{A}}-1)\oplus(u_{\mathsf{B}}-1)+1$ released by the verifiers A and B. The random bits are sent to the honest prover at high speed over specialized coaxial cables. The honest prover measures the photon and replies to verifiers A and B with the measurement outcome over fast coaxial cables. Verifiers A and B then record these signals $z_{\mathsf{A}}, z_{\mathsf{B}}$ on their timetaggers.

Provided that the recorded signals violate a Bell inequality, the time difference between when the random inputs were provided and the measurement outcome was received by A and B allows the verifiers to perform a secure ranging based on round-trip communication times. Any excess latencies during transmission (i.e., transmission slower than the speed of light), or in the measurement process at the honest prover implies a larger target region E for localizing the prover. This highlights the importance of low-latency measurements at the implemented honest prover station and high-speed classical communication. In our experiment, this is enabled with the use of Pockels cells, custom high-speed electronics, and low-latency coaxial cables. The experimental realization of a QPV protocol requires careful design and optimization to minimize latency in electronics, measurement, and transmission.

With the Bell-test protocol, much of the "quantum" hardware (such as the source and the measurement stations) need not be trusted, or even be within the secure verifier labs, allowing for more flexibility in configuration choices. A careful account of the assumptions involved in our demonstration can be found in Sect. II.

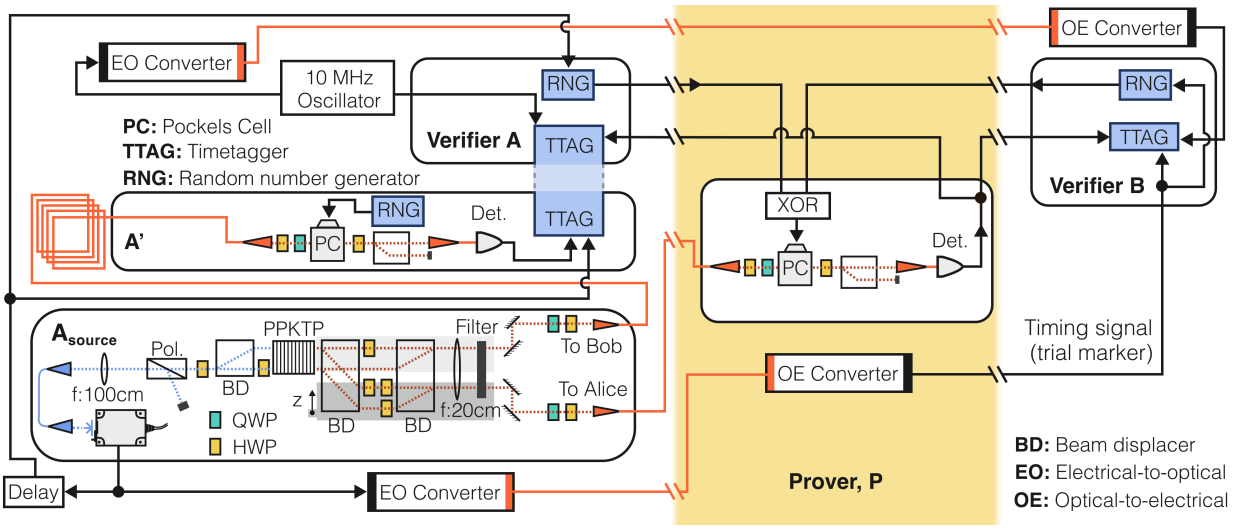

FIG. 5. Experimental diagram for the position verification demonstration. Verifiers A, A′, and B, the honest prover P as implemented, and the entangled photon source $\mathsf{A}_{\text{source}}$ are indicated. The yellow shading around the honest prover indicates the region prover hardware must be in to pass the protocol (target region E). The RNGs and timetaggers are associated with the verifiers. Electrical and optical cables carry information to and from the various experimental stations. The entangled photons are distributed via optical cables, and the measurement choice bits are transmitted via high-speed coaxial cables. Electrical cables are indicated in black, and optical cables in orange. Electrical-to-optical (OE) and optical-to-electrical (EO) converters convert information between electrical and optical encoding. These are used in the lines that transmit synchronization signals among the verifiers. $\mathsf{A}_{\text{source}}$ is the source of entangled photons, which produces entangled photons via spontaneous parametric downconversion with a periodically poled potassium titanyl phosphate (ppKTP) crystal placed inside a dual Mach-Zehnder interferometer formed with beam displacers (BD) [37]. For the measurements at A′ and P, the pair of measurement bases are implemented with quarter-waveplates (QWP) and half-waveplates (HWP), and fast switching between the bases is performed with a Pockels cell.

B. Protocol Implementation

The labs containing verifiers A and B are assumed to be trusted, in the sense that the equipment inside them is not under the influence of any adversaries. This is related to Assumptions 4 and 10 in Sect. II. Physically, this means the signals are assumed to be shielded from adversaries (in that they cannot be altered or read by adversaries) until they physically exit the labs, and that any signals coming in from the outside are similarly shielded once they enter the verifier labs.

All signals exit and enter the verifier labs via access points in the walls through which cables are laid. In all our analyses, we rely on the trusted lab assumption stated to use these physical access points to define the verifier A and B's physical locations. Here, we are modeling a situation where the verifier labs are carefully shielded from external signals, while an opening allows transmission of signals in and out of the labs. The times when challenge (response) signals exit (enter) these openings are the earliest (latest) time at which they can be influenced or read by adversaries. From the adversarial point of view, it thus suffices



FIG. 6. Photographs of the physical entry/exit points of the classical and quantum signals that serve as the verifiers A and B in our demonstration. Any hardware and cabling inside the verifier B lab is trusted. We do not need to trust the quantum hardware/measurement device, at Verifier A. See Fig. 5 for a full account of the devices at verifier A. At verifier B, the lab houses a random number generator for deciding the challenge u_{B}, and a timetagger that records classical timing information and the response from provers z_{B}. All information exiting and entering the Verifier B lab is classical.

to reduce the verifiers to the spatial points corresponding to the signal entry and exit. See Fig. 6 for photographs of the entry/exit points. Signals traveling to or from the verifier labs and crossing these physical thresholds at time 't' are considered to have been generated or detected at 't'.

The protocol proceeds as a series of trials, with the number of trials performed in a single run of the protocol determined by previous calibration data and a desired threshold for soundness error (see Sect. I). Each trial starts with the creation of an entangled pair of photons at $\mathsf{A}_{\text{source}}$. The photons are generated via spontaneous parametric downconversion. A solid-state pulsed laser at 775 nm pumps an entangled pair generation source that probabilistically down-converts the pump to generate entangled pairs of photons at 1550 nm. The laser operates at a repetition rate of 80 MHz, meaning we attempt to produce entangled photons every 12.5 ns.

The source of entangled photons is largely similar to the one employed in Ref. [37], and is based on a twin Mach-Zehnder interferometer constructed with beam displacers. The entangled photon source is tuned to produce a non-maximally entangled state, nominally $0.383|\text{HH}\rangle + 0.924|\text{VV}\rangle$, conditional on a pair of photons being spontaneously generated, which happens with a probability of approximately 1/350. This state (along with the measurement angles at the measurement stations) was chosen based on a numerical optimization maximizing the Bell violation. The optimization takes experimental losses, dark counts, and imperfect state fidelity into account. For comparison to bounds on prior entanglement by adversaries, the robustness of entanglement of the nominal state is $(0.383 + 0.924)^2 - 1 = 0.71$. Accounting for the probability of photon pair emission, this gives a robustness of entangle-

ment per trial of 2×10^{-3}. The source sends one photon ($\mathsf{Q_P}$) from each of the entangled photon pairs to P via optical fiber, and the measurement station $\mathsf{A'}$ makes measurements locally on its photon ($\mathsf{Q_{A'}}$). The $\mathsf{Q_{A'}}$ photon reaches the measurement station at $\mathsf{A'}$ after being delayed in a fiber spool. This delay ensures that the measurement basis choice at this station is space-like separated from the trial region, see Assumption 8 in Sect. II.

Following the release of entangled photons from $\mathsf{A}_{\text{source}}$, the verifier A transmits a random bit (u_A). This bit is freshly generated via a hardware random number generator at A, and its value is immediately recorded on a timetagger located at A. The bit is then encoded as an electronic non-return-to-zero (NRZ) signal with about a 1 V amplitude and transmitted via a specialized coaxial fiber that has a central conductor surrounded by a foam dielectric. The air in the foam reduces the relative permittivity of the surrounding insulator and allows these cables to support a signal propagation velocity that is between 85-87% of the speed of light [69]. In traditional low-density polyethylene (LDPE) coaxial cables, the velocity of propagation is closer to 66%. This classical signal travels towards the prover faster than the 1550 nm photon $\mathsf{Q_P}$, which travels through an optical fiber of roughly comparable length at approximately 66% of the speed of light. Some time after u_A is released, verifier B releases its bit (u_B), which is recorded on a timetagger at verifier B). In our configuration, with verifier B situated closer to the honest prover P than verifier A, this delay is set so that the bits u_A and u_B reach the (known) location of the honest prover at approximately the same time. The classical challenge bits reaching the honest prover at the same time is optimal because any discrepancy in the arrival times would mean that the honest prover would have to wait for either u_A or u_B before computing their XOR. Waiting in this way would increase protocol latency at the honest prover hardware station and the size of the realizable target region of the Bell-test protocol. In our realization, the synchronization is slightly imperfect and u_B arrives at the prover 9(2) ns after u_A. Once the bits u_A and u_B reach the prover, the prover computes the XOR of the bits $m_\mathsf{P} = (u_\mathsf{A} - 1) \oplus (u_\mathsf{B} - 1) + 1$ using an application-specific integrated circuit (ASIC). This computation takes only a few nanoseconds and does not contribute significantly to the latency at the honest prover P. The honest prover then sets the basis of the Pockels cell (see Fig. 5) so that incoming photons will be projectively measured in basis $a1 = 6.7°$ or $a2 = -29.26°$, depending on whether $m_\mathsf{P} = 1$ or $m_\mathsf{P} = 2$ respectively. Here, the angles represent the projective bases corresponding to rotations of a linear polarizer relative to the horizontal. The angles are the same as in Ref. [37]. An incoming photon $\mathsf{Q_P}$ arrives shortly after the Pockels cell is set. The photon is then measured in the appropriate basis with a superconducting nanowire single photon detector (SNSPD). The output is duplicated with a power splitter into z_A and z_B, which are sent to the verifiers over separate specialized coaxial cables. These are similar to those used for the transmission of u_A and u_B to the honest prover P.

Approximately at the same time as the photon $\mathsf{Q_P}$ is measured at the honest prover, its partner photon $\mathsf{Q_{A'}}$ is measured at the measurement station $\mathsf{A'}$ using a local random bit produced at the station $m_{\mathsf{A'}}$. The photon is measured in the basis $b1 = -6.7°$ or $b2 = 29.26°$ depending on if the value of $m_{\mathsf{A'}}$ is 1 or 2. The result of this local measurement, $o_{\mathsf{A'}}$, is recorded on the timetagger at verifier A. Finally, the returning outputs from the honest prover z_A to verifier A are recorded on verifier A's timetagger, and z_B on verifier B's timetagger. This concludes one trial in the experiment. Trials are repeated sequentially at a 250 kHz rate. The data recorded on the timetaggers, including the input random bits from the verifiers and the output bits back from any provers, constitutes the raw trial record. This raw trial record is then processed based on classical timing markers that are exchanged

between the two verifiers via a trusted classical channel. A trusted classical channel is also used by the verifiers to exchange information used to align and process the collected data.

C. Hardware and Timing

As discussed above, the experiment has three physically distinct locations, shown in, Fig. 5. One lab, which we call "A's lab", contains A, A', and $\mathsf{A}_{\text{source}}$; one contains B; and one contains P.

A number of optical and electrical cables carry information back and forth between the labs in our demonstration. Because the timetaggers that record this information are separated from each other, and recorded trial information must be reconciled to a common timebase, the timetaggers must exchange timing information between themselves. In our demonstration, timing information is derived from an electronic "sync" signal, which is divided down from the 775 nm pump laser pulsed at 80 MHz, and serves as the master clock for pulse generation. This signal is first converted to optical pulses, which are sent to the honest prover's lab over optical fiber. The optical signal is then converted back to an electronic signal and propagated to verifier B. The sync signal is also used by the prover's hardware to sample the output of the XOR circuit correctly. In principle, the challenges can serve as sync signals, in which case sending separate sync signals is not required.

Another timing signal that the experiment uses is a 10 MHz clock that serves as a timebase for the remote timetaggers. The "sync" signal serves as a trial marker, but it is not at the correct frequency to synchronize our remote timetaggers. The 10 MHz clock is generated at verifier A, split off and supplied to the local timetagger, converted to an optical signal, and transmitted via optical fibers all the way to verifier B, where it is supplied to B's timetagger.

The sync signal is divided down to 250 kHz from the electronic output monitor of the 80 MHz pump laser such that the laser emits 320 pulses during every sync period. The laser pulses then propagate through a set of beam displacers and optics to reach a periodically poled potassium titanyl phosphate crystal in two spatial modes to produce the entangled state, consisting of the Q_{P} and $\mathsf{Q}_{\mathsf{A}'}$ portions of the entangled state. The time of emission of the quantum state Q_{P} from the verifier A lab is much earlier than the emission of the classical bit u_{A}. This is to ensure that the photon traveling slowly through fiber arrives at the honest prover in time to be measured. Because the measurement station A' is in the same room as the verifier A, the fiber is looped up into a spool to delay measurement by 791(1) ns, as mentioned in a previous section. These single mode optical fibers are each of length ≈ 160 m and have a transmission efficiency of over 99%.

Other processes are also clocked off of the electrical sync signal. In addition to the sync signal that is propagated to verifier B, it is also sent to verifier A after passing through a digital delay circuit. The sync signal triggers hardware random number generators at A and B to in-part determine when they each emit a random bit (u_{A} and u_{B}) to the prover P over custom coaxial cables. The delay at verifier A is then set so that the random bits u_{A} and u_{B} arrive at the prover at approximately the same time. The relative timing of the bits is decided so that there is sufficient time for the XOR of the bits to be computed and the Pockels cell to be set to either a high or low voltage based on the result of the XOR by the time the photon Q_{P} arrives.

The electronic sync signal is also provided to the measurement station A', which measures its photon $\mathsf{Q}_{\mathsf{A}'}$ based on a random number generator that is triggered again by this sync signal. The output of this measurement is immediately recorded on a local timetagger.

Finally, the prover P transmits the measurement outcome from its measurement of photon Q_P to A and B over electrical coax cables. These signals containing the measurement outcome are recorded on timetaggers. Due to electronic errors, in about 2×10^{-6} of trials, A's and B's records of the prover's measurement outcome disagree. Besides the outcomes, a local record of the RNG bits u_A and u_B is recorded on the timetaggers every trial, along with the timetag of the sync signal to serve as a reference. Because the entire experiment is clocked on this sync signal, the timetagger records can be reconciled with one another using the trusted delay measurements. This allows the creation of a trial-by-trial timing record of all relevant events.

IV. CLASSICAL SPACETIME MEASUREMENTS

Determining the localization regions achieved by the Bell-test protocol requires making measurements on the timing signals and the physical distances separating the verifiers. Here, we provide a summary of our spacetime measurement techniques and results.

We define time $t = 0$ as when the synchronization signal exits the divider circuit, which electronically divides the monitor output of the 775 nm pump laser. Several techniques are used to estimate the timing of various events relative to this reference during a trial. For instance, to establish the time when the sync signal reaches the prover P and the far verifier station B, we measure some electrical latencies directly with an oscilloscope, measure certain electrical cable latencies using electrical reflectometry, measure optical cable latencies using optical time domain reflectometry (OTDR), and perform relative delay measurements with timetaggers.

a. Direct Oscilloscope Measurements Where appropriate, we measure the electrical latency of components in a single-pass configuration. We split pulses produced by a pulse generator with a ≈ 1 ns rise time, and measure relative delays on a 1 GHz oscilloscope employing length-matched coaxial cables. For single-pass optical measurements (always in free-space), we employ physical measurements of the path lengths.

b. Time Domain Reflectometry When single-pass measurements were not appropriate, we use electrical and OTDR to measure component latencies. Here, a pulsed source is directed through one end of an electrical or optical path such that the reflections off the far end can be timed and divided by two to compute the component latency.

The results for key events in the experiment are listed in table I and depicted in Fig. 7. While provided here for completeness, not all of these times need to be characterized or measured for the execution of the position verification protocol by the verifiers. The only times that must remain constant during a protocol run are the times when the verifiers send the challenge bits and receive the response bits (numbered 1, 2, 13 and 14). The constancy of these times can be checked with the verifiers' timetaggers during data runs. The timetaggers are assumed to be functioning correctly and outside adversarial influence (see assumptions 3 and 4). For our demonstration, we verify the timings on a few data sets, and assume that the timings remain constant during the actual data runs.

A. Physical Distances

The physical distance between the verifiers is assumed to be known and constant during the protocol, see Sect. II, assumption 5. We determine this distance with measuring tape

Index	Parameter	Value
1.	Verifier A: random u_{A} out	1291.0(5) ns
2.	Verifier B: random u_{B} out	1429.1(6) ns
3.	Verifier A: photon Q_{P} exits laser	1023.3(3) ns
4.	Verifier A: photon Q_{P} emitted	1133(2) ns
5.	Verifier A′: setting $m_{\mathsf{A}'}$ created	1738.8(4) ns
6.	Verifier A′: outcome $o_{\mathsf{A}'}$ recorded	1874.5(3) ns
7.	Prover P: u_{A} in from A	1766(1) ns
8.	Prover P: u_{B} in from B	1769(1) ns
9.	Prover P: signal into Pockels cell	1798(1) ns
10.	Prover P: photon Q_{P} arrives at Pockels cell	1831(1) ns
11.	Prover P: photon Q_{P} arrives SNSPD	1858(2) ns
12.	Prover P: signals $z_{\mathsf{A}}, z_{\mathsf{B}}$ created	1868(2) ns
13.	Verifier B: response z_{B} in	2207.7(6) ns
14.	Verifier A: response z_{A} in	2340.3(5) ns

TABLE I. A summary of the times events occur in a position verification trial, referenced to the time the synchronization signal exits the divider circuit as time $t = 0$ ns. Relevant events in the trial span about 1050 ns, if we consider the classical signal exiting the verifier A lab as the start time at verifier A. The protocol run time would increase with larger separations between the verifiers A and B. The trial rate is 4 µs. See Fig. 7 for a visualization of these key events.

by hand-measuring a set of orthogonal spans. The location of the prover hardware that we install for demonstrating successful runs of the protocol has a bearing on the choices of time delays used for the classical position verification. We measure these distances with measuring tape also. In a real-world implementation of this protocol, these distances must first be established using classical geomapping techniques before one could implement the position verification protocol.

The major contributors to uncertainty in the physical measurements is the uncertainty in the orthogonality of the various spans in the building, and the size of the exit points of the verifier labs. To estimate the effect of various uncertainties, we perform Monte-Carlo sampling based on the uncertainties of the individual spans, and the uncertainty in the solid angles connecting the various spans. Uncertainty in the solid angles shortens the best estimate for the separation between the verifiers because deviations from 90° for the angles between the orthogonal spans reduce the verifier separation. We estimate that the building is true to 1°, and employ a projected normal distribution to model angular uncertainty. The largest source of uncertainty is the 200 mm uncertainty in the size of the verifier locations (cable exit points) A and B, pictured in Fig. 6. A summary of the distance measurements is presented in table II.



FIG. 7. Spacetime diagram of the position verification protocol. Important events during one trial of the protocol (listed in table I) are marked. The purple lightcones are forward and backward lightcones from the times s_{A}, s_{B}, r_{A}, and r_{B} (labeled 1, 2, 14 and 13 respectively) when the verifiers send the challenge bits and receive the response bits (see Sect. IV B). The target region is the darker diamond that contains points 7 through 12. The events labeled 5 and 6 occur at A', which is slightly offset from A. The origin (time $t = 0$) of both this figure and table I is when the sync signal corresponding to a particular trial exits the divider board. (see Sect. III C).

B. Target Regions in the Spatial Dimensions

Using the distance and timing measurements, the verifiers can establish the target region that effectively localizes the successful prover. A formal treatment of the target region is presented in Sect. V B. In this section, we outline how the target region realized by the experimentally implemented protocol instance is calculated.

Distance	**Value**
Verifier A - Verifier B	195.1(3) m
Verifier A - Prover P (proj.)	92.8(1) m
Verifier B - Verifer A′ TTAG (proj.)	192.2(3) m
Verifier B - Verifer A′ RNG (proj.)	192.2(2) m

TABLE II. This table summarizes key measured distances and their associated uncertainties for our experimental layout. Because there are only two classical verifiers transmitting challenges in our demonstration, the only relevant distances are those projected onto the line joining the verifiers A and B, indicated by "(proj.)".

1. Experimental Parameters for Determining Target Regions

Our experiment is a series of trials, and times for trial k are described relative to the start time t_k for that trial. To compute the extent of the target region E in 3+1 dimensions, and lower-dimensional projections thereof, as in Fig. 3 of the main text, only four times are relevant. These are the times when the verifiers send classical challenge bits and the latest times they accept responses. Following the notation introduced in Sect. I, $t_k + s_{\mathsf{A}}$ and $t_k + s_{\mathsf{B}}$ are the times at which A and B emit their challenges u_{A} and u_{B} toward the prover. The verifiers agree beforehand that responses must be received by times $t_k + r_{\mathsf{A}}$ and $t_k + r_{\mathsf{B}}$ at A and B respectively. In our implementation, the verifiers time-tag when challenges are emitted, and when responses are received.

A successful prover must simultaneously satisfy the following constraints related to the timing of signals in every trial (in addition to producing a Bell violation at the verifiers).

1. Receive challenge u_{A} (released at $t_k + s_{\mathsf{A}}$) from A and reply with z_{A} to A (by $t_k + r_{\mathsf{A}}$).
2. Receive challenge u_{B} (released at $t_k + s_{\mathsf{B}}$) from B and reply with z_{B} to B (by $t_k + r_{\mathsf{B}}$).
3. Receive challenge u_{A} (released at $t_k + s_{\mathsf{A}}$) from A and reply with z_{B} to B (by $t_k + r_{\mathsf{B}}$).
4. Receive challenge u_{B} (released at $t_k + s_{\mathsf{B}}$) from B and reply with z_{A} to A (by $t_k + r_{\mathsf{A}}$).

These constraints require that the prover measures Q_{P} with low enough latency to be able to respond in time and that the messages are transmitted to and from a prover over fast enough classical channels. As explained in the main text,

- Constraint 1 enforces a sphere of radius $\frac{c\,(r_{\mathsf{A}} - s_{\mathsf{A}})}{2}$ centered around A, $\mathsf{S}(\mathsf{A}, c(r_{\mathsf{A}} - s_{\mathsf{A}}))$.
- Constraint 2 enforces a sphere of radius $\frac{c\,(r_{\mathsf{B}} - s_{\mathsf{B}})}{2}$ centered around B, $\mathsf{S}(\mathsf{B}, c(r_{\mathsf{B}} - s_{\mathsf{B}}))$.
- Constraint 3 enforces an ellipsoid of rotation with foci at the verifiers and a major axis $c(r_{\mathsf{B}} - s_{\mathsf{A}}), \mathsf{E}(\mathsf{A}, \mathsf{B}; c(r_{\mathsf{B}} - s_{\mathsf{A}}))$.
- Constraint 4 enforces an ellipsoid of rotation with foci at the verifiers and a major axis $c(r_{\mathsf{A}} - s_{\mathsf{B}})$, $\mathsf{E}(\mathsf{A}, \mathsf{B}; c(r_{\mathsf{A}} - s_{\mathsf{B}}))$.

The dimensions of these shapes are indicated in table III. The target region for each trial is the intersection of these spheres and ellipsoids, which are all rotationally symmetric about the axis joining the verifiers.

Because of the periodic nature of the trials, there is also a periodic constraint in the times that a prover must be located in these regions, which is not captured in the spatial projections.

Parameter Description	**Value**
Sphere $\mathsf{S}(\mathsf{A}, c(r_\mathsf{A} - s_\mathsf{A}))$, radius	157.3(2) m
Sphere $\mathsf{S}(\mathsf{B}, c(r_\mathsf{B} - s_\mathsf{B}))$, radius	116.7(2) m
Ellipsoid $\mathsf{E}(\mathsf{A}, \mathsf{B}; c(r_\mathsf{B} - s_\mathsf{A}))$, major axis	274.8(2) m
Ellipsoid $\mathsf{E}(\mathsf{A}, \mathsf{B}; c(r_\mathsf{A} - s_\mathsf{B}))$, major axis	273.1(2) m
Focal distance for ellipsoids (verifier separation)	195.1(3) m

TABLE III. Extents of the geometric localization regions derived from the timing and distance measurements in the experimental setup. All values are presented with their associated uncertainties, rounded to two decimal places. The units for these measurements are meters.

2. Computing Target Regions

When projecting the target regions to the spatial dimensions (to perform position localization), the aforementioned two spheres and two ellipsoids are the only relevant geometries that verifiers need to estimate. Estimating these geometries requires measurement of the physical distance between the verifiers determined via the procedures outlined in table II, and the start and end transmission times determined by the verifiers' timetagger measurements (see table I).

For the QPV protocol, the target region E (where an honest prover must be located) projected to spatial dimensions is the intersection of the two spheres and two ellipses from Constraints 1-4 above, $E = \mathsf{S}(\mathsf{A}, c\Delta\tau_1) \cap \mathsf{S}(\mathsf{B}, c\Delta\tau_2) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_3) \cap \mathsf{E}(\mathsf{A}, \mathsf{B}; c\Delta\tau_4)$. This is shown in Sect. V B.

The target regions can also be projected into lower spatial dimensions (one and two dimensions). For a one-dimensional projection, the projection of interest is along the line joining the verifiers, and for two dimensions, it is the plane roughly parallel to the ground and containing the line joining the verifiers. Because of the rotational symmetry of the regions, projection onto any plane rotated along the line joining the verifiers yields an identical target region.

When projecting onto lower-dimensional regions, the spheres and ellipsoids reduce to lines in one dimension and circles and ellipses in two dimensions. The target region E can also be computed as a projection into these lower-dimensional regions. To perform these calculations, we use the known distances between the verifiers from table VI, and the timings from table I to numerically compute the lengths, areas and volumes of the target regions of the quantum protocol (quantum target region) in one, two and three dimensions. We also use the uncertainties in these quantities to run a Monte Carlo simulation to yield an estimate of the uncertainty in the length/area/volume of the target region.

3. Comparing Classical and Quantum Target Regions

An alternative protocol where the verifiers only require an honest prover to compute a classical function based on inputs u_A and u_B can be considered, see Sect. V B. Because such a protocol requires no quantum resources on the part of the prover or the verifiers, it serves as a useful comparison to the Bell-test protocol. If the target region achievable with the

classical position verification protocol is comparable in size to the quantum target region E, we gain little by adding quantum resources to the Bell-test protocol. On the other hand, if the quantum target region is significantly smaller than the classically achievable target region, we take the Bell-test protocol to have demonstrated a quantum advantage. We note that the security of such a classical protocol is based on the impossibility of faster-than-light signaling and thus does not require restrictions on the entanglement shared between the adversaries that the Bell-test protocol requires (detailed in V E, also see Assumption 6).

We employ the following metric to quantify the advantage of using quantum resources,

$$\text{Quantum Advantage} = \frac{\text{size of classical target region}}{\text{size of quantum target region (E)}} \tag{1}$$

The "size" in the metrics may be a length, area, or volume when considering one, two or three-dimensional position verification. An ideal quantum protocol (without experimental latencies in transmission and processing) can achieve an infinite Quantum Advantage, as defined above. In general, the goal of a good quantum protocol is to maximize the Quantum Advantage.

In this work, we compare our quantum experiment to two classical protocols:

The ideal classical protocol involves the verifiers sending challenges at the speed of light, and the implemented prover replying back with zero latency, also at the speed of light. The target region is the point exactly at the midpoint of the line joining the verifier. This represents the classical protocol with the smallest possible target region.

The comparable classical protocol involves the verifiers sending challenges and receiving responses at the same times as our quantum protocol $(t_k + r_{\mathsf{A}}, t_k + r_{\mathsf{B}}, t_k + s_{\mathsf{A}}, t_k + s_{\mathsf{B}})$. The motivation here is to consider a practically implementable classical protocol. For any prover implementation, computing classical functions has lower latency than making quantum measurements. So it is reasonable that a classical protocol with the latencies we demonstrate here for the quantum protocol can be implemented.

For the classical protocols, a pair of provers can collude to pass the protocol as long as one of them is in $\Big(\mathsf{S}(\mathsf{A}, c\Delta\tau_1) \cap \mathsf{E}(\mathsf{A},\mathsf{B}; c\Delta\tau_3)\Big)$, and the other in $\Big(\mathsf{S}(\mathsf{B}, c\Delta\tau_2) \cap \mathsf{E}(\mathsf{A},\mathsf{B}; c\Delta\tau_4)\Big)$. This is shown in Sect. V B. We take this to mean that the classical target region is the union of these two regions for all our comparisons.

For **the ideal classical protocol**, the ellipsoids from Constraints 3 and 4 reduce to degenerate line segments. This is because the major axis is equal to the distance between the verifiers ($c(r_{\mathsf{A}} - s_{\mathsf{B}}) = c(r_{\mathsf{B}} - s_{\mathsf{A}}) = d_{\mathsf{AB}}$). The classical target region is the line segment joining the verifiers. Because this is an *ideal* protocol, the target region is one-dimensional. The only uncertainty associated with the size of this region is in the distance between the verifiers. The quantum advantage for the ideal classical protocol is the ratio of its target region to the quantum target region projected along the line joining the verifiers. We performed Monte Carlo simulations to estimate the uncertainty in this ratio. The results are presented in table IV.

For **the comparable classical protocol**, the target regions in one, two and three dimensions all have non-zero sizes, and we compare these to the sizes of the quantum target region. This computation is similar to that done for the ideal classical protocol. The results of these comparisons are presented in table IV, Fig. 8 and in the main text.

