

Cosmic Bell Test using Random Measurement Settings from High-Redshift Quasars

Dominik Rauch,^{1, 2, *} Johannes Handsteiner,^{1, 2} Armin Hochrainer,^{1, 2} Jason Gallicchio,³ Andrew S. Friedman,⁴ Calvin Leung,^{1, 2, 3, 5} Bo Liu,⁶ Lukas Bulla,^{1, 2} Sebastian Ecker,^{1, 2} Fabian Steinlechner,^{1, 2} Rupert Ursin,^{1, 2} Beili Hu,³ David Leon,⁴ Chris Benn,⁷ Adriano Ghedina,⁸ Massimo Cecconi,⁸ Alan H. Guth,⁵ David I. Kaiser,^{5, †} Thomas Scheidl,^{1, 2} and Anton Zeilinger^{1, 2, ‡}

¹ Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

² Vienna Center for Quantum Science & Technology (VCQ), Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

³ Department of Physics, Harvey Mudd College, Claremont, California 91711, USA

⁴ Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, California 92093, USA

⁵ Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

⁶ School of Computer, NUDT, 410073 Changsha, China

⁷ Isaac Newton Group, Apartado 321, 38700 Santa Cruz de La Palma, Spain

⁸ Fundación Galileo Galilei—INAF, 38712 Breña Baja, Spain

In this paper we present a cosmic Bell experiment with polarization-entangled photons, in which measurement settings were determined based on real-time measurements of the wavelength of photons from high-redshift quasars, whose light was emitted billions of years ago; the experiment simultaneously ensures locality. Assuming fair sampling for all detected photons, and that the wavelength of the quasar photons had not been selectively altered or previewed between emission and detection, we observe statistically significant violation of Bell's inequality by 9.3 standard deviations, corresponding to an estimated p value of $\lesssim 7.4 \times 10^{-21}$. This experiment pushes back to at least ~ 7.8 Gyr ago the most recent time by which any local-realist influences could have exploited the “freedom-of-choice” loophole to engineer the observed Bell violation, excluding any such mechanism from 96% of the space-time volume of the past light cone of our experiment, extending from the big bang to today.

Background. To Erwin Schrödinger, entanglement was “the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought” [1]. He referred to an analysis by Einstein, Podolsky, and Rosen (EPR) [2], regarding the quantum-mechanical predictions for perfect correlations in certain quantum systems. EPR made two assumptions explicit. Regarding locality, they wrote: “Since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system.” They also articulated a “reality criterion”: “If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, there exists an element of physical reality corresponding to this physical quantity.” In the light of these two assumptions and their analysis of a particular two-particle state, EPR concluded that quantum mechanics is incomplete. While the EPR reasoning is logically unassailable, Niels Bohr pointed out that the EPR assumptions need not hold for quantum observations [3].

This discussion had laid dormant for several decades, but in 1964 John Stewart Bell demonstrated that a complete theory based on the EPR premises makes predictions that are in conflict with those of quantum mechanics [4, 5]. In such local-realist theories, it is assumed that every individual system carries its own set of properties prior to measurement, which are presumed to be inde-

pendent of any possible influence from outside its past light cone. Bell concluded that in a local-realist theory the strength of correlations among measurements on different particles’ properties is limited and smaller than the predictions of quantum physics. This is expressed by Bell’s inequality.

With Bell’s result, a question that previously had been dismissed as “merely philosophical” became experimentally testable. In 1969, Clauser, Horne, Shimony, and Holt (CHSH) published their inequality as an experimentally accessible variant of Bell’s version [6]. The idea was to measure the four probabilities $p(A, B|a_i, b_j)$ of measurement results $A, B \in \{+1, -1\}$, in which Alice chooses between two measurement bases a_1 and a_2 , and likewise Bob chooses between the two measurement bases b_1 and b_2 . For systems in particular states subject to judicious choices of measurement bases, the predictions for correlations among the measurement outcomes A, B under various combinations of settings a_i, b_j differ markedly between quantum mechanics and models that satisfy EPR’s assumptions of locality and realism.

Subsequently, entangled-particle states have been shown to violate Bell’s inequality in numerous situations, consistent with the predictions of quantum theory [7–9]. Yet experiments always require sets of assumptions for their interpretation [10, 11]. In tests of local realism, these assumptions can be seen as loopholes, by which, in principle, it could be argued that local realism has not been completely ruled out [8, 12]. Closing the locality

loophole [13, 14], for example, requires that the measurement settings are changed by Alice shortly before the arrival of an entangled particle at her detector, such that no signal could inform Bob about Alice’s measurement setting or outcome before Bob completes a measurement at his own detector (and vice versa). The fair sampling assumption, on the other hand, states that the measured subset of particles is representative of the complete set. This loophole is closed if a sufficiently high fraction of the entangled pairs is detected [15–17]. Recently, several experiments have observed significant violations of Bell’s inequality while simultaneously closing both the locality and fair-sampling loopholes [18–21].

Arguably the most interesting assumption is that the choice of measurement settings is “free and random,” and independent of any physical process that could affect the measurement outcomes [5, 22–25]. As Bell himself noted, his inequality was derived under the assumption “that the settings of instruments are in some sense free variables—say at the whim of experimenters—or in any case not determined in the overlap of the backward light cones” [22]. In recent years, this “freedom-of-choice” loophole has garnered significant theoretical interest [26–35], as well as growing experimental attention [36–41].

The freedom-of-choice loophole, as usually understood, concerns events that might have transpired within the causal past of a given experiment, which a local-realist mechanism could have exploited in order to mimic the predictions from quantum mechanics [42]. In a recent pilot test [38], measurement settings for a test of Bell’s inequality were determined by real-time observation of light from Milky Way stars, thereby constraining any such local-realist mechanism to have acted no more recently than ~ 600 years ago, rather than microseconds before a given experimental run (as in previous tests [36]). The magnitude of that leap reflected how comparatively little attention had been devoted previously to experimentally addressing this loophole. Given the expansion history of the universe since the big bang, however, the pilot test [38] excluded only about one hundred-thousandth of one percent of the relevant space-time volume within the past light cone of the experiment.

In this paper, we describe a Cosmic Bell experiment that pushes the origin of the measurement settings considerably deeper into cosmic history, constraining any local-realist mechanism to have acted no more recently than 7.78 Gyr ago. Based on the arrangement of high-redshift quasars used in our experiment, these results exclude any local-realist mechanism that might have exploited the freedom-of-choice loophole from 96.0% of the space-time volume of the past light cone of the experiment, extending from the big bang to today.

Experimental implementation. Figure 1 shows a schematic of the experimental setup at the *Observatorio del Roque de los Muchachos* on the Canary Island of La Palma. A central entangled photon source was located

in a container next to the *Nordic Optical Telescope*. One entangled-photon observer, Alice, was situated in another container next to the *Telescopio Nazionale Galileo* (TNG), and Bob was stationed at the ground floor of the *William Herschel Telescope* (WHT). The quasar photons were collected by the TNG [43] and the WHT [44]. The random numbers extracted from these signals were transmitted to the observers using BNC cables. The polarization-entangled photons were distributed from the source to the receivers via free-space optical links. A more detailed schematic of the setup can be seen in Figure 2.

The entangled photon source (see Supplementary Materials [45]) was based on type-0 spontaneous parametric down-conversion (SPDC) in a Sagnac loop configuration [46, 47]. It generated fiber-coupled photon pairs at center wavelengths $\lambda_A = 850$ nm and $\lambda_B = 773.6$ nm in a state close to the maximally entangled Bell state $|\Phi^\pm\rangle = (1/\sqrt{2})(|H_A H_B\rangle \pm |V_A V_B\rangle)$, where subscripts A and B label the respective single-mode fiber for Alice and Bob. Each photon was guided to a transmitting telescope (Tx), and distributed via free-space optical channels to the receiving stations of Alice and Bob. Each station consisted of a receiving telescope for entangled photons (Rx), a polarization analyzer (POL), a control and data acquisition unit (CaDa) locked to a rubidium frequency standard, and a cosmic random number generator (CRNG). The entangled photons were guided from the Rx to the polarization analyzer, where an electro-optical modulator (EOM) performed the fast switching between the complementary measurement bases accordingly. The polarization was measured using a polarizing beam splitter followed by a single-photon avalanche diode (SPAD) in each output port.

The CRNGs at TNG (Alice) and WHT (Bob) were essentially identical. The optical path for each CRNG featured a magnified intermediate image, which enabled one to adjust the field of view with an iris in order to minimize background light. A dichroic mirror with a cutoff wavelength of 630 nm split the incoming light in a transmitted “red” and a reflected “blue” arm. Additional filters (shortpass at 620 nm in the blue arm and longpass at 637 nm in the red arm) efficiently filtered out misdirected photons whose wavelengths were near the cutoff wavelength of the dichroic mirror. Incorporating these additional filters yielded much smaller fractions of misdirected astronomical photons than in our previous experiment [38], with $f_w < 2 \times 10^{-5}$ (see the Supplemental Material [45]). Light from each arm was fed to a SPAD. Electric signals from the SPADs were processed by the CaDa, which triggered the EOM to apply the corresponding measurement settings. Alice measured linear polarization along $22.5^\circ/112.5^\circ$ (red) and $67.5^\circ/157.5^\circ$ (blue), while Bob measured linear polarization along $0^\circ/90^\circ$ (red) and $45^\circ/135^\circ$ (blue). All SPAD events, from the CRNGs and the polarization analyzers,

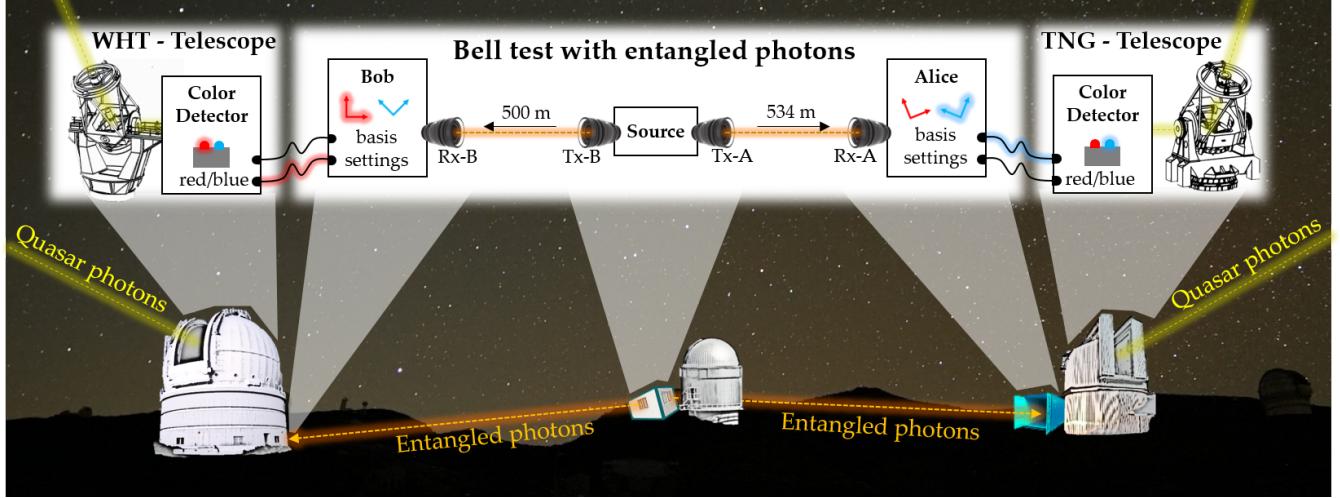


FIG. 1. The experimental stations for our Cosmic Bell test. Alice’s station received cosmic photons with the *Telescopio Nazionale Galileo* (TNG), whose primary mirror diameter is 3.58 m, while Bob’s station received cosmic photons with the *William Herschel Telescope* (WHT), whose primary mirror diameter is 4.20 m. Polarization entangled photons were sent from the source to Alice and Bob. Diameters and focal lengths of the quantum channel telescopes were Tx: $d = 70$ mm, $f = 280$ mm; Rx: $d = 140$ mm, $f_{\text{eff}} = 1470$ mm. Latitude, longitude, and elevation for the experimental sites were Alice (\mathcal{A} : 28.75410° , -17.88915° , 2375 m), Bob (\mathcal{B} : 28.760636° , -17.8816861° , 2352 m), and the Source (\mathcal{S} : 28.757189° , -17.884961° , 2385 m). The distances from Source to Bob and Alice were 500 m and 534 m, respectively.

were timestamped in the CaDa and recorded by a computer.

Using the wavelength of cosmic photons to implement the measurement settings requires the assumption that the wavelength of each photon was set at emission and has not been selectively altered or previewed between emission and detection. (Well-known processes, such as cosmological redshift and gravitational lensing, treat all photons from a given astronomical source in a uniform way, independent of the photons’ wavelength at emission [48–50].)

Within an optically linear medium, there does not exist any known physical process that can absorb and re-radiate a given photon at a different wavelength along the same line of sight, without violating the local conservation of energy and momentum. We further assume that the detected cosmic photons represent a fair sample, despite significant losses in the intergalactic and interstellar media, the Earth’s atmosphere, and the experimental setup.

Various scenarios could (in principle) lead to corrupt choices of measurement settings within our experiment. For example, local sources of photons (“noise”) rather than genuine astronomical photons could trigger the CRNGs. The most significant sources of local noise include sky glow, light pollution, and detector dark counts. The overall background was measured by pointing each telescope to a dark sky patch 10 arcseconds away from its target quasar before and after each observing period.

Space-time arrangement. Ensuring locality requires that any information leaving Alice’s quasar at the speed

of light along with her setting-determining cosmic photon could not have reached Bob before his measurement of the entangled photon is completed, and vice versa.

The projected space-time diagram in Figure 3 illustrates the situation for the first observed quasar pair (pair 1) of our experiment. The entangled photons are generated at point \mathcal{S} and travel through 12 m of optical fiber, resulting in a delay of $\tau_{\text{fiber}} \approx 50$ ns. The distance over the free-space channels is $x_{\mathcal{A}} = 534$ m to \mathcal{A} and $x_{\mathcal{B}} = 500$ m to \mathcal{B} .

To ensure the locality conditions, measurements of entangled photons must only be accepted within a certain valid time window, τ_{valid}^k , which has to be chosen such that the selection and implementation of the corresponding settings on one side remain space-like separated from the measurements on the other side. τ_{valid}^k is constrained to within a certain time window τ_{geom}^k , which depends on the time-dependent directions of the quasars relative to \mathcal{A} and \mathcal{B} . Given the moderate time dependence of τ_{geom}^k over the relatively brief observing periods (≤ 17 min), we use the shortest value per side within the observing period: $\bar{\tau}_{\text{geom}}^k = \min_t(\tau_{\text{geom}}^k)$, where $\bar{\tau}_{\text{geom}}^{\mathcal{A}} = 2.81$ μ s (2.67 μ s) and $\bar{\tau}_{\text{geom}}^{\mathcal{B}} = 1.48$ μ s (1.11 μ s) for pair 1 (2). Various delays from signal transmission through fibers and BNC cables, and to implement a given setting with the EOM, have to be subtracted from τ_{geom}^k to compute the correct validity time τ_{valid}^k . The delay until a certain setting was implemented, τ_{set}^k , was measured to be 325 ns and 430 ns for Alice and Bob, respectively. An additional buffer was used on both sides with $\tau_{\text{buffer}} = 150$ ns to account for small inaccuracies in tim-

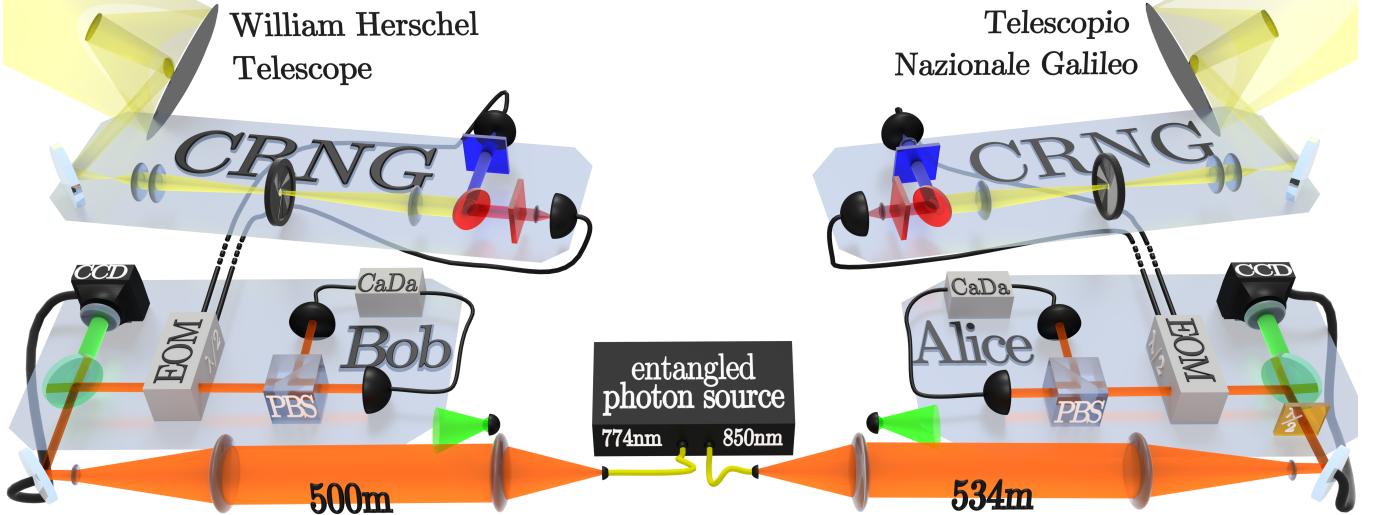


FIG. 2. A photon pair source located in the middle produced polarization-entangled photons at center wavelengths of 773.6 nm and 850 nm. The photons were separated into two spatial modes via a dichroic mirror and sent via free-space channels to the quantum receivers at Bob (773.6 nm) and Alice (850 nm). Fast steering mirrors guided the photons to the receivers using a green LED as a reference. Electro-optical modulators (EOM) rotate the measurement basis according to the input signals from the CRNGs. Polarization measurements are performed using a polarizing beam splitter (PBS) with avalanche photodiodes in each output path. Detection events are timestamped by the control and data acquisition unit (CaDa) and stored locally. Quasar light is collected by the astronomical telescopes and fed into an optical system that creates a magnified image with an iris to restrict the field of view. The quasar light is then split according to its wavelength into a “blue” and a “red” channel, whereby each channel contains additional filters to remove misdirected photons. The detector signals are used to trigger the implementation of the corresponding measurement basis at the EOM.