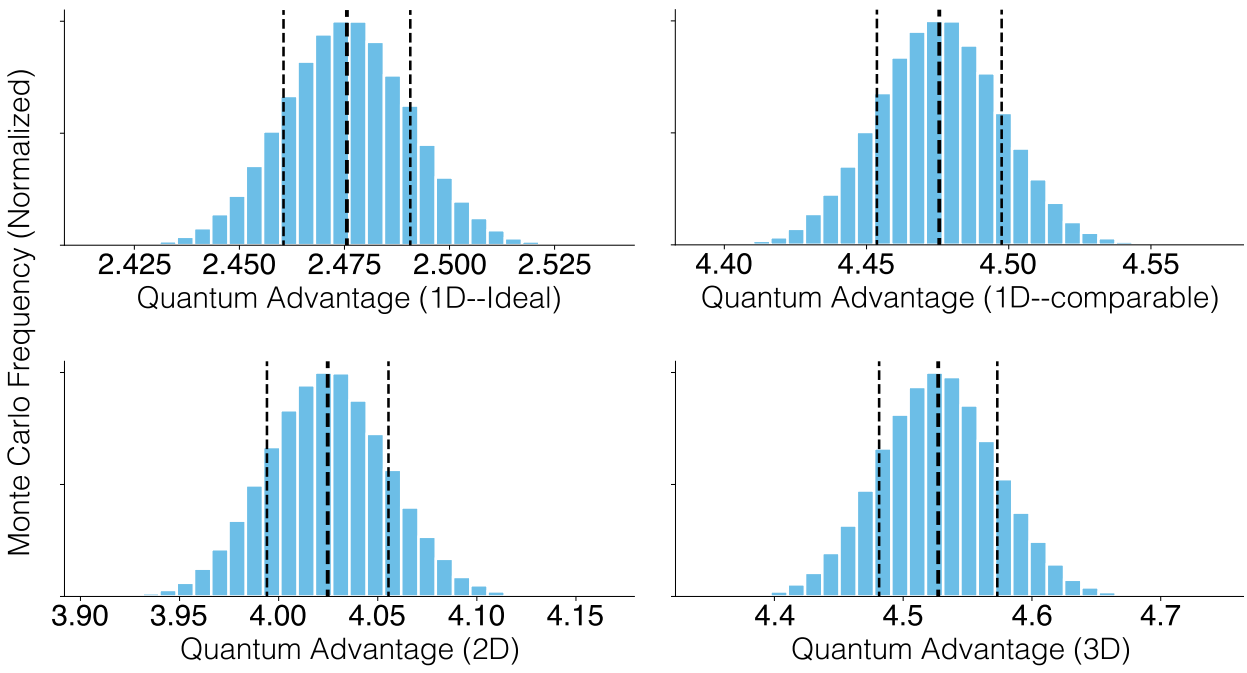

FIG. 8. Histograms showing the relative frequencies of the quantum advantage from 10^5 Monte Carlo simulations. The simulations sample the measured distances and timings based on estimated uncertainties to arrive at the quantum advantage distributions. The uncertainties on individual timing and distance measurements are modeled as Gaussian (with the exception of angular uncertainties, which use projected Gaussians). The largest source of uncertainty is the physical ambiguity in the definition of the verifier A and B positions, see Fig. 6. The mean and $\pm 1\sigma$ for each distribution are indicated with vertical black dashed lines. The top row plots the quantum advantage against ideal and comparable quantum protocols in 1 dimension, while the bottom row plots the advantage against the comparable protocol in 2 and 3 dimensions.

Dimension	**Quantum Advantage**
1D (vs Ideal Classical)	2.47(2)
1D (vs Comparable Classical)	4.48(2)
2D (vs Comparable Classical)	4.02(3)
3D (vs Comparable Classical)	4.53(5)

TABLE IV. Comparison of localization ratios and quantum advantages for QPV protocols in different dimensions. The quantum protocol achieves zero quantum advantage compared to the ideal classical protocol in 2D and 3D because the ideal classical size is zero in 2D and 3D. Reported uncertainties are one standard deviation.

V. THEORY

A. Statements of results

We prove soundness and completeness of the Bell-test protocol in Sect. II for versions of the assumptions in Sect. II, specifically versions of Assumption 6. The same configuration of verifiers and honest prover witnesses completeness regardless of which version of the assumptions is used. In all cases, for soundness $0 < \delta < 1$ and completeness $0 < \epsilon < 1$, completeness can be achieved with a number of trials of order $O(\log(1/\delta)) + O(\log(1/(1-\epsilon)))$, where the suppressed constants depend on the quality of the Bell-test on which protocol completeness is based and the class of adversaries under consideration.

Result V.1. *Protocol completeness: For all $\epsilon \in (0,1)$, the protocol of Sect. II is ϵ-complete. To witness completeness it suffices for the verifiers and honest prover to violate local realism for some Bell-test configuration with two measurement settings for the prover. For every class of adversaries under consideration and a given violation of local realism, the number of trials required for completeness ϵ and soundness δ is $O(\log(1/\delta)) + O(\log(1/(1-\epsilon)))$.*

The proof of this result is in Sect. V C. Because there is no restriction on how the entangled state for the Bell-test protocol is obtained by the verifiers and prover before each trial, completeness is preserved if a slow quantum network is used to prepare these states. The quantum link may be heralded and can support prover commitment strategies [16]. Therefore, loss in photonic links is well tolerated. See Sect. V F for how the protocol supports heralded sources and commitment.

The soundness results depend on the class of adversaries under consideration. In each case, soundness holds for all space dimensions, and the source is untrusted and may be controlled by one of the adversaries. Classical communication by the adversaries is always unconstrained. The first soundness result involves the class of adversaries with no prior entanglement.

Result V.2. *Protocol soundness against adversaries with no prior entanglement: For all $\delta > 0$, the protocol of Sect. II is δ-sound against adversaries who have no prior entanglement for each trial. The adversaries can freely use quantum computation and communication during each trial.*

The proof of Result V.2 is contained in Sects. V B, V C and V D. The first step of the proof follows the traditional path of reducing the adversary attacks to those involving two adversaries on opposite sides of the target region [70]. The Bell-test protocol requires an analysis that is independent of space dimension, takes into account the time required by the honest prover to perform its measurement task, and supports device-independence with test factors. For this we adapt the spacetime circuit analysis introduced in Ref. [50] and fully formalize the reduction to two adversaries. The subsequent analysis of adversary attacks deviates from the traditional path in order to reduce the statistics of adversarial trial records to a three-party non-signaling scenario for which we can construct test factors.

The absence of prior entanglement for the above soundness result implies a constraint on the location of the entanglement source at the trial boundary: The source must be located on the side closer to verifier A', which requires that we prevent the quantum system destined for the prover from leaking to the other side. This constraint on the location of the source is mitigated by the next soundness result and absent in the last two soundness results.

The second soundness result allows for small amounts of prior entanglement.

Result V.3. *Protocol soundness against adversaries with small amounts of prior entanglement: For all $\delta > 0$, the Bell-test protocol of Sect. II is δ-sound against adversaries with small amounts of prior entanglement when quantified in terms of robustness of entanglement. The adversaries can freely use classical and quantum computation and communication during each trial.*

This is proven in Sect. V E. The amount of prior entanglement allowed depends on the quality of the Bell test used by the verifiers and prover. The amount of prior entanglement is quantified either as maximum robustness of entanglement over trials, or as average robustness of entanglement over trials. For the experiment, the upper bounds on allowed prior entanglement are smaller than the entanglement created by the source, limiting the strength of the result. The technique used here can be viewed as a quantified strengthening of Lemma 5.3 of Ref. [45]. It is also related to the entanglement forging method of Ref. [59]. Neither our above bounds nor those based on Ref. [45] have so far led to bounds on prior entanglement of adversaries that are higher than the resources used by the verifiers and honest prover when quantum computation and communication during a trial is unrestricted.

To prove soundness against adversaries with significant amounts of prior entanglement requires increasing the complexity of the settings function $f(u_{\mathsf{A}}, u_{\mathsf{B}})$. For the next two results, l is the bit-length of u_{A} and u_{B}. The results hold with $l = 2l' + 1$ and $f(u_{\mathsf{A}}, u_{\mathsf{B}}) = g_0(u_{\mathsf{A}} \oplus u_{\mathsf{B}})$ where $g_0(x_0, \ldots x_{2l'}) = x_0 + \sum_{i=1}^{l'} x_i x_{i+l'} \mod 2$, where for this purpose we use $\{0, 1\}$ for the range of f and the set of settings choices for the prover.

Result V.4. *Protocol soundness against adversaries with logarithmic prior entanglement: There are efficiently implementable settings functions $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ such that for all $\delta > 0$, the Bell-test protocol of Sect. II is δ-sound against adversaries with prior entanglement on at most $\log(l)/2 - O(1)$ qubits, no quantum communication but arbitrary quantum computation and classical communication during each trial.*

The proof is in Sect. V G. Since quantum communication can be replaced by prior entanglement and classical communication, the soundness can be interpreted as being against adversaries with at most $\log(l)/2 - O(1)$ qubits of prior entanglement and quantum communication. To support practical applications, we fully quantify the allowed prior-entanglement bound. To maintain a given soundness and completeness, it suffices to increase the number of trials by a constant factor. We choose a factor of 6 in the examples given in Sect. V G. The suppressed constant in the bound $\log(l)/2 - O(1)$ depends on the test factor used and its gain. See the paragraphs starting at Eq. 54 in Sect. V G. The proof of Result V.4 is inspired by the proof of Prop. 6.1 in Ref. [63], but requires a different strategy for applying an ϵ-net and a different communication-complexity bound. A specific analysis of the relationship between adversarial test-factor expectations and guessing probabilities for the settings function is required to efficiently preserve completeness under all circumstances.

Instead of bounding prior entanglement and quantum communication, we can bound qubits and gates used by the adversary nearer B.

Result V.5. *Protocol soundness against adversaries with linear quantum operations: There are efficiently implementable settings functions $f(u_{\mathsf{A}}, u_{\mathsf{B}})$ such that for all $\delta > 0$, the Bell-test protocol of Sect. II is δ-sound against adversaries with arbitrary classical communication but no quantum communication, and where the adversary nearer A can use arbitrary local quantum resources, but the adversary nearer B uses $O(l/(\log(q) + 1))$ quantum operations in a basic universal set, where q is the number of qubits used by this adversary.*

The proof of this result is based on the proof of Result V.4 and given at the end of Sect. V G. Again, to support practical applications, we fully quantify the bound on quantum resources allowed by adversaries. The proof of Result V.5 is motivated by the argument in Ref. [15]. The suppressed constant in the bound $O(l/(\log(q)+1))$ depends on the size of the basic universal set of quantum operations, which includes qubit measurements.

In Refs. [15, 63], the preferred settings functions are either random or the inner product function $f(u_{\mathsf{A}}, u_{\mathsf{B}}) = \sum_i u_{\mathsf{A}i} u_{\mathsf{B}i} \mod 2$. Random settings functions are not scalably implementable. The inner product function does not satisfy the assumptions of the Bell-test protocol, see Assumptions 9 and 10. In particular, if one verifier's challenge is all zeros, the value of the inner product function is zero, which determines the prover setting. The inner product function is also a poor choice for establishing adequate lower bounds on the communication complexity of the communication scenario that plays a key role in the proof of Result V.4.

B. Spacetime analysis

For defining the target region and the restrictions on prior entanglement of adversaries, we use a version of the spacetime circuit formalism and adversary analysis introduced in Ref. [50] for QPV protocols with parties that have access to so-called random oracles for defining the settings function. The analysis of Ref. [50] applies to any number of verifiers and needs no restriction on the amount of prior entanglement of adversaries. However, the protocol involves one round consisting of a number of parallel trials, depends on quantum random oracles and does not determine a trial boundary on which to define prior entanglement. We require a different version of this spacetime analysis to identify the adversary capabilities and the trial boundary on which entanglement is restricted. For this purpose, we provide a self-contained description of the formalism and analysis. We first define the target region and the trial boundary on which adversaries' entanglement is restricted. We then reduce the problem of analyzing adversary capabilities without prior entanglement to the case of having two separate, initially unentangled adversaries that can each communicate to the other only once before committing to responses to the verifiers.

Spacetime circuits: We assume familiarity with the fundamental concepts of relativistic spacetime. Our theory applies to any time-oriented globally-hyperbolic Lorenzian manifold $\mathcal{M}$. Minkowski space is the most familiar example of such a manifold. For a comprehensive treatment of the relevant concepts see, for example, Ref. [71]. Here we give a brief overview of the concepts used here. An event or point in spacetime is identified by its position in the manifold. A worldline is a forward-directed light- or time-like path in spacetime. For our purposes, a worldline may be identified with the image of a differentiable function $\boldsymbol{r}(x)$ from an interval I on the reals to spacetime, where I may be open or closed on either end, and $d\boldsymbol{r}(x)/dx$ is forward-directed, light-like or time-like on the interior of I. The worldline is inextendible if it is not a proper subset of another worldline. Event $\boldsymbol{r}'$ causally follows event $\boldsymbol{r}$ if there is a worldline that contains both $\boldsymbol{r}$ and $\boldsymbol{r}'$, and $\boldsymbol{r}'$ follows $\boldsymbol{r}$ on the worldline. Because we assume global hyperbolicity, this defines the causal partial order on spacetime, and we write $\boldsymbol{r} \preceq \boldsymbol{r}'$ if $\boldsymbol{r}'$ causally follows $\boldsymbol{r}$. The closed forward lightcone $C_{\boldsymbol{r},+}$ at $\boldsymbol{r}$ consists of the events that causally follow $\boldsymbol{r}$, including $\boldsymbol{r}$. Similarly, the closed backward lightcone $C_{\boldsymbol{r},-}$ at $\boldsymbol{r}$ consists of the events that causally precede $\boldsymbol{r}$, including $\boldsymbol{r}$. Lightcones have the property that no worldline can exit a forward lightcone or enter a backward lightcone. This means that the intersection of an inextendible worldline with a lightcone is either empty

or a closed half-line. Backward lightcones are instances of past sets, which are defined as subsets J of spacetime such that if $\boldsymbol{r} \in J$ and $\boldsymbol{r}' \preceq \boldsymbol{r}$, than $\boldsymbol{r}' \in J$. Boundaries of past sets are instances of time slices, which we define to be closed space-like subsets of spacetime whose spacetime complement has two connected components, one consisting of the events that causally precede the time slice, and the other consisting of the events that causally follow the time slice. Space-like subsets satisfy that for every pair $\boldsymbol{r}$ and $\boldsymbol{r}'$ in the subset, neither $\boldsymbol{r} \preceq \boldsymbol{r}'$ nor $\boldsymbol{r}' \preceq \boldsymbol{r}$. Global hyperbolicity is equivalent to the existence of a time slices satisfying the property that every inextendible worldline intersects the time slice. In Minkowski space, the boundaries of light cones are examples of time slices that do not satisfy this property. A subset D of spacetime is causally convex if for every worldline that starts and ends in D, the worldline is contained in D. Equivalently, if $\boldsymbol{r} \preceq \boldsymbol{r}'$ are events in D and $\boldsymbol{r}''$ satisfies $\boldsymbol{r} \preceq \boldsymbol{r}'' \preceq \boldsymbol{r}'$, then $\boldsymbol{r}'' \in D$. A special case of a causally convex subset is a light diamond from one event $\boldsymbol{r}$ to another $\boldsymbol{r}'$, which is defined as the intersection of the forward light cone $C_{\boldsymbol{r},+}$ at $\boldsymbol{r}$ and the backward light cone $C_{\boldsymbol{r}',-}$ at $\boldsymbol{r}'$. Light diamonds are empty if $\boldsymbol{r}'$ does not causally follow $\boldsymbol{r}$. The intersection of a worldline with a light diamond forms a time-bounded interval, which may be empty.

The following is a conceptual description of the behavior of entities such as verifiers and provers in spacetime: All entities exist on worldlines. The entities can interact with communications. Communications travel on worldlines starting on one entity's worldline and ending at another's. We refer to the events where an entity sends or receives a communication as interaction vertices and assume that there are finitely many such vertices in every compact region of spacetime.

Spacetime circuits abstract the conceptional description above. Spacetime circuits do not distinguish between worldlines of communications and entities connecting vertices, and confine the dynamics of systems to the vertices. In the conceptual picture, entities may perform operations along their worldlines. In the spacetime circuit, such operations are moved forward to the next interaction vertex along the line and combined into one operation.

Spacetime circuits $\mathcal{S}$ are defined by the following: (1) A set of interaction vertices $\mathcal{V} = \mathcal{V}(\mathcal{S})$, which are events in spacetime. (2) A set of causally directed edges $\mathcal{E} = \mathcal{E}(\mathcal{S})$ representing worldlines connecting two events in spacetime, at least one of which is an interaction vertex in $\mathcal{V}$. The two events connected by an edge E are referred to as the source vertex $V_{\mathrm{src}}(E)$ and the target vertex $V_{\mathrm{tgt}}(E)$, where $V_{\mathrm{tgt}}(E)$ causally follows $V_{\mathrm{src}}(E)$. An edge E is internal if both endpoints are in the set $\mathcal{V}$. It is external if exactly one of the endpoints is in $\mathcal{V}$. (3) For each edge $E \in \mathcal{E}$ an assignment of a quantum system $Q(E)$ carried by E. (4) For each vertex $V \in \mathcal{V}$ a quantum operation $\mathcal{O}(V)$ that jointly transforms incoming quantum systems to outgoing quantum systems at V. For a vertex $V \in \mathcal{V}$, the incoming quantum systems are those associated with edges E whose target vertex is V, that is, $V_{\mathrm{tgt}}(E) = V$. The outgoing quantum systems are those associated with edges E whose source vertex is V. The quantum operation at a vertex without incoming edges is a state preparation on the joint quantum systems of the outgoing edges. The operation at a vertex with no outgoing edges discards the quantum systems of the incoming edges. We refer to the state of quantum systems of edges simply as the state of the edges.

Spacetime circuits describe the dynamics of the edge systems. Intuitively, for any time slice, the systems of edges crossing this time slice have a joint state, and the quantum operations on the vertices determine the joint state on a later time slice given the joint state on an earlier one. To treat this dynamics formally, we introduce a number of causal concepts for spacetime circuits. The vertices and edges of a spacetime circuit form a directed

acyclic graph, which induces a partial order over the union of the vertex and edge sets. A directed path in the circuit is a sequence of alternating vertices and edges $(A_k)_{k=0}^m$ with $m \geq 0$ that starts with either a vertex or an edge, where the vertices and edges along the sequence are incident on each other and in causal order. Specifically, if A_k is an edge and A_{k+1} is a vertex, then $A_{k+1} = V_{\text{tgt}}(A_k)$, and if A_k is a vertex and A_{k+1} is an edge, then $A_k = V_{\text{src}}(A_{k+1})$. Topologically, we can associate to each edge its open worldline between its source vertex and its target vertex. Then the union of the vertices and edges on a directed path is a worldline. We use directed paths to define the circuit partial order. For two vertices or edges A and B, A precedes B in the circuit partial order, written $A \leq B$, iff there is a directed path starts with A and ends with B. This partial order is weaker than the partial order induced by spacetime causality, in the sense that we can have A causally precede B in spacetime but not $A \leq B$ in the circuit partial order.

A vertex subset $\mathcal{V}'$ of $\mathcal{V}$ is causally convex if for every directed path starting and ending in $\mathcal{V}'$, all vertices on the path are in $\mathcal{V}'$. We define the sets of in-edges and out-edges of $\mathcal{V}'$ as $\text{in}(\mathcal{V}') = \{E \in \mathcal{E} : V_{\text{src}}(E) \notin \mathcal{V}', V_{\text{tgt}}(E) \in \mathcal{V}'\}$ and $\text{out}(\mathcal{V}') = \{E \in \mathcal{E} : V_{\text{src}}(E) \in \mathcal{V}', V_{\text{tgt}}(E) \notin \mathcal{V}'\}$. That is, $\text{in}(\mathcal{V}')$ includes all edges entering $\mathcal{V}'$ from outside, while $\text{out}(\mathcal{V}')$ includes all edges exiting $\mathcal{V}'$ to the outside. A causally convex set of vertices $\mathcal{V}'$ defines a spacetime subcircuit $\mathcal{S}(\mathcal{V}')$ with vertex set $\mathcal{V}'$ and edge set $\mathcal{E}'$ consisting of the edges in $\mathcal{E}$ with at least one endpoint in $\mathcal{V}$. The external edges of $\mathcal{S}(\mathcal{V}')$ consist of the union of $\text{in}(\mathcal{V}')$ and $\text{out}(\mathcal{V}')$. A causally convex set of vertices $\mathcal{V}'$ and its associated spacetime subcircuit $\mathcal{S}(\mathcal{V}')$ define a quantum operation $\mathcal{O}(\mathcal{V}')$ that maps the quantum systems of its in-edges to those of its out-edges. This quantum operation can be constructed recursively as follows: 1) Every set $\mathcal{V}'$ consisting of a single vertex V is causally convex, and its quantum operation is given by $\mathcal{O}(\mathcal{V}') = \mathcal{O}(V)$. 2) If $\mathcal{V}'$ consists of more than one vertex, choose a vertex $Z \in \mathcal{V}'$ that is maximal in the circuit partial order restricted to $\mathcal{V}'$. That is, no other vertex in $\mathcal{V}'$ lies after Z along a directed path. Then the reduced set $\mathcal{V}'' = \mathcal{V}' \setminus \{Z\}$ is causally convex. Suppose that we have constructed the quantum operation $\mathcal{O}(\mathcal{V}'')$ of $\mathcal{V}''$. The in-edges of Z partition into $\mathcal{Z}_1 = \text{in}(\{Z\}) \cap \text{out}(\mathcal{V}'')$ consisting of the in-edges of Z that are out-edges of $\mathcal{V}''$, and $\mathcal{Z}_2 = \text{in}(\{Z\}) \cap \text{in}(\mathcal{V}')$ consisting of the in-edges of Z that are in-edges of $\mathcal{V}'$. The in-edges of $\mathcal{V}'$ consist of the in-edges of $\mathcal{V}''$ and the edges of $\mathcal{Z}_2$. The quantum operation of $\mathcal{V}'$ is constructed by adjoining the quantum systems of edges in $\mathcal{Z}_2$ to those of the edges $\text{in}(\mathcal{V}'')$, applying $\mathcal{O}(\mathcal{V}'')$ to the latter, and then applying the $\mathcal{O}(Z)$ to both the systems of edges in $\mathcal{Z}_2$ and the systems of edges in $\mathcal{Z}_1$ that appeared at the output of $\mathcal{O}(\mathcal{V}'')$.

The spacetime-circuit notion of a time slice and its state is captured by an antichain of edges and the joint state of their systems. A set of edges $\mathcal{A}$ is an antichain if no pair of edges in $\mathcal{A}$ is comparable in the circuit partial order, that is, there is no directed path connecting two edges in $\mathcal{A}$. For an antichain of edges $\mathcal{A}$, consider the set $\mathcal{V}'$ of vertices V for which there exists an edge E in $\mathcal{A}$ such that $V \leq E$. Intuitively, $\mathcal{V}'$ consists of the vertices that lie in the causal past of the antichain $\mathcal{A}$. Then the set $\mathcal{V}'$ is causally convex. To see this, consider a directed path from vertex $V_1 \in \mathcal{V}'$ to vertex $V_2 \in \mathcal{V}'$, and let V_3 be a vertex on the path. By definition of the circuit partial order, $V_3 \leq V_2$, and by construction of $\mathcal{V}'$, $V_2 \leq E$ for some edge $E \in \mathcal{A}$. Therefore, $V_3 \leq E$, which implies that $V_3 \in \mathcal{V}'$. The state of the systems of the antichain is determined by the joint state of the quantum systems on the in-edges of $\mathcal{V}'$. The quantum operation $\mathcal{O}(\mathcal{V}')$ transforms this joint state into a joint state of the quantum systems of its out-edges. Because the source vertices of the edges in $\mathcal{A}$ are in $\mathcal{V}'$, the edges in $\mathcal{A}$ are among the out-edges of $\mathcal{V}'$. The state of $\mathcal{A}$ is obtained by tracing out the systems on the out-edges of $\mathcal{V}'$ that are not in $\mathcal{A}$.

For a spacetime circuit $\mathcal{S}$ with vertex set $\mathcal{V}$ and a subset D of spacetime, the intersection $\mathcal{V} \cap D$ consists of all vertices in $\mathcal{V}$ that lie within the region D. If D is causally convex in spacetime, then $\mathcal{V} \cap D$ is causally convex in the spacetime circuit $\mathcal{S}$. This is because the circuit partial order on the vertices of $\mathcal{V}$ is weaker than the partial order induced by spacetime causality. Consequently, If V and V' are vertices in $\mathcal{V} \cap D$ and V'' is a vertex of $\mathcal{V}$ satisfying $V \leq V'' \leq V'$ in the circuit partial order for $\mathcal{S}$, then $V \preceq V'' \preceq V'$ in the causal partial order, so $V'' \in D$.

Spacetime circuit for verifiers and provers: We model verifiers and provers as a spacetime circuit $\mathcal{S}$ that spans all trials of the Bell-test QPV protocol. We consider a generic trial with start time t. We specify an event with the notation $\boldsymbol{r} = (\boldsymbol{x}, s)$, where $\boldsymbol{x}$ is the space location, and s is the time. For $\mathsf{F} = \mathsf{A}, \mathsf{B}$, let $C_{\mathsf{F},+}$ be the forward lightcone from $\boldsymbol{r}_{\mathsf{F},s} = (\boldsymbol{x}_{\mathsf{F}}, t + s_{\mathsf{F}})$ and $C_{\mathsf{F},-}$ the backward lightcone from $\boldsymbol{r}_{\mathsf{F},r} = (\boldsymbol{x}_{\mathsf{F}}, t + r_{\mathsf{F}})$, where $t + s_{\mathsf{F}}$ and $t + r_{\mathsf{F}}$ are the times at which the verifier F sends a challenge and receives a response, respectively. $C_{\mathsf{F},+}$ is the region within which the challenge sent by F is visible to provers, and $C_{\mathsf{F},-}$ is the region within which it is possible for a prover to send a response to F. Let $D_{\mathsf{F}} = C_{\mathsf{F},+} \cap C_{\mathsf{F},-}$ be the light diamond from $\boldsymbol{r}_{\mathsf{F},s}$ to $\boldsymbol{r}_{\mathsf{F},r}$, representing the region in which a valid challenge–response interaction with the verifier F could occur. The target region for the trial is defined as $E = D_{\mathsf{A}} \cap D_{\mathsf{B}}$, which we assume to be non-empty. The target region for the Bell-test protocol is the union of the target regions for each trial. We analyze spacetime circuits in $d + 1$-dimensional spacetime, where $d \geq 1$. For an illustration of the relevant regions in a $1 + 1$-dimensional Minkowski cross-section containing the verifiers in the experiment, see Fig. 9. The labeling of the regions is introduced before Eq. (1), when we reduce the available adversary strategies to those achievable by two adversaries.

For each trial, the trial boundary on which entanglement of adversaries is restricted is defined as the time slice that is formed by the boundary T of the set $\mathcal{M}_{\text{in}} = (C_{\mathsf{A},-} \cup C_{\mathsf{B},-}) \cap (\mathcal{M} \setminus (C_{\mathsf{A},+} \cup C_{\mathsf{B},+})$, where $\mathcal{M}$ is the spacetime manifold. See Fig. 9 for a visualization of T in the $1 + 1$-dimensional cross-section containing the verifiers in the experiment. We refer to T as the trial boundary. The set $\mathcal{M}_{\text{in}}$ is a past set. $\mathcal{M}_{\text{in}}$ can be described as the set of events at which the trial's challenges are not yet visible, but from which it is possible to send a response to at least one of the verifiers A or B. All events relevant to the trial occur in the union $C_{\mathsf{A},-} \cup C_{\mathsf{B},-}$. If $\mathcal{M}'_{\text{in}}$ is the region defined in the same way as $\mathcal{M}_{\text{in}}$ but for the next trial, then $\mathcal{M}'_{\text{in}}$ contains $C_{\mathsf{A},-} \cup C_{\mathsf{B},-}$ of the current trial. This ensures that all events of the current trial occur before the trial boundary of the next one, and, similarly, all events of previous trials occur before the current boundary T. The restriction on adversarial entanglement is formulated in terms of the spacetime circuit's edges that exit $\mathcal{M}_{\text{in}}$ and is defined in terms of the systems listed at (3) after introducing the adversary spacetime circuit.