TABLE I. For Alice and Bob’s side ($k = \{\mathcal{A}, \mathcal{B}\}$), we list the QSO Simbad identifiers, azimuth (az_k) (clockwise from due North) and altitude (alt_k) above horizon at the start of the observing periods, and redshift (z) and lookback time to emission (t_{lb}) for quasars observed in pairs 1 and 2, beginning at UTC 2018-01-11 00:20:00 (pair 1) and 2018-01-11 01:21:00 (pair 2). Pair 1 was observed for a total of 17 min, pair 2 for 12 min. τ_{valid}^k is the time the detector setting was valid, taking into account delays and safety margins (see Figure 3). The last three columns show the measured CHSH parameter, as well as the p value and the number of standard deviations ν by which our local-realist model can be rejected (see the Supplemental Material [45]).

Pair	Side	ID	az_k°	alt_k°	z	t_{lb} [Gyr]	τ_{valid}^k [μs]	S_{exp}	p -value	ν
1	\mathcal{A}	QSO B0350-073	233	38	0.964	7.78	2.34	2.65	7.4×10^{-21}	9.3
	\mathcal{B}	QSO J0831+5245	35	57	3.911	12.21	0.90			
2	\mathcal{A}	QSO B0422+004	246	38	0.268	3.22	2.20	2.63	7.0×10^{-13}	7.1
	\mathcal{B}	QSO J0831+5245	21	64	3.911	12.21	0.53			

ing and distance measurements (for details please refer to the Supplemental Material [45]). The final validity time we used is then $\tau_{\text{valid}}^k = \bar{\tau}_{\text{geom}}^k - \tau_{\text{set}}^k - \tau_{\text{buffer}}$.

For pairs 1 and 2, measurement settings at Bob’s station were determined based on observations of a quasar with redshift $z_B = 3.911$ [51], corresponding to a look-back time to the emission of that light $t_{\text{lb}}^B = 12.21$ Gyr ago. Measurement settings at Alice’s station were determined based on observations of quasars with $z_A = 0.964$ [52] (pair 1) and $z_A = 0.268$ [53] (pair 2), corresponding to $t_{\text{lb}}^A = 7.78$ Gyr and 3.22 Gyr ago, respectively. (See Table I.) These times may be compared with the age of our observable universe since the big bang, $t_{\text{lb}} = 13.80$ Gyr [54]. We consider possible implications of inhomogeneities along the lines of sight to these objects,

such as gravitational lensing effects, in the Supplemental Material [45].

Figure 4 depicts the past light cone of our experiment (gray) together with the past light cones of quasar emission events \mathcal{Q}_A (blue) and \mathcal{Q}_B (red) for the quasars of pair 1. The past light cones from \mathcal{Q}_A and \mathcal{Q}_B for this pair last intersected $t_{\text{lb}}^{AB} = 13.15$ Gyr ago, less than 650 million years after the big bang. (For pair 2, the past light cones most recently intersected $t_{\text{lb}}^{AB} = 12.47$ Gyr ago.) This is the most recent time by which a correlation between the two quasars could have occurred or been orchestrated. The space-time 4-volume contained within the union of the past light cones from \mathcal{Q}_A and \mathcal{Q}_B constitutes just 4.0% (pair 1) and 36.5% (pair 2) of the 4-volume within the past light cone of our experi-

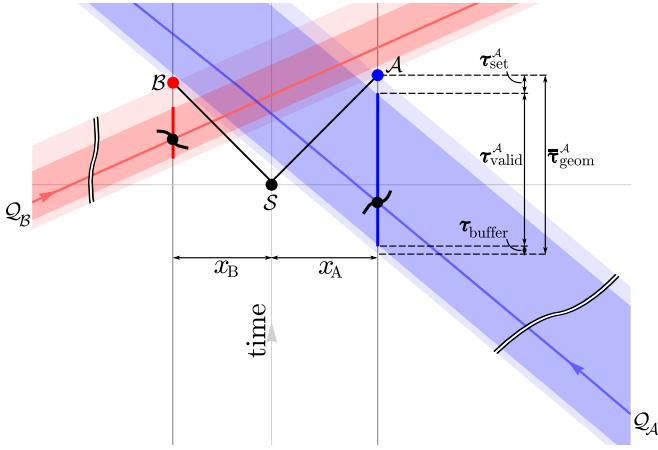


FIG. 3. (1+1)D space-time diagram for pair 1, with the origin at the source \mathcal{S} of entangled pair creation (black dot) and a spatial projection axis chosen to minimize its distance to Alice and Bob. After a short fiber delay (too small to see), entangled photons are sent via free-space channels (black lines) to be measured by Alice and Bob at events A and B . Galaxy symbols indicate examples of measurements of valid settings from quasar photons emitted far away at space-time events \mathcal{Q}_A and \mathcal{Q}_B . Ensuring locality limits settings to the shaded regions. Delays to implement each setting and an added safety buffer shorten the validity time windows that were actually used to the darker shaded regions.

ment. (See Supplemental Material [45].) Events associated with any local-realist mechanism that could have affected detector settings and measurement outcomes of our experiment would need to lie within the past light cones of \mathcal{Q}_A and/or \mathcal{Q}_B , and hence are restricted to have acted no more recently than $t_{lb}^A = 7.78$ Gyr or 3.22 Gyr ago for pairs 1 and 2, respectively.

Analysis and results. We performed two Cosmic Bell tests with the quasars listed in Table I, for a total measurement time of 17 min (pair 1) and 12 min (pair 2). In the analysis of our acquired data, we follow the assumption of fair sampling and fair coincidences [12]. Thus, our data can be postselected for coincidence events at Alice's and Bob's stations. We correct for the clock drift as in Ref. [55] and identify coincidences within a time window of 2.66 ns. We then check for correlations between measurement outcomes $A, B \in \{+1, -1\}$ for particular settings choices $(a_i, b_j), i, j \in \{1, 2\}$ using the Clauser-Horne-Shimony-Holt (CHSH) inequality [6]:

$$S \equiv |E_{11} + E_{12} + E_{21} - E_{22}| \leq 2, \quad (1)$$

where $E_{ij} = 2p(A = B|a_i b_j) - 1$ and $p(A = B|a_i b_j)$ is the probability of Alice and Bob obtaining the same measurement outcome for the joint settings (a_i, b_j) . While four probabilities can arithmetically add up to 4, local-realistic correlations cannot exceed an S value of 2 and the quantum-mechanical limit is $2\sqrt{2}$ [56].

As can be seen from Table I, the measured S_{exp} values are 2.65 and 2.63 for pairs 1 and 2, which clearly ex-

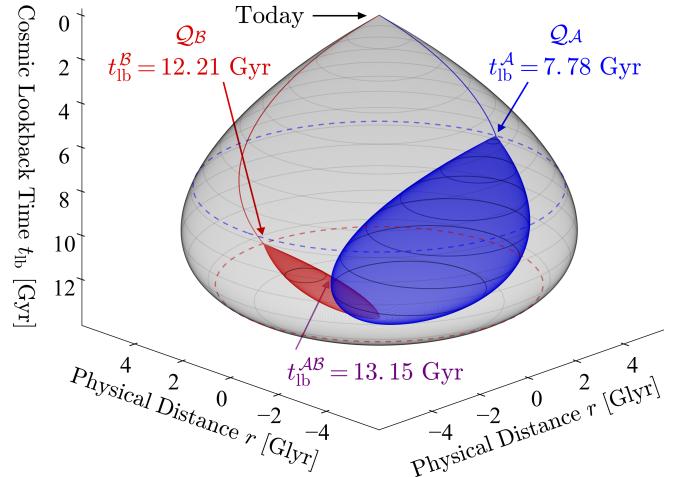


FIG. 4. (2+1)D space-time diagram for pair 1, showing the past light cone of our experiment (gray) and of the quasar emission events \mathcal{Q}_A (blue) and \mathcal{Q}_B (red), extending back to the big bang, 13.80 Gyr ago. The quasars in pair 1 emitted the light that we observed during our experiment $t_{lb}^A = 7.78$ Gyr and $t_{lb}^B = 12.21$ Gyr ago. The past light cones from \mathcal{Q}_A and \mathcal{Q}_B last intersected $t_{lb}^{AB} = 13.15$ Gyr ago. The shapes of the light cones reflect the changing rate of cosmic expansion since the big bang. To be consistent with our data, any local-realist mechanism would need to have affected detector settings and measurement outcomes of our experiment from within the past light cones of events \mathcal{Q}_A , \mathcal{Q}_B , or their overlap, a space-time region that consists of only 4.0% of the physical space-time volume contained within the past light cone of our experiment. Such a local-realist scenario would need to have been set in motion at least 7.78 Gyr ago.

ceed the local-realist bound of 2. However, not all of our settings were determined by genuine cosmic photons. A certain fraction of settings ϵ_k on each side ($k \in \{\mathcal{A}, \mathcal{B}\}$) was produced by some kind of local process, including sky glow, ambient light, and detector dark counts. We therefore consider such settings to be “corrupt” and assume that a local-realist mechanism could have exploited them to produce maximal CHSH correlations, with $S = 4$. Such a (hypothetical) mechanism could produce CHSH correlations as large as $S = 2(1 - \epsilon_A - \epsilon_B) + 4(\epsilon_A + \epsilon_B)$ [38, 40, 57].

In our analysis we account for such “corrupt” settings as well as unequal (biased) frequencies for various combinations of detector settings (a_i, b_j) , and possible “memory effects” by which a local-realist mechanism could exploit knowledge of settings and outcomes of previous trials (see Supplemental Material [45]). From this detailed treatment, we find that correlations at least as large as observed in our data could have been produced by a local-realist mechanism only with probabilities $p \leq 7.4 \times 10^{-21}$ for pair 1 and $p \leq 7.0 \times 10^{-13}$ for pair 2, corresponding to experimental violations of the Bell-CHSH bound by at least 9.3 and 7.1 standard deviations, respectively.

Conclusions. For each Cosmic Bell test reported here, we assume fair sampling and close the locality loophole. We also constrain the freedom-of-choice loophole with detector settings determined by extragalactic events, such that any local-realist mechanism would need to have acted no more recently than 7.78 Gyr or 3.22 Gyr ago for pairs 1 and 2, respectively—more than six orders of magnitude deeper into cosmic history than the experiments reported in Ref. [38]. This corresponds to excluding such local-realist mechanisms from 96.0% (pair 1) and 63.5% (pair 2) of the relevant space-time regions, compared to $\sim 10^{-5}\%$ of the relevant space-time region as in Ref. [38] (see Supplemental Materials [45]).

We have therefore dramatically limited the space-time regions from which local-realist mechanisms could have affected the outcome of our experiment to early in the history of our universe. To constrain such models further, one could use other physical signals to set detector settings, such as patches of the cosmic microwave background radiation (CMB), or even primordial neutrinos or gravitational waves, thereby constraining such models all the way back to the big bang—or perhaps even earlier, into a phase of early-universe inflation [31, 38]. Such extreme tests might ultimately prove relevant to the question of whether quantum entanglement undergirds the emergence of space-time itself. (For a recent review, see Ref. [58]).

Note Added. After we completed our experiment, a similar experiment was conducted by another group, the results of which are reported in Ref. [59].

Acknowledgements. The authors would like to thank Cecilia Fariña, Émilie L’Homé, Karl Kolle, Neil O’Mahony, Jürg Rey, Fiona Riddick and the whole team at the WHT as well as Emilio Molinari, Giovanni Mainella, Carlos Gonzalez and the whole team at the TNG for their tremendous support of our experiment. We also thank Thomas Augusteijn, Carlos Perez and all the staff at the NOT for their support, which did not decrease even after our container crashed into their telescope in a storm. We are also grateful for the encouraging support of Cesare Barbieri. In addition, we are grateful to Brian Keating, Hien Nguyen, Paul Schechter, and Gary Cole for helpful discussions. This work was supported by the Austrian Academy of Sciences (OEAW), by the Austrian Science Fund (FWF) with SFB F40 (FO-QUS) and FWF project CoQuS No. W1210-N16, the Austrian Federal Ministry of Education, Science and Research (BMBWF) and the University of Vienna via the project QUESS. This work was also supported by NSF INSPIRE Grant no. PHY-1541160. Portions of this work were conducted in MIT’s Center for Theoretical Physics and supported in part by the U.S. Department of Energy under Contract No. DE-SC0012567. C.L. was supported by the U.S. Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

* dominik.rauch@oeaw.ac.at

† dikaiser@mit.edu

‡ anton.zeilinger@univie.ac.at

- [1] E. Schrödinger, “Discussion of Probability Relations between Separated Systems,” *Math. Proc. Cambridge Phil. Soc.* **31**, 555–563 (1935).
- [2] A. Einstein, B. Podolsky, and N. Rosen, “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” *Phys. Rev.* **47**, 777–780 (1935).
- [3] N. Bohr, “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?” *Phys. Rev.* **48**, 696–702 (1935).
- [4] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics* **1**, 195–200 (1964).
- [5] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, 1987).
- [6] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed Experiment to Test Local Hidden-Variable Theories,” *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [7] J. F. Clauser and A. Shimony, “Bell’s theorem. Experimental tests and implications,” *Rep. Prog. Phys.* **41**, 1881–1927 (1978).
- [8] J.-Å. Larsson, “Loopholes in Bell inequality tests of local realism,” *J. Phys. A* **47**, 424003 (2014), arXiv:1407.0363 [quant-ph].
- [9] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, “Bell Nonlocality,” *Rev. Mod. Phys.* **86**, 419–478 (2014), arXiv:1303.2849 [quant-ph].
- [10] W. V. Quine, “Main Trends in Recent Philosophy: Two Dogmas of Empiricism,” *Philos. Rev.* **60**, 20 (1951).
- [11] P. M. Duhem, *The Aim and Structure of Physical Theory* (Princeton University Press, 1954).
- [12] J.-Å. Larsson, M. Giustina, J. Kofler, B. Wittmann, R. Ursin, and S. Ramelow, “Bell-inequality violation with entangled photons, free of the coincidence-time loophole,” *Phys. Rev. A* **90**, 032107 (2014), arXiv:1309.0712 [quant-ph].
- [13] A. Aspect, J. Dalibard, and G. Roger, “Experimental Test of Bell’s Inequalities Using Time-Varying Analyzers,” *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
- [14] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, “Violation of Bell’s Inequality under Strict Einstein Locality Conditions,” *Phys. Rev. Lett.* **81**, 5039–5043 (1998), arXiv:9810080 [quant-ph].
- [15] 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 (London)* **409**, 791–794 (2001).
- [16] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, “Bell violation using entangled photons without the fair-sampling assumption,” *Nature (London)* **497**, 227–230 (2013), arXiv:1309.0712v2.
- [17] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat, “Detection-Loophole-Free Test of Quantum Nonlocality, and Applications,” *Phys. Rev. Lett.* **111**, 130406 (2013), arXiv:1306.5772.