We now determine the adversary spacetime circuit, which by definition satisfies that there are no vertices in the trial's target region E. Later we generalize the analysis by allowing a limited class of vertices in the target region E to prove that honest provers must employ quantum devices in the target region to successfully pass the protocol. The trial circuit $\mathcal{C}$ is defined as the spacetime circuit associated with the vertices in the causally convex set $(C_{\mathsf{A},-} \cup C_{\mathsf{B},-}) \cap (C_{\mathsf{A},+} \cup C_{\mathsf{B},+})$. The trial record is determined by the trial circuit $\mathcal{C}$ and the state on the in-edges of $\mathcal{C}$. The vertices of $\mathcal{C}$ partition into verifier vertices and adversary vertices, where verifier vertices include any operations that are performed by verifier-owned devices and needed to implement the verifiers' protocol actions. The edges connecting verifiers and adversaries are restricted by Assumption 4. In particular, the only edges from verifiers to adversaries are those carrying the challenges. In addition, there are

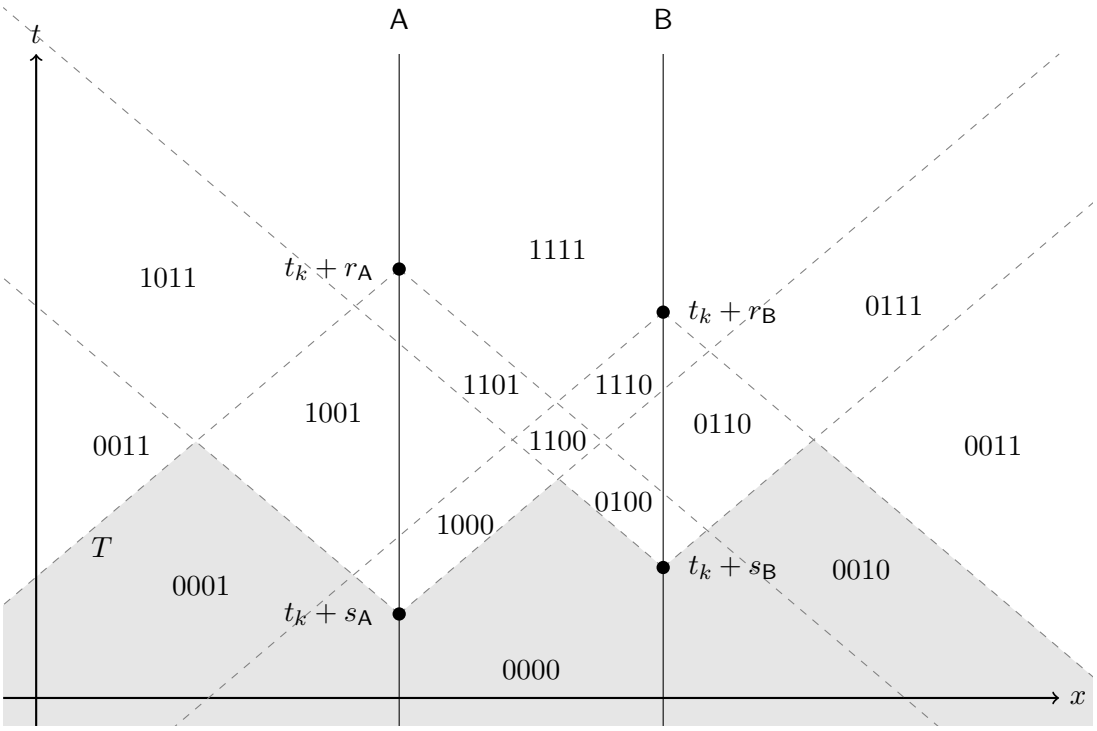

FIG. 9. Schematic $1+1$-dimensional spacetime cross-section of a trial. The trial boundary T is the boundary of the gray-shaded region, which is the $1+1$-dimensional cross-section of the region $\mathcal{M}_{\text{in}}$ defined in the text. The kth trial is shown and t_k is the trial's start time according to the protocol. The regions are labeled by 4-bit strings according to the convention introduced before Eq. (1), where bits 1 and 2 determine whether the points are causally after $t_k + s_\mathsf{A}$ and $t_k + s_\mathsf{B}$, respectively; bits 3 and 4 determine whether the points are not causally before $t_k + r_\mathsf{A}$ and $t_k + r_\mathsf{B}$, respectively. The region labeled with 0011 appears in two places, and regions labeled with 1010 and 0101 do not appear in this cross-section.

no edges from adversaries to verifiers that precede an edge from verifiers to adversaries. This restriction implies that no directed path starting and ending at adversary vertices can pass through a verifier vertex. Therefore, the set of adversary vertices is causally convex. Let $\mathcal{A}$ be the spacetime subcircuit associated with the adversary vertices.

Reduction to two adversaries and one round of communication: Such a reduction is given for a protocol with idealized timing in Ref. [70] for one space dimension. Here we provide the reduction for any number of space dimensions. To analyze the capabilities of the adversary spacetime circuit $\mathcal{A}$, we partition the vertices of the circuit into causally convex subsets determined by the locations of vertices relative to the verifiers' sending and receiving vertices. For every event $\boldsymbol{r}$ in spacetime, we assign a 4-bit string $l(\boldsymbol{r}) = b_1 b_2 b_3 b_4$ defined as follows: $b_1 = 1$ if $\boldsymbol{r} \in C_{\mathsf{A},+}$, $b_2 = 1$ if $\boldsymbol{r} \in C_{\mathsf{B},+}$, $b_3 = 1$ if $\boldsymbol{r} \notin C_{\mathsf{A},-}$, and $b_4 = 1$ if $\boldsymbol{r} \notin C_{\mathsf{B},-}$, with $b_j = 0$ if the condition for $b_j = 1$ is not satisfied. Each bit encodes whether the event $\boldsymbol{r}$ lies inside or outside an associated forward or backward lightcone. We form bit strings by concatenating the bit values. For two such bit strings $\boldsymbol{b}$ and $\boldsymbol{b}'$, we define a partial order on bitstrings by bit-wise comparison and write $\boldsymbol{b} \leq \boldsymbol{b}'$ if each bit of $\boldsymbol{b}$ is less than or equal to the corresponding bit of $\boldsymbol{b}'$. This partial order reflects the causal ordering of events: If an event $\boldsymbol{r}$ causally precedes another event $\boldsymbol{r}'$, then $l(\boldsymbol{r}) \leq l(\boldsymbol{r}')$.

For each 4-bit string $\boldsymbol{b}$, the set $\mathcal{M}_{\boldsymbol{b}} = \{\boldsymbol{r} : l(\boldsymbol{r}) = \boldsymbol{b}\}$ is causally convex. This is because $\mathcal{M}_{\boldsymbol{b}}$ is the intersection of four causally convex sets specified by the values of the bits of $\boldsymbol{b}$. Since the intersection of causally convex sets is also causally convex, $\mathcal{M}_{\boldsymbol{b}}$ inherits this property. The set $\mathcal{M}_{1100}$ is the target region E. This region is the intersection of the four closed lightcones $C_{\mathsf{A},\pm}$ and $C_{\mathsf{B},\pm}$, and is therefore also closed. For a subset B of 4-bit strings, let $\mathcal{M}_B = \cup_{\boldsymbol{b}\in B}\mathcal{M}_{\boldsymbol{b}}$. In general, a subset I of a partially ordered set P is called order-convex if for every $x, y \in I$ with $x \leq y$ and $z \in P$ with $x \leq z \leq y$, we have $z \in I$. In particular, causally convex sets of interaction vertices in a spacetime circuit are order-convex for the circuit partial order. If B is order-convex, then the set $\mathcal{M}_B$ is causally convex. To see this, consider arbitrary events $\boldsymbol{r}, \boldsymbol{r}' \in \mathcal{M}_B$ and an event $\boldsymbol{r}''$ in $\mathcal{M}$ with $\boldsymbol{r} \preceq \boldsymbol{r}'' \preceq \boldsymbol{r}'$. We have to show that $\boldsymbol{r}'' \in \mathcal{M}_B$. Be the relationship between the causal ordering and the partial order of bit strings, $l(\boldsymbol{r}) \leq l(\boldsymbol{r}'') \leq l(\boldsymbol{r}')$. Since B is order-convex, $l(\boldsymbol{r}'') \in B$ which implies that $\boldsymbol{r}'' \in \mathcal{M}_B$. We adopt the convention that the bit symbol "$*$" stands for either 0 or 1, and a bit string that has $*$'s denotes the set of bit strings obtained by replacing the $*$'s with 0 or 1. With this convention, the strings $\boldsymbol{b}$ with alphabet $\{0, 1, *\}$ denote order-convex subsets of bitstrings and consequently, the corresponding sets $\mathcal{M}_{\boldsymbol{b}}$ are causally convex. The regions that appear in the cross-section shown in Fig. 9 are labeled by bitstrings according to this convention. The regions $\mathcal{M}_{1010}$ and $\mathcal{M}_{0101}$ do not appear in the figure.

Let $\mathcal{V}$ be the set of vertices of $\mathcal{A}$. For an order-convex set B of 4-bit strings, $\mathcal{M}_B$ is causally convex in spacetime, from which it follows that $\mathcal{V} \cap \mathcal{M}_B$ is causally convex in $\mathcal{A}$. Since the set of vertices $\mathcal{V}$ of $\mathcal{A}$ are causally convex in the trial circuit $\mathcal{C}$, $\mathcal{V} \cap \mathcal{M}_B$ is also causally convex in $\mathcal{C}$. We define $\mathcal{A}_B = \mathcal{S}(\mathcal{V} \cap \mathcal{M}_B)$ to be the spacetime subcircuit of $\mathcal{A}$ and of $\mathcal{C}$ determined by the vertices in $\mathcal{V} \cap \mathcal{M}_B$. We partition $\mathcal{A}$ into the following vertex-disjoint spacetime subcircuits:

$$\begin{aligned}
\mathcal{A}_{\text{in}} &= \mathcal{A}_{\{00*0,000*\}},\\
\mathcal{A}_{\mathsf{A},\text{start}} &= \mathcal{A}_{\{100*,1010\}},\\
\mathcal{A}_{\mathsf{B},\text{start}} &= \mathcal{A}_{\{01*0,0101\}},\\
\mathcal{A}_{\mathsf{A},\text{end}} &= \mathcal{A}_{1101},\\
\mathcal{A}_{\mathsf{B},\text{end}} &= \mathcal{A}_{1110},\\
\mathcal{A}_{\text{out}} &= \mathcal{A}_{**11}.
\end{aligned} \tag{1}$$

We let $\mathcal{A}_{\mathsf{A}}$ be the part of the adversary that consists of $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{A},\text{end}}$, and $\mathcal{A}_{\mathsf{B}}$ the part that consists of $\mathcal{A}_{\mathsf{B},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$:

$$\begin{aligned}
\mathcal{A}_{\mathsf{A}} &= \mathcal{A}_{\{100*,1010,1101\}},\\
\mathcal{A}_{\mathsf{B}} &= \mathcal{A}_{\{01*0,0101,1110\}}.
\end{aligned} \tag{2}$$

Because $\mathcal{A}$ contains no vertices in the target region, the bit-string suffixes in the above list do not include 1100. The partition of $\mathcal{A}$ is shown in Fig. 10 with arrows indicating where there can be edges connecting the parts. $\mathcal{A}_{\text{in}}$ is the part of the adversary in $\mathcal{M}_{\text{in}}$ and consists of the vertices of $\mathcal{A}$ that do not yet have access to the challenges but are in principle able to send a response to at least one of the verifiers. $\mathcal{A}_{\text{out}}$ consists of the adversary vertices that can no longer respond to either of the verifiers. The edges from $\mathcal{A}_{\mathsf{A},\text{start}}$ to $\mathcal{A}_{\mathsf{B},\text{end}}$ and from $\mathcal{A}_{\mathsf{B},\text{start}}$ to $\mathcal{A}_{\mathsf{A},\text{end}}$ allow for one-round communication between $\mathcal{A}_{\mathsf{A}}$ and $\mathcal{A}_{\mathsf{B}}$. There is no edge connecting $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$ or connecting $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$, as the two parts are space-like separated from each other.

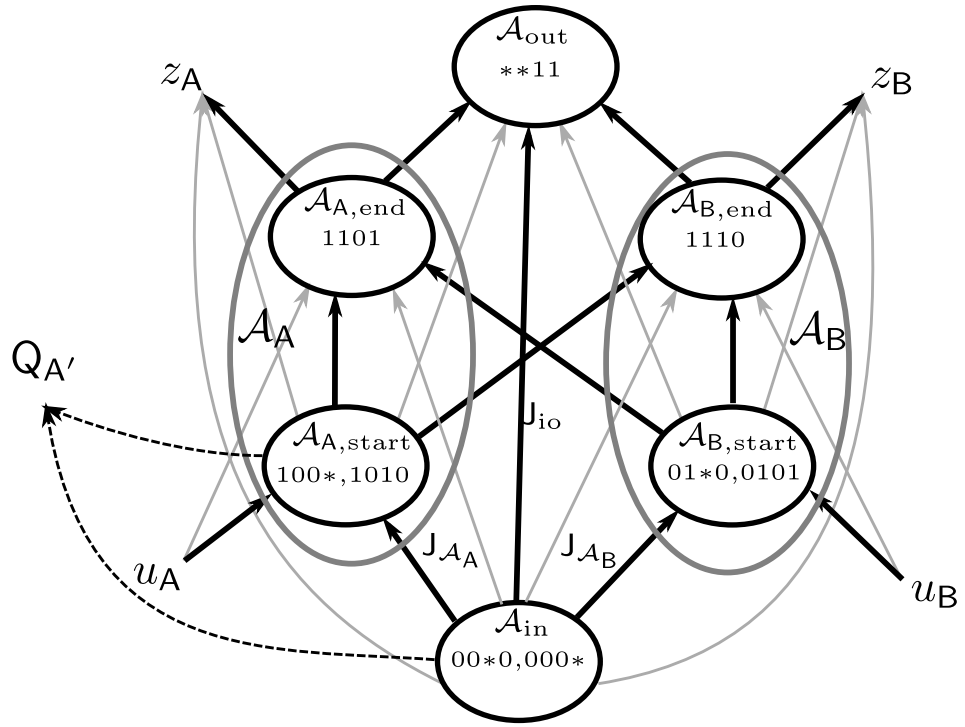

FIG. 10. Adversary spacetime circuit. The diagram depicts the adversary circuits and their interconnections. The circuit receives challenges u_{A} and u_{B} coming from the verifiers and delivers responses z_{A} and z_{B} to the verifiers. The adversary circuit partitions into six subcircuits indicated by the black ellipses. Each subcircuit enacts a quantum operation from its in-edges to its out-edges. The binary strings indicate the associated spacetime regions as described in the text. The entangled quantum systems prepared by the source are $\mathsf{Q}_{\mathsf{A}'}$ and Q_{P}. Because the source is treated as part of the adversary, Q_{P} is carried on the internal edges of the adversary circuit, while $\mathsf{Q}_{\mathsf{A}'}$ may be sent to A' from $\mathcal{A}_{\text{in}}$ or from $\mathcal{A}_{\mathsf{A},\text{start}}$, where we show the edges that alternatively or jointly carry $\mathsf{Q}_{\mathsf{A}'}$ with dashed arrows. By Assumption 8, $\mathcal{A}_{\mathsf{A},\text{end}}$, $\mathcal{A}_{\mathsf{B},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$ cannot send anything to A' before completion of A''s measurement process. We therefore do not need to include edges from these subcircuits for producing $\mathsf{Q}_{\mathsf{A}'}$. The part of the circuit directly relevant to the protocol consists of two groups indicated by gray ellipses and labeled by $\mathcal{A}_{\mathsf{A}}$ on the left and $\mathcal{A}_{\mathsf{B}}$ on the right. The black arrows depict the principal circuit connections. The light-gray edges may be present in the spacetime circuit but are removed by passing them through one of the bypassed adversary circuits in the abstract analysis of adversary capabilities.

The restrictions on prior entanglement are expressed in terms of the joint state of the systems carried by the outgoing edges of $\mathcal{A}_{\text{in}}$, where we exclude direct edges to $\mathcal{A}_{\text{out}}$. These direct edges do not contribute to the current trial of the protocol. The systems carried by edges that may contribute relevant entanglement partition into

1) the system on the edge from $\mathcal{A}_{\text{in}}$ that may contribute to $\mathsf{Q}_{\mathsf{A}'}$ and the system $\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ carried by edges from $\mathcal{A}_{\text{in}}$ to $\mathcal{A}_{\mathsf{A}}$,
2) the system $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ carried by edges from $\mathcal{A}_{\text{in}}$ to $\mathcal{A}_{\mathsf{B}}$. (3)

After we eliminate the indirect edges in the abstract treatment of the circuit below, these edges go to A' and $\mathcal{A}_{\mathsf{A},\text{start}}$, and $\mathcal{A}_{\mathsf{B},\text{start}}$, respectively. In our analysis, we choose to restrict

the entanglement between $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_\mathsf{A}}$ and $\mathsf{J}_{\mathcal{A}_\mathsf{B}}$. Except for the untrusted measurement apparatus of A', the verifiers may be assumed to be classical on exiting $\mathcal{M}_{\text{in}}$. Restrictions on communication by the measurement apparatus are described in Assumption 8.

Because of the absence of adversary vertices in the target region $\mathcal{M}_{1100}$, the vertices of $\mathcal{A}_\mathsf{A}$ and $\mathcal{A}_\mathsf{B}$ are causally convex in the circuit partial order, making $\mathcal{A}_\mathsf{A}$ and $\mathcal{A}_\mathsf{B}$ spacetime subcircuits of $\mathcal{A}$. In particular, because it would violate causality, there is no directed path that starts at a vertex in $\mathcal{A}_\mathsf{A}$, enters $\mathcal{A}_\mathsf{B}$, and returns to $\mathcal{A}_\mathsf{A}$, or vice versa. Consequently, in the absence of the light-grey arrows in Fig. 10, the adversary actions in a trial can be viewed as follows: 1) An initial state preparation by $\mathcal{A}_{\text{in}}$, which is in the causal past of the trial, followed by local processing by $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$. $\mathcal{A}_{\text{in}}$ and $\mathcal{A}_{\mathsf{A},\text{start}}$ may contribute to the system $\mathsf{Q}_{\mathsf{A}'}$ to be measured by A'. 2) A single round of communication consisting of one message from $\mathcal{A}_{\mathsf{A},\text{start}}$ to $\mathcal{A}_{\mathsf{B},\text{end}}$, and another from $\mathcal{A}_{\mathsf{B},\text{start}}$ to $\mathcal{A}_{\mathsf{A},\text{end}}$. 3) Responses sent from $\mathcal{A}_{\mathsf{A},\text{end}}$ to A and from $\mathcal{A}_{\mathsf{B},\text{end}}$ to B. After completing the trial, the adversaries pass a quantum state to $\mathcal{A}_{\text{out}}$ for use in the next trial. The actions of the adversaries can therefore be represented by two adversaries $\mathcal{A}_\mathsf{A}$ and $\mathcal{A}_\mathsf{B}$ who prepare an initial joint state in their past, receive challenges locally, and communicate once in one round of communication before committing to responses to the verifiers. The light-grey arrows in Fig. 10 will be eliminated without weakening the capabilities of the adversary in the abstract analysis of the circuit in Sect. V C.

Projection of target region: A single trial's target region in spacetime is $\mathcal{M}_{1100}$. For visualization, it is helpful to project this region onto the space dimension of an inertial frame. In the experiment, both verifiers are stationary relative to the ground, so we use the inertial frame in which the verifiers are at rest. To determine the spatial projection of $\mathcal{M}_{1100}$, we find the set of space positions $\boldsymbol{x}$ whose stationary worldlines intersect $\mathcal{M}_{1100}$. The stationary worldline for space position $\boldsymbol{x}$ is given by $P = \{(\boldsymbol{x}, t)\}_{t\in\mathbb{R}}$. The distances from $\boldsymbol{x}$ to the verifiers A and B are denoted by l_A and l_B. Let t_1, t_2, t_3, and t_4 be the times at which the worldline P enters or leaves the lightcones $C_{\mathsf{A},+}$, $C_{\mathsf{B},+}$, $C_{\mathsf{A},-}$ and $C_{\mathsf{B},-}$. Then, with units where the speed of light is $c = 1$, we have

$$\begin{aligned} t_1 &= t + s_\mathsf{A} + l_\mathsf{A} \\ t_2 &= t + s_\mathsf{B} + l_\mathsf{B}, \\ t_3 &= t + r_\mathsf{A} - l_\mathsf{A}, \\ t_4 &= t + r_\mathsf{B} - l_\mathsf{B}. \end{aligned} \tag{4}$$

To ensure that the stationary worldline P intersects the target region $\mathcal{M}_{1100}$, the condition $\max(t_1, t_2) \leq \min(t_3, t_4)$ must be satisfied. From the equivalent four inequalities $t_j \leq t_k$ for $j \in \{1, 2\}$ and $k \in \{3, 4\}$, we get

$$\begin{aligned} &t_1 \leq t_3 \leftrightarrow 2l_\mathsf{A} \leq r_\mathsf{A} - s_\mathsf{A}, \\ &t_2 \leq t_4 \leftrightarrow 2l_\mathsf{B} \leq r_\mathsf{B} - s_\mathsf{B}, \\ &t_1 \leq t_4 \leftrightarrow l_\mathsf{A} + l_\mathsf{B} \leq r_\mathsf{B} - s_\mathsf{A}, \\ &t_2 \leq t_3 \leftrightarrow l_\mathsf{A} + l_\mathsf{B} \leq r_\mathsf{A} - s_\mathsf{B}. \end{aligned} \tag{5}$$

The first two inequalities constrain $\boldsymbol{x}$ to lie within spheres centered at the verifiers A and B, while the last two constrain $\boldsymbol{x}$ to ellipsoids with foci at the verifiers. Thus, the projection of $\mathcal{M}_{1100}$ onto space in the verifiers' inertial frame is the intersection of these spheres and ellipsoids.

Classically achievable target region: Classical position verification with the two verifiers A and B involves the verifiers' sending challenges u_A and u_B and the prover responding to both verifiers with the value of $g(u_\mathsf{A}, u_\mathsf{B})$ for some fixed function g. In the spacetime circuit picture, to succeed with probability 1 it is necessary and sufficient to have one prover vertex in the region where u_A and u_B are visible and there is sufficient time to send $g(u_\mathsf{A}, u_\mathsf{B})$ to A, and another prover vertex in the region where u_A and u_B are visible and there is sufficient time to send $g(u_\mathsf{A}, u_\mathsf{B})$ to B. These constraints are related to the analysis of adversaries for classical position verification in Ref. [72]. If either of these conditions is not met, the prover has a nonzero probability of failure. The first condition requires a prover vertex in $C_{\mathsf{A},+} \cap C_{\mathsf{B},+} \cap C_{\mathsf{A},-}$, while the second requires another prover vertex in $C_{\mathsf{A},+} \cap C_{\mathsf{B},+} \cap C_{\mathsf{B},-}$. We identify the union of these two spacetime regions as the effective target region E_c for classical position verification. Although this target region includes the locations of the verifiers themselves, it serves as a useful reference for comparing with the Bell-test protocol's target region E. The projection of E_c onto space in the verifiers' inertial frame can be determined in the same way as the projection of E. To determine the projection of the prover vertex in $C_{\mathsf{A},+} \cap C_{\mathsf{B},+} \cap C_{\mathsf{A},-}$, the inequalities that need to be satisfied are

$$\begin{aligned} &t_1 \leq t_3 \leftrightarrow 2l_\mathsf{A} \leq r_\mathsf{A} - s_\mathsf{A}, \\ &t_2 \leq t_3 \leftrightarrow l_\mathsf{B} + l_\mathsf{A} \leq r_\mathsf{A} - s_\mathsf{B}. \end{aligned} \tag{6}$$

To determine the projection of the prover vertex in $C_{\mathsf{A},+} \cap C_{\mathsf{B},+} \cap C_{\mathsf{B},-}$, the inequalities to be satisfied are

$$\begin{aligned} &t_1 \leq t_4 \leftrightarrow l_\mathsf{B} + l_\mathsf{A} \leq r_\mathsf{B} - s_\mathsf{A}, \\ &t_2 \leq t_4 \leftrightarrow 2l_\mathsf{B} \leq r_\mathsf{B} - s_\mathsf{B}. \end{aligned} \tag{7}$$

Both of these regions are intersections of an ellipsoid and a sphere. The projection of the target region E_c is the union of the two.

C. Reduction to three-party non-signaling and connection to two-party local realism

In the previous section, we reduced the adversary capabilities to two spacetime circuits $\mathcal{A}_\mathsf{A}$ and $\mathcal{A}_\mathsf{B}$ with limited inputs and restricted communication. We now assume that there is no entanglement between $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_\mathsf{A}}$ and $\mathsf{J}_{\mathcal{A}_\mathsf{B}}$, which implies that $\mathcal{A}_\mathsf{A}$ and $\mathcal{A}_\mathsf{B}$ share no prior entanglement. We relax this assumption first for weak entanglement in Sect. V E and then for significant amounts of entanglement in Sect. V G. The goal of this section is to characterize the probability distributions of the trial record achievable by the adversary in terms of a three-party non-signaling scenario. Our use of test factors dictates that we characterize these probability distributions for any joint state of all entities' edges crossing the trial boundary T defining the start of the trial. This joint state includes the trial records of previous trials as well as all other information that can influence the trial. In the spacetime analysis, this accounts for all possible dependencies on the past. To check that a test factor W has expectation at most 1 for all probability distributions achievable by the adversary, it suffices to consider extremal joint states as expectations cannot increase under convex combinations.

Before proceeding, we eliminate redundant edges in the high-level description of the adversary spacetime circuit depicted in Fig. 10. These are the edges associated with connections shown in light-gray. For this purpose, we detach the circuit from its specific spacetime realization and consider it as an abstract quantum circuit with processing vertices

$\mathcal{A}_{\text{in}}, \mathcal{A}_{\mathsf{A},\text{start}}, \mathcal{A}_{\mathsf{A},\text{end}}, \mathcal{A}_{\mathsf{B},\text{start}}, \mathcal{A}_{\mathsf{B},\text{end}}, \mathcal{A}_{\text{out}}$ that are directionally connected with inputs and outputs as shown in Fig. 10. In this abstract circuit, each of the redundant edges bypasses one or more intervening vertices. We can re-route the redundant edges by passing them through the intervening vertices, re-defining the process at these intervening vertices to copy the system carried by the redundant edge from the vertex input to the vertex output. The modified circuit contains no bypassing edges, and it preserves the power of the adversary.

The connection in Fig. 10 with system J_{io} in Fig. 10 is not re-routed because it does not contribute to the actions of the adversary during the trial. Entanglement between it and the systems on other edges into $\mathcal{A}_{\mathsf{A}}$ and $\mathcal{A}_{\mathsf{B}}$ is allowed, provided there is no entanglement between $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ and $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ after tracing out J_{io}. The test factor W does not depend on J_{io}, since the trial record is independent of this system. In addition to J_{io}, we also need to consider the classical state of the verifiers on edges exiting the causal past $\mathcal{M}_{\text{in}}$ of the trial. The classical state of the verifiers may include shared randomness that is kept secret from and therefore independent of other entities before the trial and that can help generate measurement settings and challenges. Such randomness is used to satisfy the independence requirements of Assumption 10. Accordingly, and because the reduced state on $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ and $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ is separable, the adversary's expectation of the test factor W is maximized by some pure product state over the systems $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ and $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$. We now fix such a product state. This state may be assumed to be known by every component of the adversary.

The first half of the adversary strategy, executed by $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$, can be described as follows: $\mathcal{A}_{\mathsf{A},\text{start}}$ transforms its input systems $\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ into output systems $\mathsf{R}_{\mathcal{A}_{\mathsf{A}}}$ and $\mathsf{R}_{\mathcal{A}_{\mathsf{A}}\to\mathcal{A}_{\mathsf{B}}}$, where $\mathsf{R}_{\mathcal{A}_{\mathsf{A}}}$ includes $\mathsf{Q}_{\mathsf{A}'}$. This transformation may depend on the classical challenge u_{A}, as it is an input to $\mathcal{A}_{\mathsf{A},\text{start}}$. The resulting system $\mathsf{R}_{\mathcal{A}_{\mathsf{A}}}$ is then passed on to $\mathcal{A}_{\mathsf{A},\text{end}}$ with the component $\mathsf{Q}_{\mathsf{A}'}$ being passed on to A', and $\mathsf{R}_{\mathcal{A}_{\mathsf{A}}\to\mathcal{A}_{\mathsf{B}}}$ is passed on to $\mathcal{A}_{\mathsf{B},\text{end}}$. Because the quantum inputs to $\mathcal{A}_{\mathsf{A}}$ and $\mathcal{A}_{\mathsf{B}}$ are in a product state, the quantum system $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$, originally transmitted from $\mathcal{A}_{\text{in}}$ to $\mathcal{A}_{\mathsf{B},\text{start}}$, can instead be treated as being internally prepared by $\mathcal{A}_{\mathsf{B},\text{start}}$ in a pure state. The operation by $\mathcal{A}_{\mathsf{B},\text{start}}$ is equivalent to transforming $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ into quantum systems $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}$ and $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$. The joint state of these two systems may depend on the classical challenge u_{B}, which is an input to $\mathcal{A}_{\mathsf{B},\text{start}}$. The systems $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}$ and $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$ are then passed on to $\mathcal{A}_{\mathsf{B},\text{end}}$ and $\mathcal{A}_{\mathsf{A},\text{end}}$, respectively. The communicated systems are shown in Fig. 11 with the bypassing edges removed. Without loss of generality, we may assume that the challenge u_{A} is separately passed on by $\mathcal{A}_{\mathsf{A},\text{start}}$ to both $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$, and similarly, u_{B} is passed on by $\mathcal{A}_{\mathsf{B},\text{start}}$. The relevant second half of the strategy, which determines the classical challenge responses z_{A} and z_{B} recorded by the verifiers, can be treated as measurements performed by $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$. Each of these measurements may depend on both challenges u_{A} and u_{B}. In the following, we modify the adversary strategy while improving the corresponding expectation of the test factor W, and show that for the modified strategy the trial record follows a three-party non-signaling distribution.