- [18] 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 (London)* **526**, 682–686 (2015), arXiv:1508.05949 [quant-ph].
- [19] 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,” *Phys. Rev. Lett.* **115**, 250401 (2015), arXiv:1511.03190.
- [20] 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,” *Phys. Rev. Lett.* **115**, 250402 (2015), arXiv:1511.03189.
- [21] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, “Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes,” *Phys. Rev. Lett.* **119**, 010402 (2017), arXiv:1611.04604 [quant-ph].
- [22] J. S. Bell, “The Theory of Local Beables,” *Epistemological Lett.* **9**, 86–96 (1976).
- [23] A. Shimony, M. A. Horne, and J. F. Clauser, “Comment on ‘The theory of local beables’,” *Epistemological Lett.* **13**, 97–102 (1976).
- [24] J. S. Bell, “Free variables and local causality,” *Epistemological Lett.* **15**, 79 (1977).
- [25] C. H. Brans, “Bell’s Theorem Does Not Eliminate Fully Causal Hidden Variables,” *Int. J. Theo. Phys.* **27**, 219–226 (1988).
- [26] J. Kofler, T. Paterek, and Č. Brukner, “Experimenter’s freedom in Bell’s theorem and quantum cryptography,” *Phys. Rev. A* **73**, 022104 (2006), quant-ph/0510167.
- [27] M. J. W. Hall, “Local Deterministic Model of Singlet State Correlations Based on Relaxing Measurement Independence,” *Phys. Rev. Lett.* **105**, 250404 (2010), arXiv:1007.5518 [quant-ph].
- [28] M. J. W. Hall, “Relaxed Bell inequalities and Kochen-Specker theorems,” *Phys. Rev. A* **84**, 022102 (2011), arXiv:1102.4467 [quant-ph].
- [29] J. Barrett and N. Gisin, “How Much Measurement Independence Is Needed to Demonstrate Nonlocality?” *Phys. Rev. Lett.* **106**, 100406 (2011), arXiv:1008.3612 [quant-ph].
- [30] M. Banik, M. Rajjak Gazi, S. Das, A. Rai, and S. Kunkri, “Optimal free will on one side in reproducing the singlet correlation,” *J. Phys. A* **45**, 205301 (2012), arXiv:1204.3835 [quant-ph].
- [31] J. Gallicchio, A. S. Friedman, and D. I. Kaiser, “Testing Bell’s Inequality with Cosmic Photons: Closing the Setting-Independence Loophole,” *Phys. Rev. Lett.* **112**, 110405 (2014), arXiv:1310.3288 [quant-ph].
- [32] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, “Arbitrarily Small Amount of Measurement Independence Is Sufficient to Manifest Quantum Nonlocality,” *Phys. Rev. Lett.* **113**, 190402 (2014), arXiv:1407.5634 [quant-ph].
- [33] G. Pütz and N. Gisin, “Measurement dependent locality,” *New J. Phys.* **18**, 055006 (2016), arXiv:1510.09087 [quant-ph].
- [34] M. J. W. Hall, “The significance of measurement independence for Bell inequalities and locality,” in *At the Frontier of Spacetime: Scalar-Tensor Theory, Bell’s Inequality, Mach’s Principle, Exotic Smoothness*, edited by T. Asselmeyer-Maluga (Springer, Switzerland, 2016) Chap. 11, pp. 189–204, arXiv:1511.00729 [quant-ph].
- [35] S. Pironio, “Random ‘choices’ and the locality loophole,” arXiv:1510.00248 [quant-ph].
- [36] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. K. Langford, T. Jennewein, and A. Zeilinger, “Violation of local realism with freedom of choice,” *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19708–19713 (2010), arXiv:0811.3129.
- [37] D. Aktas, S. Tanzilli, A. Martin, G. Pütz, R. Thew, and N. Gisin, “Demonstration of Quantum Nonlocality in the Presence of Measurement Dependence,” *Phys. Rev. Lett.* **114**, 220404 (2015), arXiv:1504.08332 [quant-ph].
- [38] J. Handsteiner, A. S. Friedman, D. Rauch, J. Gallicchio, B. Liu, H. Hosp, J. Kofler, D. Bricher, M. Fink, C. Leung, A. Mark, H. T. Nguyen, I. Sanders, F. Steinlechner, R. Ursin, S. Wengerowsky, A. H. Guth, D. I. Kaiser, T. Scheidl, and A. Zeilinger, “Cosmic Bell Test: Measurement Settings from Milky Way Stars,” *Phys. Rev. Lett.* **118**, 060401 (2017), arXiv:1611.06985 [quant-ph].
- [39] C. Wu, B. Bai, Y. Liu, X. Zhang, M. Yang, Y. Cao, J. Wang, S. Zhang, H. Zhou, X. Shi, X. Ma, J.-G. Ren, J. Zhang, C.-Z. Peng, J. Fan, Q. Zhang, and J.-W. Pan, “Random Number Generation with Cosmic Photons,” *Phys. Rev. Lett.* **118**, 140402 (2017), arXiv:1611.07126 [quant-ph].
- [40] C. Leung, A. Brown, H. Nguyen, A. S. Friedman, D. I. Kaiser, and J. Gallicchio, “Astronomical random numbers for quantum foundations experiments,” *Phys. Rev. A* **97**, 042120 (2018), arXiv:1706.02276 [quant-ph].
- [41] C. Abellán *et al.* (BIG Bell Test Collaboration), “Challenging local realism with human choices,” *Nature* **557**, 212–216 (2018), arXiv:1805.04431 [quant-ph].
- [42] One may also consider retrocausal models, in which the relevant hidden variables affect the measurement outcomes from the future [60–62].
- [43] The Italian Telescopio Nazionale Galileo (TNG) operated on the island of La Palma by the Fundación Galileo Galilei of the INAF (Istituto Nazionale di Astrofisica) at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.
- [44] The William Herschel Telescope (WHT) operated on the island of La Palma by the Isaac Newton Group of Telescopes (ING) at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.
- [45] See Supplemental Material (SM), which includes Refs. [63–86], for more information about the experimental setup and data analysis.
- [46] F. Steinlechner, M. Gilaberte, M. Jofre, T. Scheidl, J. P. Torres, V. Pruneri, and R. Ursin, “Efficient herald-

- ing of polarization-entangled photons from type-0 and type-II spontaneous parametric downconversion in periodically poled KTiOPO₄,” *J. Opt. Soc. America B* **31**, 2068 (2014).
- [47] T. Kim, M. Fiorentino, and F. N. C. Wong, “Phase-stable source of polarization-entangled photons using a polarization Sagnac interferometer,” in *Conference on Lasers and Electro-Optics and 2006 Quantum Electronics and Laser Science Conference, CLEO/OELS* (IEEE, Long Beach, CA, 2006) pp. 1–5, arXiv:quant-ph/0509219.
- [48] P. J. E. Peebles, *Principles of Physical Cosmology* (Princeton University Press, 1993).
- [49] S. Weinberg, *Cosmology* (Oxford University Press, 2008).
- [50] R. D. Blandford and R. Narayan, “Cosmological applications of gravitational lensing,” *Ann. Rev. Astron. Astrophys.* **30**, 311–358 (1992).
- [51] D. Downes, R. Neri, T. Wiklind, D. J. Wilner, and P. A. Shaver, “Detection of CO (4-3), CO (9-8), and Dust Emission in the Broad Absorption Line Quasar APM 08279+5255 at a Redshift of 3.9,” *Astrophys. J. Lett.* **513**, L1–L4 (1999), arXiv:astro-ph/9810111.
- [52] I. Pâris, P. Petitjean, É. Aubourg, A. D. Myers, A. Streblyanska, B. W. Lyke, S. F. Anderson, É. Armengaud, J. Bautista, M. R. Blanton, M. Blomqvist, J. Brinkmann, J. R. Brownstein, W. N. Brandt, É. Burtin, K. Dawson, S. de la Torre, A. Georgakakis, H. Gil-Marín, P. J. Green, P. B. Hall, J.-P. Kneib, S. M. LaMassa, J.-M. Le Goff, C. MacLeod, V. Mariappan, I. D. McGreer, A. Merloni, P. Noterdaeme, N. Palanque-Delabrouille, W. J. Percival, A. J. Ross, G. Rossi, D. P. Schneider, H.-J. Seo, R. Tojeiro, B. A. Weaver, A.-M. Weijmans, C. Yèche, P. Zarrouk, and G.-B. Zhao, “The Sloan Digital Sky Survey Quasar Catalog: Fourteenth data release,” *Astron. Astrophys.* **613**, A51 (2018), arXiv:1712.05029 [astro-ph.GA].
- [53] M. S. Shaw, R. W. Romani, G. Cotter, S. E. Healey, P. F. Michelson, A. C. S. Readhead, J. L. Richards, W. Max-Moerbeck, O. G. King, and W. J. Potter, “Spectroscopy of the Largest Ever γ -Ray-selected BL Lac Sample,” *Astrophys. J.* **764**, 135 (2013), arXiv:1301.0323 [astro-ph.HE].
- [54] P. A. R. Ade *et al.* (Planck Collaboration), “Planck 2015 results. XIII. Cosmological parameters,” *Astron. Astrophys.* **594**, A13 (2016), arXiv:1502.01589 [astro-ph.CO].
- [55] T. Scheidl, R. Ursin, A. Fedrizzi, S. Ramelow, X. S. Ma, T. Herbst, R. Prevedel, L. Ratschbacher, J. Kofler, T. Jennewein, and A. Zeilinger, “Feasibility of 300 km quantum key distribution with entangled states,” *New J. Phys.* **11**, 085002 (2009), arXiv:1007.4645.
- [56] B. S. Cirelson, “Quantum generalizations of Bell’s inequality,” *Lett. Math. Phys.* **4**, 93–100 (1980).
- [57] J. Kofler, M. Giustina, J.-Å. Larsson, and M. W. Mitchell, “Requirements for a Loophole-Free Photonic Bell Test using Imperfect Setting Generators,” *Phys. Rev. A* **93**, 032115 (2016), arXiv:1411.4787 [quant-ph].
- [58] M. Van Raamsdonk, “Lectures on Gravity and Entanglement,” in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1–26, 2015* (2017) pp. 297–351, arXiv:1609.00026 [hep-th].
- [59] 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**, in press (2018).
- [60] O. Costa De Beauregard, “S-matrix, Feynman zigzag and Einstein correlation,” *Phys. Lett. A* **67**, 171–174 (1978).
- [61] N. Argaman, “Bell’s theorem and the causal arrow of time,” *Am. J. Phys.* **78**, 1007–1013 (2010), arXiv:0807.2041 [quant-ph].
- [62] H. Price and K. Wharton, “Disentangling the Quantum World,” *Entropy* **17**, 7752–7767 (2015), arXiv:1508.01140 [quant-ph].
- [63] P. Sarkar, J. Yadav, B. Pandey, and S. Bharadwaj, “The scale of homogeneity of the galaxy distribution in SDSS DR6,” *M. Not. Roy. Astron. Soc.* **399**, L128–L131 (2009), arXiv:0906.3431 [astro-ph.CO].
- [64] C. Marinoni, J. Bel, and A. Buzzi, “The scale of cosmic isotropy,” *J. Cosmol. Astropart. Phys.* **10**, 036 (2012), arXiv:1205.3309 [astro-ph.CO].
- [65] A. S. Friedman, D. I. Kaiser, and J. Gallicchio, “The shared causal pasts and futures of cosmological events,” *Phys. Rev. D* **88**, 044038 (2013), arXiv:1305.3943 [astro-ph.CO].
- [66] M. G. Mingaliev, Y. V. Sotnikova, R. Y. Udvovitskiy, T. V. Mufakharov, E. Nieppola, and A. K. Erkenov, “RATAN-600 multi-frequency data for the BL Lacertae objects,” *Astron. Astrophys.* **572**, A59 (2014), arXiv:1410.2835 [astro-ph.GA].
- [67] W. F. Kern and J. R. Bland, *Solid Mensuration, With Proofs* (Wiley, 1938).
- [68] E. Egami, G. Neugebauer, B. T. Soifer, K. Matthews, M. Ressler, E. E. Becklin, T. W. Murphy, Jr., and D. A. Dale, “APM 08279+5255: Keck Near- and Mid-Infrared High-Resolution Imaging,” *Astrophys. J.* **535**, 561–574 (2000), arXiv:astro-ph/0001200.
- [69] S. Oya, Y. Minowa, H. Terada, M. Watanabe, M. Hattori, Y. Hayano, Y. Saito, M. Ito, T.-S. Pyo, H. Takami, and M. Iye, “Spatially Resolved Near-Infrared Imaging of a Gravitationally Lensed Quasar, APM 08279+5255, with Adaptive Optics on the Subaru Telescope,” *Publ. Astron. Soc. Japan* **65**, 9 (2013).
- [70] M. Rauch, “The Lyman Alpha Forest in the Spectra of QSOs,” *Ann. Rev. Astron. Astrophys.* **36**, 267–316 (1998), arXiv:astro-ph/9806286.
- [71] P. McDonald, U. Seljak, S. Burles, D. J. Schlegel, D. H. Weinberg, R. Cen, D. Shih, J. Schaye, D. P. Schneider, N. A. Bahcall, J. W. Briggs, J. Brinkmann, R. J. Brunner, M. Fukugita, J. E. Gunn, Ž. Ivezić, S. Kent, R. H. Lupton, and D. E. Vanden Berk, “The Ly α Forest Power Spectrum from the Sloan Digital Sky Survey,” *Astrophys. J. Suppl.* **163**, 80–109 (2006), arXiv:astro-ph/0405013.
- [72] C. R. Benn and S. L. Ellison, “Brightness of the night sky over La Palma,” *New Astron. Rev.* **42**, 503 (1998), arXiv:astro-ph/9909153.
- [73] I. A. B. Meinel, “OH Emission Bands in the Spectrum of the Night Sky,” *Astrophys. J.* **111**, 555 (1950).
- [74] K. Z. Stanek, T. W.-S. Holoién, C. S. Kochanek, A. B. Davis, G. Simonian, U. Basu, N. Goss, J. F. Beacom, B. J. Shappee, J. L. Prieto, D. Bersier, S. Dong, P. R. Wozniak, J. Brimacombe, D. Szczygiel, and G. Pojmanski, “ASAS-SN Photometry of QSO BZB J0424+0036,”

- [The Astronomer's Telegram 6866 \(2015\)](#).
- [75] L. Carrasco, G. Escobedo, A. Porras, E. Recillas, V. Chavushyan, A. Carraminana, and D. Mayya, “NIR brightening of the blazar BZB J0424+0036,” [The Astronomer's Telegram 5712 \(2014\)](#).
- [76] L. Carrasco, A. Porras, E. Recillas, J. Leon-Tavares, V. Chavushyan, and A. Carraminana, “An ongoing NIR Flare of the QSO BZB J0424+0036,” [The Astronomer's Telegram 6865 \(2015\)](#).
- [77] L. Carrasco, A. Porras, E. Recillas, J. Leon-Tavares, V. Chavushyan, and A. Carraminana, “NIR brightening of the QSO HB89 0422+004,” [The Astronomer's Telegram 6971 \(2015\)](#).
- [78] D. L. King, “Atmospheric Extinction at the Roque de los Muchachos Observatory, La Palma,” [RGO/La Palma technical note 31 \(1985\)](#).
- [79] F. G. Saturni, D. Trevese, F. Vagnetti, M. Perna, and M. Dadina, “A multi-epoch spectroscopic study of the BAL quasar APM 08279+5255. II. Emission- and absorption-line variability time lags,” [Astron. Astrophys. 587, A43 \(2016\)](#), arXiv:1512.03195 [astro-ph.GA].
- [80] B. Sbarufatti, A. Treves, R. Falomo, J. Heidt, J. Kotilainen, and R. Scarpa, “ESO Very Large Telescope Optical Spectroscopy of BL Lacertae Objects. II. New Redshifts, Featureless Objects, and Classification Assessments,” [Astron. J. 132, 1–19 \(2006\)](#), arXiv:astro-ph/0601506.
- [81] A. Treves and S. Panzeri, “The upward bias in measures of information derived from limited data samples,” [Neural Computation 7, 399–407 \(1995\)](#).
- [82] L. E. Bassham III, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, E. B. Barker, S. D. Leigh, M. Levenson, M. Vangel, D. L. Banks, *et al.*, “Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications,” [Special Publication \(NIST SP\) 800-22 Rev 1a \(2010\)](#).
- [83] R. D. Gill, “Time, Finite Statistics, and Bell’s Fifth Position,” in *Proceedings of Foundations of Probability and Physics - 2, Math. Modelling in Phys., Engin., and Cogn. Sc.*, Vol. 5 (Växjö University Press, 2003) p. 179, arXiv:quant-ph/0301059.
- [84] R. D. Gill, “Statistics, Causality and Bell’s Theorem,” [Statist. Sci. 29, 512–528 \(2014\)](#), arXiv:1207.5103 [stat.AP].
- [85] P. Bierhorst, “A robust mathematical model for a loophole-free Clauser-Horne experiment,” [J. Phys. A 48, 195302 \(2015\)](#), arXiv:1312.2999 [quant-ph].
- [86] D. Elkouss and S. Wehner, “(Nearly) optimal P values for all Bell inequalities,” [NPJ Quantum Information 2, 16026 \(2016\)](#), arXiv:1510.07233 [quant-ph].

Supplemental Material

Cosmic Bell Test using Random Measurement Settings from High-Redshift Quasars

Causal Alignment

As described in Ref. [38], the time-dependent locations of astronomical sources on the sky relative to our ground-based experimental site complicates the enforcement of the space-like separation conditions needed to address both the locality and freedom-of-choice loopholes. For example, the photon from quasar emission event \mathcal{Q}_A must be received by Alice's cosmic-photon receiving telescope (Rx-CP) before that photon's causal wavefront reaches either the Rx-CP or the entangled-photon receiving telescope (Rx-EP) on Bob's side, and vice versa.

In this section we first present our main result for the causal-alignment windows, $\tau_{\text{valid}}^k(t)$ (for sides $k, \ell = \{\mathcal{A}, \mathcal{B}\}$), within which settings chosen by astronomical photons remain valid, and then derive the various terms in our expression. As shown in Figure 3 of the Main paper, we parameterize $\tau_{\text{valid}}^k(t)$ as

$$\tau_{\text{valid}}^k(t) = \bar{\tau}_{\text{geom}}^k(t) - \tau_{\text{set}}^k - \tau_{\text{buffer}}^k, \quad (1)$$

where $\tau_{\text{geom}}^k(t)$ arises from the geometrical arrangement of the quasars relative to the locations of relevant instrumentation on Earth, and $\bar{\tau}_{\text{geom}}^k(t) \equiv \min_t [\tau_{\text{geom}}^k(t)]$ is the minimum value of $\tau_{\text{geom}}^k(t)$ during an observing window. The term τ_{set}^k indicates the time required to electronically output a bit and implement the detector setting, while τ_{buffer}^k accommodates total delays due to atmosphere, telescope optics, and detector response. Negative validity times for either k would indicate an instantaneous configuration that was out of “causal alignment,” in which at least one setting would be invalid for the purposes of closing the locality loophole.

For $\tau_{\text{geom}}^k(t)$, we find

$$\begin{aligned} \tau_{\text{geom}}^k(t) &= \frac{1}{c} \hat{\mathbf{n}}_k(t) \cdot (\mathbf{r}_k - \mathbf{m}_\ell) + \frac{n}{c} \left[|\mathbf{m}_k - \mathbf{s}| - |\mathbf{m}_\ell - \mathbf{s}| \right] \\ &\quad - \frac{\gamma_k}{c} |\mathbf{r}_k - \mathbf{m}_k|, \end{aligned} \quad (2)$$

where \mathbf{r}_k and \mathbf{m}_k are the spatial 3-vectors for the locations of the cosmic receiving telescopes (Rx-CP) and entangled-particle detectors (Rx-EP) for side k , respectively; \mathbf{s} is the spatial 3-vector for the location of the source of entangled particles; and c is the speed of light in vacuum. The time-dependent unit vector $\hat{\mathbf{n}}_k(t)$ points toward the quasar used to set detector settings on side k , and is computed using astronomical ephemeris calculations. Additionally, n is the index of refraction of air and γ_k , which acts like an index of refraction, parameterizes the group velocity delay through fiber optics and/or electrical cables connecting the telescope and entangled photon detector on side k .

We work with a space-time metric signature $(+, -, -, -)$, so that space-time events represented by four-vectors \mathcal{A}^μ and \mathcal{B}^μ will be space-like separated if $(\mathcal{A}^\mu - \mathcal{B}^\mu)^2 < 0$. We represent spatial and temporal intervals of cosmological magnitude—such as the interval between emission of light from a distant quasar and its detection on Earth today—in terms of a spatially flat Friedmann-Lemaître-Robertson-Walker (FLRW) line element because on length-scales greater than $\mathcal{O}(100)$ Mpc, our universe has been measured to be homogeneous [63], isotropic [64], and spatially flat [54] to high accuracy. We have

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= c^2 dt^2 - R_0^2 a^2(t) \left[d\chi^2 + \chi^2 d\Omega_{(2)}^2(\theta, \phi) \right], \end{aligned} \quad (3)$$

where t is cosmic time, equal to the proper time recorded by a freely falling observer at the origin of the spatial coordinate system, χ is a (dimensionless) comoving distance, and $d\Omega_{(2)}^2(\theta, \phi) = d\theta^2 + \sin^2 \theta d\phi^2$ is the line-element for a unit 2-sphere. In this section we ignore possible complications from inhomogeneities, such as gravitational-lensing effects.

In Eq. (3), $a(t)$ is the (dimensionless) cosmic scale factor and the constant $R_0 \equiv c/H_0$ has dimensions of length. We use $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ as the present value of the Hubble parameter [54], corresponding to a Hubble time $t_H = H_0^{-1} = 14.43$ Gyr and hence $R_0 = 14.43$ Glyr. Physical distances at a given cosmic time, $r(t)$, are related to comoving distances by $r(t) = R_0 a(t) \chi$. The observed redshift, z , for astronomical objects arising from cosmic expansion is given by

$$1 + z = \frac{a(t_0)}{a(t_e)}, \quad (4)$$

where t_0 is the present time and t_e is the time of emission. We set $a(t_0) = 1$ and take $t = 0$ to be the time of the hot big bang (following any primordial phase of inflation, if inflation occurred). We further assign the origin of the spatial coordinates to be the center of the Earth. Errors introduced by treating the rotating Earth as an inertial frame are less than one part in 10^6 , and are easily accommodated within τ_{buffer} .

Quasar emission event \mathcal{Q}_k occurred a long time ago, in a galaxy far, far away. Hence it is convenient to introduce (dimensionless) conformal time, $d\eta \equiv H_0 dt/a(t)$, to take into account the cosmic expansion between the emission and detection of the cosmic photons. Then Eq. (3) becomes

$$ds^2 = R_0^2 a^2(\eta) \left[d\eta^2 - d\chi^2 - \chi^2 d\Omega_{(2)}^2 \right], \quad (5)$$

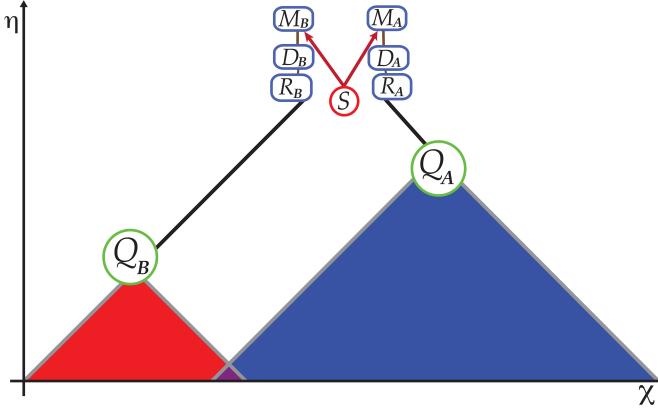


FIG. 1. Schematic space-time diagram (not to scale) of our cosmic Bell test, in (dimensionless) conformal time η versus comoving distance χ . In these coordinates, null geodesics appear as 45° diagonals. On each side $k \in \{A, B\}$, light from quasar emission event \mathcal{Q}_k is received on Earth at event \mathcal{R}_k and used to determine the detector setting at event \mathcal{D}_k . Meanwhile, spacelike-separated from events \mathcal{Q}_k , \mathcal{R}_k , and \mathcal{D}_k , the source S emits a pair of entangled particles, which are measured at events \mathcal{M}_k .

and (radial) null geodesics correspond to $d\eta = d\chi$. In these coordinates, the 4-vector corresponding to the receipt of a quasar photon at detector k on Earth may be written $\mathcal{R}_k^\mu = (\eta_{r_k}, r_k)$, and the 4-vector for the emission of the photon at the quasar is

$$\mathcal{Q}_k^\mu = (\eta_{q_k}, \chi_{r_k} + (\eta_{r_k} - \eta_{q_k}) \hat{\mathbf{n}}_k), \quad (6)$$

where χ_{r_k} is the (comoving) spatial location of the quasar photon detector Rx-CP on side k , and $\hat{\mathbf{n}}_k(t)$ is the unit spatial 3-vector pointing from the center of the Earth toward the quasar. The entangled pair is emitted from the source at $\mathcal{S}^\mu = (\eta_s, \chi_s)$. See Figure 1.

The entangled-pair emission event should be space-like separated from the arrival of the quasar photons, which requires $(\mathcal{Q}_k - \mathcal{S})^2 < 0$. Assuming $(\eta_{r_k} - \eta_s), |\chi_{r_k} - \chi_s| \ll (\eta_{r_k} - \eta_{q_k})$, this implies

$$-\hat{\mathbf{n}}_k \cdot (\chi_{r_k} - \chi_s) < \eta_{r_k} - \eta_s. \quad (7)$$

(We further note that in the limit $(\eta_{r_k} - \eta_{q_k}) \gg |\chi_{r_k}|$, the spherical waves emitted from the quasar arrive at the Earth as plane waves to a very good approximation.) The quantities in Eq. (7) all refer to Earthbound events, and hence we may transform back to coordinates more convenient for describing a given experimental trial. Upon recalling that $a(t_0) = 1$, we have $\chi_{r_k} = R_0^{-1}\mathbf{r}_k$, where \mathbf{r}_k is the present (physical) spatial location of the quasar-photon detector Rx-CP as reckoned from the center of the Earth, and likewise $\chi_s = R_0^{-1}\mathbf{s}$. We also have

$$\eta_{r_k} - \eta_s = H_0 \int_{t_s}^{t_{r_k}} \frac{dt'}{a(t')}. \quad (8)$$

The cosmic scale factor varies imperceptibly during the course of the experiment, so we may expand $a(t_0 + \delta t) = a(t_0) + \dot{a}(t_0)\delta t + \mathcal{O}[(\delta t)^2]$, given that $t_H^{-1} = H_0 = \dot{a}(t_0)/a(t_0)$, and overdots denote derivatives with respect to cosmic time, t . During a given experimental trial, $\delta t = t_{r_k} - t_s$ is typically a fraction of a second, so we have $(\eta_{r_k} - \eta_s) = H_0(t_{r_k} - t_s) + \mathcal{O}(10^{-17})$. Then Eq. (7) becomes

$$-\frac{\hat{\mathbf{n}}_k}{c} \cdot (\mathbf{r}_k - \mathbf{s}) < t_{r_k} - t_s. \quad (9)$$

Each measurement on an entangled particle should be completed before the causal wavefront from the quasar emission on the other side arrives, which requires $(\mathcal{Q}_k - \mathcal{M}_\ell)^2 < 0$. Representing $\mathcal{M}_k^\mu = (\eta_{m_k}, \chi_{m_k})$, using Eq. (6) for \mathcal{Q}_k^μ , and proceeding as above to relate (η_{m_k}, χ_{m_k}) to $(t_{m_k}, \mathbf{r}_{m_k})$, we find

$$(t_{r_k} - t_{m_\ell}) + \frac{\hat{\mathbf{n}}_k}{c} \cdot (\mathbf{r}_k - \mathbf{m}_\ell) > 0. \quad (10)$$

Meanwhile, each quasar photon must be received, processed, and converted to a stable detector setting before the entangled photon from the source arrives. To calculate $\tau_{\text{geom}}^k(t)$, we consider only the spatial arrangement of the various events, since we separately accommodate additional delays (from telescope optics, electronics, cables, and the like) with the factors τ_{set}^k and τ_{buffer}^k . For $\tau_{\text{geom}}^k(t)$, we may therefore parameterize

$$t_{d_k} - t_{r_k} = \frac{\gamma_k}{c} |\mathbf{r}_k - \mathbf{m}_k|, \quad (11)$$

where t_{d_k} is the time when the setting for the entangled-photon detector on side k is set. Eq. (11) takes into account the fact that the quasar-photon reception and the detector-setting event can occur at different spatial locations. For measurements on the future light cone of the entangled-particle emission, we can use the particles' travel time from the source to the detectors to write

$$t_{m_k} = t_s + \frac{n}{c} |\mathbf{m}_k - \mathbf{s}|. \quad (12)$$

For the setting to be valid, meanwhile, it must be set before the measurement, $t_{d_k} < t_{m_k}$. Then we may rearrange Eqs. (11) and (12) to write

$$t_{r_k} - t_s < \frac{n}{c} |\mathbf{m}_k - \mathbf{s}| - \frac{\gamma_k}{c} |\mathbf{r}_k - \mathbf{m}_k|. \quad (13)$$

Similarly substituting Eq. (12) for t_{m_ℓ} into Eq. (10) we have

$$\frac{n}{c} |\mathbf{m}_\ell - \mathbf{s}| - \frac{\hat{\mathbf{n}}_k}{c} \cdot (\mathbf{r}_k - \mathbf{m}_\ell) < t_{r_k} - t_s. \quad (14)$$

We want the most conservative limit on $\tau_{\text{geom}}^k(t)$. Subtracting the lefthand side of Eq. (9) from the lefthand side of Eq. (14) yields $c^{-1}(n + \cos\theta)|\mathbf{m}_\ell - \mathbf{s}| \geq 0$, since

the index of refraction satisfies $n \geq 1$, and $|\cos \theta| \leq 1$, where θ is the angle between $\hat{\mathbf{n}}_k$ and $(\mathbf{m}_\ell - \mathbf{s})$. Therefore the lefthand side of Eq. (14) provides the tighter lower bound on $(t_{r_k} - t_s)$. The validity time is determined by the difference between the upper and lower bounds on $(t_{r_k} - t_s)$: subtracting the lefthand side of Eq. (14) from the righthand side of Eq. (13) yields our expression for $\tau_{\text{geom}}^k(t)$ in Eq. (2).

For our experiment, we may set $\mathbf{m}_k \simeq \mathbf{r}_k$ for each side, and accommodate measured delays for signal propagation and processing within the factors τ_{set}^k . Then we may compute values for $\tau_{\text{geom}}^k(t)$ using the coordinates for the various experimental stations shown in Table I. We used $\tau_{\text{buffer}} = 150$ ns on each side, as well as $\tau_{\text{set}}^A = 325$ ns and $\tau_{\text{set}}^B = 430$ ns. Incorporating these values for τ_{set}^k and τ_{buffer} as in Eq. (1) yielded the final values of τ_{valid}^k that we used, shown in Table II.

Component	Lat. [°]	Lon. [°]	Elev. [m]	diam _k [m]
Rx-CP \mathcal{A}	28.75410 [°]	-17.88915 [°]	2375	3.58
Source S	28.757189 [°]	-17.884961 [°]	2385	
Rx-CP \mathcal{B}	28.760636 [°]	-17.8816861 [°]	2352	4.20

TABLE I. Latitude, Longitude, and Elevation for the cosmic-photon receiving telescopes (Rx-CP) for Alice (\mathcal{A}) and Bob (\mathcal{B}), and for the Source (S). Also shown are the aperture diameters (diam_k) for the telescopes.

Excluded Spacetime Regions

By using light from distant quasars to determine detector settings, we may constrain the space-time region within which any putative local-realist mechanism could have engineered the observed correlations among measurements on the entangled particles. In this section we consider the space-time regions excluded from such local-realist scenarios.

Following the discussion in Ref. [65], we may relate the measured redshift for a given astronomical object to the conformal time at which the light we receive on Earth was emitted by the object, η_q . We take $\eta = 0$ to correspond to the time of the hot big bang. We may also compute the lookback time to the emission event (in cosmic time), t_{lb} , reckoned from the present, t_0 .