We first show that the quantum messages $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$ from $\mathcal{A}_{\mathsf{B},\text{start}}$ can be eliminated without decreasing the expectation of the test factor W for the adversary strategy. Given that the initial state of $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ is a fixed pure state and considering the causal structure shown in Fig. 11, the state of $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$ depends on u_{B} but is independent of the other systems involved in the trial. $\mathcal{A}_{\mathsf{B},\text{start}}$ can therefore prepare it as a partial trace over an ancillary system of a u_{B}-dependent pure state $|\psi_{u_{\mathsf{B}}}\rangle$. $\mathcal{A}_{\mathsf{B},\text{start}}$ can modify $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}$ so that it includes the ancillary system and let $\mathcal{A}_{\mathsf{B},\text{end}}$ remove or ignore it. We can therefore assume that $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$ is in state $|\psi_{u_{\mathsf{B}}}\rangle$ when the challenge is u_{B}. Instead of $\mathcal{A}_{\mathsf{B},\text{start}}$ preparing and passing on $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}}$ and $\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}}}$, we can have $\mathcal{A}_{\mathsf{A},\text{start}}$ prepare independent systems $\mathsf{R}_{\mathcal{A}_{\mathsf{B}},u_{\mathsf{B}}'}\mathsf{R}_{\mathcal{A}_{\mathsf{B}}\to\mathcal{A}_{\mathsf{A}},u_{\mathsf{B}}'}$ for each

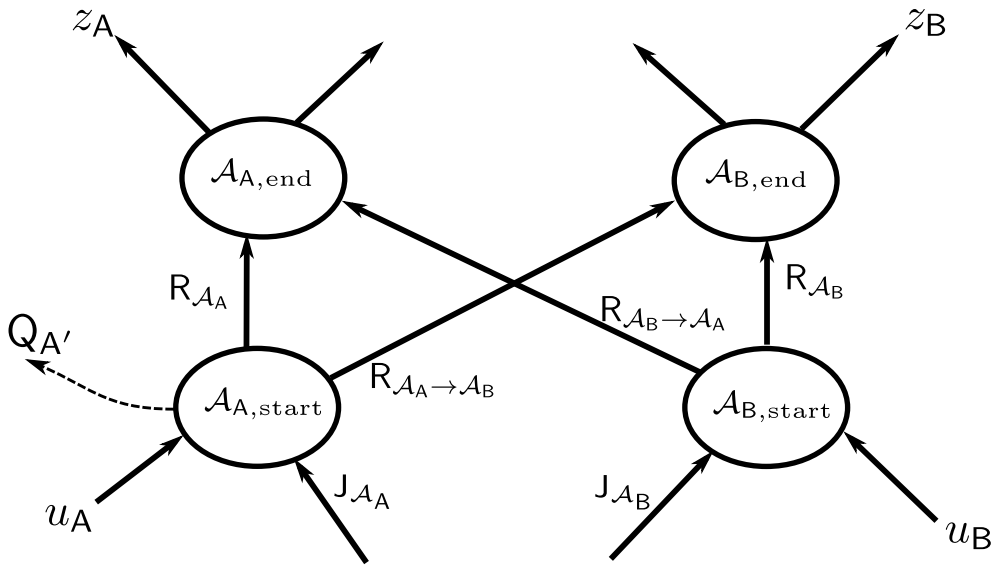

FIG. 11. Adversary spacetime circuit with bypassing edges removed and communicated systems labeled. The dashed arrow may contribute to $\mathsf{Q}_{\mathsf{A}'}$, see Fig. 10.

possible value $u_{\mathsf{B}}{}'$ of U_{B}, where the state of $\mathsf{R}_{\mathcal{A}_\mathsf{B},u_{\mathsf{B}}{}'}\mathsf{R}_{\mathcal{A}_\mathsf{B}\to\mathcal{A}_\mathsf{A},u_{\mathsf{B}}{}'}$ is $|\psi_{u_{\mathsf{B}}{}'}\rangle$. In addition to passing on the systems $\mathsf{R}_{\mathcal{A}_\mathsf{A}}$ and $\mathsf{R}_{\mathcal{A}_\mathsf{A}\to\mathcal{A}_\mathsf{B}}$, $\mathcal{A}_{\mathsf{A},\mathrm{start}}$ passes on all systems $\mathsf{R}_{\mathcal{A}_\mathsf{B}\to\mathcal{A}_\mathsf{A},u_{\mathsf{B}}{}'}$ to $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and all systems $\mathsf{R}_{\mathcal{A}_\mathsf{B},u_{\mathsf{B}}{}'}$ to $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. Since both $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ know the actual challenge u_{B} used in the trial, they can discard all systems $\mathsf{R}_{\mathcal{A}_\mathsf{B},u_{\mathsf{B}}{}'}$ and $\mathsf{R}_{\mathcal{A}_\mathsf{B}\to\mathcal{A}_\mathsf{A},u_{\mathsf{B}}{}'}$ for $u_{\mathsf{B}}{}' \neq u_{\mathsf{B}}$. After this step, the joint state at $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ and $\mathsf{Q}_{\mathsf{A}'}$ relevant for the measurements producing the responses and the verifier measurement outcomes is identical to that in the original strategy. The trial-record probability distribution is therefore unchanged, and the edges leaving $\mathcal{A}_{\mathsf{B},\mathrm{start}}$ only carry the classical message consisting of u_{B} to $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$.

The action by which $\mathcal{A}_{\mathsf{A},\mathrm{start}}$ creates the system $\mathsf{R}_{\mathcal{A}_\mathsf{A}}\mathsf{R}_{\mathcal{A}_\mathsf{A}\to\mathcal{A}_\mathsf{B}}$ may depend on u_{A}. Next we show that the adversary circuit can be modified so that this creation of $\mathsf{R}_{\mathcal{A}_\mathsf{A}}\mathsf{R}_{\mathcal{A}_\mathsf{A}\to\mathcal{A}_\mathsf{B}}$ becomes independent of u_{A}, without reducing the adversary's expectation $\mathrm{Exp}(W)$ of the test factor W. Let $\bar{W}_{u_\mathsf{A}} = \mathrm{Exp}(W|U_{\mathsf{A}} = u_{\mathsf{A}})$ be the expectation of W conditional on the challenge u_{A}. We have $\mathrm{Exp}(W) = \sum_{u_\mathsf{A}} \mathbb{P}(U_{\mathsf{A}} = u_{\mathsf{A}})\bar{W}_{u_\mathsf{A}}$, which is a convex combination of the $\bar{W}_{u_\mathsf{A}}$. It follows that $\max(\bar{W}_{u_\mathsf{A}}) \geq \mathrm{Exp}(W)$. Let $u_{\mathsf{A}0} = \arg\max_{u_\mathsf{A}:\mathbb{P}(u_\mathsf{A})>0} \bar{W}_{u_\mathsf{A}}$. We now modify the adversary circuit so that $\mathcal{A}_{\mathsf{A},\mathrm{start}}$ always prepares $\mathsf{R}_{\mathcal{A}_\mathsf{A}}\mathsf{R}_{\mathcal{A}_\mathsf{A}\to\mathcal{A}_\mathsf{B}}$ as if the challenge from A were $U_{\mathsf{A}} = u_{\mathsf{A}0}$, regardless of the actual u_{A} used in the trial. $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ modify their measurements according to those that would have been performed if the challenge from B were $u_{\mathsf{B}}{}'$, where they jointly choose $u_{\mathsf{B}}{}'$ from the probability distribution of the challenges from B conditional on $U_{\mathsf{A}} = u_{\mathsf{A}0}$ and $M_{\mathsf{P}} = m_{\mathsf{P}} = f(u_{\mathsf{A}}, u_{\mathsf{B}})$, with m_{P} being the actual prover setting and $u_{\mathsf{A}}, u_{\mathsf{B}}$ the actual challenges in this trial. To see that this conditional probability distribution is well-defined, we note that by Assumption 10, M_{P} and U_{A} are independent. Since m_{P} is the actual prover setting, $\mathbb{P}(m_{\mathsf{P}}) > 0$. Similarly, $\mathbb{P}(U_{\mathsf{A}} = u_{\mathsf{A}0}) > 0$ by construction. Due to independence, the joint probability is simply the product $\mathbb{P}(m_{\mathsf{P}}, u_{\mathsf{A}0}) = \mathbb{P}(m_{\mathsf{P}})\mathbb{P}(u_{\mathsf{A}0})$, which is strictly positive. This ensures that the conditioning event $(U_{\mathsf{A}} = u_{\mathsf{A}0}, M_{\mathsf{P}} = m_{\mathsf{P}})$ has non-zero probability and the conditional distribution for choosing $u_{\mathsf{B}}{}'$ is well-defined. The method for choosing $u_{\mathsf{B}}{}'$ defines the random variable $U_{\mathsf{B}}{}'$ and ensures that $f(u_{\mathsf{A}0}, U_{\mathsf{B}}{}') = M_{\mathsf{P}} = f(U_{\mathsf{A}}, U_{\mathsf{B}})$. The joint choice of $u_{\mathsf{B}}{}'$ requires joint randomness, which can be shared by $\mathcal{A}_{\mathsf{A},\mathrm{start}}$ as part of their classical message to $\mathcal{A}_{\mathsf{B},\mathrm{end}}$.

Although $\mathcal{A}_{\mathsf{A},\text{start}}$ does not know m_{P}, $\mathcal{A}_{\mathsf{A},\text{start}}$ can generate the needed shared randomness for each possible value of M_{P} and pass on the shares to $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$. $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$ have both learned m_{P} and can choose from the shared randomness accordingly.

From now on, we abbreviate events and conditions of the form $R = r$ for RVs R by leaving out "$R =$" inside probability and expectation expressions, provided that this does not introduce ambiguity. We claim that the expectation of W after the modifications of $\mathcal{A}_{\mathsf{A}}$ and $\mathcal{A}_{\mathsf{B}}$ is $\bar{W}_{u_{\mathsf{A}0}}$. To show this, let $O_{\mathsf{A}'}{}'$ be the verifier A' outcome RV and $Z_{\mathsf{A}}{}'$ and $Z_{\mathsf{B}}{}'$ the response RVs with the modifications. We compute the expectation of $W(M_{\mathsf{A}'}, O_{\mathsf{A}'}{}', M_{\mathsf{P}}, Z_{\mathsf{A}}{}', Z_{\mathsf{B}}{}')$ below. By construction, and because $M_{\mathsf{A}'}$ is conditionally independent of U_{A} and U_{B} given M_{P} (see Assumption 10), the state being measured to obtain $O_{\mathsf{A}'}{}'$, $Z_{\mathsf{A}}{}'$ and $Z_{\mathsf{B}}{}'$ and the measurements being performed are the same as if $U_{\mathsf{A}} = u_{\mathsf{A}0}$ and $U_{\mathsf{B}} = u_{\mathsf{B}}{}'$, where $M_{\mathsf{P}} = f(u_{\mathsf{A}}, u_{\mathsf{B}}) = f(u_{\mathsf{A}0}, u_{\mathsf{B}}{}')$. We first verify this equivalence by calculating the probability distribution of the settings and outcomes as follows. By assumption and then by construction,

$$\begin{aligned}
\mathbb{P}(m_{\mathsf{A}'}, u_{\mathsf{B}}{}'|m_{\mathsf{P}}) &= \mathbb{P}(m_{\mathsf{A}'}|m_{\mathsf{P}})\mathbb{P}(u_{\mathsf{B}}{}'|m_{\mathsf{P}}) \\
&= \mathbb{P}(m_{\mathsf{A}'}|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}})\mathbb{P}(U_{\mathsf{B}} = u_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}}) \\
&= \mathbb{P}(m_{\mathsf{A}'}, U_{\mathsf{B}} = u_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}}). \qquad (8)
\end{aligned}$$

We apply this identity to determine

$$\begin{aligned}
\mathbb{P}(m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}') &= \sum_{u_{\mathsf{B}}{}'} \mathbb{P}(m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', u_{\mathsf{B}}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}') \\
&= \sum_{u_{\mathsf{B}}{}'} \mathbb{P}(o_{\mathsf{A}'}{}', z_{\mathsf{A}}{}', z_{\mathsf{B}}{}'|m_{\mathsf{A}'}, u_{\mathsf{B}}{}', m_{\mathsf{P}})\mathbb{P}(m_{\mathsf{A}'}, u_{\mathsf{B}}{}'|m_{\mathsf{P}})\mathbb{P}(m_{\mathsf{P}}) \\
&= \sum_{u_{\mathsf{B}}{}'} \mathbb{P}(O_{\mathsf{A}'} = o_{\mathsf{A}'}{}', Z_{\mathsf{A}} = z_{\mathsf{A}}{}', Z_{\mathsf{B}} = z_{\mathsf{B}}{}'|m_{\mathsf{A}'}, U_{\mathsf{A}} = u_{\mathsf{A}0}, U_{\mathsf{B}} = u_{\mathsf{B}}{}', m_{\mathsf{P}}) \\
&\qquad \times \mathbb{P}(m_{\mathsf{A}'}, U_{\mathsf{B}} = u_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}})\mathbb{P}(m_{\mathsf{P}}) \\
&= \sum_{u_{\mathsf{B}}{}'} \mathbb{P}(m_{\mathsf{A}'}, O_{\mathsf{A}'} = o_{\mathsf{A}'}{}', Z_{\mathsf{A}} = z_{\mathsf{A}}{}', Z_{\mathsf{B}} = z_{\mathsf{B}}{}', U_{\mathsf{B}} = u_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}}) \\
&\qquad \times \mathbb{P}(m_{\mathsf{P}}) \\
&= \mathbb{P}(m_{\mathsf{A}'}, O_{\mathsf{A}'} = o_{\mathsf{A}'}{}', Z_{\mathsf{A}} = z_{\mathsf{A}}{}', Z_{\mathsf{B}} = z_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}, m_{\mathsf{P}})\mathbb{P}(m_{\mathsf{P}}) \\
&= \mathbb{P}(m_{\mathsf{A}'}, O_{\mathsf{A}'} = o_{\mathsf{A}'}{}', m_{\mathsf{P}}, Z_{\mathsf{A}} = z_{\mathsf{A}}{}', Z_{\mathsf{B}} = z_{\mathsf{B}}{}'|U_{\mathsf{A}} = u_{\mathsf{A}0}), \qquad (9)
\end{aligned}$$

where in the first two lines, we calculate probabilities involving the RVs $O_{\mathsf{A}'}{}', Z_{\mathsf{A}}{}', Z_{\mathsf{B}}{}'$ according to the modified adversary circuit, and in the last step we used $\mathbb{P}(m_{\mathsf{P}}) = \mathbb{P}(m_{\mathsf{P}}|U_{\mathsf{A}} = u_{\mathsf{A}0})$ by independence of M_{P} and U_{A} according to Assumption 10. We can now calculate the

expectation of W for the modified adversary circuit.

$$\begin{aligned}
&\mathrm{Exp}(W(M_{\mathsf{A}'}, O_{\mathsf{A}'}{}', M_{\mathsf{P}}, Z_{\mathsf{A}}{}', Z_{\mathsf{B}}{}')) \\
&= \sum_{m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}'} W(m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}')\mathbb{P}(m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}') \\
&= \sum_{m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}'} W(m_{\mathsf{A}'}, o_{\mathsf{A}'}{}', m_{\mathsf{P}}, z_{\mathsf{A}}{}', z_{\mathsf{B}}{}')\mathbb{P}(m_{\mathsf{A}'}, O_{\mathsf{A}'} = o_{\mathsf{A}'}{}', m_{\mathsf{P}}, Z_{\mathsf{A}} = z_{\mathsf{A}}{}', Z_{\mathsf{B}} = z_{\mathsf{B}}{}' | U_{\mathsf{A}} = u_{\mathsf{A}0}) \\
&= \sum_{m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}}} W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}})\mathbb{P}(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}} | U_{\mathsf{A}} = u_{\mathsf{A}0}) \\
&= \bar{W}_{u_{\mathsf{A}0}}, \qquad (10)
\end{aligned}$$

where we used Eq. (9) for the second identity and a change of variables for the third.

Because of Assumption 10, the incoming adversary state including that of the source and quantum systems being shared is independent of the verifier settings and challenge distribution. Consequently, with the above modifications to the adversary circuit, we have arrived at a situation where $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ and A' receive a joint quantum state ρ that does not depend on the challenges u_{A}, u_{B} or the measurement setting $m_{\mathsf{A}'}$ used by A' in the trial. The challenges and the measurement setting constitute classical inputs to $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ and A', and the measurements performed by $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ now depend only on $u_{\mathsf{B}}{}'$. We can further modify the measurements so that they depend only on the actual prover setting m_{P}. For this additional modification, consider a particular m_{P}. Then we have, for the above modified adversary circuit and every $u_{\mathsf{B}}{}'$ for which $f(u_{\mathsf{A}0}, u_{\mathsf{B}}{}') = m_{\mathsf{P}}$

$$\mathrm{Exp}(W|m_{\mathsf{P}}, u_{\mathsf{B}}{}') = \mathrm{Exp}(W|u_{\mathsf{B}}{}') \leq \max_{u_{\mathsf{B}}{}''}\{\mathrm{Exp}(W|U_{\mathsf{B}}{}' = u_{\mathsf{B}}{}'') : f(u_{\mathsf{A}0}, u_{\mathsf{B}}{}'') = m_{\mathsf{P}}\}, \qquad (11)$$

where we define conditional expectations and probabilities to be zero whenever the conditioning event has zero probability. The adversary's expectation of W therefore improves if given $m_{\mathsf{P}} = f(u_{\mathsf{A}}, u_{\mathsf{B}})$, $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ always perform the measurements for an input $u_{\mathsf{A}0}, u_{\mathsf{B}}{}''$ that maximizes $\mathrm{Exp}(W|u_{\mathsf{A}0}, u_{\mathsf{B}}{}'')$ subject to $f(u_{\mathsf{A}0}, u_{\mathsf{B}}{}'') = m_{\mathsf{P}}$. With this additional modification, $u_{\mathsf{B}}{}''$ is determined by m_{P}, which implies that the joint randomness used for the first modification can now be omitted.

Instead of providing both u_{A} and u_{B} as inputs, it suffices to provide just m_{P}. The local measurement settings now depend only on the local inputs: $m_{\mathsf{A}'}$ for A', m_{P} for $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, and m_{P} for $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. Because the measurement settings are local to A', $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, and the joint state ρ does not depend on the inputs, we can define the conditional distribution of the three measurement outcomes, namely, the outcome $O_{\mathsf{A}'}$ of A', the response Z_{A} from $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, and the response Z_{B} from $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. This conditional distribution can be defined not only for the case where the inputs for $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ are identical, but also when their inputs are different. Consequently, we constructed a modified adversary strategy that yields an expectation of W at least as high as the original, and for which the probability distribution of the trial record is derived from the full set of conditional distributions according to a three-party, necessarily non-signaling, scenario with a joint quantum state for the three parties, $k_{\mathsf{A}'}$ settings and $c_{\mathsf{A}'}$ outcomes for the first party, A', and k_{P} settings and c_{P} outcomes for each of the second and third parties, $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. The expectation of W is evaluated according to a distribution of inputs where the second and third party's inputs are always the same. It follows that to ensure that W is a test factor against adversaries without prior

entanglement, it suffices to verify that its maximum expected value over all three-party non-signaling distributions does not exceed 1, where the three inputs are $m_{\mathsf{A}'}$, m_{P} and m_{P} with probability distribution $\mathbb{P}(m_{\mathsf{A}'}, m_{\mathsf{P}})$.

Existence of test factors for protocol completeness: The Bell-test protocol requires finding test factors against adversaries that ensure completeness of the protocol for an implementable honest prover. It suffices to consider non-negative candidate functions $W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}})$ that can be computed from the trial record and do not depend directly on u_{A} and u_{B}. To take advantage of the reduction to the three-party quantum scenario derived above, we write W in terms of a function $W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, b; z_{\mathsf{B}}, b')$ according to

$$W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}}) = W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, m_{\mathsf{P}}; z_{\mathsf{B}}, m_{\mathsf{P}}). \tag{12}$$

The reduction to three parties implies that the conditional distributions accessible by adversary strategies for the protocol can be written as

$$\mathbb{P}(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) = \mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b = m_{\mathsf{P}}, b' = m_{\mathsf{P}}), \tag{13}$$

where the conditional distributions $\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')$ satisfy both quantum and non-signaling constraints for three parties according to the configuration described in the previous paragraph. We use b and b' to refer to the inputs of parties $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ in the extended three party scenario where $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ may have different inputs determining their measurements. For a review of the quantum and non-signaling constraints, see Ref. [73]. We do not take advantage of the quantum constraints and instead consider all non-signaling conditional distributions. Let $\nu(m_{\mathsf{A}'}, m_{\mathsf{P}})$ be the probability distribution of $m_{\mathsf{A}'}, m_{\mathsf{P}} = f(u_{\mathsf{A}}, u_{\mathsf{B}})$ used in the Bell-test protocol, where $\nu(m_{\mathsf{A}'}, m_{\mathsf{P}}) > 0$ for each $m_{\mathsf{A}'}, m_{\mathsf{P}}$. Both conditional and unconditional probability distributions are denoted by $\mu, \nu, \ldots$ and treated as functions of their arguments, for which we avoid the use of RV conventions. The candidate function W' is a test factor against adversaries if for every three-party non-signaling conditional distribution $\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')$, we have

$$\mathrm{Exp}(W') = \sum_{m_{\mathsf{A}'}, b, o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}} W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, b; z_{\mathsf{B}}, b)\nu(m_{\mathsf{A}'}, b)\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b) \leq 1. \tag{14}$$

The construction and optimization of functions W' satisfying this condition is described in Sect. V D.

The three-party non-signaling constraints are sufficient for constructing test factors for the general Bell-test protocol. In the case where the prover has only two settings, that is, $k_{\mathsf{P}} = 2$ which is relevant for the experimental implementation, a significant simplification can be obtained by taking advantage of a connection to two-party local realism and standard Bell tests. We use this simplification to prove completeness. For this, we consider only test factors W' that are symmetric under the exchange of parties $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, that is,

$$W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, b; z_{\mathsf{B}}, b') = W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{B}}, b'; z_{\mathsf{A}}, b) \tag{15}$$

for all values of the arguments. As a result, the expectation of W' is unchanged if we symmetrize the non-signaling conditional distribution $\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')$ under the exchange of parties $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$ according to

$$\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b') = \frac{1}{2}\left(\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b') + \mu(o_{\mathsf{A}'}, z_{\mathsf{B}}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b', b)\right). \tag{16}$$

Because the exchanged conditional distribution is also non-signaling, μ' is the mixture of two non-signaling conditional distributions. Since mixtures of non-signaling conditional distributions are non-signaling, μ' is non-signaling.

For any three-party non-signaling distribution $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')$ that is symmetric under the exchange of parties $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, the two marginal conditional distributions $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b)$ for A' and $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, and $\mu'(o_{\mathsf{A}'}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b')$ for A' and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, are identical. In the terminology of Ref. [74], this means that the marginal $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b)$ is 2-shareable. According to "Result 1" in this reference, any such 2-shareable probability distribution with $k_{\mathsf{P}} = 2$ is local realistic. For the marginal conditional distribution $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b)$, the two parties are A' and $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ with inputs $m_{\mathsf{A}'}$ and b, respectively. In the experiment, we directly sample according to $\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b)$ with probability $\nu(m_{\mathsf{A}'}, b)$. From this we can in principle simulate samples from $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b)$ by swapping z_{A} with z_{B} with probability 1/2 before recording the trial results. Because of symmetry of W', it is not necessary to do this explicitly. That is, the value of W' is not affected by explicitly performing the probabilistic swap.

We do not directly observe different inputs b, b' to $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, but the analysis above shows that there exists a three-party non-signaling distribution consistent with the observations. By non-signaling of the distribution μ', $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b, b') = \mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b, b)$. That is, the observed marginal for A' and $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ conditional on their inputs does not depend on which b' is given as input to $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. Consequently, although we cannot directly observe input combinations with $b \neq b'$, for the three-party non-signaling distribution implied by the best adversary strategies obtained above, after symmetrization the observable conditional marginal for A' and $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ is the 2-shareable and hence local-realistic marginal conditional distribution $\mu'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, b)$ for all such adversary strategies.

Ideally, and with high probability in the experiment, the honest prover's responses satisfy $z_{\mathsf{A}} = z_{\mathsf{B}}$. Let $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ be the achieved conditional distribution for the honest prover, and $\sigma'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) = (\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) + \sigma(o_{\mathsf{A}'}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}}))/2$ its symmetrized marginal. The honest prover behaves so that $\sigma'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ violates some two-party Bell inequality with $k_{\mathsf{P}} = 2$ values for input m_{P}. Therefore, σ' is outside the local-realistic polytope. Given the conditional distribution $\sigma'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ and the input distribution $\nu(m_{\mathsf{A}'}, m_{\mathsf{P}})$, a test factor $W_{\mathrm{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, m_{\mathsf{P}})$ against local realism can be constructed that witnesses the Bell violation and has positive gain [47]. Here, $W_{\mathrm{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, m_{\mathsf{P}})$ is non-negative and satisfies that its expectation is at most 1 under any local realistic predictions. If we set

$$W'(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, b; z_{\mathsf{B}}, b') = \frac{1}{2}(W_{\mathrm{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, b) + W_{\mathrm{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{B}}, b')) \tag{17}$$

then W' is symmetric under the exchange of parties $\mathcal{A}_{\mathsf{A},\mathrm{end}}$ and $\mathcal{A}_{\mathsf{B},\mathrm{end}}$. Therefore, the argument presented in the paragraphs above applies. Given this and in view of the conditions satisfied by the test factor W_{LR} against local realism, we conclude that W' is a test factor against three-party non-signaling conditional distributions, where the three inputs have probability distribution $\nu(m_{\mathsf{A}'}, b)\delta_{b,b'}$. Here, $\delta_{b,b'}$ is the Kronecker delta. Consequently, W' satisfies Eq. (14) and so is a test factor against adversary strategies.

Because W_{LR} has positive gain for the symmetrized marginal $\sigma'(o_{\mathsf{A}'}, z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ introduced above, W' has positive gain g for the honest distribution in the experiment, but expectation at most 1 for adversary strategies. Therefore, the completeness of the Bell-test protocol is ensured, provided that the number of trials is large enough. Since the gain g is the expectation of $\log(W')$ for the honest distribution and the p-value against adversaries

after a finite number of trials is given by $1/(\prod_k w'_k)$ with w'_k being the value taken by W' in the k'th trial, an estimate for the number of trials required for soundness δ is $\bar{n} = \log(1/\delta)/g$. If $n \geq \bar{n}$, then the average $\log(p)$-value for the protocol is $ng \geq \log(1/\delta)$. For completeness $\epsilon \in (0,1)$, the probability of passing needs to be at least $1-\epsilon$. The probability of passing is the probability that the $\log(p)$-value for the protocol is at least $\log(1/\delta)$. Since the honest distribution is close to i.i.d. and the number of trials n is large, the $\log(p)$-value is expected to be close to Gaussian with variance n times the variance of $\log(W')$. We use this as a heuristic to choose the number of trials required to ensure soundness and completeness given a reference honest distribution in the experimental protocol, see Sect. VI.

To obtain a mathematical bound on the number of trials required for completeness ϵ, we may conservatively assume that at each trial, conditional on the past, the expected value of $\log(W')$ is at least some value g that can be estimated before the experiment based on the expected honest distribution. Optimized test factors W' generally satisfy that $\log(W')$ is bounded. Whenever necessary, this boundedness can be enforced by mixing W' with the trivial test factor 1. Suppose that $\log(W') \in [\lambda_l, \lambda_u]$ and let $c = |\lambda_u - \lambda_l|$. Let W'_k be W' evaluated on the k'th trial record and $M_k = \sum_{j=1}^{k} \log\bigl(W'_j\bigr)$. According to the Azuma-Hoeffding inequality,

$$\mathbb{P}(M_n < ng - \Delta) \leq e^{-\Delta^2/(2nc^2)}. \tag{18}$$

For completeness ϵ, we set $\Delta = c\sqrt{2n\log(1/(1-\epsilon))}$. To determine the number of trials required, we solve for the minimum n satisfying

$$\begin{aligned} \log(1/\delta) &\leq ng - \Delta \\ &= ng - c\sqrt{2n\log(1/(1-\epsilon))} \\ &= \sqrt{ng}(\sqrt{ng} - c\sqrt{2\log(1/(1-\epsilon))/g}). \end{aligned} \tag{19}$$

This inequality is satisfied if $\sqrt{ng} = \sqrt{\log(1/\delta)} + c\sqrt{2\log(1/(1-\epsilon))/g}$, because then both $\sqrt{ng}$ and $\sqrt{ng} - c\sqrt{2\log(1/(1-\epsilon))/g}$ are greater than or equal to $\sqrt{\log(1/\delta)}$. Consequently, it suffices to set

$$n = 2\bigl(\log(1/\delta)/g + 2c^2\log(1/(1-\epsilon))/g^2\bigr), \tag{20}$$

which shows that for completeness ϵ with soundness δ, $O(\log(1/\delta)) + O(\log(1/(1-\epsilon)))$ trials suffice. The constants depend on the performance of the honest prover and are finite provided that g is positive.

The connection to the hypothesis test against local realism implies that the Bell-test protocol is robust against experimental imperfections such as loss. In our experimental implementation, both the verifier A' and the prover P have two measurement settings, where under each setting there are two possible measurement outcomes. That is, $k_{\mathsf{A}'} = k_{\mathsf{P}} = 2$ and $c_{\mathsf{A}'} = c_{\mathsf{P}} = 2$. According to Ref. [75], it is possible to violate a two-party Bell inequality in this configuration for any photon loss less than 1/3 for each party. In our experiment, the source is near A''s measurement apparatus, so A' can have substantially less loss than the honest prover. If A' has no loss, then any photon loss between the source and the honest prover of less than 1/2 can be tolerated [76].

The test-factor construction described in Sect. V D below for position verification uses a test factor for two-party local realism as a starting point but exploits the greater flexibility

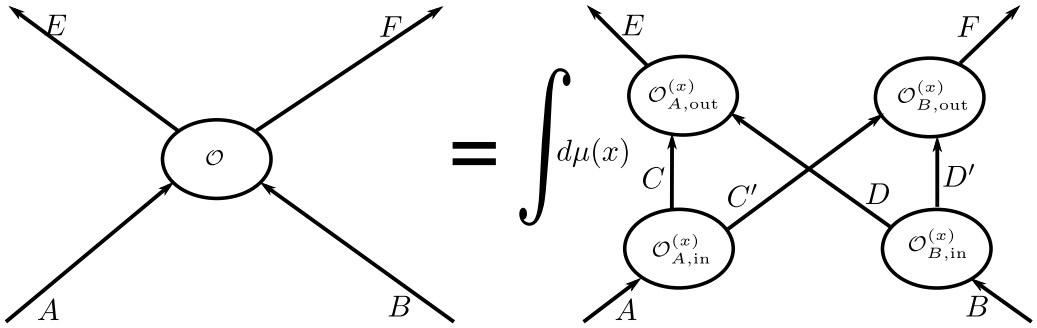

FIG. 12. Removable quantum operations.

of the three-party non-signaling constraints. In general, the test factors constructed directly for the non-signaling constraints without taking advantage of the simplification based on local realism need not be expressable in terms of Eq. (17).

Requirement for quantum vertices in the target region: The Bell-test QPV protocol fails if the honest prover uses only classical devices. Here, a device is considered classical if it decoheres its inputs in a fixed product basis and produces separable outputs for each product basis input. We claim that no prover, either honest or adversarial, who uses only classical devices in the target region can successfully pass the protocol. Consider adversary circuits that include vertices in the target region $\mathcal{M}_{1100}$. According to our conventions, such vertices change the adversary into an honest prover. Here we show that if the joint action of the vertices in $\mathcal{M}_{1100}$ satisfies the definition of being removable given below, then the adversary can not increase their expectation of the test factors used in the protocol over an adversary with no such vertices. We also show that if the vertices act classically as formally defined below, then they are removable. This is the sense in which an honest prover must have a quantum presence in the target region to pass the protocol.

The $\mathcal{A}_{1100}$ subcircuit is now not empty and is part of the adversary circuit $\mathcal{A}$ and needs to be added to the list of components of $\mathcal{A}$ in Eq. (1). $\mathcal{A}_{1100}$ receives inputs from $\mathcal{A}_{\text{in}}$, $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$ and passes outputs to $\mathcal{A}_{\text{out}}$, $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$. As in the analysis of $\mathcal{A}$ given above, to simplify the circuit, without loss of generality, we can pass the input from $\mathcal{A}_{\text{in}}$ through $\mathcal{A}_{\mathsf{A},\text{start}}$ or $\mathcal{A}_{\mathsf{B},\text{start}}$ and pass the output to $\mathcal{A}_{\text{out}}$ through $\mathcal{A}_{\mathsf{A},\text{end}}$ or $\mathcal{A}_{\mathsf{B},\text{end}}$. With this simplification, the quantum operation of $\mathcal{A}_{1100}$ can be viewed as an operation with two inputs, from $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$, and two outputs, to $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$.

Consider a general two-input two-output quantum operation $\mathcal{O} : \rho_{AB} \mapsto \mathcal{O}(\rho_{AB})_{EF}$, where A, B are the input quantum systems and E, F the output quantum systems. For $\mathcal{A}_{1100}$, the operation is implemented by $\mathcal{A}_{1100}$, and the inputs and outputs are the edges from $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$, and to $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$. We call $\mathcal{O}$ removable if 1) there exist quantum systems C, C', D, D', a probability distribution $\mu(x)$ over a classical variable x, and the following x-dependent quantum operations: $\mathcal{O}^{(x)}_{A,\text{in}}$ from A to C and C', $\mathcal{O}^{(x)}_{B,\text{in}}$ from B to D and D', $\mathcal{O}^{(x)}_{A,\text{out}}$ from C and D to E, and $\mathcal{O}^{(x)}_{B,\text{out}}$ from C' and D' to F. And 2) these quantum operations satisfy

$$\mathcal{O} = \int d\mu(x) \left(\mathcal{O}^{(x)}_{A,\text{out}} \otimes \mathcal{O}^{(x)}_{B,\text{out}}\right) \circ \left(\mathcal{O}^{(x)}_{A,\text{in}} \otimes \mathcal{O}^{(x)}_{B,\text{in}}\right). \tag{21}$$

The right-hand side is the replacement for the removable operation $\mathcal{O}$. The situation is illustrated in Fig. 12.

To make the correspondence between $\mathcal{A}_{1100}$ and $\mathcal{O}$, let A be the system coming in from $\mathcal{A}_{\mathsf{A},\text{start}}$, B the system coming in from $\mathcal{A}_{\mathsf{B},\text{start}}$, E the system going to $\mathcal{A}_{\mathsf{A},\text{end}}$ and F the system going to $\mathcal{A}_{\mathsf{B},\text{end}}$. To show that an adversary with removable $\mathcal{A}_{1100}$ can do no better than an adversary with no vertices in the target region, with $\mathcal{O}$ the quantum operation implemented by $\mathcal{A}_{1100}$, let $\mathcal{O}^{(x)}_{A,\text{in}}, \mathcal{O}^{(x)}_{A,\text{out}}, \mathcal{O}^{(x)}_{B,\text{in}}$ and $\mathcal{O}^{(x)}_{B,\text{out}}$ be the operations for the replacement in Eq. (21). We modify the other components of the adversary circuit to eliminate $\mathcal{A}_{1100}$. For this purpose, we modify $\mathcal{A}_{\text{in}}$ and its outgoing edges so that $\mathcal{A}_{\text{in}}$ randomly selects x with probability distribution $\mu(x)$ and sends x to both $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$ by adding it to its outgoing edges. After completing their existing quantum operations, $\mathcal{A}_{\mathsf{A},\text{start}}$ and $\mathcal{A}_{\mathsf{B},\text{start}}$ apply $\mathcal{O}^{(x)}_{A,\text{in}}$ to A and $\mathcal{O}^{(x)}_{B,\text{in}}$ to B, respectively, producing systems C, C' and D, D'. They then send C and D to $\mathcal{A}_{\mathsf{A},\text{end}}$, and C' and D' to $\mathcal{A}_{\mathsf{B},\text{end}}$, along with x. Upon receiving these systems and x, $\mathcal{A}_{\mathsf{A},\text{end}}$ and $\mathcal{A}_{\mathsf{B},\text{end}}$ apply $\mathcal{O}^{(x)}_{A,\text{out}}$ to C and D, and $\mathcal{O}^{(x)}_{B,\text{out}}$ to C' and D', which after averaging over x, reproduces the outputs they previously obtained from $\mathcal{A}_{1100}$ in the original circuit.

Removable quantum operations include quantum routers that forward subsystems from the inputs to the outputs according to a fixed routing rule independent of the inputs, and semiclassical operations that decohere the inputs in a product basis and map each basis element to a separable state. To implement such a semiclassical operation in the form given on the right-hand side of Eq. (21), decoherence is implemented by operations $\mathcal{O}_{A,\text{in}}$ and $\mathcal{O}_{B,\text{in}}$ that do not depend on the classical variable x introduced below. These operations implement the decoherence in their respective bases and copy their decohered states to their outputs. Let the basis elements of A and B be denoted by $|a\rangle_A$ and $|b\rangle_B$, where a and b range over finite index sets. The semiclassical operation produces, for each product state $|a\rangle_A |b\rangle_B$, a separable state. We express this separable state as an explicit mixture $\sum_{x_{a,b}} \mu_{a,b}(x_{a,b}) |\phi(x_{a,b})\rangle_E |\psi(x_{a,b})\rangle_F$, where for each a, b, $\mu_{a,b}(x_{a,b})$ is a probability distribution on a finite set depending on a, b. We let the variable x be a list $(x_{a,b})_{a,b}$, where each $x_{a,b}$ is independently distributed with probability distribution $\mu_{a,b}(x_{a,b})$. The operations $\mathcal{O}^{(x)}_{A,\text{out}}$ and $\mathcal{O}^{(x)}_{B,\text{out}}$ are then designed to independently prepare the states $|\phi(x_{a,b})\rangle_E$ and $|\psi(x_{a,b})\rangle_F$, given the inputs $|a\rangle_C |b\rangle_D$ and $|a\rangle_{C'} |b\rangle_{D'}$, respectively. After integrating over the probability distribution of x, this reproduces the original semiclassical operation. Such semiclassical operations include probabilistic classical computations and Markov processes that decohere the inputs in a product basis and perform probabilistic classical computation given the input basis elements to produce a random output, which is then encoded in a fixed output product basis.

D. Test-factor construction

Test factors for detecting adversaries without prior entanglement are constructed to maximize the gain per trial for the actual honest prover. This requires estimates of the conditional probabilities $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}} | m_{\mathsf{A}'}, m_{\mathsf{P}})$ for the protocol with the actual prover. The quality of the estimates does not affect the validity of the test factors constructed but may affect the probability of success for the honest prover. We obtain such estimates from calibration data obtained before runs of the protocol. Ideally, the prover always responds with $z_{\mathsf{A}} = z_{\mathsf{B}}$. However, in the experiment we observe trials where $z_{\mathsf{A}} \neq z_{\mathsf{B}}$. The method for estimating $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}} | m_{\mathsf{A}'}, m_{\mathsf{P}})$ is described in Sect. VI. Here, we assume that $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}} | m_{\mathsf{A}'}, m_{\mathsf{P}})$

is given. The associated trial-wise probability distribution is

$$\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, m_{\mathsf{P}}) = \sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})\mu_{\text{set}}(m_{\mathsf{A}'}, m_{\mathsf{P}}), \tag{22}$$

where $\mu_{\text{set}}(m_{\mathsf{A}'}, m_{\mathsf{P}})$ is the probability of $m_{\mathsf{A}'}$ and $m_{\mathsf{P}} = f(u_{\mathsf{A}}, u_{\mathsf{B}})$, which is determined by the joint distribution of settings and challenges. See Assumption 10. Let $\mathcal{P}$ be the polytope of conditional distributions $\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')$ that are three-party non-signaling. See [77] for a detailed treatment of the geometry and extremal structure of $\mathcal{P}$. For such a $\mu \in \mathcal{P}$, define

$$\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, b, b') = \mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, b, b')\mu_{\text{set}}(m_{\mathsf{A}'}, b)\delta_{b,b'}. \tag{23}$$

We can maximize the gain by solving the following problem:

Gain optimization problem:

Given: Conditional probabilities $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$.

Maximize over W:

$$\text{Exp}_\sigma(\log(W)) = \sum_{o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, m_{\mathsf{P}}} \log(W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}}))\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, m_{\mathsf{P}}). \tag{24}$$

Constraints:

$$\begin{aligned} W &\geq 0, \\ \text{Exp}_\mu(W) &= \sum_{o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, m_{\mathsf{P}}} W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}})\mu(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}, m_{\mathsf{A}'}, m_{\mathsf{P}}, m_{\mathsf{P}}) \\ &\leq 1 \quad \text{for all } \mu \in \mathcal{P}. \end{aligned} \tag{25}$$

The gain optimization problem requires maximizing a concave function over the convex set $\mathcal{P}$ of three-party non-signaling distributions. For the case of two settings per party and two outcomes for each setting as in our experimental implementation, this can be done directly. However, to simplify the calculations, we restricted the test factor W to be symmetric under the exchange of z_{A} and z_{B}. In our experiment, the probability of $z_{\mathsf{A}} \neq z_{\mathsf{B}}$ is very low, on the order of 10^{-6}. If it is zero, then the probability distribution of $(o_{\mathsf{A}'}, z_{\mathsf{A}}, m_{\mathsf{A}'}, m_{\mathsf{P}})$ conditional on $z_{\mathsf{A}} = z_{\mathsf{B}}$ for three-party non-signaling is local realistic, a consequence of the results of Ref. [74] discussed above. With this as motivation, we took advantage of existing strategies for optimizing the gain of test factors in two-party Bell tests against local realism [47]. From calibration data, we estimated the actual prover's probability distribution of $(o_{\mathsf{A}'}, z_{\mathsf{A}}, m_{\mathsf{A}'}, m_{\mathsf{P}})$ conditional on $z_{\mathsf{A}} = z_{\mathsf{B}}$, and then determined the gain-optimizing test factor $W_{\text{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, m_{\mathsf{P}})$ against local realism given this probability distribution. We then considered test factors W of the form

$$W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}}) = \delta_{z_{\mathsf{A}}, z_{\mathsf{B}}} W_{\text{LR}}(o_{\mathsf{A}'}, m_{\mathsf{A}'}; z_{\mathsf{A}}, m_{\mathsf{P}}) + \lambda(1 - \delta_{z_{\mathsf{A}}, z_{\mathsf{B}}}), \tag{26}$$

where λ was chosen to have the largest value for which the resulting W is a test factor against three-party non-signaling, that is, for which W satisfies the constraints of the gain optimization problem. Although these test factors are not gain-optimal, they are robust against variations in the small probability that $z_{\mathsf{A}} \neq z_{\mathsf{B}}$. Empirical tests indicate that for the probability distributions encountered in our experiments, near-optimal gain is achieved.

E. Test factors against small amounts of prior entanglement

We can exclude small amounts of prior entanglement shared by adversaries, depending on the expectation of the chosen test factor that can be achieved by the actual prover. For general protocols, such bounds can be derived from Lemma 5.3 of Ref. [45]. This lemma obtains a lower bound of order δ'/δ on the total dimension of the entangled quantum systems required for adversaries to pass the protocol if the protocol is δ-sound when adversaries start with separable states and the soundness is changed to $\delta' > \delta$ to account for prior entanglement. This bound is obtained by replacing the adversary's entangled state before the protocol with the completely mixed state and disallowing quantum communication during the protocol. Here we develop quantified bounds that apply trial-by-trial or on average over a protocol by taking advantage of properties of test factors. For bounds that overcome this limitation by increasing the complexity of the settings function, see Sect. V G. We quantify entanglement in terms of the robustness of entanglement [78], see also Ref. [79]. For a density matrix ρ, we write $R(\rho)$ for the robustness of entanglement. $R(\rho)$ is an entanglement monotone. For separable states, $R(\rho) = 0$, and for entangled states, $R(\rho) > 0$. In particular, for a bipartite pure entangled state ρ with Schmidt coefficients a_j, we have $R(\rho) = \left(\sum_j a_j\right)^2 - 1$. The significance of robustness of entanglement is that for every bipartite entangled state ρ in finite dimensions, there exist separable states σ and τ such that $\rho = (1+\xi)\sigma - \xi\tau$ with $\xi = R(\rho)$, and no such representation exists with $0 \leq \xi < R(\rho)$. These properties of robustness of entanglement can be found in Ref. [78]. We also remark that for a bipartite pure state ρ, the Rényi entropy of order $1/2$ of its reduced density matrix is given by $S_{1/2} = \log(R(\rho) + 1)$, which establishes a direct connection between the robustness of entanglement and a conventional entropic measure.

Consider a test factor W for the Bell-test protocol against adversaries without prior entanglement. Let $\bar{w}_{\min}$ be the average over settings and challenges of the minimum value of W conditional on the settings and challenges, that is

$$\bar{w}_{\min} = \sum_{m_{\mathsf{A}'}, m_{\mathsf{P}}} \mu_{\text{set}}(m_{\mathsf{A}'}, m_{\mathsf{P}}) \min_{o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}} W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}}). \tag{27}$$

Since the probability distribution of settings and challenges is fixed, the expectation of W is at least $\bar{w}_{\min}$. Without loss of generality, we may assume that $\bar{w}_{\min} < 1$, as the test factor is otherwise useless for the protocol. If the entangled state ρ of the two subsystems $\mathsf{Q}_{\mathsf{A}'}\mathsf{J}_{\mathcal{A}_{\mathsf{A}}}$ and $\mathsf{J}_{\mathcal{A}_{\mathsf{B}}}$ on leaving the causal past $\mathcal{M}_{\text{in}}$ of the trial satisfies $R(\rho) = \xi' \leq \xi$, then we represent the state as an affine combination of separable states σ and τ as $(1+\xi')\sigma - \xi'\tau$. By construction, the expectation of W for each of σ and τ is between $\bar{w}_{\min}$ and 1. By linearity over the affine combination of the expectation of W, the expectation of W for state ρ is at most $(1 + \xi') - \xi'\bar{w}_{\min} = 1 + \xi'(1 - \bar{w}_{\min}) \leq 1 + \xi(1 - \bar{w}_{\min})$. Let $\bar{w}_u = 1 + \xi(1 - \bar{w}_{\min})$. This implies that $W/\bar{w}_u$ is a test factor against adversaries whose prior entanglement has robustness bounded from above by ξ. By using the test factor $W/\bar{w}_u$ for each trial, we can effectively reject such adversaries.

We can also test the hypothesis that adversaries have prior entanglement with average robustness below some threshold, averaged over all trials. For this, suppose that at trial k, conditional on the past, the adversaries have prior entanglement with robustness Ξ_k. We treat Ξ_k as a random variable determined by the past of trial k. Let W_k denote the instantiation of the test factor W on the record of the k'th trial. Conditional on the past,

Ξ_k takes on a specific, determined value ξ_k, and the expectation of W_k given this past is at most $1 + \xi_k(1 - \bar{w}_{\min}) \leq e^{\xi_k(1-\bar{w}_{\min})}$. Equivalently, the past-conditional expectation of $W_k e^{-\Xi_k(1-\bar{w}_{\min})}$ is at most 1. By the law of total expectation, the product Z of $W_k e^{-\Xi_k(1-\bar{w}_{\min})}$ over n trials also has expectation at most 1. By Markov's inequality, we have $\mathbb{P}(Z \geq 1/\delta) \leq \delta \mathrm{Exp}(Z)$, where $\delta > 0$ is the desired soundness parameter. Rearranging the inequality in this expression, we conclude that the probability that

$$\frac{1}{n(1-\bar{w}_{\min})}\left(\sum_k \log(W_k) - \log(1/\delta)\right) \geq \frac{1}{n}\sum_k \Xi_k \tag{28}$$

is at most δ. We can therefore take the left-hand side of this inequality as a lower bound on the average robustness of entanglement that the adversaries must share prior to each trial in order to pass the protocol with probability at least $1 - \delta$. We modify the protocol so that it succeeds when the lower bound on the average robustness certified at soundness parameter δ exceeds a threshold chosen ahead of time. Let r_{th} be the chosen threshold for average robustness of entanglement per trial. Then, to pass the protocol we require $\frac{1}{(1-\bar{w}_{\min})n}(\sum_k \log(W_k) - \log(1/\delta)) \geq r_{\mathrm{th}}$. We can rewrite this inequality as $\prod_k W_k e^{-r_{\mathrm{th}}(1-\bar{w}_{\min})} \geq 1/\delta$. Therefore, it suffices to replace the trial-independent test factor in Protocol 1 by $We^{-r_{\mathrm{th}}(1-\bar{w}_{\min})}$ to test against adversaries with prior entanglement.

Before the run of the protocol, the verifiers determine the values of the robustness threshold r_{th} for which the actual prover has a high probability of passing the protocol given enough trials. Suppose that the test factor W against adversaries without prior entanglement has gain $g = \mathrm{Exp}(\log(W))$ for the actual prover, and define $\bar{w}_{\min}$ as above. For n trials with the actual prover, the expectation of the left-hand side in Eq. 28 is $g/(1-\bar{w}_{\min}) - \log(1/\delta)/(n(1-\bar{w}_{\min}))$. The term multiplying $\log(1/\delta)$ goes to zero as $n \to \infty$, so given sufficiently many trials, it is possible to reject adversaries with average robustness of entanglement per trial up to $g/(1-\bar{w}_{\min})$.

Any test factor W for QPV against adversaries without prior entanglement can be used in the above expressions to also test against adversaries with prior entanglement. For the experiment, we modify W to minimize the number of trials required to ensure that the actual prover can pass the protocol with chosen threshold r_{th} and with sufficiently high probability. The modification consists of replacing W with $W' = \lambda W + (1 - \lambda)$ and minimizing the required number of trials over $\lambda \in [0, 1/(1 - w_{\min})]$, where $w_{\min}$ is the minimum value that W can take. See Sect. VI A for details.

F. Heralded sources and commitment to overcome transmission loss

Significant advantages of test factors include that they permit early stopping if thresholds on the product of trial-wise test factors are reached, and that they can be adapted based on past trials if desired. The latter property enables setting the next trial's test factor to 1, if conditions for performing the next trial are not satisfied. A test factor of 1 is equivalent to ignoring the contribution of the next trial. A common situation where it is desirable to not perform the next trial is if heralded sources are used. In this case, provided the herald is committed early enough, such as before the trial boundary or at A' before the measurement is applied, a test factor of 1 can be used for trials where the herald indicates that no entanglement has been generated. Heralded sources can be used to overcome the problem of loss in transmission. For example, the protocol of Ref. [80], known as the DLCZ protocol, can

extend entanglement links to arbitrary distances with the help of local quantum memories, and a recent experiment employing high-fidelity entanglement over 100 km using a single-photon-interference protocol [41] demonstrates that heralded entanglement can already be established over large distances between atomic memories with fairly long storage times of around 300 ms.

For bounded distances, the commitment method of Refs. [16, 81] can be used to avoid the need for quantum memory and allow for transmission loss γ arbitrarily close to 1 if the detectors have efficiency 1. The probability of a successful trial with commitment is effectively $1-\gamma$. In terms of distance l, we usually have $1-\gamma = e^{-\lambda l}$ where λ is a loss rate. Consequently, commitment can enable a significant increase in distance without the use of quantum repeaters, albeit at the cost of an exponential increase in the number of trials required for the protocol. When using commitment, the prover, before entering the target region, decides whether they are ready to commit to the upcoming trial. Commitment is determined by whether or not they detected the presence of the photon. They send a commitment bit with value 1 for "commit" to verifier A', before they enter the target region, ensuring that it will be received by verifier A' or A early enough. The verifiers must certify that commitment happened before the target region, for which it suffices if they receive the commitment in the complement of the set that causally depends on the target region. For simplicity, we assume that verifier A and A' are co-located and receive the commitment before they exit region 1001 in Fig. 9. The verifiers set the test factor to 1 in post-processing if the prover indicated not being ready. To avoid the requirement for local memory in the case that the source is near A', the verifier measurement is performed immediately after their photon arrives. This requires replacing the causal separation assumption of Assumption 8 with physical security. In particular, A' must prevent the measurement apparatus from communicating measurement results beyond A''s laboratory during the period from when the setting is applied to the end of the trial. Alternatively, A' can make their own commitment measurement before generating and applying their measurement setting. We remark that the quantum measurement apparatus used for deciding whether to commit does not need to be trusted.

Refs. [16, 81] prove that the commitment method works in some generality for QPV protocols that use simple correct/incorrect decisions for each trial. To verify security of the commitment method for Bell-test QPV, it is necessary to extend the analysis of adversary capabilities to account for the fact that adversary $\mathcal{A}_{\mathsf{A},\mathrm{start}}$, but not $\mathcal{A}_{\mathsf{A},\mathrm{end}}$, $\mathcal{A}_{\mathsf{B},\mathrm{start}}$ or $\mathcal{A}_{\mathsf{B},\mathrm{end}}$, can control the commitment bit received by A'. The reduction to three-party non-signaling in Sect. V C needs no modification, because the commitment bit can be added to $\mathsf{Q}_{\mathsf{A}'}$ as an additional classical bit, and $\mathsf{Q}_{\mathsf{A}'}$ is included in $\mathsf{R}_{\mathcal{A}_\mathsf{A}}$. The resulting three-party state including the commitment bit is independent of the challenges and the verifier setting. Conditioning on the commitment bit having value 1 results in a conditional three-party state, which again results in a three-party quantum non-signaling scenario. In particular the test factor has expectation bounded above by 1 conditional on the commitment bit. Setting the test factor to 1 if the commitment bit is 0 preserves this bound. In general, any conditioning performed by A' before applying the measurement setting and independent of this setting performs similarly.

G. Security against adversaries with significant quantum resources

In order to obtain security against adversaries who possess larger amounts of entanglement than can be handled by the results of Sect. V E, we use quantum position verification protocols with more complex settings functions. Ref. [63] considers settings functions which accept two length-l bit strings x and y as input, which are transmitted to the prover by the left and right verifiers, respectively. Ref. [63] Prop. 6.1 establishes a lower bound of $l^{1/2}/8$ on the dimension of the memory registers needed for adversaries to defeat the soundness of their protocols for concrete settings functions and sufficiently small noise levels. The memory registers include classical registers such as those needed for classical communication. They also show non-constructively the existence of settings functions requiring dimension at least $2^{l/2-5}$ under the same assumptions. See also Ref. [82]. The proofs are based on ϵ-net arguments and on results in communication complexity. Note that in this section, ϵ denotes the parameter associated with the ϵ-net technique, not the completeness parameter of the protocol.

Here, we use a similar approach to analyze the Bell-test protocol for the case where provers have two setting choices and the verifier settings and challenges are independent and uniformly distributed. We show that there are simple and concrete settings functions such that the adversary closer to verifier B needs quantum systems of dimension $d = \Omega(l^{1/2})$ for prior entanglement to defeat soundness of the protocol. See Eq. (55) for an explicit inequality relating l and d. Whereas Ref. [63] only provided a lower bound on the total size of both classical and quantum registers that cheating adversaries must use, we provide a lower bound specifically on the size of the adversaries' quantum registers, thus allowing for free classical computation and communication. Crucially, our results require no restriction on the noise level in the quantum systems used by the verifiers and honest prover.

In our security analysis, we assume that the adversaries possess an initial entangled quantum state. We do not directly consider the case in which the adversaries use quantum communication with one another, since in this setting quantum communication can be simulated with teleportation. To achieve this, it suffices to extend the initial quantum state with one Bell pair for each qubit to be communicated. We let d denote the local dimension of the entangled quantum state. Following Ref. [63], we approximate the states obtained after the initial measurements made by the adversary nearer verifier B with the states of an ϵ-net. The approximation demonstrates that the adversary nearer verifier B can reduce the size of the classical messages to the other adversary to the size of the ϵ-net, and conditional on these messages, we recover a three-party quantum non-signaling scenario. We then take advantage of the properties of test factors to deduce the existence of a solution to a communication-complexity problem. By applying bounds on such solutions, we obtain a lower bound on d.

An alternative approach to proving bounds on adversary resource requirements is given in Ref. [15]. In this approach the number of quantum gates and measurements required by successful adversaries is bounded from below. Since the sequence of gates, measurements and measurement outcomes can be efficiently encoded, it is possible to avoid the use of ϵ-net approximations. Ref. [15] effectively obtains a linear lower bounds on the number of gates and measurements in terms of l. Unlike Ref. [63], there is no restriction on classical communication. In this section, we also extend these bounds to Bell-test QPV. For this purpose, in the proof of the bounds on dimension of the entangled quantum states below, we can directly substitute gate sequences for ϵ-nets to obtain a linear bound on quantum

gates and measurements. This bound is established at the end of the section for the adversary closer to verifier B for the Bell-test protocol. The other adversary's local quantum resources are unconstrained, and classical communication is free. See Eq. (64) for an explicit bound relating l to the number of gates and qubits used for a standard universal gate set.

In this section, we use simplified notation to refer to the parties and states. We refer to verifier A''s quantum measurement apparatus as $\mathcal{V}$ and to the adversaries as $\mathcal{B}, \mathcal{C}$, where $\mathcal{B}$ is the adversary nearer $\mathcal{V}$. The adversary strategy consists of two phases, where in the first phase they receive the challenges and measure and manipulate their shared entangled state, and in the second phase they receive the other adversary's classical message and send their responses to the verifiers. The quantum system destined for $\mathcal{V}$ is V and the quantum systems of $\mathcal{B}$ and $\mathcal{C}$ are B and C. Without loss of generality, the initial state is pure on systems VBC, where VB is in initially in the possession of $\mathcal{B}$, and C is in the possession of $\mathcal{C}$. V must be sent to $\mathcal{V}$ by $\mathcal{B}$ because the measurement station must produce its output before $\mathcal{C}$ has a chance to send anything to it. Let $|\psi\rangle_{VBC}$ be the initial state. We let d be the dimension of the support of the reduced state on system C, which is also the dimension of the support of system VB. During the first phase of the trial, $\mathcal{B}$ and $\mathcal{C}$ receive their challenges x and y. Conditional on challenge x, $\mathcal{B}$ applies a quantum instrument to VB with outcome b, and $\mathcal{C}$ applies another quantum instrument to C and obtains outcome c. System V goes to the verifier to be measured by the verifier's measurement device, which is not trusted and therefore may be assumed to receive both x and b along with system V. Systems B and C are stored for the second phase of the trial during which $\mathcal{B}$ and $\mathcal{C}$ make measurements to determine the values to be sent to the respective verifiers. (See Figure 13.) We consider the transmission of subsystem V as an unrestricted quantum communication by adversary $\mathcal{B}$. By increasing the instrument outcome spaces, we can, without loss of generality, assume that the operations conditional on outcomes are pure, meaning that as operations on density matrices they have just one Kraus operator X and act as $|\phi\rangle \mapsto X|\phi\rangle$ on the Hilbert space. The probability of the corresponding instrument outcome is $\langle\phi| X^\dagger X |\phi\rangle$.

Before proceeding, we show that if $d = 2^{2^l}$, then the adversaries can perfectly simulate any pair of verifier measurement station and honest prover sharing a pair of qubits from a source. For some protocols with trusted devices, such perfect attacks are impossible, as shown in Ref. [83]. The following is a simple instance of a "garden-hose" attack [9] adapted to our scenario. In our case, the attack is enabled by adversary $\mathcal{B}$'s control of the quantum state that is measured by $\mathcal{V}$. The attack strategy requires that $\mathcal{B}$ and $\mathcal{C}$ share 2^l pairs of qubits $\{P_x\}_x$, each in the state that is intended to be shared between the verifier measurement station and the honest prover. In the first phase of the trial, when $\mathcal{B}$ receives challenge x, $\mathcal{B}$ sends their share of P_x to $\mathcal{V}$ and x to $\mathcal{C}$. When $\mathcal{C}$ receives challenge y, for each possible x', $\mathcal{C}$ applies measurement $f(x', y)$ to their share of $P_{x'}$, using the same positive operator valued measure (POVM) as would be used by the honest prover being simulated. The measurement outcome is recorded as $r(x', y)$. $\mathcal{C}$ sends to $\mathcal{B}$ the challenge y along with a string of length 2^l whose symbol in position x' has value $r(x', y)$. In the second phase, $\mathcal{C}$, having learned x, responds with $r(x, y)$ and $\mathcal{B}$ does the same, using the information received from the first phase.