We parameterize the Friedmann equation governing the evolution of $a(t)$ in terms of the function

$$E(a) \equiv \frac{H(a)}{H_0} = \sqrt{\Omega_\Lambda + \Omega_k a^{-2} + \Omega_M a^{-3} + \Omega_R a^{-4}}, \quad (15)$$

where $H(a)$ is the Hubble parameter for a given scale factor $a = a(t)$, and we again use the best-fit value $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ [54]. The $\Omega_i \equiv \rho_i/\rho_c$ are the present-day ratios of the energy densities of dark energy (ρ_Λ), cold matter (ρ_M), and radiation (ρ_R) to the criti-

cal density, $\rho_c = 3H_0^2/(8\pi G)$, where G is Newton's gravitational constant. (The quantity ρ_M includes contributions from both baryonic matter and cold dark matter.) We also define the total fractional density of dark energy, cold matter, and radiation ($\Omega_T \equiv \Omega_\Lambda + \Omega_M + \Omega_R$), and the fractional density associated with spatial curvature ($\Omega_k \equiv 1 - \Omega_T$). We assume that ρ_Λ arises from a genuine cosmological constant with equation of state $w = p/\rho = -1$, and hence $\Omega_\Lambda a^{-3(1+w)} = \Omega_\Lambda$, which is consistent with observations [54]. We adopt the best-fit cosmological parameters from Ref. [54],

$$\vec{\Omega} = (\Omega_\Lambda, \Omega_M, \Omega_R) = (0.6911, 0.3089, 9.16 \times 10^{-5}), \quad (16)$$

consistent with $|\Omega_k| < \mathcal{O}(10^{-3})$. Here $\Omega_R = \Omega_M/(1 + z_{\text{eq}})$, with the redshift for matter-radiation equality given by $z_{\text{eq}} = 3371$ [54].

Redshifts for the three quasars we observed are listed in Table II. For QSO J0831+5245, which was observed for both quasar pairs, we use a reported host galaxy redshift of $z = 3.9114 \pm 0.0003$ from Ref. [51]. For QSO B0350-073, we use the reported redshift of $z = 0.9635389 \pm 0.00011$ from the Sloan Digital Sky Survey (SDSS) Quasar Catalog Fourteenth Data Release [52]. For QSO B0422+004, reported redshifts include $z = 0.31$ [66] and $z = 0.268$ [53]. Neither reported redshift included an uncertainty, so we conservatively adopt the smaller value, $z = 0.268$. Given the small redshift uncertainties for QSO J0831+5245 and QSO B0350-073, and that no redshift uncertainties were reported for QSO B0422+004, we assume that all redshift uncertainties are negligible.

The conformal time for the emission event from a distant quasar at redshift z may be written [65]

$$\eta_k(z) = \int_0^{1/(1+z)} \frac{da}{a^2 E(a)}, \quad (17)$$

upon using $a_e = 1/(1+z)$ from Eq. (4) (and recalling our convention that $a(t_0) = 1$). Using $d\eta = H_0 dt/a(t)$, we may similarly compute the (cosmic time) lookback time to the emission event from today as

$$t_{\text{lb}}^k(z) = \int_{1/(1+z)}^1 \frac{da}{a H(a)}, \quad (18)$$

again using $a_e = 1/(1+z)$. In Table II we list the quasars used in pairs 1 and 2 for $k = \mathcal{A}, \mathcal{B}$, their measured redshifts, z , and the corresponding values of $\eta_k(z)$ and $t_{\text{lb}}^k(z)$ for each quasar. The present age of the universe corresponds to $\eta_0 = \eta(z=0) = 3.20$, and the lookback time to the hot big bang is $t_{\text{lb}}(\infty) = 13.80$ Gyr.

Neglecting (for the moment) any possible effects from inhomogeneities along the lines of sight between the quasar emission events and our receipt of the cosmic photons on Earth, we assume that any local-realist mechanism that could have engineered the observed violations

Pair	Side	Simbad ID	RA°	DEC°	α°	az $_{\text{k}}^{\circ}$	alt $_{\text{k}}^{\circ}$	z	η_k	t_{lb}^k (Gyr)	F_{excl}	τ_{valid}^k [μs]
1	\mathcal{A}	QSO B0350-073	58.127300	-7.183976	83.81	233	38	0.964	2.46	7.78	0.960	2.34
	\mathcal{B}	QSO J0831+5245	127.923750	52.754860		35	57	3.911	1.56	12.21		0.90
2	\mathcal{A}	QSO B0422+004	66.195175	0.601758	72.84	246	38	0.268	2.95	3.22	0.635	2.20
	\mathcal{B}	QSO J0831+5245	127.923750	52.754860		21	64	3.911	1.56	12.21		0.53

TABLE II. Quasars whose light was used to determine detector settings for Alice (\mathcal{A}) and Bob (\mathcal{B}) for pairs 1 and 2. For each quasar, we list its QSO ID number from the Simbad database, celestial coordinates, angular separations (α) of each pair, azimuth (clockwise from due North) and altitude above horizon at the start of each observing run, and redshift (z). We also list the (dimensionless) conformal time of quasar emission (η_k), the lookback times (in Gyr) to each quasar emission event (t_{lb}^k), as well as the fraction of the physical 4-volume of the past light cone of our experiment, extending back to the big bang, from which any local-realist mechanism that might account for the measured violations of the Bell-CHSH inequality is excluded (F_{excl}). Finally, we list the validity time τ_{valid}^k from Eq. (1), which gives the minimum time that detector settings are valid for side $k = \{\mathcal{A}, \mathcal{B}\}$ during each experimental run, taking into account various delays and safety margins.

of the Bell-CHSH inequality must have acted within the past lightcone of either quasar emission event. Only within those spacetime regions could the local-realist mechanism have altered or previewed the bit that we would later receive on Earth, and shared that information (at or below the speed of light) with other elements of our experimental apparatus, such as the source of entangled particles or the detectors on the other side of our experiment [31, 38]. For pair 1, any such local-realist mechanism is constrained to have acted no more recently than $t_{\text{lb}} = 7.78$ Gyr ago, while for pair 2 the constraint is $t_{\text{lb}} = 3.22$ Gyr ago.

We may further characterize the space-time region within which any local-realist mechanism could have acted in order to produce the observed violations of the Bell-CHSH inequality. That region consists of the union of the past lightcones from the quasar emission events utilized for a given experimental run, $V_Q^{(4)}$, which we may compare with the space-time 4-volume of the past lightcone of the experiment itself, $V_{\text{exp}}^{(4)}$. To calculate $V_Q^{(4)}$, we must consider the 4-volume of the past light cone from each emission event and subtract the 4-volume of those light cones' intersection.

We calculate the 4-volume contained within the past light cone of a quasar emission event by integrating the invariant volume element $dV = \sqrt{-g} d^4x$ over the region bounded by past-directed null geodesics extending from the quasar emission event, where $g = \det[g_{\mu\nu}(x)]$ is the determinant of the space-time metric. As we saw above, null geodesics take the form $d\eta = d\chi$ in the coordinates of Eq. (5). Taking the spatial origin to lie along the worldline of quasar k , the 4-volume of the past light cone from emission event \mathcal{Q}_k may be evaluated as

$$V^{(4)}(\eta_k) = R_0^4 \int \Theta(\eta_k - \eta - \chi) a^4(\eta) \chi^2 \sin \theta d\eta d\chi d\theta d\phi \\ = \frac{4\pi R_0^4}{3} \int_0^{\eta_k} d\eta a^4(\eta) (\eta_k - \eta)^3, \quad (19)$$

where $\Theta(x)$ is the Heaviside step function. In a FLRW universe, the most recent (conformal) time at which the past light cones from emission events \mathcal{Q}_A and \mathcal{Q}_B overlap

is given by [65]

$$\eta_{AB} = \frac{1}{2} (\eta_A + \eta_B - \chi_L). \quad (20)$$

Here χ_L is the comoving spatial distance between the worldlines of quasars A and B , which, in a spatially flat universe, is given by

$$\chi_L = \sqrt{\chi_A^2 + \chi_B^2 - 2\chi_A\chi_B \cos \alpha}, \quad (21)$$

where α is the angle between the quasars as seen from Earth, and [65]

$$\chi_k(z) = \int_{1/(1+z)}^1 \frac{da}{a^2 E(a)}. \quad (22)$$

Along any $\eta = \text{constant}$ surface, with $0 \leq \eta \leq \eta_{AB}$, the past light cones from emission events \mathcal{Q}_A and \mathcal{Q}_B appear as three-dimensional spheres that partially overlap. In Euclidean space, the (spatial) three-volume of the intersection region of two spheres of radii r_1 and r_2 , with distance between their centers d , is given by [67]

$$V_I^{(3)} = \frac{\pi}{12d} (r_1 + r_2 - d)^2 [d^2 + 2d(r_1 + r_2) - 3(r_1 - r_2)^2]. \quad (23)$$

Upon substituting $d \rightarrow \chi_L$, $r_1 \rightarrow \eta_A - \eta$, and $r_2 \rightarrow \eta_B - \eta$, making use of Eq. (20) for η_{AB} , and performing some straightforward algebra, we find the space-time 4-volume of the intersection region of the past light cones from emission events \mathcal{Q}_A and \mathcal{Q}_B to be

$$V_I^{(4)}(\eta_A, \eta_B, \alpha) \\ = 4\pi R_0^4 \int_0^{\eta_{AB}} d\eta a^4(\eta) \left[\frac{1}{3} (\eta_{AB} - \eta)^3 \right. \\ \left. + (\eta_{AB} - \eta)^2 \left(\frac{\chi_L^2 - (\eta_A - \eta_B)^2}{4\chi_L} \right) \right]. \quad (24)$$

The union of the past light cones from emission events \mathcal{Q}_A and \mathcal{Q}_B therefore has the 4-volume

$$V_Q^{(4)}(\eta_A, \eta_B, \alpha) = V^{(4)}(\eta_A) + V^{(4)}(\eta_B) - V_I^{(4)}(\eta_A, \eta_B, \alpha), \quad (25)$$

while the 4-volume of the past light cone of our experiment is given by $V_{\text{exp}}^{(4)} = V^{(4)}(\eta_0)$.

For an experimental run using a pair of quasars with redshifts z_A , z_B , and relative angle α , the space-time region that is *excluded* from playing any role in an explanation based on a local-realist mechanism is given by $V_{\text{exp}}^{(4)} - V_Q^{(4)}$. As a fraction of $V_{\text{exp}}^{(4)}$, this may be written

$$F_{\text{excl}} = 1 - \left(\frac{V_Q^{(4)}(\eta_A, \eta_B, \alpha)}{V^{(4)}(\eta_0)} \right). \quad (26)$$

The relative angle (as seen from Earth) between the quasars in pair 1 is $\alpha = 83.81^\circ$, and for the quasars of pair 2 is $\alpha = 72.84^\circ$. Given the redshifts for each quasar listed in [Table II](#), we find $F_{\text{excl}} = 0.960$ for pair 1, and $F_{\text{excl}} = 0.635$ for pair 2. See [Figure 2](#) and [Figure 3](#).

We may compare these values for F_{excl} with the corresponding values for our Vienna pilot test involving Milky Way stars [\[38\]](#). We again use Eq. (4) to relate redshift to cosmic scale factor. Since the motions of Milky Way stars are dominated by local peculiar velocities independent of Hubble expansion, astronomers do not measure cosmic-expansion redshifts for Milky Way stars. Nonetheless, we may compute effective values of z with which to parameterize the various emission times. For the nearby sources used in Ref. [\[38\]](#), we may Taylor expand

$$a(t_e) = a(t_0) - H_0(t_0 - t_e) + \mathcal{O}([(t_0 - t_e)/t_H]^2), \quad (27)$$

where, as above, $H_0 = \dot{a}(t_0)/a(t_0)$ is the present value of the Hubble parameter, $t_H = H_0^{-1} = 14.43$ Gyr, and we scale $a(t_0) = 1$. Neglecting uncertainties on the measured distances to the Milky Way stars that we used in our pilot test, we have $(t_0 - t_A) = 604$ years and $(t_0 - t_B) = 1930$ years for run 1, and $(t_0 - t_A) = 577$ years and $(t_0 - t_B) = 3624$ years for run 2. Using Eqs. (4) and (27), these correspond to effective redshifts $z_A = 4.19 \times 10^{-8}$ and $z_B = 1.32 \times 10^{-7}$ for run 1, and $z_A = 4.00 \times 10^{-8}$ and $z_B = 2.51 \times 10^{-7}$ for run 2. Given $\alpha_1 = 119^\circ$ for run 1 and $\alpha_2 = 112^\circ$ for run 2, the times when the past light cones from emission events A and B most recently overlapped were $t_{AB} = 2409$ years ago (run 1) and $t_{AB} = 4039$ years ago (run 2), corresponding to $z_{AB} = 1.67 \times 10^{-7}$ (run 1) and $z_{AB} = 2.80 \times 10^{-7}$ (run 2). Repeating the calculation as above, we then find $F_{\text{excl}} = 1.38 \times 10^{-7}$ for run 1, and $F_{\text{excl}} = 1.45 \times 10^{-7}$ for run 2. In other words, the Vienna pilot test excluded about one hundred-thousandth of one percent of the relevant space-time volume, compared to the exclusion of 96.0% (pair 1) and 63.5% (pair 2) achieved in the present experiment.

Effects of Inhomogeneities along the Line of Sight

Our discussion to this point has assumed that the quasar photons were emitted from point-like astronom-

ical objects. In reality, quasars are large, messy objects; a given photon may be subject to complicated interactions involving optically thick atmospheres before escaping from the quasar. To address such scenarios, we consider the “effective emission time” to be the latest time η_k that any local interactions associated with the quasar could have altered the wavelength of the photon. Any corrections arising from strong electromagnetic fields, plasma effects, or related atmospheric phenomena near the quasar would affect a precise calculation of η_k by some small quantity (with cosmic-time values for the corrections $\Delta t \ll t_{\text{lb}}^k$). Compared to our simple estimates of $\eta_k(z)$ and $t_{\text{lb}}^k(z)$ based on the measured redshifts of the quasars, any such corrections would be indiscernible, given $t_{\text{lb}}^k(z) \sim \mathcal{O}(1-10)$ Gyr for the quasars used in pairs 1 and 2.

Atmospheric or related interactions at the quasar could introduce delays between the arrival at Earth of the causal waveform from the emission event of a given photon and the receipt of that photon on Earth. In principle, a local-realist mechanism could therefore exploit information about the wavelength of the incoming quasar photon prior to its detection, in order to engineer the observed violations of the Bell-CHSH inequality. However, any such advanced signal about the incoming quasar photon would need to be correlated with the detector-setting photon itself, and therefore the information carried by the advanced signal must also have originated within the past light cone of the quasar emission event \mathcal{Q}_k , with effective emission time $\eta_k(z)$.

Similar considerations apply to the case in which photons from a given quasar are subject to strong gravitational lensing en route from the quasar to Earth. For example, it is known that light from quasar QSO J0831+5245 (which we used to determine Bob’s settings in pairs 1 and 2) is lensed by a large, intervening mass [\[68, 69\]](#), producing multiple images of the original quasar as seen from Earth. The multiple images arise from different paths that quasar photons take between the lens and Earth. In the case of this particular quasar, it has been estimated that the distinct paths correspond to a difference in arrival times at Earth of as much as 1 day [\[69\]](#).

Given the delay in arrival times, it is possible (in principle) that a local-realist mechanism could receive information from a short-path photon about the wavelength of a long-path photon before the latter arrives at Earth, and exploit that advanced signal to engineer the observed violations of the Bell-CHSH inequality. Much like the case of atmospheric delays at the quasar itself, however, any information of value to the local-realist mechanism would need to have originated within the past light cone of the emission event, with effective emission time, $\eta_k(z)$. Such scenarios are therefore constrained to the same space-time regions described above. See [Figure 4](#).