Returning to the analysis of adversaries with constrained entanglement dimension, let $B_{x,b}$ and $C_{y,c}$ be the operators applied by the quantum instruments of $\mathcal{B}$ and $\mathcal{C}$ upon challenges x and y and outcomes b and c, respectively. We have $\sum_b B^\dagger_{x,b} B_{x,b} = \mathbb{1}$ and similarly for $C_{y,c}$. To simplify notation, we treat states $|\phi\rangle$ of system VBC as $d' \times d$ dimensional matrices ϕ, where $d' \geq d$, VB is represented by the row space, and C by the column space. In this

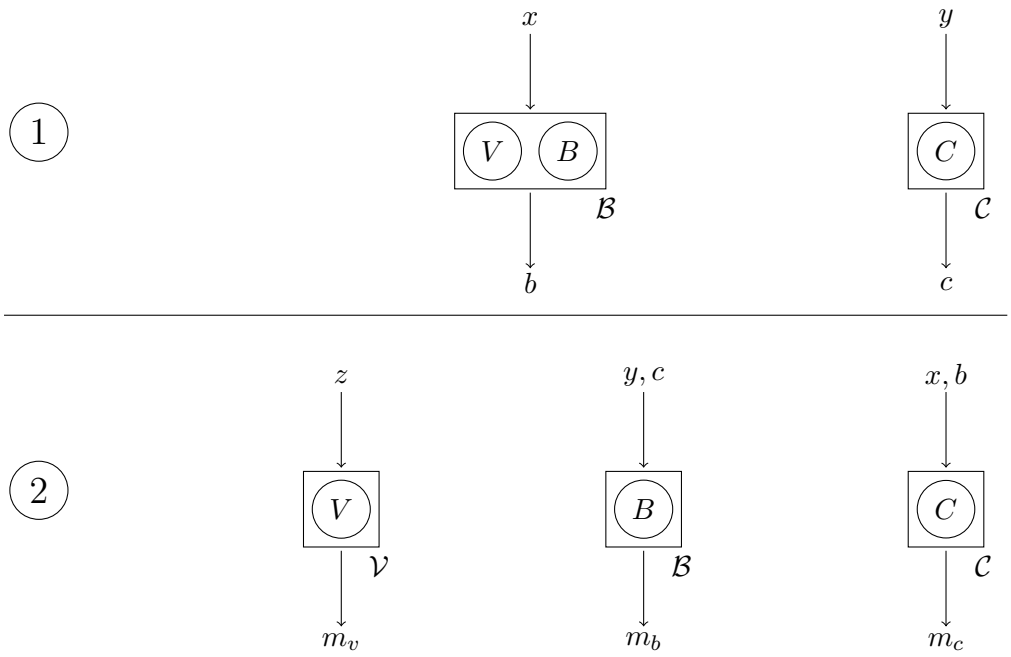

FIG. 13. A theoretical cheating model. In the first phase, the systems V, B and C are measured by the adversarial parties $\mathcal{B}$ and $\mathcal{C}$. In the second phase, the same systems are re-measured by the first verifier's measurement apparatus $\mathcal{V}$ and the parties $\mathcal{B}$ and $\mathcal{C}$.

representation, $B_{x,b}$ are matrices of size $d' \times d'$ acting on the left, and $C_{y,c}$ are matrices of size $d \times d$ whose transpose is acting on the right. That is, conditional on x, b, y, c, the normalized state obtained after the instruments act is $\psi_{x,b,y,c} = B_{x,b}\psi C_{y,c}^T/|B_{x,b}\psi C_{y,c}^T|$, where ψ is the initial entangled state, which is known to $\mathcal{B}$ and $\mathcal{C}$. Since the support of the reduced density matrix of the initial state on VB has dimension at most d, we assume that the matrix ψ is non-zero only on the first d rows. Here $|\cdot|$ denotes the Hilbert-Schmidt norm $|M| = \sqrt{\mathrm{tr}(M^\dagger M)} = \sqrt{\mathrm{tr}(MM^\dagger)}$, which is the usual Hilbert-space norm for states represented by matrices. In principle, the quantum instruments of $\mathcal{C}$ could map the d-dimensional support $\mathcal{H}$ of ψ's reduced density matrix on system C to arbitrary places in a higher dimensional space. However, for a given y, c the reduced density matrix of state $\psi C_{y,c}^T$ still has support of at most d dimensions. We can therefore ensure that the output of $C_{y,c}$ is in $\mathcal{H}$ by applying an isometry $V_{y,c}$ to the output that returns the output to $\mathcal{H}$. A standard such isometry is obtained from the polar decomposition $C_{y,c} = V_{y,c}^\dagger \tilde{C}_{y,c}$, where $\tilde{C}_{y,c}$ is positive semidefinite Hermitian on H. $\mathcal{C}$ knows which isometry was applied and can reverse it in the second phase before performing their final measurement, or simply integrate it into the final measurement. In order to reduce the real dimension of the matrices $\psi C_{y,c}^T$ to d^2, we further modify $C_{y,c}$ by post-processing with a unitary depending on y, c applied to system C such that $\psi C_{y,c}^T$ restricted to the first d rows is a Hermitian matrix, where $\psi C_{y,c}^T$ is by construction zero on the other rows. This can be done by replacing $C_{y,c}$ with $U^{\dagger T} V^T C_{y,c}$, where $\psi C_{y,c}^T$ has singular value decomposition $UDV^\dagger$. Here U may be considered as a $d \times d$ unitary matrix acting on the first d rows when multiplied on the left.

Let W be a test factor used for Bell-test QPV that has expectation at most 1 for the three-party quantum non-signaling scenario derived in Sect. V C. The goal is to show that if the entanglement dimension of the entangled state shared by the adversaries at the trial

boundary is at most d, then the expectation of the test factor W for the adversaries exceeds 1 by at most a factor $1/\kappa$, where $\kappa \in (0,1)$ and κ can be made close to 1 by increasing l. To protect against such adversaries it then suffices to replace the test factor W by $W' = \kappa W$. The basic idea of the proof is that the communication from adversary $\mathcal{C}$ to $\mathcal{B}$ can be compressed to consist of only one of L symbols, where L is determined by d. At the same time, $\mathcal{B}$ is able to guess the value of the prover setting $f(x,y)$ knowing only their challenge and the compressed symbol from $\mathcal{C}$, with a guessing advantage linearly related to $1/\kappa - 1$. If l is large enough and the function $f(x,y)$ has sufficient complexity, a communication-complexity analysis shows that $1/\kappa - 1$ must be small. The proof is obtained in a series of steps. We begin by fixing d and the excess expectation $\epsilon' = \mathrm{Exp}(W) - 1$ of W achieved by the adversaries. The steps are as follows:

1. **Construction of an ϵ-net:** We construct an ϵ-net $\mathcal{N}$ of L normalized states such that for each state $\psi C_{y,c}^T$ there is a member $\phi_{y,c}$ of the ϵ-net within distance ϵ of $\psi C_{y,c}^T / |\psi C_{y,c}^T|$. Later, we express ϵ as a positive multiple of ϵ' with coefficient determined by the values of W. The number L of states in the ϵ-net is monotonically increasing in d and $1/\epsilon$.
2. **New adversary strategy:** For each x, b, we replace the states $\psi_{x,b,y,c}$ by the states $B_{x,b}\phi_{y,c}/|B_{x,b}\phi_{y,c}|$ to obtain a new, albeit artificial strategy for the adversaries. With this replacement the new probability distribution μ' of the trial record and x, b, y, c is within total-variation distance 2ϵ of the original probability distribution μ.
3. **Limited decrease of test-factor expectation:** We show that the decrease in the expectation of the test factor for the new strategy compared to the original one is bounded by a linear multiple of ϵ. By choosing ϵ small enough, we can ensure that the test factor expectation for the new strategy still exceeds 1 by a positive multiple of ϵ.
4. **Guessing probability and communication scenario:** For each x, b and each relevant $\phi \in \mathcal{N}$, we reduce the new strategy to a three-party quantum non-signaling scenario corresponding to the events with $\phi_{y,c} = \phi$. We show that if, conditional on x, b and ϕ, the test factor expectation exceeds 1 by a given amount δ, then the settings distribution for $\mathcal{B}$ and $\mathcal{C}$ in the non-signaling scenario must be biased away from $1/2$, making it possible for $\mathcal{B}$ to guess $f(x,y)$ knowing only x, b and ϕ. The guessing probability then also applies to the original strategy's state and probability distribution, with a controlled decrease due to the small total-variation distance previously established. We find that by averaging over all x, b and $\phi \in \mathcal{N}$, adversary $\mathcal{B}$ can guess $f(x,y)$ with guessing probability at least $1/2 + \epsilon$, being assisted only by initial shared entanglement with $\mathcal{C}$ and a message from an L-letter alphabet from $\mathcal{C}$. The alphabet consists of the members of $\mathcal{N}$ encoded classically, and the message replaces the original message to $\mathcal{B}$ from $\mathcal{C}$.
5. **Communication-complexity bound:** We identifying the strategy for $\mathcal{B}$ to guess $f(x,y)$ with a simultaneous message passing protocol for $f(x,y)$ with non-trivial guessing probability. We show that the guessing probability for this protocol can be bounded in terms of L, l and the maximum entry of the binary Fourier transform of the communication matrix of $f(x,y)$. For semi-bent functions $f(x,y) = g_0(x \oplus y)$, the maximum entry is small, leading to a favorable bound.
6. **Test-factor modification:** From the bounds obtained, we determine explicit factors κ depending on l satisfying that $W' = \kappa W$ is a test factor against adversaries with $d = O(l^{1/2})$. Furthermore, we show how to modify the strategy to protect against adversaries whose average over the trials of the logarithm of the entanglement dimension

is bounded by a bound that can be made arbitrarily large by increasing l.

7. **Gate-complexity bound:** Finally, we bypass the ϵ-net construction to show that if adversary $\mathcal{B}$ uses fewer than q qubits and $G = O(l/(\log_2(q) + 1)$ gates, then the adversaries pass the protocol with probability at most the soundness parameter.

Construction of the ϵ-net: Consider $\epsilon > 0$. We make use of an ϵ-net $\mathcal{N}$ on Hermitian matrices of dimension $d \times d$. The ϵ-net serves to approximate the Hermitian matrices $\psi C_{y,c}^T$. The real dimension of the space of such matrices is d^2. The space of Hermitian matrices is a real inner product space under the trace inner product that induces the Hilbert-Schmidt norm. We therefore identify this space with $\mathbb{R}^{d^2}$. The distance between two $d \times d$ Hermitian matrices ϕ and ϕ' may be expressed explicitly as $|\phi - \phi'| = \sqrt{\mathrm{tr}((\phi - \phi')^\dagger(\phi - \phi'))}$ and can be written as a standard Euclidean distance in the basis constructed from the Gell-Mann matrices. There exists a set $\mathcal{N}$ consisting of $L \leq (1 + 2/\epsilon)^{d^2}$ Hermitian matrices ϕ with $|\phi| = 1$ such that for every y, c with $|\psi C_{y,c}^T| > 0$ there is a member ϕ of $\mathcal{N}$ within distance ϵ of the normalized state $\psi C_{y,c}^T/|\psi C_{y,c}^T|$. This follows from Ref. [84], Proposition 4.2.10. This proposition states that for every subset K of $\mathbb{R}^{d^2}$, there is an ϵ-net $\mathcal{N}$ for K of size $\mathrm{vol}(K + (\epsilon/2)B_{d^2})/\mathrm{vol}((\epsilon/2)B_{d^2})$. The ϵ-net is a set of points in K such that every member of K is within ϵ of one of these points. Here, $\mathrm{vol}(\cdot)$ denotes volume, and B_{d^2} is the unit ball in $\mathbb{R}^{d^2}$. For our application, we let K be the unit sphere in $\mathbb{R}^{d^2}$, and observe that $K + (\epsilon/2)B_{d^2}$ is contained in the ball of radius $1 + \epsilon/2$. The ratio of volumes in the proposition is then the ratio of the volume of a ball of radius $1 + \epsilon/2$ to the volume of a ball of radius $\epsilon/2$. Since the normalized states $\psi C_{y,c}^T/|\psi C_{y,c}^T|$ are in the unit sphere, $\mathcal{N}$ satisfies the claimed properties.

New adversary strategy: Next, we reduce the space consisting of the input y and the instrument outcomes c for $\mathcal{C}$ to L possible outcomes without significantly decreasing the verifiers' test factor expectation after the second phase of the trial. For each y, c such that $|\psi C_{y,c}^T| > 0$, choose $\phi_{y,c} \in \mathcal{N}$ such that $\left|\phi_{y,c} - \psi C_{y,c}^T/|\psi C_{y,c}^T|\right| \leq \epsilon$. If $|\psi C_{y,c}^T| = 0$, we set $\phi_{y,c} = 0$.

We consider $\phi_{y,c} \in \mathcal{N}$ both as a state and as a symbol from an alphabet of L letters. The cases where $\phi_{y,c} = 0$ have zero probability of occurring and therefore do not contribute. Let $\mu(b, c|x, y) = |B_{x,b}\psi C_{y,c}^T|^2$ be the probability of outcomes b, c conditional on x and y. When $\mu(b, c|x, y) > 0$, the resulting normalized state is $\psi_{x,b,y,c} = B_{x,b}\psi C_{y,c}^T/\sqrt{\mu(b, c|x, y)}$. We set $\psi_{x,b,y,c} = 0$ if $\mu(b, c|x, y) = 0$. We modify the first phase of the trial by artificially replacing $\psi_{x,b,y,c}$ with

$$\phi_{x,b,y,c} = B_{x,b}\phi_{y,c}/|B_{x,b}\phi_{y,c}|, \tag{29}$$

where we set $\phi_{x,b,y,c} = 0$ if $\mu(b, c|x, y) = 0$. It is not necessary for this replacement to be physically realizable, but it enables the reduction to a communication-complexity problem.

We claim that for each pair of x and y, the expectation over b, c of $|\phi_{x,b,y,c} - \psi_{x,b,y,c}|^2$ is bounded by $4\epsilon^2$. The claim is established by the following sequence of inequalities.

$$\begin{aligned}
&\sum_{b,c} \mu(b, c|x, y)|\phi_{x,b,y,c} - \psi_{x,b,y,c}|^2 \\
&\quad = \sum_{b,c} \left| \frac{|B_{x,b}\psi C_{y,c}^T|}{|B_{x,b}\phi_{y,c}|} B_{x,b}\phi_{y,c} - B_{x,b}\psi C_{y,c}^T \right|^2
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{b,c} \left| \frac{|B_{x,b}\psi C_{y,c}^T| - |\psi C_{y,c}^T||B_{x,b}\phi_{y,c}|}{|B_{x,b}\phi_{y,c}|} B_{x,b}\phi_{y,c} + |\psi C_{y,c}^T| B_{x,b}\phi_{y,c} - B_{x,b}\psi C_{y,c}^T \right|^2 \\
&\leq \sum_{b,c} 2 \left| \frac{|B_{x,b}\psi C_{y,c}^T| - |\psi C_{y,c}^T||B_{x,b}\phi_{y,c}|}{|B_{x,b}\phi_{y,c}|} B_{x,b}\phi_{y,c} \right|^2 + 2\left| |\psi C_{y,c}^T| B_{x,b}\phi_{y,c} - B_{x,b}\psi C_{y,c}^T \right|^2 \quad (30) \\
&= 2\sum_{b,c} \left| |B_{x,b}\psi C_{y,c}^T| - |\psi C_{y,c}^T||B_{x,b}\phi_{y,c}| \right|^2 + 2\sum_{b,c} \left| |\psi C_{y,c}^T| B_{x,b}\phi_{y,c} - B_{x,b}\psi C_{y,c}^T \right|^2 \\
&\leq 4\sum_{b,c} \left| |\psi C_{y,c}^T| B_{x,b}\phi_{y,c} - B_{x,b}\psi C_{y,c}^T \right|^2 \quad (31) \\
&= 4\sum_{b,c} \operatorname{tr}\Big(B_{x,b} \big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big) \big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big)^\dagger B_{x,b}^\dagger \Big) \\
&= 4\sum_{b,c} \operatorname{tr}\Big(B_{x,b}^\dagger B_{x,b} \big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big) \big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big)^\dagger \Big) \\
&= 4\sum_{c} \operatorname{tr}\Big(\big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big) \big(|\psi C_{y,c}^T|\phi_{y,c} - \psi C_{y,c}^T \big)^\dagger \Big) \\
&= 4\sum_{c} |\psi C_{y,c}^T|^2 \left| \phi_{y,c} - \psi C_{y,c}^T / |\psi C_{y,c}^T| \right|^2 \\
&\leq 4\epsilon^2 \sum_{c} \operatorname{tr}\big(\psi (C_{y,c}^\dagger C_{y,c})^T \psi^\dagger \big) \\
&= 4\epsilon^2. \quad (32)
\end{aligned}
$$

To obtain line (30), we used $|u+v|^2 \leq 2|u|^2 + 2|v|^2$. For line (31) we applied the triangle inequality in the form $||u| - |v|| \leq |u - v|$ to the first summand's terms.

Next we claim that the probability distribution of the trial record according to the adversary strategy changes by at most 2ϵ in total-variation distance after we make the state replacement of the previous paragraph. For the purpose of this proof, we change the notation for verifier measurement settings and measurement outcomes used in other sections. Let z be the verifier measurement setting. The probability distribution of z, x, b, y, c is not changed by the replacement. Let m_v, m_b, m_c be the measurement outcomes obtained by $\mathcal{V}, \mathcal{B}, \mathcal{C}$ in the second phase of the trial, where the same measurement POVMs are used before and after the replacements. Let μ, μ' the distributions of $z, x, b, y, c, m_v, m_b, m_c$ before and after the replacement. By construction, we have $\mu'(x, b, y, c) = \mu(x, b, y, c)$. Since total-variation distance satisfies the data-processing inequality, it suffices to bound the total-variation distance of the trial record by the total-variation distance between μ and μ'. We have

$$
\begin{aligned}
2\mathrm{TV}(\mu, \mu') &= \sum_{z,x,b,y,c,m_v,m_b,m_c} |\mu(z, x, b, y, c, m_v, m_b, m_c) - \mu'(z, x, b, y, c, m_v, m_b, m_c)| \\
&= \sum_{z,x,b,y,c} \mu(z, x, b, y, c) \\
&\qquad \times \sum_{m_v,m_b,m_c} |\mu(m_v, m_b, m_c | z, x, b, y, c) - \mu'(m_v, m_b, m_c | z, x, b, y, c)|. \quad (33)
\end{aligned}
$$

Consider a general POVM with POVM operators P_i for i in an index set I and normalized

states $|\sigma\rangle$ and $|\sigma'\rangle$. Write $|\Delta\rangle = |\sigma\rangle - |\sigma'\rangle$. We can estimate

$$\begin{aligned}\sum_i |\langle\sigma| P_i |\sigma\rangle - \langle\sigma'| P_i |\sigma'\rangle| &= \sum_i |(\langle\sigma| - \langle\sigma'|)P_i|\sigma\rangle - \langle\sigma'| P_i(|\sigma'\rangle - |\sigma\rangle)| \\ &\leq \sum_i (|\langle\Delta| P_i |\sigma\rangle| + |\langle\sigma'| P_i |\Delta\rangle|) \\ &\leq 2||\Delta\rangle|, \end{aligned} \tag{34}$$

where, for the last line we applied the following inequality twice: For every normalized state $|\rho\rangle$,

$$\begin{aligned}\sum_i |\langle\Delta| P_i |\rho\rangle| &= \sum_i |\langle\Delta| P_i^{1/2} P_i^{1/2} |\rho\rangle| \\ &\leq \sum_i \langle\Delta| P_i |\Delta\rangle^{1/2} \langle\rho| P_i |\rho\rangle^{1/2} \\ &\leq \left(\sum_i \langle\Delta| P_i |\Delta\rangle\right)^{1/2} \left(\sum_k \langle\rho| P_k |\rho\rangle\right)^{1/2} \\ &= \langle\Delta|\Delta\rangle^{1/2} \langle\rho|\rho\rangle^{1/2} = ||\Delta\rangle|. \end{aligned} \tag{35}$$

Here we applied the Cauchy-Schwarz inequality twice, first to the inner product of states in the first line, and then to the sum in the second line after recognizing the sum as an inner product. Returning to Eq. (33) and substituting the inequality of Eq. (34) we obtain

$$\begin{aligned}2\mathrm{TV}(\mu,\mu') &\leq 2 \sum_{z,x,b,y,c} \mu(z,x,b,y,c)|\phi_{x,b,y,c} - \psi_{x,b,y,c}| \\ &= 2 \sum_{x,b,y,c} \mu(x,b,y,c)|\phi_{x,b,y,c} - \psi_{x,b,y,c}| \\ &\leq 2\left(\sum_{x,b,y,c} \mu(x,b,y,c)\right)^{1/2} \left(\sum_{x,b,y,c} \mu(x,b,y,c)|\phi_{x,b,y,c} - \psi_{x,b,y,c}|^2\right)^{1/2} \\ &\leq 4\epsilon, \end{aligned} \tag{36}$$

where we used the Cauchy-Schwarz inequality to obtain the third line and Eq. (32) to obtain the last line.

Limited decrease of test-factor expectation: We can now obtain a bound on how much the expectation of the test factor W can be reduced by the above artificial state replacement. The test factor W depends only on z, $f(x,y)$ and m_v, m_b, m_c. Let $W_{\max}(z) = \max_{f,m_v,m_b,m_c} W(z,f,m_v,m_b,m_c)$ and $W_{\min}(z) = \min_{f,m_v,m_b,m_c} W(z,f,m_v,m_b,m_c)$. Let $\Delta_{\max} = \max_z(W_{\max}(z) - W_{\min}(z))$. Since $\mu(z) = \mu'(z)$, and the total-variation distance between μ and μ' is at most 2ϵ, it follows that

$$\mathrm{Exp}_{\mu'}(W) \geq \mathrm{Exp}_{\mu}(W) - 2\epsilon\Delta_{\max}. \tag{37}$$

We provide the explicit calculation to establish this inequality. In general, let $w(z,r)$ be a real-valued function of variables z and r. Let $w_{\max}(z) = \max_r w(z,r)$ and $w_{\min}(z) =$

$\min_r w(z,r)$. Let $\delta_{\max} = \max_z(w_{\max}(z) - w_{\min}(z))$. Let ν and ν' be probability distributions such that $\nu(z) = \nu'(z)$ and $\mathrm{TV}(\nu,\nu') \leq \epsilon'$. Because the distribution of z is the same for ν and ν', we have $\mathrm{TV}(\nu,\nu') = \mathrm{Exp}_\nu(\mathrm{TV}(\nu(\cdot|z),\nu'(\cdot|z)))$. Let $R_z = \{r : \nu'(r|z) \geq \nu(r|z)\}$ and $\bar{R}_z$ the complement of R_z in the domain for the variable r. Then

$$\begin{aligned}
\mathrm{Exp}_{\nu'}(w) - \mathrm{Exp}_\nu(w) &= \sum_z \nu(z) \sum_r (\nu'(r|z) - \nu(r|z)) w(z,r) \\
&= \sum_z \nu(z) \left(\sum_{r \in R_z} (\nu'(r|z) - \nu(r|z)) w(z,r) - \sum_{r \in \bar{R}_z} (\nu(r|z) - \nu'(r|z)) w(z,r) \right) \\
&\geq \sum_z \nu(z) \left(\sum_{r \in R_z} (\nu'(r|z) - \nu(r|z)) w_{\min}(z) - \sum_{r \in \bar{R}_z} (\nu(r|z) - \nu'(r|z)) w_{\max}(z) \right) \\
&= \sum_z \nu(z) \left(\mathrm{TV}(\nu(\cdot|z),\nu'(\cdot|z)) w_{\min}(z) - \sum_{r \in \bar{R}_z} \mathrm{TV}(\nu(\cdot|z),\nu'(\cdot|z)) w_{\max}(z) \right) \\
&= -\sum_z \nu(z) \mathrm{TV}(\nu(\cdot|z),\nu'(\cdot|z)) (w_{\max}(z) - w_{\min}(z)) \\
&\geq -\sum_z \nu(z) \mathrm{TV}(\nu(\cdot|z),\nu'(\cdot|z)) \delta_{\max} \\
&= -\mathrm{TV}(\nu,\nu') \delta_{\max} \\
&\geq -\epsilon' \delta_{\max}. \qquad (38)
\end{aligned}$$

Guessing probability and communication scenario: Let $\bar{\Delta} = \sum_z \mu(z)(W_{\max}(z) - W_{\min}(z))$ be the average over z of $(W_{\max}(z) - W_{\min}(z))$. Suppose that $\mathrm{Exp}_\mu(W) \geq 1 + \epsilon(3\bar{\Delta} + 2\Delta_{\max})$. Then, in view of Eq. (37), $\mathrm{Exp}_{\mu'}(W) \geq 1 + 3\epsilon\bar{\Delta}$. This expectation can be expressed as the average over $\phi \in \mathcal{N}$ and x, b of the expectation of W with respect to μ' conditional on $\phi_{y,c} = \phi$ and x, b. Consider a fixed $\phi \in \mathcal{N}$ and suppose that conditional on $\phi_{y,c} = \phi$ and x, b, the expectation with respect to μ' of W is $1 + \delta(\phi, x, b)$. Fix ϕ, x, b and consider the case $\delta(\phi, x, b) \geq 0$. Conditional on $\phi_{y,c} = \phi$ and x, b, the situation in the modified second phase of the trial is that $\mathcal{V}, \mathcal{B}, \mathcal{C}$ share a fixed quantum state $\phi_{x,b,y,c} = B_{x,b}\phi / |B_{x,b}\phi|$. This is a three-party scenario with a fixed shared quantum state, and measurements are made based on additionally provided classical inputs z, y, c with a conditional probability distribution that is the same as the original conditional distribution $\mu(z, y, c | \phi_{y,c} = \phi, x, b)$. $\mathcal{V}$'s quantum measurement device receives z and $\mathcal{B}$ and $\mathcal{C}$ both get y, c, where in this scenario, $\mathcal{V}, \mathcal{B}, \mathcal{C}$ may be assumed to know the given values ϕ, x, b before receiving the additional input z and y, c, respectively. The test factor W depends only on z, $f(x, y)$ and the measurement outcomes. We claim that with the addition of shared classical randomness with a specific distribution for $\mathcal{B}$ and $\mathcal{C}$, this scenario is quantum non-signaling with respect to the setting choice z at $\mathcal{V}$ and identical settings choices $f = f(x, y)$ for $\mathcal{B}$ and $\mathcal{C}$. Here we take $\{0, 1\}$ as the set of values for $f = f(x, y)$. The value of z received by $\mathcal{V}$ is uniform and independent of y, c, and $\mathcal{V}$ applies a z-conditional POVM producing output m_v, where the POVM does not depend on y, c. Consequently there is no signaling from $\mathcal{V}$ to $\mathcal{BC}$ in the sense that the probability distributions of m_b, m_c and y, c observed by $\mathcal{BC}$ do not depend on the value of z. Similarly, the probability distribution of z and m_v do not depend on the value of y, c.

To obtain the correct non-signaling conditions for $\mathcal{B}$ and $\mathcal{C}$ separately, we express $f(x,y)$ in terms of additional shared randomness that does not depend on the value of $f(x,y)$. The shared randomness consists of auxilliary variables y_0, c_0, y_1, c_1 that are independent of y, c and all other pre-existing variables, and shared between $\mathcal{B}$ and $\mathcal{C}$. The probability of y_0, c_0 is defined as the probability of $y = y_0, c = c_0$ conditional on $f(x,y) = 0$, which is given by $\mu(y = y_0, c = c_0 | f(x,y) = 0, \phi_{y,c} = \phi, x, b)$. The variables y_1, c_1 are independent of y_0, c_0 and have the distribution of $y = y_1, c = c_1$ conditional on $f(x,y) = 1$. The auxiliary variables can be added to the quantum state ϕ in the form of an extra, independent mixed state σ shared between $\mathcal{B}$ and $\mathcal{C}$ that is perfectly correlated between $\mathcal{B}$ and $\mathcal{C}$. $\mathcal{B}$ and $\mathcal{C}$ on input $f = f(x,y)$ can simulate the original scenario by each of them assigning $y = y_f, c = c_f$ and using these values as before. This can be interpreted as $\mathcal{B}$ and $\mathcal{C}$ each locally applying an f-conditional POVM to the extended shared quantum state consisting of ϕ and σ, producing outcomes m_b and m_c. The scenario is now defined consistently also for different inputs f and f' for $\mathcal{B}$ and $\mathcal{C}$. If the joint distribution of z and $f = f(x,y)$ is uniform conditional on $\phi_{y,c} = \phi$ and x, b, the conditional expectation of W is at most 1 as W is a test factor for such non-signaling scenarios with uniform settings z and f. Since, in this section, we are assuming that verifier measurement settings and challenges are uniform and independent, the distribution of z alone is uniform and independent of $f(x,y)$, even conditional on $\phi_{y,c} = \phi$ and x, b. That is $\mu'(z | f(x,y) = f, \phi_{y,c} = \phi, x, b) = \mu'(z) = \mu(z)$. We can therefore obtain a lower bound on the guessing probability of the value of $f = f(x,y)$ by $\mathcal{B}$ given only the symbol ϕ in addition to x, b and the conditional expectation of W. To obtain the lower bound, we instead show that if the guessing probability for f is less than $1/2 + \epsilon'$, then the expectation of W can exceed 1 by at most $\epsilon' \bar{\Delta}$. Suppose that the guessing probability is less than $1/2 + \epsilon'$, that is, for each $f \in \{0,1\}$, we have $\mu'(f | \phi_{y,c} = \phi, x, b) < 1/2 + \epsilon'$. Let ν be the distribution μ' conditioned on $\phi_{y,c} = \phi, x, b$ that governs the constructed non-signaling scenario. Since $W_{\min}(z) \leq \mathrm{Exp}_\nu(W | z, f) \leq W_{\max}(z)$ and by independence of z, we obtain

$$\begin{aligned}
\mathrm{Exp}_\nu(W|z) &= \sum_f \nu(f|z) \mathrm{Exp}_\nu(W|z,f) \\
&= \sum_f \nu(f) \mathrm{Exp}_\nu(W|z,f) \\
&= \sum_f \frac{1}{2} \mathrm{Exp}_\nu(W|z,f) + \sum_f \left(\nu(f) - \frac{1}{2} \right) \mathrm{Exp}_\nu(W|z,f) \\
&< \sum_f \frac{1}{2} \mathrm{Exp}_\nu(W|z,f) + \epsilon' (W_{\max}(z) - W_{\min}(z)). \qquad (39)
\end{aligned}$$

The average over z of the first summand on the last line is bounded by 1 since it is the average of W with respect to the uniform settings distribution. Averaging Eq. (39) over z gives

$$\mathrm{Exp}_\nu(W) < 1 + \epsilon' \bar{\Delta}. \qquad (40)$$

It follows that if $\delta(\phi, x, b) \geq 0$ and $\mathrm{Exp}_\nu(W) = 1 + \delta(\phi, x, b)$, then there is a value of f for which $\nu(f) = \mu'(f = f(x,y) | \phi_{y,c}, x, b) \geq 1/2 + \delta(\phi, x, b)/\bar{\Delta}$, that is, the guessing probability for f is at least $1/2 + \delta(\phi, x, b)/\bar{\Delta}$. Let $\bar{\delta}$ be the average over ϕ and x, b of $\delta(\phi, x, b)$. Since the guessing probability is always at least $1/2$, the average of the guessing probability over

ϕ, x, b with distribution μ' is lower bounded by

$$\begin{aligned}\mathrm{Exp}_{\mu'}(1/2 + \max(\delta(\phi, x, b)/\bar{\Delta}, 0)) &\geq \mathrm{Exp}_{\mu'}(1/2 + \delta(\phi, x, b)/\bar{\Delta}) \\ &= \mathrm{Exp}_{\mu'}(1/2 + (W-1)/\bar{\Delta}) \\ &\geq 1/2 + 3\epsilon, \end{aligned} \tag{41}$$

where we applied $\mathrm{Exp}_{\mu'}(W) \geq 1 + 3\epsilon\bar{\Delta}$. Because the total-variation distance between μ and μ' is bounded by 2ϵ, the average guessing probability for f for $\mathcal{B}$ when the variables have their original distribution μ and $\mathcal{B}$ learns only $\phi = \phi_{y,c}$ from $\mathcal{C}$ is reduced by at most 2ϵ to at least $1/2 + \epsilon$. This holds because the guessing probability is the expectation of the binary random variable describing whether the guess is correct or not, and the total-variation distance is the maximum change in expectation of such binary random variables.