One scenario in which gravitational lensing *could* affect

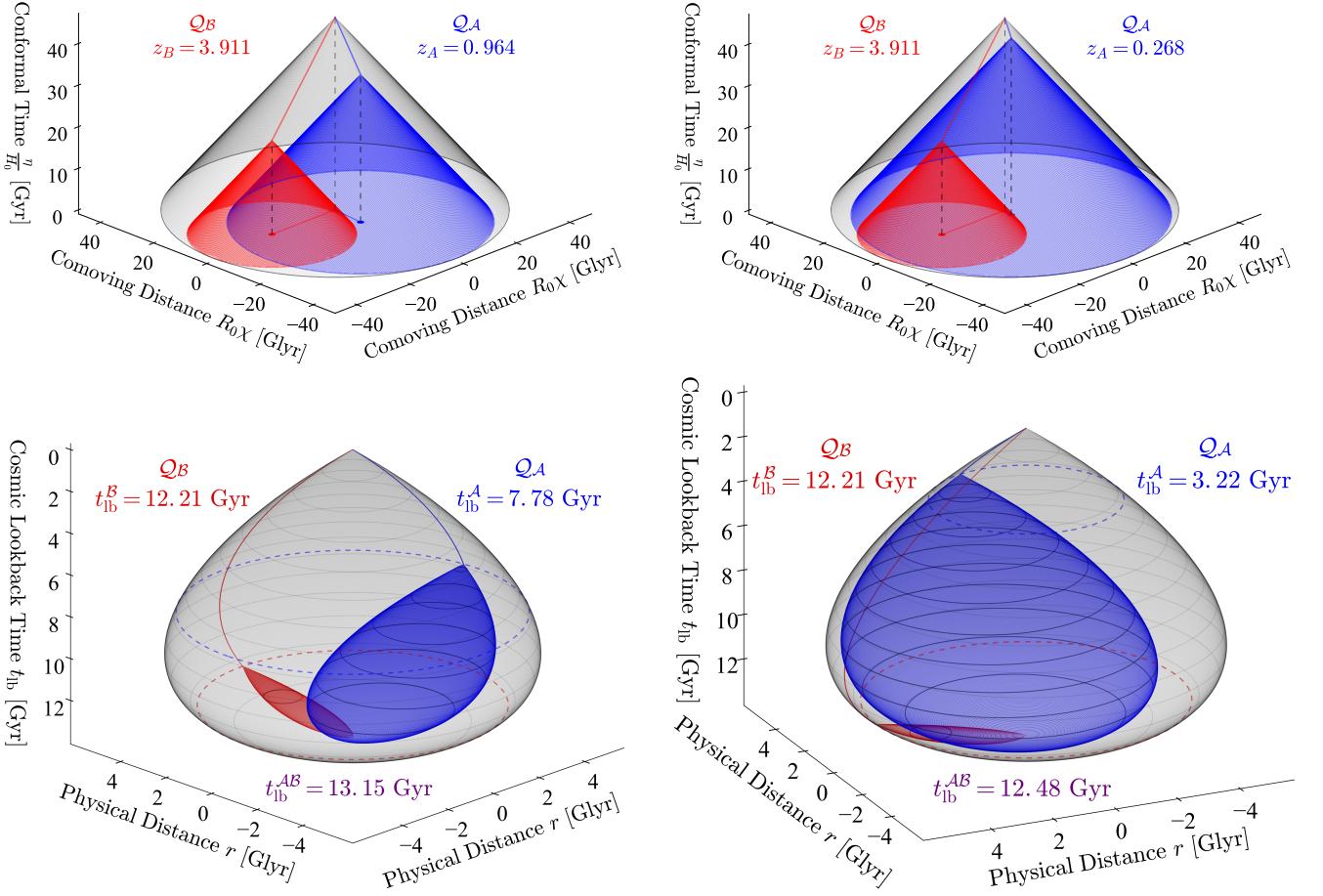


FIG. 2. (2 + 1)D space-time diagrams (with one spatial dimension suppressed) for the quasars in pairs 1 (*left*) and 2 (*right*), showing the past light cones from quasar emission events Q_A (blue), Q_B (red), and the past light cone of our experiment (gray). Top row figures show (rescaled) conformal time η/H_0 and comoving distance $R_0\chi$; bottom row figures show cosmic lookback time t_{lb} and physical distance $r = R_0a(t)\chi$. In conformal coordinates, the big bang singularity is mapped to the surface $\eta = 0$, much as the Earth's poles in a Mercator projection are depicted as lines of comparable length to the Earth's equator. The angle between the red and blue vectors in the $\eta = 0$ plane in the conformal diagrams (top row) represents the quasar pair's angular separation on the sky, given by $\alpha = 83.81^\circ$ and $\alpha = 72.84^\circ$ for pairs 1 and 2, respectively. The bottom-row plots show the lookback times for each emission event (t_{lb}^k) and the lookback time to when their past light cones most recently intersected (t_{lb}^{AB}). Note that the current age of the universe is $t_0 = 13.80$ Gyr. The teardrop shape of the past light cones when plotted in (t, r) coordinates arises from the varying expansion rate of the universe over time, given by the scale factor $a(t)$. The paths traveled by the quasar photons along the surface of the gray past light cone of the experiment are shown in red and blue. In each figure, a local-realistic mechanism could have exploited information from the red and/or blue regions (and their overlap) to engineer the observed violations of the Bell-CHSH inequality. The gray regions outside of the red and blue regions are *excluded* from any such local-realistic explanation. Assuming negligible uncertainties for the reported redshifts, these excluded regions amount to $F_{\text{excl}} = 96.0\%$ (pair 1) and 63.5% (pair 2) of the total space-time 4-volume within the past light cone of our experiment, spanning all of cosmic history since the big bang.

our conclusions would arise if the wavelength of a quasar photon could somehow be measured without altering the photon's wavelength or trajectory. (No such measurement is possible according to quantum mechanics, but a local-realistic mechanism, by design, is meant to be distinct from quantum mechanics.) If an “in-flight” measurement were possible, then a local-realistic mechanism could potentially measure the wavelength of a quasar photon as it approaches the gravitational lens, and arrange for information about that photon's wavelength to arrive at Earth

via a short path from the lens, rather than a long path. In such a scenario, the most recent time by which the local-realistic mechanism would need to have acted would be bounded by the time the quasar photons reach the lens, which is more recent than the emission-time from the quasar. (We do not consider a scenario in which the local-realistic mechanism could *alter* the wavelength of the quasar photon without changing its trajectory; any such mechanism would violate the conservation of energy and momentum.)

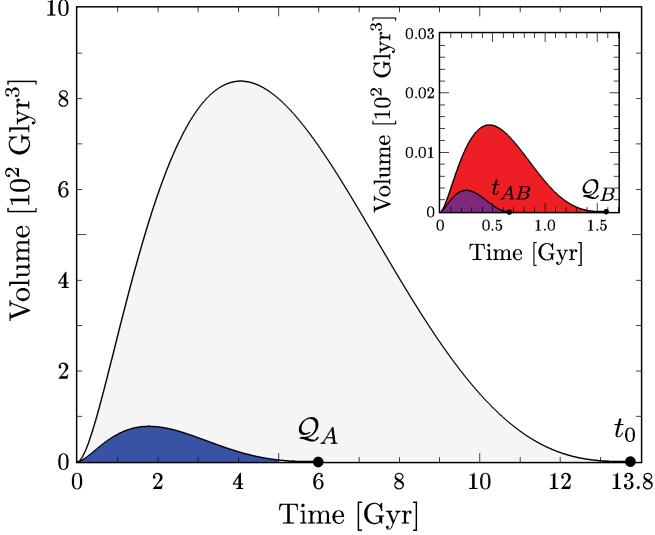


FIG. 3. Spatial volumes as a function of cosmic time t for the relevant past light cones for pair 1. The gray area shows the full past light cone of our experiment, and the blue area shows the past light cone of the emission event Q_A at $t_A = 6.02$ Gyr ($t_{lb}^A = 7.78$ Gyr ago). The spacetime volume of this region is only 4.0% of the full past light cone. The insert shows emission event Q_B , at $t_B = 1.59$ Gyr ($t_{lb}^B = 12.21$ Gyr ago). The space-time volume of its past light cone (red) is only 0.020% of the full past light cone. The past light cones of Q_A and Q_B last intersected at $t_{AB} = 0.65$ Gyr ($t_{lb}^{AB} = 13.15$ Gyr ago), and the space-time volume of their intersection (purple) is only 0.0023% of the total.

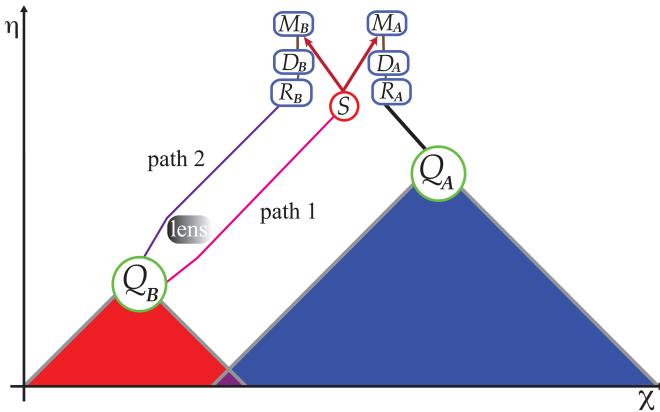


FIG. 4. A scenario in which light from quasar emission event Q_B is lensed by a large, intervening mass between the quasar and Earth. The lens can produce multiple images of the quasar as seen from Earth. In the scenario shown, photons following path 1 arrive at Earth before photons that follow path 2. However, to be valuable to a local-realist mechanism, the information that arrives along path 1 must be correlated with information that originates within the past light cone of the effective emission event Q_B (red or purple regions).

In the case of the lensed quasar in our study, the intervening lens has an estimated redshift $z_{lens} \simeq 3$ [68]. This corresponds to a lookback time from Earth of $t_{lb}(z_{lens}) \simeq 11.6$ Gyr ago, compared to the quasar emission time $t_{lb}^B = 12.21$ Gyr ago. The time at which photons arrive at the lens is considerably earlier than the emission times from either of the quasars with which this quasar was paired ($t_{lb}^A = 7.78$ Gyr for pair 1 and $t_{lb}^A = 3.22$ Gyr for pair 2), so even such an in-flight measurement scenario would have no impact on our overall conclusions. Likewise, because the fraction of the relevant space-time 4-volume is dominated by the volume of the past light cone of the more recent emission event, re-calculating F_{excl} using $z_{lens} \simeq 3$ rather than $z_B = 3.911$ yields virtually no change compared to the values computed above: $F_{\text{excl},\text{lens}} = 0.960$ for pair 1, and $F_{\text{excl},\text{lens}} = 0.635$ for pair 2.

Photons from distant quasars can be affected in other ways between emission and detection, beyond gravitational lensing. In particular, the intergalactic medium can affect the quasar spectra observed on Earth. Quasar sources typically have strong emission at the Lyman- α line, which, in the laboratory frame, corresponds to $\lambda_{\text{emit}}^\alpha = 121.6$ nm (deep in the ultraviolet). However, en route, photons from high-redshift quasars encounter clouds of neutral hydrogen gas, which preferentially absorb photons at the Lyman- α wavelength — for quasar photons that have been redshifted during their travel to 121.6 nm by the time they encounter the gas cloud. Photons from very distant quasars can encounter multiple gas clouds en route to Earth, resulting in a dense “Lyman- α forest” of absorption lines at wavelengths shorter than $\lambda_{\text{obs}}^\alpha = (1+z)\lambda_{\text{emit}}^\alpha$, where z is the redshift of the original quasar emission [70, 71].

In our experiment, Alice’s receiving station observed quasars at redshifts $z_A = 0.964$ (pair 1) and $z_A = 0.268$ (pair 2). Hence the Lyman- α forest affected quasar photons that would have been received with wavelengths $\lambda_{\text{obs}} < 238.8$ nm (pair 1) and $\lambda_{\text{obs}} < 154.2$ nm (pair 2). But, as discussed further below, the detectors of our cosmic random number generators (CRNGs) were largely insensitive to $\lambda_{\text{obs}} < 400$ nm. Hence selective absorption by the Lyman- α forest would have had no observable effect for either of the “quasar A” sources in our experiment. This also means that effects from transmission through the intergalactic medium could not have introduced correlations between the detected photons from quasars A and B, because any effects on the quasar-A photons would have fallen outside the sensitivity range of our detectors.

On the other hand, at Bob’s receiving station we used a more distant quasar, with $z_B = 3.911$. The photons from this distant quasar certainly could have been affected by selective absorption within the Lyman- α forest, for observed wavelengths $\lambda < \lambda_{\text{obs}}^\alpha = (1+z_B)\lambda_{\text{emit}}^\alpha = 597.2$ nm. Our CRNGs used dichroic filters to distinguish

“red” from “blue” astronomical photons on either side of $\lambda_{\text{filter}} = 630 \text{ nm}$. Therefore the entire range of photons from quasar B that could have been affected by the Lyman- α forest falls within the “blue” channel. Akin to the gravitational-lensing scenario described above, one may imagine some local-realist “conspiracy” that used the Lyman- α forest to purposefully alter the spectra that would be received at Bob’s station (by suppressing “blue” photons), and/or that could have “alerted” other elements of our experimental apparatus with a preview of the patterns in the upcoming sequence of “red” and “blue” detections at Bob’s station.

However, given our detectors’ sensitivity for $\lambda \geq 400 \text{ nm}$, the *closest* gas cloud that could have affected the observable portion of the spectrum from quasar B would be at some redshift z_{cloud} such that $400 \text{ nm} = (1 + z_{\text{cloud}})\lambda_{\text{emit}}^{\alpha}$, or $z_{\text{cloud}} = 2.29$, corresponding to a lookback time at which the quasar photons encountered that cloud of $t_{\text{lb}}^{\text{cloud}} = 10.93 \text{ Gyr}$ ago. This falls considerably longer ago than the lookback times to the emission events from quasar A : $t_{\text{lb}}^A = 7.78 \text{ Gyr}$ ago for pair 1, $t_{\text{lb}}^A = 3.22 \text{ Gyr}$ ago for pair 2. Therefore, any “conspiracy” that might have occurred as recently as $t_{\text{lb}}^{\text{cloud}} = 10.93 \text{ Gyr}$ ago is consistent with our overall conclusion: we have constrained any such local-realist mechanisms to have occurred no more recently than t_{lb}^A . Likewise, we may calculate the excluded space-time volume fraction under the assumption that some “conspiracy” occurred at a gas cloud at $z_{\text{cloud}} = 2.29$, by substituting $z_B \rightarrow z_{\text{cloud}}$. Repeating the calculation as above, we then find $F_{\text{excl}} = 0.958$ rather than 0.960 for pair 1, whereas for pair 2, we find $F_{\text{excl}} = 0.635$, unchanged from our original calculation.

Quasar Selection

In anticipation of our limited observing opportunities at the observatory, we searched for quasar pairs whose space-time arrangement would exclude the largest fraction of the 4-volume of the past light cone of our experiment, while maintaining a sufficiently high signal-to-noise ratio to yield a statistically significant result, given constraints on telescope time. (Following Refs. [38, 40, 57], we discuss requirements for signal-to-noise below.) We started with a database of the $\sim 62\,000$ quasars from the Simbad database that had a Sloan r'-band magnitude brighter than 19. We cross-referenced this to the SDSS DR14 database, taking the most conservative of their redshift estimates (or other redshifts from Simbad and the literature, where SDSS redshifts were not available). We considered only quasars that were the brightest for their distance within each 5° by 5° patch of sky. This left $\sim 4\,000$ quasars. For each pair of these, and for every minute of allotted telescope time, we simulated the number of quasar and skyglow photons that each telescope’s detectors would record when pointed at the relevant in-

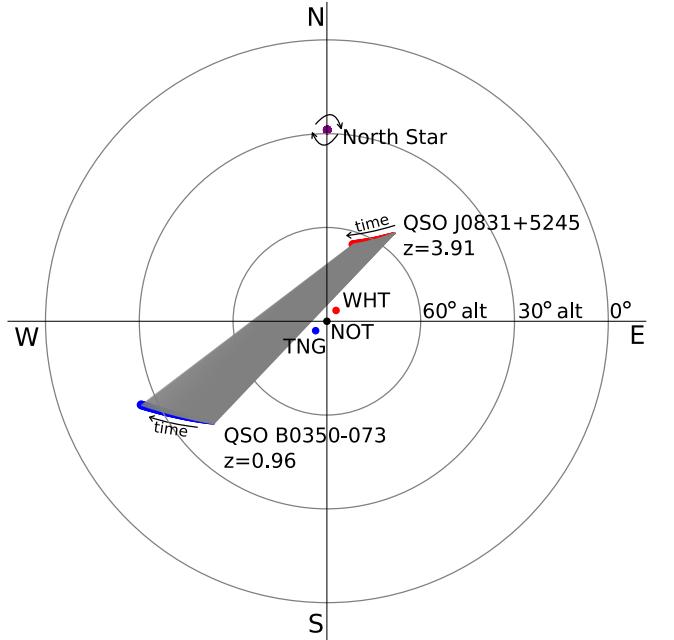


FIG. 5. Overhead view of the orientation of quasars in the sky with respect to the source at NOT and the receivers at TNG and WHT on 2018-01-11 between 00:00 and 01:30 UTC, ending with the pair 1 measurement.

stantaneous elevation angle $> 25^\circ$ and observed through the associated airmass. Figure 5 shows the path on the sky that the quasars of pair 1 took during a period of 1.5 hours on one of our observing nights; our experiment with pair 1 was conducted near the end of that window.

To estimate the rate at which entangled photon coincidences would accumulate, we required that the relevant signal-to-noise threshold be exceeded in each of the red/blue detector-setting ports for both quasars, and that detectors be triggered while both quasars were in causal alignment with respect to the experimental stations such that $\tau_{\text{valid}}^k(t) > 0$ for both sides $k = A, B$. This ensured that we only included entangled-photon coincidences while closing the locality loophole and ensuring the signal-to-noise was sufficiently high.

For a given experimental visibility, entangled-photon coincidence rate, and quasar-photon rate, we estimated the statistical significance we could achieve during the time window while all these conditions were met. Then we chose the highest-redshift pairs with the largest observable angular separations that our simulations predicted could yield significant results in the time allotted.

For the best of these candidates, we further required each object to have published spectra and verified redshifts (either SDSS DR14 or elsewhere in the literature). We manually vetted these to ensure they were legitimate quasar spectra and accurately estimated redshifts, and not, for example, misidentified stellar spectra of $r' \sim 13$ -19 magnitude stars (which the SDSS algorithms some-

times misclassified as extragalactic objects with redshifts $z > 0.1$ [52]). These cases were non-negligible as contaminants to the subset of the SDSS DR14 database that our software preferentially selected, so the manual checks were required before choosing final target quasars to observe. For each of the four initially scheduled time slots on the telescopes, we performed this procedure to select $\sim 10\text{--}20$ vetted targets, yielding several possible pairs that would be optimal at the beginning of the scheduled observation window. The best pairs balanced the trade-off between the largest excluded cosmological 4-volume and the highest statistical significance that could be expected to accumulate within the relevant time window.

Although we were originally scheduled for observation windows over the course of four consecutive nights, due to bad weather and technical problems, we were only able to conduct experiments during the last of our scheduled sessions (early in the morning of 11 January 2018).

Experimental Details

We used Type-0 spontaneous parametric down conversion in a 30 mm periodically poled KTiOPO₄ (ppKTP) crystal placed in a Sagnac-interferometer loop to produce entangled photon pairs. The crystal was bi-directionally pumped by a grating-stabilized 405 nm laser to produce pairs of horizontally polarized down-converted photons at wavelengths of 773.6 nm and 850 nm with a spectral bandwidth of ≈ 2.5 nm FWHM. The entangled state was rotated near the maximally entangled Bell state $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ with fiber polarization controller paddles. The relative phase was adjusted by the polarization of the pump beam by optimizing the Bell test visibility. We locally measured heralding efficiencies of 31% and 41%, which differed by about 1% before and after the experiment, confirming that these parameters remained stable over the duration of the measurement. For pair 1 (pair 2) the duty cycle of Alice's and Bob's measurements - i.e. the temporal sum of used valid setting intervals divided by the total measurement time per run - were 0.32% (1.17%) and 1.62% (0.96%), respectively, resulting in a duty-cycle for valid coincidence detections between Alice and Bob of 5.2×10^{-5} (1.1×10^{-4}).

Dichroic Selection of CRNGs

As emphasized in our previous work [38, 40], it is critical to select appropriate dichroic filters for generating astronomical random bits on the basis of their wavelength. The filters should be chosen to minimize the predictability of the random bits generated. This means minimizing cross-talk between the “red” and “blue” detection channels, and choosing the infrared cutoff of our red band to

be opaque to skyglow, which at the Roque de los Muchachos observatory increases rapidly with increasing wavelength over the range $\sim 700 - 1000$ nm [72] due to transitions in the rovibrational states of OH radicals which are abundant in the upper atmosphere [73].

We employed a system of four dichroic elements which define our detection bands as shown in Figure 7. First, a longpass dichroic beamsplitter (BS) is used to reflect incoming light with $\lambda \lesssim 635$ nm and send it towards our “blue” detector. After the dichroic beamsplitter, we place additional shortpass (SP1: $\lambda < 620$ nm) and long-pass filters (LP1: $\lambda > 637$ nm) in the blue and red output ports to reject photons near the transition wavelength, which may go either way at the dichroic beamsplitter. This way, the fraction of red photons detected in the blue arm and vice versa is negligible ($f_w < 2 \times 10^{-5}$). This represents a significant improvement over our previous experiment, in which the wrong-way fractions for each detector were $f_w \sim \mathcal{O}(10^{-2})$ [38]. Finally, after the longpass filter in the transmitted (red) output port, we define the long-wavelength cutoff of our red band with a shortpass filter (SP2) at ≈ 745 nm, a wavelength chosen to optimize the trade-off between rejection of infrared skyglow and transmission of quasar photons.

In order to make the selection of the dichroic elements BS, SP1, LP1, and SP2, we began with a hand-prepared list of available filtersets (BS, SP1, LP1) whose wavelengths were compatible as well as a list of available filters SP2. Then, for every combination of (BS, SP1, LP1) and SP2, we computed the signal-to-noise ratio in the red and blue channels up to an unknown overall constant which varies from quasar to quasar. We chose the final filter combination that yielded a strong signal-to-noise ratio for all three quasars observed. For every observation except that of QSO B0422+004, our first-principles computation of the red-blue imbalance agreed with the measured count rates to within 11%. For the violently-variable BL Lac object QSO B0422+004, our observation was redder than predicted by our model by a factor of ≈ 1.6 . This is consistent with recent photometric observations, which report a brightening of QSO B0422+004 in the V band by a factor of 2.5 between December 9, 2014, and December 29, 2015 [74]. Similar dramatic variations in brightness were also observed in the 1.2 – 2.2 micron J, H, and K bands between February 2013 and January 2015 [75–77].

The optical efficiency of our CRNGs is shown in Table III. It varies from quasar to quasar due to their diverse spectral shape and different observing altitudes. For all computations, we employ a spectral model, formulated in Ref. [38], which takes as input a quasar spectrum (counts per second per wavelength), its observation altitude, and choices for the dichroic elements BS, SP1, LP1, and LP2. Quasar spectra are corrected for atmospheric extinction with tabulated values for atmospheric reddening at the Roque de los Muchachos [78]. Both quasar and skyglow spectra are weighted by the transmission curves of the

Path	% Transmit (red)	% Reflect(blue)
Through atmosphere	95-96	86-89
Through all optics	38-39	20-30
Detector quantum efficiency	75-76	30-44

TABLE III. The probability that a quasar photon in our transmit/reflect wavelength band is transmitted through the atmosphere, through the telescope/CRNG collection optics, and the probability of registering as an electronic pulse at the APD.

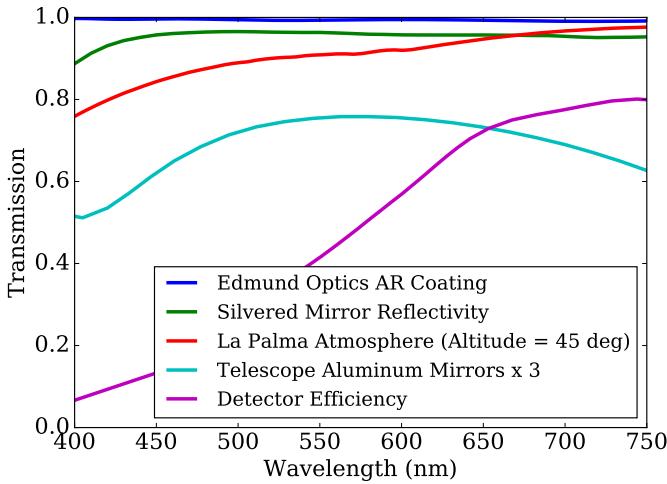


FIG. 6. Transmission spectra of optical elements in our Cosmic RNGs. Both the WHT and TNG have three aluminum mirrors, whose cumulative transmission is plotted here.

telescopes' aluminum mirror coatings, our lenses' anti-reflection coatings, and our ID120 detectors' quantum efficiency curve. These constituent curves are plotted in Figure 6, and the weighted spectra are plotted in Figure 7.

Data Analysis

In this section we analyze the data from the two experimental runs, which were each conducted early in the morning of 11 January 2018. We make the assumptions of fair sampling and fair coincidences [12]. Thus, for testing local realism, all data can be postselected to coincidence events between Alice's and Bob's measurement stations. These coincidences were identified using a time window of 2.66 ns.

As in Ref. [38], we denote by N_{ij}^{AB} the number of coincidences in which Alice had outcome $A \in \{+, -\}$ under setting a_i (with $i = 1, 2$) and Bob had outcome $B \in \{+, -\}$ under setting b_j (with $j = 1, 2$). Then the number of all coincidences for settings combination $a_i b_j$

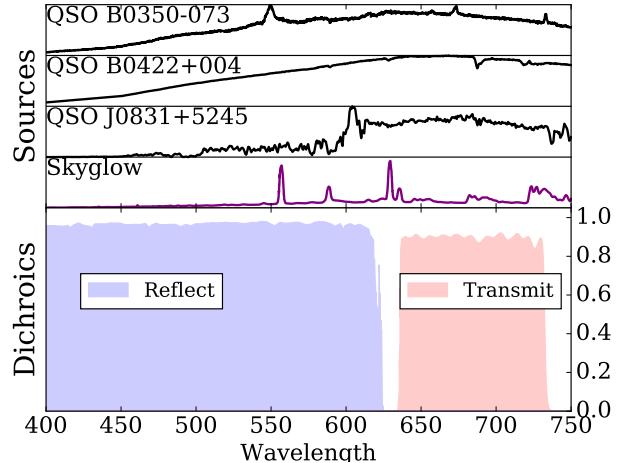


FIG. 7. Top: The number distribution of photons from each source, weighted by the transmission of optical elements and the detector efficiency. The quasar spectra are reddened by Rayleigh scattering. Bottom: Our system of four dichroic elements splits incoming photons cleanly into a reflect (blue) channel and a transmit (red) channel. The red channel is shortpassed to reject a large fraction of Meinel rovibrational emissions which grow rapidly at wavelengths longer than ~ 700 nm. Quasar spectra were taken from: QSO J0831+5245 [79], QSO B0350-073 (SDSS DR14, [52]), and QSO B0422+004 [80]. The skyglow spectrum is from Ref. [72].

is given by

$$N_{ij} \equiv \sum_{A,B=+,-} N_{ij}^{AB}, \quad (28)$$

and the total number of all recorded coincidences is

$$N \equiv \sum_{i,j=1,2} N_{ij}. \quad (29)$$

A point estimate of the joint setting probabilities is given by

$$q_{ij} \equiv p(a_i b_j) = \frac{N_{ij}}{N}. \quad (30)$$

We may then test whether the probabilities q_{ij} can be factorized, that is, whether they can be written (approximately) as

$$p_{ij} \equiv p(a_i)p(b_j), \quad (31)$$

where

$$\begin{aligned} p(a_i) &\equiv \frac{N_{i1} + N_{i2}}{N} = q_{i1} + q_{i2}, \\ p(b_j) &\equiv \frac{N_{1j} + N_{2j}}{N} = q_{1j} + q_{2j}. \end{aligned} \quad (32)$$

We may also estimate the conditional probabilities for correlated outcomes in which both Alice and Bob observe

the same result:

$$p(A = B|a_i b_j) = \frac{N_{ij}^{++} + N_{ij}^{--}}{N_{ij}}. \quad (33)$$

If we define the quantity

$$\begin{aligned} C &\equiv -p(A = B|a_1 b_1) - p(A = B|a_1 b_2) \\ &\quad - p(A = B|a_2 b_1) + p(A = B|a_2 b_2), \end{aligned} \quad (34)$$

then the Bell-CHSH inequality for local-realist models [6] takes the form $C \leq 0$ (if one neglects the “freedom-of-choice” loophole [57]). We may also construct the correlation functions

$$E_{ij} \equiv 2p(A = B|a_i b_j) - 1, \quad (35)$$

in terms of which we may construct the quantity

$$S \equiv |E_{11} + E_{12} + E_{21} - E_{22}|. \quad (36)$$

The Bell-CHSH inequality may then be written $S \leq 2$. The quantities C and S are related as $S = 2|C + 1|$.

The measured coincidence counts N_{ij}^{AB} for pair 1 are

	$ij \setminus AB$	++	+-	-+	--
Pair 1 :	11	145	956	1229	291
	12	487	1618	2417	370
	21	440	1514	1399	206
	22	3229	418	593	2321

(37)

For pair 1, we have $N = 17,633$ total coincidence counts. For pair 2, we find

	$ij \setminus AB$	++	+-	-+	--
Pair 2 :	11	71	1007	1102	168
	12	654	1394	1975	494
	21	315	771	702	186
	22	1975	108	165	1333

(38)

For pair 2, we have $N = 12,420$ total coincidence counts.

Using the measured coincidence counts in Eqs. (38) and (37), we find the values for the Bell-CHSH parameters C and S of Eqs. (34) and (36) as shown in Table IV. There we also show the visibility fraction, defined as

$$\mathcal{V} \equiv \frac{S}{S_{\max}^{QM}}, \quad (39)$$

where $S_{\max}^{QM} = 2\sqrt{2}$, the Tsirelson bound [56], is the maximum value that the quantity S may attain according to quantum mechanics.

Statistical Independence of Settings Choices

We first consider whether the joint settings frequencies q_{ij} for the data from pairs 1 and 2 may be factorized. For

Pair	C	S	\mathcal{V}
1	0.3229	2.6457	0.935
2	0.3140	2.6281	0.929

TABLE IV. Values for the Bell-CHSH parameters and the visibility fraction for each experimental run.

Pair	Side	ID	n_r	s_r	n_b	s_b
1	\mathcal{A}	QSO B0350-073	288	2094	350	2774
1	\mathcal{B}	QSO J0831+5245	358	9320	408	5064
2	\mathcal{A}	QSO B0422+004	640	3970	684	3224
2	\mathcal{B}	QSO J0831+5245	389	10908	347	6213

TABLE V. Average signal and background count rates from the quasars measured during our experiment. All rates are reported in counts per second. While the fluctuations in signal rates varied significantly due to atmospheric seeing, fluctuations in the measured noise rates were consistent with \sqrt{N} error over the course of a 5-minute noise measurement.

Pair	$p(a_1)$	$p(a_2)$	$p(b_1)$	$p(b_2)$
1	0.4261	0.5739	0.3505	0.6495
2	0.5527	0.4473	0.3480	0.6520

TABLE VI. Values for the single-side settings probabilities.

Pair	quantity	11	12	21	22
1	q_{ij}	0.1486	0.2774	0.2018	0.3721
	p_{ij}	0.1493	0.2767	0.2011	0.3728
2	q_{ij}	0.1890	0.3637	0.1589	0.2883
	p_{ij}	0.1923	0.3604	0.1556	0.2916

TABLE VII. Values for the joint settings frequencies q_{ij} and inferred joint probabilities p_{ij} .

the individual settings probabilities $p(a_i)$ and $p(b_j)$ defined in Eq. (32), we find the values shown in Table VI. The joint frequencies q_{ij} and the inferred joint probabilities p_{ij} are shown in Table VII. Note that for each dataset, we find $\sum_{ij} q_{ij} = \sum_{ij} p_{ij} = 1.00$, as required.

We expect that for each dataset, $q_{ij} \simeq p_{ij}$. If we did not find $q_{ij} \simeq p_{ij}$, then (in principle) there could have been some common cause that established correlations among the various setting choices; the choices ij would not be independent. To test for any such violation of independence among the joint settings ij , we may conduct a Pearson’s χ^2 test by calculating the statistic

$$\chi^2 = N \sum_{ij} \frac{(q_{ij} - p_{ij})^2}{p_{ij}}, \quad (40)$$

and then computing the p -value corresponding to this value of χ^2 [38]. For pair 1, we find $\chi^2 = 0.1504$, which implies that (under the assumption of independent setting choices) there is a purely statistical chance of $p = 0.698$ that the observed frequencies q_{ij} (or data that deviate even more from the inferred p_{ij}) would be obtained. For pair 2, we find $\chi^2 = 2.405$, corresponding to a statistical chance $p = 0.121$. Given that each of

these p values is larger than typical thresholds (such as $p < 0.05$), we conclude that there is no strong statistical support to refute the hypothesis that the settings ij were independent of each other, for both pairs 1 and 2.