Communication-complexity bound: We have shown that if the adversarial expectation of W exceeds 1 by at least $\epsilon(3\bar{\Delta} + 2\Delta_{\max})$ and the adversary $\mathcal{C}$, instead of communicating y, c to $\mathcal{B}$, sends only $\phi = \phi_{y,c}$ from the L members of the ϵ-net $\mathcal{N}$, then $\mathcal{B}$, knowing x, b, can guess the value of $f = f(x, y)$ with an average probability of at least $1/2 + \epsilon$. Thus ϵ is the guessing advantage of $\mathcal{B}$ for f in this modified adversary protocol. To make the guess, $\mathcal{B}$ takes advantage of complete knowledge of the shared entangled state and the quantum instruments used to compute the necessary conditional probabilities according to μ. This adversary protocol is equivalent to an entanglement-assisted protocol in the simultaneous message passing (SMP) model, where on uniformly random inputs x and y, $\mathcal{B}$ and $\mathcal{C}$ send messages to a third party $\mathcal{D}$, sharing with each other ahead of time only a pure quantum state on a bipartite Hilbert space of total dimension $d \times d$. $\mathcal{B}$'s message, obtained before transmitting the subsystem destined for $\mathcal{V}$, is from an alphabet of arbitrary size and consists of x, b, and $\mathcal{C}$'s message is from an alphabet of size L and consists of $\phi = \phi_{y,c} \in \mathcal{N}$. $\mathcal{D}$ uses the two messages to guess the value of $f(x, y)$, succeeding with probability at least $1/2 + \epsilon$. The dimension of the entangled state in this SMP protocol is $d \times d$ because ψ's reduced density matrix on C has dimension d. We consider functions $f(x, y)$ of the form $g(x \oplus y)$ with x and y bit strings of length l. For such functions, we show below how to determine lower bounds on L for such protocols. The conventional communication complexity for such protocols is $\log_2(L)$. The best bounds for functions satisfying our assumptions on the settings function are obtained with balanced, semi-bent functions [85–87]. Balance ensures that the settings function satisfies the assumptions stated in Sect. II when the verifier setting z and challenges x, y are uniform and independent. A simple example of an efficiently implementable, balanced, semi-bent function for $l = 2k + 1$ is the exclusive OR of one bit with the split inner product on $2k$ additional bits defined as $g_0(u) = u_0 + \sum_{j=1}^{k} u_j u_{j+k}$ mod 2.

The analysis of the lower bound on L given ϵ and d for the SMP protocol described in the previous paragraph consists of three parts. In the first part the problem is reduced to estimating a trace of the sandwiched product of the communication matrix defined below with a factorization of a guessing matrix with entries given in Eq. (47), where the factors' Hilbert-Schmidt norm can be bounded. In the second part the guessing probability is related to the maximum Fourier coefficient of the communication matrix and the Hilbert-Schmidt norms of the factors of the guessing matrix. In the third part, we instantiate the communication matrix with a semi-bent function of $x \oplus y$ taking advantage of the small value of the maximum Fourier coefficient of such functions' communication matrix. The techniques used are from Refs. [88–90] and inspired by the factorization norm analyses of Ref. [91].

We switch to column vector notation for quantum states and transition from guessing bits to guessing values in $\{-1,1\}$. The communication matrix M for the SMP protocol is defined by its entries $M_{x,y} = (-1)^{f(x,y)} \in \{-1,1\}$. The steps of the SMP protocol are as follows:

1. $\mathcal{B}$ and $\mathcal{C}$ share a pure quantum state ψ in a bipartite Hilbert space $\mathcal{H}_B \otimes \mathcal{H}_C$. To establish the general result below, the dimension of this Hilbert space is unconstrained.
2. $\mathcal{B}$ receives a uniformly random input x and $\mathcal{C}$ a uniformly random input y.
3. $\mathcal{B}$ applies POVM $(B_{x,i})_{i=1}^{L'}$ to $\mathcal{H}_B$ and sends outcome i to $\mathcal{D}$, while $\mathcal{C}$ applies POVM $(C_{y,j})_{j=1}^{L}$ to $\mathcal{H}_C$ and sends outcome j to $\mathcal{D}$. Here $\mathcal{B}$ can include classical information about x in the POVM labels so that it can be extracted from the outcomes, but $\mathcal{C}$ is limited to L outcomes, which restricts $\mathcal{C}$'s ability to convey information about y. Here we overloaded notation to denote the POVM operators, which are obtained from the instrument operators in the QPV protocol according to $B_{x,i} = B_{x,b_i}^\dagger B_{x,b_i}$ for a labeling of the outcomes b for $\mathcal{B}$, and $C_{y,j} = \sum_{c:\phi_{y,c}=\phi_j} {C_{y,c}}^\dagger C_{y,c}$ for some indexing ϕ_j of the members of $\mathcal{N}$.
4. $\mathcal{D}$ produces the guess $R_{ij} \in \{-1,1\}$ of $M_{x,y}$.

The probability of outcome i,j given x,y is

$$\nu(i,j|x,y) = \psi^\dagger B_{x,i} \otimes C_{y,j}\psi = ((B_{x,i}\otimes \mathbb{1})\psi)^\dagger((\mathbb{1}\otimes C_{y,j})\psi), \tag{42}$$

where we used $B_{x,i} = B_{x,i}^\dagger$, since POVM elements are Hermitian. Let $1/2 + s_{x,y}$ be the probability that $\mathcal{D}$ correctly guesses $M_{x,y}$ on input x,y and $s = 2^{-2l}\sum_{x,y} s_{x,y}$ its expectation over x,y. We have

$$2s_{x,y} = \sum_{i,j} \nu(i,j|x,y) M_{x,y} R_{i,j}. \tag{43}$$

Write vectors in $\mathcal{H}_B \otimes \mathcal{H}_C$ in a fixed orthogonormal basis $(e_k)_{k=1}^{d'}$, where d' is the dimension of $\mathcal{H}_B \otimes \mathcal{H}_C$. For $\varphi \in \mathcal{H}_B \otimes \mathcal{H}_C$, φ_k is the k'th coefficient of this vector in this basis. Let $\mathbf{B}$ be the matrix of dimension $2^l \times (L'd')$ whose entries are

$$\boldsymbol{B}_{x,ik} = ((B_{x,i}\otimes\mathbb{1})\psi)_k^*, \tag{44}$$

where ik denotes the pair (i,k) and indexes the columns with $i \in \{1,\ldots,L'\}$ and $k \in \{1,\ldots,d'\}$. Let $\boldsymbol{C}$ be the matrix of dimension $Ld' \times 2^l$ whose entries are

$$\boldsymbol{C}_{jk',y} = ((\mathbb{1}\otimes C_{y,j})\psi)_{k'}. \tag{45}$$

Intuitively, $\boldsymbol{B}$ is the matrix of row vectors that catalogs the action of $\mathcal{B}$'s measurements on the shared state for every input x, while $\boldsymbol{C}$ is the matrix of column vectors cataloging $\mathcal{C}$'s measurements for every input y. Their form enables the use of matrix multiplication to express the guessing advantages. Let $\boldsymbol{R}$ be the $L'd' \times Ld'$ matrix whose entries are

$$\boldsymbol{R}_{ik,jk'} = R_{ij}\delta_{k,k'}. \tag{46}$$

We have

$$\begin{aligned}\sum_{i,j}\nu(i,j|x,y)R_{i,j} &= \sum_{i,j} R_{i,j}\sum_k ((B_{x,i}\otimes\mathbb{1})\psi)_k^*((\mathbb{1}\otimes C_{y,j})\psi)_k \\ &= \sum_{ik,jk'} ((B_{x,i}\otimes\mathbb{1})\psi)_k^* R_{i,j}\delta_{k,k'}((\mathbb{1}\otimes C_{y,j})\psi)_{k'} \\ &= (\boldsymbol{BRC})_{x,y}. \end{aligned} \tag{47}$$

Therefore

$$2s = 2^{-2l}\operatorname{tr}\big(M^T(\boldsymbol{BRC})\big) = 2^{-2l}\operatorname{tr}\big(\boldsymbol{C}M^T(\boldsymbol{BR})\big), \tag{48}$$

where $M^T = M^\dagger$ since M is real. It follows that $2|s|$ is upper-bounded by 2^{-2l} times the product of the Hilbert-Schmidt norms $|\boldsymbol{C}|$ and $|\boldsymbol{BR}|$ of $\boldsymbol{C}$ and $\boldsymbol{BR}$ and the maximum singular value $\|M\|$ of M. This is a consequence of the Cauchy-Schwarz inequality for the Hilbert-Schmidt inner product, the properties of Schatten p-norms, and the fact that the Hilbert-Schmidt norm is the Schatten 2-norm. Schatten p-norms are unitarily invariant, submultiplicative, and non-increasing under multiplication by a matrix of maximum singular value less than 1. For example, see Ref. [92], Chapter IV. By the Cauchy-Schwarz inequality, $|\operatorname{tr}\big(\boldsymbol{C}M^T(\boldsymbol{BR})\big)| \le |\boldsymbol{C}||M^T(\boldsymbol{BR})|$, and by the last property, $|M^T(\boldsymbol{BR})| \le \|M\||\boldsymbol{BR}|$. We have

$$\begin{aligned}
|\boldsymbol{C}|^2 &= \operatorname{tr}\big(\boldsymbol{C}^\dagger\boldsymbol{C}\big) \\
&= \sum_{y=1}^{2^l}\sum_{j=1}^{L}\sum_{k'=1}^{d'} |\boldsymbol{C}_{jk',y}|^2 \\
&= \sum_{y=1}^{2^l}\sum_{j=1}^{L} \psi^\dagger(\mathbb{1}\otimes C_{y,j}^2)\psi \\
&\le \sum_{y=1}^{2^l}\sum_{j=1}^{L} \psi^\dagger(\mathbb{1}\otimes C_{y,j})\psi \\
&= \sum_{y=1}^{2^l} \psi^\dagger(\mathbb{1}\otimes\mathbb{1})\psi \\
&= 2^l,
\end{aligned} \tag{49}$$

where we used the fact that the POVM element $C_{y,j}$ is Hermitian and, because $0 \le C_{y,j} \le \mathbb{1}$, satisfies $C_{y,j}^2 \le C_{y,j}$. For $|\boldsymbol{BR}|$ we get

$$\begin{aligned}
|\boldsymbol{BR}|^2 &= \operatorname{tr}\big((\boldsymbol{BR})^\dagger\boldsymbol{BR}\big) \\
&= \sum_{x=1}^{2^l}\sum_{j=1}^{L}\sum_{k'=1}^{d'}\left|\sum_{i=1}^{L'}\sum_{k=1}^{d'} \boldsymbol{R}_{ik,jk'}((B_{x,i}\otimes\mathbb{1})\psi)_k^*\right|^2 \\
&= \sum_{x=1}^{2^l}\sum_{j=1}^{L}\sum_{k'=1}^{d'}\left|\sum_{i=1}^{L'} R_{i,j}((B_{x,i}\otimes\mathbb{1})\psi)_{k'}^*\right|^2 \\
&= \sum_{x=1}^{2^l}\sum_{j=1}^{L}\sum_{k'=1}^{d'}\sum_{i=1}^{L'}\sum_{i'=1}^{L'} R_{i,j}R_{i',j}\big(\psi^\dagger(B_{x,i}\otimes\mathbb{1})\big)_{k'}((B_{x,i'}\otimes\mathbb{1})\psi)_{k'} \\
&= \sum_{x=1}^{2^l}\sum_{j=1}^{L}\psi^\dagger\left(\sum_{i=1}^{L'} R_{ij}(B_{x,i}\otimes\mathbb{1})\right)^2\psi \\
&= \sum_{x=1}^{2^l}\sum_{j=1}^{L}\left|\sum_{i=1}^{L'} R_{ij}(B_{x,i}\otimes\mathbb{1})\psi\right|^2,
\end{aligned} \tag{50}$$

where we used that the matrix $\boldsymbol{A} = \boldsymbol{BR}$ has entries $\boldsymbol{A}_{x,jk'} = \sum_{i=1}^{L'} \sum_{k=1}^{d'} \boldsymbol{R}_{ik,jk'} ((B_{x,i} \otimes \mathbb{1})\psi)_k^*$ to obtain the second line. Let $V_j = \sum_{i=1}^{L'} R_{ij} B_{x,i}$. Let I be the set of indeces i such that $R_{ij} = 1$ and let $\bar{I}$ be the complement of I in $\{1, \ldots, L'\}$. Define $V_{j+} = \sum_{i \in I} R_{ij} B_{x,i} = \sum_{i \in I} B_{x,i}$ and $V_{j-} = -\sum_{i \in \bar{I}} R_{ij} B_{x,i} = \sum_{i \in \bar{I}} B_{x,i}$. Then V_{j+} and V_{j-} are positive semidefinite, $V_{j+} - V_{j-} = V_j$ and $V_{j+} + V_{j-} = \mathbb{1}$. Therefore the maximum eigenvalue of V_j is at most 1, and the minimum eigenvalue is at least -1. It follows that $\sum_{i=1}^{L'} R_{ij}(B_{x,i} \otimes \mathbb{1}) = V_j \otimes \mathbb{1}$ has maximum singular value at most 1. Continuing Eq. (50) gives

$$\begin{aligned} |\boldsymbol{BR}|^2 &= \sum_{x=1}^{2^l} \sum_{j=1}^{L} |(V_j \otimes \mathbb{1})\psi|^2 \\ &\leq \sum_{x=1}^{2^l} \sum_{j=1}^{L} 1 = 2^l L. \end{aligned} \tag{51}$$

Applying the bounds obtained to Eq. (48) gives

$$2s \leq \|M\| 2^{-2l} \sqrt{2^l} \sqrt{2^l L} = \|M\| 2^{-l} \sqrt{L}. \tag{52}$$

For functions $f(x, y)$ of the form $f(x, y) = g(x \oplus y)$, the matrix M can be diagonalized with the Fourier transform. For odd l, the function $f(x, y) = g_0(x \oplus y)$ satisfies that its communication matrix M_{g_0} has maximum absolute eigenvalue $2^{(l+1)/2}$, see Ref. [87]. Because M_{g_0} is symmetric, the maximum absolute eigenvalue directly corresponds to its maximum singular value, yielding $\|M_{g_0}\| = 2^{(l+1)/2}$. Therefore, for odd l and g_0,

$$2s \leq 2^{-(l-1)/2} \sqrt{L}, \tag{53}$$

which demonstrates an exponential suppression of the guessing advantage as a function of l, regardless of the state shared between $\mathcal{B}$ and $\mathcal{C}$ in the SMP protocol. Together with the previous analysis of the relationship between test-factor expectation and guessing advantage for adversary protocols in Bell-test QPV, this implies that given an upper bound on the dimension of prior entanglement shared by adversaries, it suffices to increase l to reduce the excess test-factor expectation achievable by adversaries. Once the excess is made small enough, the modification in the next step can be used to adjust the test factor to preserve protocol soundness.

Test-factor modification: Returning to the test-factor estimates, the ϵ-net analysis of adversary strategies showed that if the test-factor expectation according to the original adversarial distribution μ satisfies $\mathrm{Exp}_\mu(W) \geq 1 + \epsilon(3\bar{\Delta} + 2\Delta_{\max})$, then the adversary's guessing advantage s is at least ϵ. On the other hand, through the communication-complexity analysis, we obtained the upper bound of Eq. (53) on s given l and L. Combining these bounds and substituting the size of the ϵ-net for L, we find that the following inequality must be satisfied:

$$\epsilon \leq 2^{-(l-3)/2} (1 + 2/\epsilon)^{d^2/2}. \tag{54}$$

This imposes a lower bound on d given l and ϵ. For asymptotic bounds, we rewrite and simplify the bound. It is necessary to satisfy $\epsilon 2^{(l-3)/2} \leq (3/\epsilon)^{d^2/2}$. Taking logarithms gives

$$(l - 3) - 2\log_2(1/\epsilon) \leq d^2 \log_2(3/\epsilon). \tag{55}$$

In the next two paragraphs, we show that for a given test factor and actual prover, ϵ and the number of trials required for a given soundness can be chosen to be constant independent of l and d. Therefore, d must be $\Omega(l^{1/2})$ for the inequality to be satisfied. We can interpret this as requiring that the adversaries share at least $\log_2(l)/2 - O(1)$ qubits to defeat soundness of the protocol.

Suppose we wish for soundness against adversaries with support of dimension at most d for $\mathcal{C}$'s density matrix of the prior entanglement. This requires choosing l large enough and adjusting the test factor by multiplying it by a factor of $\kappa < 1$. Let W be a test factor against unentangled adversaries with positive gain $\mathrm{Exp}(\log(W)) = g$ for the actual prover. The expectation of W for this prover is therefore at least e^g. In the derivation above, we assumed that the adversarial expectation satisfies $\mathrm{Exp}_\mu(W) \geq 1 + \epsilon(3\bar{\Delta} + 2\Delta_{\max})$ to obtain Eq. (54), which determines a lower bound d_0 on d as a function of l and ϵ, where d_0 is the smallest integer d for which Eq. (54) is satisfied. By inspection, d_0 is monotonically increasing with both l and ϵ. This indicates that a higher adversarial expectation $\mathrm{Exp}_\mu(W)$ requires a strictly larger prior-entanglement dimension d. For a fixed expectation $\mathrm{Exp}_\mu(W)$, tolerating a higher dimension of prior entanglement necessitates increasing the number of challenge bits l. Let $\Delta_5 = 3\bar{\Delta} + 2\Delta_{\max}$. If we fix ϵ and multiply W by $\kappa = 1/(1 + \epsilon\Delta_5)$ to obtain $W' = \kappa W$, then W' is a test factor against adversaries with pre-shared entangled systems having dimension $d \leq d_0$. The value of ϵ is constrained by the need for the actual prover to still have positive gain for W', which is satisfied by ensuring $e^g > (1 + \epsilon\Delta_5)$. We therefore first choose ϵ to ensure sufficient gain for W', then, given a threshold d_0, choose l large enough to ensure soundness against adversaries with prior-entanglement dimension $d \leq d_0$.

The gain of the actual prover for W' is $-\log(1 + \epsilon\Delta_5) + g$. To ensure that the actual prover retains a gain of $g/6$, we choose ϵ so that $\log(1 + \epsilon\Delta_5) = 5g/6$. It then suffices to increase the number of trials by a factor of 6 for the same soundness when using W'. We now have

$$\epsilon = (e^{5g/6} - 1)/\Delta_5. \tag{56}$$

Given the chosen ϵ and for any fixed dimension d, there exists a threshold l_0 such that the inequality in Eq. (54) is violated for all $l > l_0$. Therefore, by choosing l large enough, adversaries with arbitrarily large d and no limit on classical communication can be rejected with high probability.

For example, with $d = 2$ and a typical test factor in our experiment, the gain is $g = 2.3223 \times 10^{-6}$ and $\Delta_5 = 1.6748$. This determines $\epsilon = 1.1555 \times 10^{-6}$ according to Eq. (56). The minimum odd l for which Eq. (54) is not satisfied with $d = 2$ is then $l = 127$. For $d = 3$ and $d = 4$ this requires $l = 229$ and $l = 375$, respectively. For a test factor suitable for an extremely effective combination of verifiers and prover who can achieve maximal CHSH violation with a Bell state, we found $g = 0.032075$ and $\Delta_5 = 2.7615$. This requires $l = 49$ for $d = 2$ and $l = 141$ for $d = 4$.

We can also use a version of the strategy used in Sect. V E to protect against adversaries with bounded average prior entanglement over the trials, with entanglement quantified as $\log(d)$. For a proof of principle, we use the following conservative version of Eq. (54):

$$\epsilon \leq 2^{-(l-3)/2}(1 + 2/\epsilon)^{d^2/2} d, \tag{57}$$

for integers $d \geq 1$. Extending the definition of the right-hand side to all real $d \geq 0$, let d'_ϵ be the real solution for d satisfying equality in Eq. (57), where we keep l fixed but can vary

$\epsilon > 0$. For $\epsilon > 0$, let $d_\epsilon = \max(2, d'_\epsilon)$, and for $\epsilon \leq 0$, we set $d_\epsilon = 1$. Then d_ϵ is a lower bound on the dimension required by the adversaries to achieve the assumed expectation of W. The maximum $\max(2, d'_\epsilon)$ takes into account that by the results for adversaries without prior entanglement, the dimension d is at least 2 to achieve an expectation $\mathrm{Exp}(W) \geq 1+\epsilon\Delta_5$ with $\epsilon > 0$ in the above derivation. Suppose that at trial i, conditional on the past, $\mathrm{Exp}(W_k) = 1 + \epsilon'_k\Delta_5$, where W_k is the instantiation of W at the k'th trial. Let $\epsilon_k = \max(0, \epsilon'_k)$. If d_k is the entanglement dimension of the entanglement shared by the adversaries before the k'th trial, then $d_k \geq d_{\epsilon_k}$. Because the expectation of W cannot exceed $\bar{W}_{\max} = \mathrm{Exp}(W_{\max}(z))$, the relevant range for ϵ_k is between 0 and $\epsilon_{\max} = (\bar{W}_{\max} - 1)/\Delta_5$.

To apply the method of Sect. V E, we need a lower bound of the form $\log(d_\epsilon) \geq \lambda\epsilon$ for some $\lambda > 0$. We have $\log(d_0) = 0$. If we can determine $\gamma > 0$ such that $d_\epsilon \geq 1 + \gamma\epsilon$, then by concavity of the logarithm

$$\log(d_\epsilon) \geq \log(1 + \gamma\epsilon) \geq \big(\log(1 + \gamma\epsilon_{\max})/\epsilon_{\max}\big)\epsilon \tag{58}$$

for $\epsilon \in [0, \epsilon_{\max}]$. We can then set $\lambda = \log(1 + \gamma\epsilon_{\max})/\epsilon_{\max}$ to satisfy $\log(d_\epsilon) \geq \lambda\epsilon$. To determine such a γ, we first show $\tilde{d} = d'_\epsilon/\epsilon$ is decreasing as a function of ϵ. Equivalently, on the interval $[0, \epsilon']$, the graph of $\epsilon \mapsto d'_\epsilon$ sits above the straight line from $(0, 0)$ to $(\epsilon', d'_{\epsilon'})$, yielding the lower bound

$$d'_\epsilon \geq (d'_{\epsilon_{\max}}/\epsilon_{\max})\epsilon \tag{59}$$

for the relevant range $[0, \epsilon_{\max}]$ of ϵ. To show the monotonicity of $\tilde{d}$, from the bound in Eq. (57), we have

$$1 = 2^{-(l-3)/2}(1 + 2/\epsilon)^{\epsilon^2\tilde{d}^2/2}\tilde{d}. \tag{60}$$

Consider the function $g(\epsilon) = (1+2/\epsilon)^{\epsilon^2}$, which converges to 1 as $\epsilon \to 0$. To determine its behavior for all $\epsilon > 0$, we analyze its logarithmic derivative by defining $h(\epsilon) = \epsilon^{-1}\frac{d}{d\epsilon}\log(g(\epsilon)) = 2\log(1 + 2/\epsilon) - 2/(\epsilon+2)$. Setting $y = 2/\epsilon \in (0, \infty)$, we have $h(\epsilon) = 2\log(1 + y) - y/(1+y)$. We have $\log(1 + y) = \int_0^y dz/(1 + z) \geq \int_0^y dz/(1 + y) = y/(1 + y)$. Therefore $h(\epsilon) > 0$, which guarantees that $g(\epsilon)$ is monotonically increasing for all positive ϵ. Because the right-hand side of Eq. (60) is fixed to be 1, the growth of $g(\epsilon)$ requires that $\tilde{d}$ strictly decrease as ϵ increases, as promised.

To determine the desired linear lower bound on d_ϵ, let ϵ_2 satisfy $d'_{\epsilon_2} = 2$. If $\epsilon_2 > \epsilon_{\max}$, let $\epsilon_2 = \epsilon_{\max}$. The definition of d_ϵ implies that $d_\epsilon \geq 1 + \gamma\epsilon$ for $\epsilon \in [0, \epsilon_{\max}]$, where γ is the minimum of $(d'_{\epsilon_{\max}} - 1)/\epsilon_{\max}$ and $(d'_{\epsilon_2} - 1)/\epsilon_2 = 1/\epsilon_2$.

We now return to the resulting inequality $\log(d_\epsilon) \geq \lambda\epsilon$ with $\lambda = \log(1 + \gamma\epsilon_{\max})/\epsilon_{\max}$ for $\epsilon \leq \epsilon_{\max}$. We claim that λ is increasing in l. For this it suffices to show that γ is increasing in l. To satisfy Eq. (60) at a given ϵ, $\tilde{d}$ and therefore d'_ϵ must increase as l increases. If we fix $d'_{\epsilon_2} = 2$ and increase l, then ϵ_2 must decrease. Thus, both of the terms whose minimum determines γ increase with l, and therefore so does γ.

The expectation of the test-factor product over n trials is bounded above by

$$\begin{aligned}\prod_{k=1}^{n}(1+\epsilon_k\Delta_5) &\le \prod_{k=1}^{n} e^{\epsilon_k\Delta_5}\\ &= \prod_{k=1}^{n} e^{(\Delta_5/\lambda)\lambda\epsilon_k}\\ &\le \prod_{k=1}^{n} e^{(\Delta_5/\lambda)\log(d_k)} \end{aligned} \tag{61}$$

Let $\bar{q} = \sum_k \log(d_k)/n$ be the average of $\log(d_k)$ over the trials. As in Sect. V E, to reject adversaries with $\bar{q} \le r_{\rm th}$, it suffices to use the modified test factor $W' = e^{-(\Delta_5/\lambda)r_{\rm th}}W$, where W be a test factor against unentangled adversaries. For a given target average $r_{\rm th}$, we can increase l to increase λ until the gain of the actual prover for W' is a constant fraction of the original gain g for W. For maintaining a gain of at least $g/6$, it suffices to choose l large enough to satisfy $r_{\rm th}(\Delta_5/\lambda) \le 5g/6$. Sect. V E.

Gate-complexity bound: For the last result of this section, we obtain a bound on the number of gates and measurements that adversary $\mathcal{C}$ needs to perform to increase the expectation of the test factor W used for excluding unentangled adversaries by a specified amount. We assume that the $\mathcal{C}$'s quantum system C consists of $q = \log_2(d)$ qubits, which may carry arbitrary prior entanglement with $\mathcal{B}$'s quantum system B. For simplicity of expressions, we assume that q is a power of 2. Upon receiving input y, $\mathcal{C}$ executes one- and two-qubit unitary and measurement operations from a finite set of such operations where the measurements have finitely many outcomes. The outcomes determine the value c in the protocol as described in Fig. 13.

We can assign to each operation and its outcome (if any) a binary code of length bounded by $E\log(q)$. The constant E depends on the gate set used. If the gate set consists of the controlled-NOT, Hadamard, T-gate, and one-qubit logical-basis measurements, the binary code for each operation and its outcome can be shorter than $2\log_2(q)+2$. One such encoding is to use the first two bits to indicate which of the four gates in the set is applied. If it is a two qubit gate, we use $2\log_2(q)$ bits to encode which two of the q qubits are acted on. If it is a one-qubit gate, we use $\log_2(q)$ bits to indicate which one. If it is a measurement, we use 1 more bit to record the outcome.

Let $\log_2(L_g)$ be the maximum length of the encoded sequence of operations with outcomes for adversary $\mathcal{C}$ in the adversary protocol. Assuming that the gate set used is the specific one as illustrated above with a fixed-length encoding where each operation uses exactly $2\log_2(q)+2$ bits, the maximum number of gates that $\mathcal{C}$ uses in the protocol is then $\log_2(L_g)/(2\log_2(q)+2)$. The state $\psi C^T_{y,c}$ is determined by the encoded sequence, so the number of such states that can occur is at most L_g. In the ϵ-net argument above, we can let $\phi_{y,c} = \psi C^T_{y,c}/|\psi C^T_{y,c}|$, which is now one of L_g many possible states. Since $\phi_{y,c}$ does not involve an approximation, the chain of distances used in the ϵ-net argument can be bypassed. In effect, we now have $\mu = \mu'$, and then, following the statement right after Eq (40), we can conclude that if $\mathrm{Exp}_\mu(W) \ge 1+\epsilon\bar{\Delta}$, the above described SMP protocol has guessing probability at least $1/2+\epsilon$. Combining this with Eq. (53), we obtain the bound

$$\epsilon \le 2^{-(l-3)/2}\sqrt{L_g}. \tag{62}$$

Given a choice of $\epsilon > 0$, for soundness against adversaries where $\mathcal{C}$ uses less than $G = \log_2(L_g)/(2\log_2(q)+2)$ gates on at most q qubits, it suffices to choose l large enough to ensure that

$$\epsilon \geq 2^{-(l-3)/2} 2^{G(\log_2(q)+1)} \tag{63}$$

and adjust the test factor by a scaling factor of $\kappa = 1/(1+\epsilon\bar{\Delta})$. Equivalently, we choose l such that

$$l \geq 2G(\log_2(q)+1) + 2\log_2(1/\epsilon) + 3, \tag{64}$$

which is linear in the number G of quantum gates used by adversary $\mathcal{C}$ and the logarithm of the number of qubits q.

To preserve positive gain for the actual prover, we choose ϵ small enough. We repeat the construction used against adversaries with bounded prior entanglement and no constraints on local quantum computation. Given the test factor W against adversaries with no prior entanglement and its positive gain g for the actual prover, the expectation of W for this prover is at least e^g. As indicated above, we set $W' = \kappa W$ with $\kappa = 1/(1+\epsilon\bar{\Delta})$ for a test factor against adversaries where $\mathcal{C}$ uses at most G gates on q qubits. To ensure that the actual prover retains a gain of $g/6$ in this setting, we again choose ϵ such that $\log\big(1+\epsilon\bar{\Delta}\big) = 5g/6$. Accordingly, we set $\epsilon = (e^{5g/6}-1)/\bar{\Delta}$. For the typical test factor and gain in our experiment given above, we have $\epsilon = 6.0704 \times 10^{-6}$. According to Eq. (64), we can protect against adversaries with up to $G = 120$ gates and measurements on $q \leq 8$ qubits with $l = 999$.

VI. ANALYSIS OF EXPERIMENTAL RESULTS

The data analysis consists of three main steps: parameter determination, calibration prior to each protocol instance, and execution of each instance. For the first step, parameter determination, we used data collected on September 19, 2024 and use this data to estimate the distribution of the trial records for the implemented prover. We refer to a prover for whom the trial records are i.i.d. according to the estimated distribution as the reference prover. This prover is honest. The second and third steps—calibration and analysis—were performed using data collected on September 20 and October 7, 2024. On each day, the data was saved across a varying number of files, with each file containing one minute of data. The trial rate in our experiment is approximately $250,000$ trials per second.

A. Parameter Determination

The first step of data analysis is to determine key parameters necessary for evaluating the protocol. These parameters include:

1. The *soundness error* δ,
2. The *target success probability* ϵ for the reference prover,
3. The *number of trials* n (the amount of data) to be used per protocol instance, and
4. The *success threshold* r_{th} on the certified average robustness of entanglement pre-shared by adversaries for implementing a successful attack.

We fixed the soundness error to $\delta = 2^{-64}$, which is considered sufficiently small for practical security purposes. In addition, we set the target success probability to $\epsilon = 0.97725$, corresponding to a two-sigma confidence level.

The procedure for determining the remaining parameters is as follows. In the case where the adversaries under consideration do not share prior entanglement, we examine the trade-off between the number of trials n and the soundness error δ (that is, the p-value threshold). We choose n such that for the reference prover, the observed p-value falls below δ with probability at least $\epsilon = 0.97725$. To consider adversaries with bounded prior entanglement, we analyze how n affects the achievable success threshold r_{th}, which serves as a lower bound on the entanglement per trial required for the adversaries to succeed. In this case, n and r_{th} are selected jointly to ensure that for the reference prover, the resulting p-value falls below δ with probability at least $\epsilon = 0.97725$. To estimate these probabilities, we proceed as follows: First, we use data collected on September 19, 2024 to estimate the conditional distribution of measurement outcomes given the settings. We refer to this data as the calibration data for parameter determination. Second, we determine a high-gain test factor for the estimated distribution. And finally, we analyze the statistical behavior of the logarithm of the test factor for the reference prover.

We begin by estimating the conditional distribution $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ from calibration data. The calibration data can be summarized by counts $n(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}; m_{\mathsf{A}'}, m_{\mathsf{P}})$. These counts are the number of trials for which the settings $(m_{\mathsf{A}'}, m_{\mathsf{P}})$ and outcomes $(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}})$ were recorded. A representative example of the counts for each such combination of values for the trials obtained in one 5-minute run is shown in Tab. V, and this example is used as the calibration data for parameter determination. Ideally, we always have $z_{\mathsf{A}} = z_{\mathsf{B}}$, but in practice, due to technical imperfections, z_{A} and z_{B} occasionally differ. Such differences can also arise from adversarial manipulations, where an adversary may sacrifice the probability of returning identical responses to improve performance conditional on a match. In the experiment, the observed frequency of such differences was normally less than $d = 2 \times 10^{-6}$. Because differences occur rarely, we regularize the four conditional probabilities $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ given settings $(m_{\mathsf{A}'}, m_{\mathsf{P}})$ with $z_{\mathsf{A}} \neq z_{\mathsf{B}}$ by setting all of them to the fixed value $d/4 = 0.5 \times 10^{-6}$ for the purpose of constructing test factors and choosing protocol parameters. When constructing test factors, we constrain the test factor to take a constant value for all settings and outcomes with $z_{\mathsf{A}} \neq z_{\mathsf{B}}$. Consequently, all such events contribute identically to the logarithm of the test factor, and so the choice of a uniform distribution over these events in the regularization does not affect the gain calculation.

TABLE V. Calibration counts $n(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}; m_{\mathsf{A}'}, m_{\mathsf{P}})$ for settings $(m_{\mathsf{A}'}, m_{\mathsf{P}})$ and outcomes $(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}})$. All settings and outcomes are binary with two possible values denoted by $1, 2$. The first four columns correspond to cases where $z_{\mathsf{A}} = z_{\mathsf{B}}$, while the last column aggregates all cases where $z_{\mathsf{A}} \neq z_{\mathsf{B}}$. These counts are based on 5 minutes of data collected on September 19, 2024. Random resampling suggests that this amount of data is sufficient for a reliable estimation of the underlying conditional distribution.

		$(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}})$				
		$(1,1,1)$	$(2,1,1)$	$(1,2,2)$	$(2,2,2)$	Others
$(m_{\mathsf{A}'}, m_{\mathsf{P}})$	$(1,1)$	18,764,031	2,339	2,390	4,794	16
	$(2,1)$	18,751,159	9,481	1,647	5,474	29
	$(1,2)$	18,752,299	1,527	6,879	5,730	32
	$(2,2)$	18,745,211	14,655	12,333	364	35

Instead of directly determining the conditional distribution $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ by fit-

ting the observed counts, we first restrict the counts to the events where $z_{\mathsf{A}} = z_{\mathsf{B}}$. We define the restricted counts as $\tilde{n}(o_{\mathsf{A}'}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}}) = n(o_{\mathsf{A}'}, o_{\mathsf{P}}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}})$ for each combination of settings $(m_{\mathsf{A}'}, m_{\mathsf{P}})$ and outcomes $(o_{\mathsf{A}'}, o_{\mathsf{P}})$. The counts $\tilde{n}(o_{\mathsf{A}'}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}})$ determine empirical conditional frequencies

$$\tilde{f}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) = \tilde{n}(o_{\mathsf{A}'}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}}) / \sum_{o_{\mathsf{A}'}, o_{\mathsf{P}}} \tilde{n}(o_{\mathsf{A}'}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}}). \tag{1}$$

If the probability that $z_{\mathsf{A}} \neq z_{\mathsf{B}}$ is zero, the counts and frequencies are sampled according to some conditional distribution $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$, which must be in the set $\mathcal{Q}_{O_{\mathsf{A}'}, O_{\mathsf{P}}|M_{\mathsf{A}'}, M_{\mathsf{P}}}$ of two-party non-signaling distributions that satisfy Tsirelson's bounds. Here we do not consider other quantum constraints. The set $\mathcal{Q}_{O_{\mathsf{A}'}, O_{\mathsf{P}}|M_{\mathsf{A}'}, M_{\mathsf{P}}}$, as characterized in Sect. IX. A of Ref. [93], forms a convex polytope. We fit $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) \in \mathcal{Q}_{O_{\mathsf{A}'}, o_{\mathsf{P}}|M_{\mathsf{A}'}, M_{\mathsf{P}}}$ to the frequencies $\tilde{f}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ by maximum likelihood. That is, we obtain the estimate $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ as the unique optimal solution of the following constrained maximum-likelihood problem:

$$\underset{\mu(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})}{\arg\max} \quad \sum_{o_{\mathsf{A}'}, o_{\mathsf{P}}, m_{\mathsf{A}'}, m_{\mathsf{P}}} \tilde{n}(o_{\mathsf{A}'}, o_{\mathsf{P}}; m_{\mathsf{A}'}, m_{\mathsf{P}}) \log \mu(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) \tag{2}$$

$$\text{subject to} \quad \mu(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) \in \mathcal{Q}_{O_{\mathsf{A}'}, O_{\mathsf{P}}|M_{\mathsf{A}'}, M_{\mathsf{P}}} \tag{3}$$

The estimate obtained is shown in Tab. VI. We then determine the conditional probabilities $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ as follows:

$$\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) = \delta_{z_{\mathsf{A}}, z_{\mathsf{B}}}(1-d)\tilde{\sigma}(o_{\mathsf{A}'}, O_{\mathsf{P}} = z_{\mathsf{A}}|m_{\mathsf{A}'}, m_{\mathsf{P}}) + (1 - \delta_{z_{\mathsf{A}}, z_{\mathsf{B}}})\, d/4. \tag{4}$$

TABLE VI. The settings-conditional distribution of outcomes $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ obtained by maximum likelihood using the calibration data in Tab. V. This probability distribution is used solely for determining key parameters in the position verification protocol and not intended to make a statement about the actual probability distribution during the calibration or analysis phase of each instance of the protocol. All displayed values are rounded to 10 digits after the decimal point.

		$(o_{\mathsf{A}'}, o_{\mathsf{P}})$			
		$(1,1)$	$(2,1)$	$(1,2)$	$(2,2)$
$(m_{\mathsf{A}'}, m_{\mathsf{P}})$	$(1,1)$	0.9994906521	0.0001264110	0.0001261772	0.0002567597
	$(2,1)$	0.9991117479	0.0005053152	0.0000885493	0.0002943876
	$(1,2)$	0.9992478451	0.0000802128	0.0003689843	0.0003029579
	$(2,2)$	0.9985476132	0.0007804446	0.0006526840	0.0000192581

Next, following the discussion in Sect. V D, we construct a test factor for the two-party Bell test using the strategy developed in Ref. [47]. This test factor is optimal for rejecting local realism [47] with respect to the settings-conditional distribution of outcomes $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ shown in Tab. VI, given a uniform settings distribution $\mu_{\text{set}}(m_{\mathsf{A}'}, m_{\mathsf{P}})$ defined as $\mu_{\text{set}}(m_{\mathsf{A}'}, m_{\mathsf{P}}) = 1/4$ for all $m_{\mathsf{A}'}, m_{\mathsf{P}} \in \{1, 2\}$. Building on this, we construct a family of test factors of the form given in Eq. (26), parameterized by λ. By maximizing λ, we obtain a well-performing test factor, as presented in Tab. VII, for the task of position verification against adversaries without prior entanglement. From now on, we refer to this

TABLE VII. Trial-wise test factor $W(m_{\mathsf{A}'}, o_{\mathsf{A}'}, m_{\mathsf{P}}, z_{\mathsf{A}}, z_{\mathsf{B}})$ for basic position verification, constructed using the calibration data in Tab. V. The four columns shown correspond to cases where $z_{\mathsf{A}} = z_{\mathsf{B}}$. For all other cases with $z_{\mathsf{A}} \neq z_{\mathsf{B}}$, which are not displayed in the table, the test factor takes a constant value of 0.9118409194, independent of the settings and outcomes. All displayed values are rounded to 10 digits after the decimal point.

		$(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}})$			
		$(1,1,1)$	$(2,1,1)$	$(1,2,2)$	$(2,2,2)$
$(m_{\mathsf{A}'}, m_{\mathsf{P}})$	$(1,1)$	1.0000133425	0.8853069445	0.8825655759	1.0836976498
	$(2,1)$	1.0000098326	0.9751727669	0.8016191273	1.0926205335
	$(1,2)$	1.0000079803	0.7988759064	0.9699591897	1.0846655877
	$(2,2)$	0.9999688446	1.0248059103	1.0300176352	0.7390162290

task as basic position verification and to the corresponding protocol as the basic protocol. In addition to its role in parameter determination below, this test factor could also be used to analyze future data for basic position verification.

Finally, we analyze the statistical behavior of the logarithm of the obtained test factor. According to our estimates of $\sigma(o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ in Eq. (4), $\tilde{\sigma}(o_{\mathsf{A}'}, o_{\mathsf{P}}|m_{\mathsf{A}'}, m_{\mathsf{P}})$ in Tab. VI, and the uniform settings distribution used in our implementation, the base-2 logarithm of the trial-wise test factor shown in Tab. VII, $\log_2(W)$, has a mean of $g_w = 3.79135 \times 10^{-6}$ and a variance of $v_w = 1.13029 \times 10^{-5}$. For the completeness analysis, we assume the reference prover so that the trials are i.i.d. and distributed according to the above estimates. We consider the case where the test factor W_k used in each trial k is identical to the one presented in Tab. VII. In view of the central limit theorem, we approximate the sum $\sum_{k=1}^{n} \log_2(W_k)$ over n i.i.d. trials as normally distributed with mean ng_w and variance nv_w. For a given soundness error δ, the condition $\sum_{k=1}^{n} \log_2(W_k) \geq \log_2(1/\delta)$ must be satisfied to ensure a p-value below δ. Accordingly, the success probability of obtaining a p-value less than δ after n trials is heuristically estimated to be $p_{\text{succ}} = Q\left(-\left(ng_w - \log_2(1/\delta)\right)/\sqrt{nv_w}\right)$, where the function Q is the tail distribution function of the standard normal distribution. To ensure $p_{\text{succ}} \geq \epsilon$, the number of trials n must be above a certain threshold. By varying the values of δ and ϵ, we obtain the trade-off curves in Fig. 14. The trade-off curves show the value of δ that can be achieved at three different success probabilities as a function of the running time of the basic protocol. For the chosen values $\delta = 2^{-64}$ and $\epsilon = 0.97725$, the figure shows that 2 minutes of data is sufficient for basic position verification.

Similarly, we analyze the behavior of the test factor adapted for position verification in the presence of adversaries with bounded prior entanglement. Following Sect. V E, we consider trial-wise test factors of the form $W' = \lambda W + (1 - \lambda)$, where $\lambda \in [0, 1/(1 - w_{\min})]$ and $w_{\min}$ is the minimum value of the test factor W for basic position verification. Let $\bar{w}'_{\min}$, defined analogously to $\bar{w}_{\min}$ in (27), denote the average of the minimum values of the test factor W' conditioned on the measurement settings and challenges. Let W'_k represent the instantiation of the test factor W' used in the k'th trial. Given n trials and a soundness error δ, the certified lower bound on the average robustness of prior entanglement is given by

$$\frac{1}{n(1 - \bar{w}'_{\min})} \left(\sum_k \log(W'_k) - \log\left(1/\delta\right) \right). \tag{5}$$

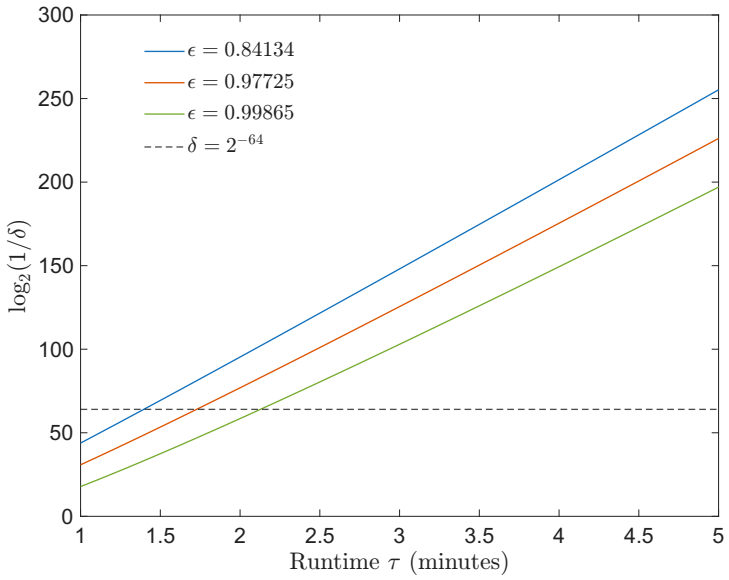

FIG. 14. Required runtime of the basic protocol (x-axis) to achieve a target soundness error δ (y-axis) with high probability. For a desired soundness error δ to be achieved with success probability at least ϵ, the experiment must run for at least the duration τ indicated on the x-axis. The three curves correspond to confidence levels of one-sigma ($\epsilon = 0.84134$), two-sigma ($\epsilon = 0.97725$), and three-sigma ($\epsilon = 0.99865$), respectively. In our experiment, we generate approximately $250,000$ trials per second. For parameter determination, we assume that the trial generation rate is exact.

Thus, to ensure that the certified lower bound exceeds the threshold r_{th}, we must have

$$\sum_k \log(W'_k) \geq n r_{\text{th}}(1 - \bar{w}'_{\text{min}}) + \log(1/\delta). \tag{6}$$

To characterize the statistical behavior of this condition, we compute the mean and variance of the random variable $\log(W')$. As before, we resort to the central limit theorem to estimate the success probability for the reference prover. For each desired success probability ϵ and each fixed value of the threshold r_{th}, we optimize over the parameter λ that characterizes the test factor W'. This optimization allows us to determine the minimum value of n required to satisfy the condition in Eq. (6) with success probability ϵ. This process also yields the corresponding optimized test factor W', which could be used to analyze future data for position verification against adversaries with prior entanglement. Varying the values of ϵ and r_{th} yields the requirement curve illustrated in Fig. 15. Based on this figure, we set $r_{\text{th}} = 8 \times 10^{-6}$. For the chosen values of $\delta = 2^{-64}$ and $\epsilon = 0.97725$, the figure shows that 4 minutes of data is sufficient for position verification against adversaries with average robustness of prior entanglement at most r_{th}.

B. Calibration and Analysis

With the protocol parameters obtained above, we executed multiple instances of the protocol using data collected on September 20, 2024, for the case of adversaries without prior entanglement, and data collected on October 7, 2024, for the case involving adversaries with limited prior entanglement. On each day, trials were performed and recorded in successive

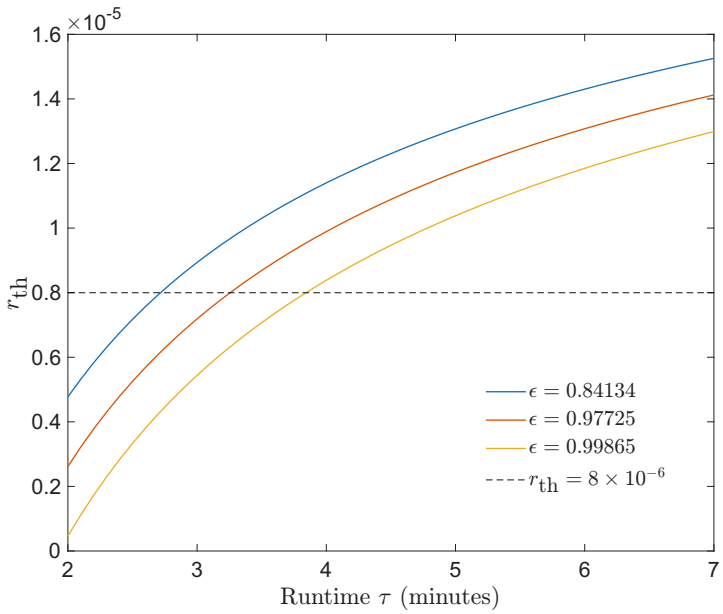

FIG. 15. Required runtime (x-axis) to ensure, with high probability, that the certified average robustness of entanglement pre-shared by adversaries reaches the target success threshold r_{th} (y-axis). The certification is performed with soundness error $\delta = 2^{-64}$. For a desired success threshold r_{th} to be reached with success probability at least ϵ, the experiment must run for at least the duration τ indicated on the x-axis. The three curves correspond to confidence levels of one-sigma ($\epsilon = 0.84134$), two-sigma ($\epsilon = 0.97725$), and three-sigma ($\epsilon = 0.99865$), respectively.

1-minute intervals, resulting in separate 1-minute data files with about 15 million trials each. The data was blinded until after the protocol parameters and analysis strategies were fixed. The protocol implementation and analysis are summarized in Protocol 2. Because the performance of the Bell test underlying the protocol varied in time, we optimized the test factor to be used for each instance of the protocol based on the 10 1-minute files immediately preceding the instance that are free of detection errors as described in the next paragraph. We refer to this data as the calibration data for the instance. For each instance, we used the calibration data to estimate the settings-conditional distribution of outcomes and to construct the corresponding test factor, following the approach described in Sect. VI A.

In our experiment, some 1-minute recording intervals experienced detector errors, an event that happened at most a few times per day. Such errors were recorded with the corresponding data files. At the beginning of each day, the first 10 1-minute files free of detector errors were used for calibration of the first protocol instance of the day. The next 2 or 4 files, regardless of whether or not detector errors happened, made up the first protocol instance, depending on whether the adversaries had no prior entanglement or bounded prior entanglement, respectively. Subsequent instances were processed sequentially using the next 2 or 4 files if available; otherwise the instance used all the remaining data files. For each such instance, calibration data consisted of the previous 10 files without detector errors.

Nominally, each 2-minute interval consists of 30 million trials, and each 4-minute interval consists of 60 million trials. However, due to clock synchronization and other uncertainties in our experiment, the actual number of recorded trials deviates from these nominal values. To ensure the validity of the analysis, we fixed the number of trials n processed to exactly

30 million for each 2-minute interval and 60 million for each 4-minute interval. If the actual number of trials exceeded n, the data were truncated and the excess trials were discarded. If it fell short, artificial trials were appended to reach n trials. In either case—whether testing for adversaries with prior entanglement or without—the appended artificial trials were analyzed using the trivial test factor that assigns a value of 1 to all possible measurement settings and outcomes, which is a valid choice. Consequently, the specific settings and outcomes assigned to each artificial trial are irrelevant. Since the logarithm of 1 is 0, these artificial trials did not contribute to the accumulated sum of the logarithms of the trial-wise test factors and thus did not influence the final analysis result. Therefore, in practice, it is not necessary to physically append artificial trials if the actual number of trials is less than n. This approach ensures the correctness of the analysis regardless of

fluctuations in the actual number of trials available for processing.

Protocol 2: Protocol Implementation and Analysis.

Input : Sequence of 1-minute data files collected in a day.

Given :

$\delta = 2^{-64}$ — soundness error (i.e., p-value threshold).

$\epsilon = 0.97725$ — target success probability for the reference prover.

n — number of analysis trials: 3×10^7 (nominally, 2-minute data) or 6×10^7 (nominally, 4-minute data), depending on the adversary model.

$r_{\text{th}} = 8 \times 10^{-6}$ — threshold on the certified average robustness of entanglement for adversaries with bounded prior entanglement.

Output: P — binary success flag: $P = 1$ (success) or $P = 0$ (failure).

Identify the 10th file in the sequence that is free of detector errors, and set the current file to the one immediately following it;

foreach *protocol instance* **do**

  Set $P \leftarrow 0$;

  Initialize the running number of calibration files: $K_{\text{calib}} \leftarrow 0$;

  Initialize the running calibration counts for all settings $z = (m_{\mathsf{A}'}, m_{\mathsf{P}})$ and outcomes $c = (o_{\mathsf{A}'}, z_{\mathsf{A}}, z_{\mathsf{B}})$: $\{\text{Data}_{\text{calib}}(cz)\}_{cz} \leftarrow 0$;

  while *a previous* 1*-minute file exists* **do**

    if *file has no detector error* **then**

      Load file; increment $K_{\text{calib}} \leftarrow K_{\text{calib}} + 1$;

      Increment $\{\text{Data}_{\text{calib}}(cz)\}_{cz}$ with counts from the file;

    end

    if $K_{\text{calib}} = 10$ **then**

      break; `// Load up to 10 files for calibration`

    end

  end

  Use $\{\text{Data}_{\text{calib}}(cz)\}_{cz}$ to estimate the settings-conditional distribution of outcomes for the purpose of test-factor construction (see Sect. VI A);

  Construct test factor W (for adversaries without prior entanglement) or W' (for adversaries with bounded prior entanglement) as in Sect. VI A;

  Load the trials from the next 2 or 4 1-minute files, if available. Otherwise, load from all remaining files. Truncate to n trials if necessary;

  Compute the accumulated sum of the logarithms of test factors S_{run}. For adversaries with prior entanglement, return the certified lower bound r_{lb} on the average robustness of entanglement according to Eq. (5) with the replacement of $\sum_k \log(W'_k)$ with S_{run};

  Calculate the success threshold $S_{\text{th}} = \log(1/\delta)$ for adversaries without prior entanglement and $S_{\text{th}} = \log(1/\delta) + nr_{\text{th}}(1 - \bar{w}'_{\min})$ for adversaries with bounded prior entanglement;

  if $S_{run} \geq S_{th}$ **then**

    Set $P \leftarrow 1$; `// Protocol succeeded`

  end

  Update the current file to the file immediately following the ones used for this protocol instance, if this is not the last instance;

end

The analysis procedure follows the implementation outlined in Protocol 2. Before unblinding the protocol instance data and conducting the final analysis, we validated that the protocol including the parameter determination functioned correctly using data collected on September 19, 2024. We then proceeded with the final analysis. Using 474 1-minute data files collected on September 20, 2024, we performed basic position verification, resulting in 232 instances, 8 of which failed. For position verification against adversaries with bounded prior entanglement, we used the 420 1-minute data files collected on October 7, 2024 to perform 103 instances. All but the final instance succeeded; the last one failed due to insufficient remaining data. Histograms of the analysis results are presented in Fig. 16. After completing the analysis of the protocol instances on the unblinded data, for diagnostic purposes, we reversed the data assignment for the two types of analyses. In the reversed analysis, using the data from September 20, 2024, we obtained 116 instances of position verification against adversaries with bounded prior entanglement, with 10 instances failing. Using the data from October 7, 2024, we obtained 205 instances of basic position verification, all of which succeeded. The results of the reversed analysis are shown in Fig. 17.

In the analysis described above, we certified a lower bound r_{lb} on the average robustness of prior entanglement to quantify the resources required for a successful adversarial attack. Alternatively, we can formulate this as a hypothesis-testing task: we aim to reject the null hypothesis that the average robustness of entanglement pre-shared by adversaries is bounded above by a specific threshold r_{th}. For this purpose, we introduce a modified test factor $W'' = W' e^{-r_{\mathrm{th}}(1-\bar{w}'_{\mathrm{min}})}$, where W' is the test factor used for certifying the average robustness, and $\bar{w}'_{\mathrm{min}}$ is the corresponding averaged minimum value (see Sect. VI A for details on determining W' and $\bar{w}'_{\mathrm{min}}$). Let W''_k denote the instantiation of the modified test factor for the k'th trial. Following the discussion below Eq. (28), we compare the accumulated product over all trials, $\prod_k W''_k$, against $1/\delta$ to determine whether the protocol instance successfully rejects the null hypothesis. When the number of available trials meets or exceeds the required number of trials n for the instance, we analyze exactly n trials. In this case, the rejection condition $\prod_k W''_k \geq 1/\delta$ is equivalent to satisfying Eq. (6). If the number of available trials is less than n, we virtually append artificial trials for which we set $W' = 1$. For each such artificial trial, $\mathrm{Exp}(W'') = \mathrm{Exp}(W' e^{-r_{\mathrm{th}}(1-\bar{w}'_{\mathrm{min}})}) \leq 1$ for adversaries with bounded average robustness, so the analysis according to the discussion below Eq. (28) in Sect. V E remains valid. Among the 103 instances using the data from October 7, 2024, there were 4 instances where the available number of analysis trial was less than n. The logarithm of the product $\prod_k W''_k$ for each instance using the data from October 7, 2024 is shown in the bottom panel of Fig. 4 of the main text.

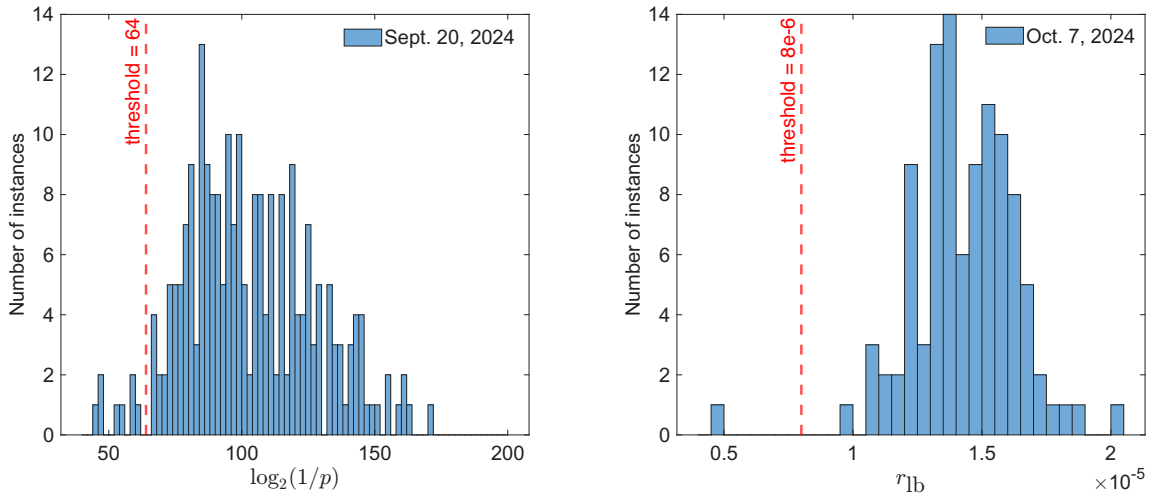

FIG. 16. Histograms of the final analysis results. The left panel shows the results for basic position verification, while the right is for position verification against adversaries with bounded prior entanglement. The histograms show the number of instances binned according to the $\log(p)$-value and the certified lower bound r_{lb} on the average robustness of prior entanglement, respectively. These values are determined by the accumulated sum of the logarithms of the trial-wise test factors used in each instance. The value of r_{lb} is given by Eq. (5) for each instance.

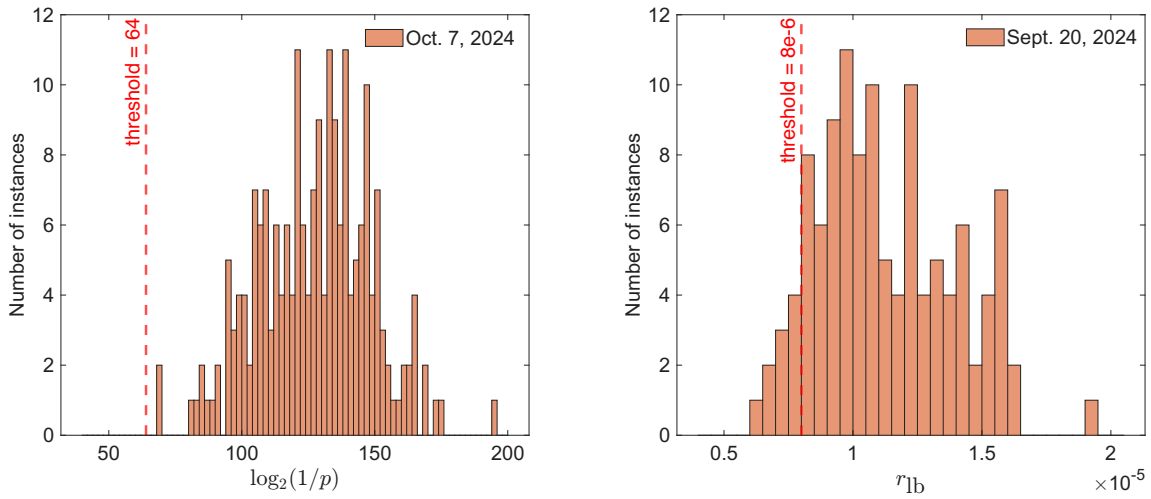

FIG. 17. Histograms of the reversed analysis results. The left panel shows the results for basic position verification, while the right is for position verification against adversaries with bounded prior entanglement. The plotted quantities are defined in the caption of Fig. 16.