Testing for Violations of No-Signaling

Another statistical test to consider is whether our data are consistent with the assumption of no-signaling. To be consistent with the principle of no-signaling, the local outcome probabilities must not depend on the setting of the distant detector, which is to say [38]:

$$\begin{aligned} p(A = +|a_i b_j) &= p(A = +|a_i b_{j'}), \\ p(B = +|a_i b_j) &= p(B = +|a_{i'} b_j). \end{aligned} \quad (41)$$

The analogous expressions for the ‘–’ outcomes follow from $p(A = -|a_i b_j) = 1 - p(A = +|a_i b_j)$. Point estimates for these conditional probabilities may be estimated from the measured coincidence counts. We define

$$\begin{aligned} N_{ij}^{A=+} &\equiv N_{ij}^{++} + N_{ij}^{+-}, \\ p(A = +|a_i b_j) &\equiv \frac{N_{ij}^{A=+}}{N_{ij}^{A=+} + N_{ij}^{A=-}}, \end{aligned} \quad (42)$$

with analogous expressions for measurement outcomes $A = -$ and $B = \pm$. Point estimates for the various probabilities are shown in Table VIII.

We may assess whether these conditional probabilities show any statistical evidence of the violation of no-signaling by performing pooled two-proportion z -tests for each relevant pair, such as $p(A = +|a_1 b_1)$ and $p(A = +|a_1 b_2)$ [38], which yields the p -values listed in Table IX. As can be seen in Table IX, nearly all of the single-sided outcomes from pairs 1 and 2 are consistent with the hypothesis of no-signaling. The only anomalously low p -value is for $p(A = +|a_2 b_j)$ for pair 2, which yields $p = 0.023 < 0.05$. However, for 8 independent tests, the probability that one shows a p -value at least as bad as 0.023 is $1 - (1 - 0.023)^8 = 0.170$, so the overall suite of tests is consistent with the hypothesis of no-signaling.

Predictability of Settings

In this section we consider imperfections in the experiment that can lead to an excess predictability [57] of the setting choices. Such excess predictability ϵ quantifies the fraction of runs in which one could predict a specific setting better than would be inferred from the overall bias of the setting choices, given all possible knowledge about the setting-generation process that could be available at the emission event for the entangled particles and thus at the distant measurement events. Loosely speaking, ϵ quantifies the fraction of runs in which the assumptions of locality and freedom-of-choice fail to hold.

In general, we consider a given trial “corrupt” if any one of three possibilities occurred at either Alice’s or Bob’s detector: it (1) involved a noise photon rather than a cosmic photon (where noise photons could arise from either detector dark counts or skyglow), or (2) involved a cosmic photon that was misdirected by a dichroic filter, or (3) involved a cosmic photon that was previewed by the local-realist model for the purpose of considering a dichroic-mirror error, but was passed over because the cosmic photon already had the desired color.

Given these considerations, we parameterize the excess predictabilities for each detector setting as [38]

$$\begin{aligned} \epsilon_{a_i} &= \frac{1}{r_{a_i}} \left(n_{a_i} + f_{i' \rightarrow i}^{(A)} s_i^{(A)} \right), \\ \epsilon_{b_j} &= \frac{1}{r_{b_j}} \left(n_{b_j} + f_{j' \rightarrow j}^{(B)} s_j^{(B)} \right), \end{aligned} \quad (43)$$

where the rate of detected photons at Alice’s detector is given by

$$r_{a_i} = (1 - f_{i \rightarrow i'}^{(A)}) s_i^{(A)} + f_{i' \rightarrow i}^{(A)} s_{i'}^{(A)} + n_{a_i}, \quad (44)$$

with a comparable expression for Bob’s detector, r_{b_j} . Here n_{a_i} is the measured rate of noise photons (dark counts and skyglow), which may be quantified by pointing the receiving telescope marginally away from its quasar target; $f_{i \rightarrow i'}^{(A)}$ is the fraction of cosmic photons whose color (if correctly identified) would have led to setting choice a_i , but which were misdirected by the dichroic mirror toward the wrong port, yielding setting $a_{i'}$; and $s_i^{(A)}$ is the detected rate of cosmic photons which have a color that (when correctly identified) yield setting choice a_i . Because the fractions of misdirected cosmic photons were so small (as noted above), with $f_w < 2 \times 10^{-5}$ for both red-to-blue and blue-to-red at Alice’s and Bob’s stations, we may neglect the effects from nonzero f_w . This simplifies our analysis compared to Ref. [38], in which the various wrong-way fractions were as large as $f_w = \mathcal{O}(10^{-2})$. For pairs 1 and 2, we find the signal rates and noise rate shown in Table V.

As in Refs. [38, 57], we assume that a local-realist model could exploit each “corrupt” trial so as to produce measurement outcomes that exceed the usual Bell-CHSH inequalities of $C \leq 0$ or $S \leq 2$. In particular, we make the conservative assumption that predictable setting events do not occur simultaneously at both detectors, so that the total fraction of corrupt joint settings is simply the sum of the corrupt settings on each side:

$$\epsilon_{ij} \equiv \epsilon_{a_i} + \epsilon_{b_j}. \quad (45)$$

If any value calculated as in Eq. (45) exceeds 1, then the corresponding ϵ_{ij} is set to 1. We further assume that the local-realist model can maximally exploit each corrupted trial, so that the maximum value of C that

Pair	$p(A = + 11)$	$p(A = + 12)$	$p(A = + 21)$	$p(A = + 22)$	$p(B = + 11)$	$p(B = + 21)$	$p(B = + 12)$	$p(B = + 22)$
1	0.4201	0.4303	0.5490	0.5559	0.5242	0.5167	0.5936	0.5825
2	0.4591	0.4534	0.5502	0.5817	0.4996	0.5152	0.5820	0.5976

TABLE VIII. Point estimates for the conditional probabilities for measurement outcomes at Alice’s and Bob’s detectors, with $p(k = +|ij) \equiv p(k = +|a_i b_j)$.

Pair	p for $p(A = + a_1 b_j)$	p for $p(A = + a_2 b_j)$	p for $p(B = + a_i b_1)$	p for $p(B = + a_i b_2)$
1	0.395	0.503	0.562	0.234
2	0.653	0.023	0.308	0.156

TABLE IX. Probabilities p that the observed data (or worse) are obtained under the null hypothesis of no-signaling.

the local-realist model could attain by exploiting excess predictabilities of detector settings would be [57]

$$C \leq \epsilon, \quad (46)$$

where

$$\epsilon \equiv \max_{ij} \epsilon_{ij} = \max_i \epsilon_{a_i} + \max_j \epsilon_{b_j}. \quad (47)$$

Following Refs. [38, 40, 57], to ensure that a sufficient number of genuine quasar photons are detected compared to skyglow, dark current, and misdirected photons, the inequality of Eq. (46) places a constraint on the visibility fraction, such that we require

$$\epsilon < \mathcal{V}\sqrt{2} - 1 \quad (48)$$

for all times during the experiment, where the visibility fraction \mathcal{V} is defined in Eq. (39). Eq. (48), in turn, sets a constraint on the signal-to-noise ratio on each of the four settings ports. Given the values of \mathcal{V} shown in Table IV, the constraint of Eq. (48) becomes $\epsilon < 0.322$ (pair 1) and $\epsilon < 0.314$ (pair 2). As shown in Table XI, all values for the excess predictabilities ϵ_{ij} easily satisfy this constraint, for both pairs of quasars.

Meanwhile, each of the corrupt fractions ϵ_{a_i} and ϵ_{b_j} has some statistical uncertainty, $\sigma_{\epsilon_{a_i}}$ and $\sigma_{\epsilon_{b_j}}$, solely due to fluctuating sky noise during the measurement run. Since the total number of runs is recorded, the only unknown is in the total number of runs that were conducted with noise photons in our CRNGs over our measurement period. This is estimated by measuring the average noise count rates before and after the measurement period at each telescope and using the higher of the two count rates to estimate the total number of corrupt runs. In each of our noise measurements we find that the estimated unbiased mean-square error in the count rates is consistent with Poisson noise.

We temporarily drop the labels for Alice and Bob, and assume that the rates r_i and n_i are independent (which follows from our assumption of fair sampling for all detected photons). Then we find

$$\sigma_{\epsilon_i}^2 = \sigma_{n_i}^2 / \bar{r}_i^2 + \mathcal{O}(f_w). \quad (49)$$

If we further assume that Alice’s and Bob’s predictabilities are independent, then we find

$$\sigma_{\epsilon_{ij}} = \sqrt{\sigma_{\epsilon_{a_i}}^2 + \sigma_{\epsilon_{b_j}}^2}, \quad (50)$$

with an estimated uncertainty on ϵ , as defined in Eq. (47), of

$$\sigma_\epsilon = \sqrt{\sigma_{\max_i \epsilon_{a_i}}^2 + \sigma_{\max_j \epsilon_{b_j}}^2}. \quad (51)$$

Values of $\epsilon_{a_i} \pm \sigma_{\epsilon_{a_i}}$ and $\epsilon_{b_j} \pm \sigma_{\epsilon_{b_j}}$ for pairs 1 and 2 are shown in Table X, and values of the excess predictabilities for joint settings, $\epsilon_{ij} \pm \sigma_{\epsilon_{ij}}$, are shown in Table XI.

Pair	$\epsilon_{a_1} \pm \sigma_{\epsilon_{a_1}}$	$\epsilon_{a_2} \pm \sigma_{\epsilon_{a_2}}$	$\epsilon_{b_1} \pm \sigma_{\epsilon_{b_1}}$	$\epsilon_{b_2} \pm \sigma_{\epsilon_{b_2}}$
1	0.1441 $\pm 1.21 \times 10^{-3}$	0.1334 $\pm 0.88 \times 10^{-3}$	0.0653 $\pm 0.46 \times 10^{-3}$	0.0342 $\pm 0.13 \times 10^{-3}$
2	0.1326 $\pm 0.46 \times 10^{-3}$	0.1679 $\pm 0.54 \times 10^{-3}$	0.0537 $\pm 0.93 \times 10^{-3}$	0.0342 $\pm 0.26 \times 10^{-3}$

TABLE X. Values for the fractions of “corrupt” detector settings at each detector.

Pair	$\epsilon_{11} \pm \sigma_{\epsilon_{11}}$	$\epsilon_{12} \pm \sigma_{\epsilon_{12}}$	$\epsilon_{21} \pm \sigma_{\epsilon_{21}}$	$\epsilon_{22} \pm \sigma_{\epsilon_{22}}$
1	0.2095 $\pm 1.30 \times 10^{-3}$	0.1783 $\pm 1.22 \times 10^{-3}$	0.1987 $\pm 0.99 \times 10^{-3}$	0.1676 $\pm 0.89 \times 10^{-3}$
2	0.1862 $\pm 1.04 \times 10^{-3}$	0.1667 $\pm 0.53 \times 10^{-3}$	0.2216 $\pm 1.07 \times 10^{-3}$	0.2021 $\pm 0.60 \times 10^{-3}$

TABLE XI. Values for the excess predictabilities for various joint detector settings.

Since the exact number of runs is known and recorded, we consider the probabilities ϵ_{ij} to be known (to within some uncertainty $\sigma_{\epsilon_{ij}}$), but the actual number of corrupt trials to be subject to statistical fluctuations. In other words, the occurrence of a corruption in any trial is taken to be an independent random event, which has probability ϵ_{ij} that depends on the settings pair ij . We assume that for “uncorrupt” trials, the local-realist model has no information about what the settings pair will be beyond the joint settings probabilities q_{ij} [38, 57].

In general, the total rates (r_{a_i}, r_{b_j}) and noise rates (n_{a_i}, n_{b_j}) can vary during the course of an observing

period. For example, when observing a given quasar over a substantial period of time, the receiving telescope will collect its light through varying amounts of airmass, as the quasar rises above or moves toward the horizon, thereby affecting the noise rate. In practice, however, our observing periods for both pairs 1 and 2 were brief enough (≤ 17 min) that the measured values for r_{a_i}, r_{b_j} , and n_{a_i}, n_{b_j} did not change substantially; the largest variation among all four detector settings yielded a difference $\Delta\epsilon_{a_1} = 8.6\%$ in the excess predictability. Therefore we adopt the conservative approach of assuming constant values of each ϵ_{a_i} and ϵ_{b_j} during a given experimental run, and use the largest values for each detector setting. Such an approach will (modestly) underestimate the statistical significance of our results.

Statistical Predictability of Random Bits

Throughout our analysis, we assume that the bits within the sequence gathered from a given quasar are independent of each other. That is, we assume that there is no sequence of bits within the bitstream that contains any information about any future bit. If this independence did not hold, then (in principle) a local-realist mechanism would be able to exploit any excess predictability (beyond the natural bias) to engineer a measured violation of Bell's inequality. Although complete independence among the bits within each bitstream can never be rigorously proven, we can calculate bounds for the available information. As we assume fair-sampling for the cosmic photons, we included all detection events, instead of postselecting for those that satisfied the requirements of $\tau_{\text{valid}}^k(t)$ to yield a valid setting. For the calculation of the mutual information $\hat{I}(m)$ we adopt the approach developed in Ref. [40]: We calculate the mutual information between every bit and every sequence of the $m = 17$ preceding bits, correcting the biased estimator for the finite size of our bitstream. The value m has been chosen such that it is strictly larger than $\lfloor \log_2(L) - 7 \rfloor$ for each measurement file, where L is the total number of detection events [81, 82]. The results are presented in Table XII. The calculated values of $\hat{I}(m)$ are 2–3 orders of magnitude smaller than the values of the excess predictabilities ϵ_{a_i} and ϵ_{b_j} that we already incorporate in our data analysis. These small values of $\hat{I}(m)$ therefore make no quantitative impact on our analysis or conclusions.

Statistical Significance of Bell Violation

As discussed in Ref. [38], there exist several different approaches to estimating the statistical significance for Bell experiments. The result of any such statistical analysis is a p -value, that is, a bound for the probability that the null hypothesis—in our case, local realism with

Pair/Side	Quasar	L	$\hat{I}(m)$
1/A	QSO B0350-073	5668580	2.0×10^{-4}
1/B	QSO J0831+5245	9010082	1.6×10^{-4}
2/A	QSO B0422+004	6338028	5.1×10^{-5}
2/B	QSO J0831+5245	13336320	2.7×10^{-4}

TABLE XII. We characterize the statistical predictability of our random bitstreams generated by each quasar by computing the mutual information of each bit with the $m = 17$ bits preceding it. This measure corresponds to the probability of guessing the bit correctly given knowledge of the prior 17 bits. The number of lookback bits m for which our estimator is valid is set by L , the length of the bitstream.

excess predictability of the detector settings ϵ_{ij} , biased detector-setting frequencies q_{ij} , fair sampling, fair coincidences, *and any other additional assumptions*—could have produced the experimentally observed data by a random, statistical variation.

Until recently, it was typical in the literature to make several simplifying assumptions when calculating the p -value for a Bell test, such as that each trial was independent and identically distributed, that the local-realist mechanism could not make any use of “memory” of the settings and outcomes of previous trials, and that excess predictabilities ϵ_{ij} could be neglected. Under those assumptions, one typically applied Poisson statistics for single coincidence counts. As emphasized in Ref. [38], however, such an approach assumes that the measured coincidence counts N_{ij}^{AB} are equal to their expected values, only to contradict that assumption by calculating the probability that the N_{ij}^{AB} could have values differing by several standard deviations from their expected values. More recent, sophisticated analyses do not make such assumptions, and represent significant improvements over previous methods [57, 83–86]. However, even these newer approaches are not optimal for our particular experimental arrangement. For example, they are not optimized for experiments with unequal (biased) settings probabilities, and/or they yield non-tight upper bounds for p that can dramatically underestimate the statistical significance.

For these reasons, in Ref. [38] we presented an *ab initio* calculation of the p -value tailored more specifically to experiments like ours. We follow the same steps here, and refer readers to the more detailed description of the calculation in Ref. [38]. First we construct the quantity

$$W \equiv \sum_{ij} \frac{N_{ij}^{\text{win}}}{q_{ij}(1 - \epsilon_{ij})}, \quad (52)$$

with $N_{ij}^{\text{win}} \equiv [N_{11}^{A \neq B}, N_{12}^{A \neq B}, N_{21}^{A \neq B}, N_{22}^{A=B}]$. If we estimate the conditional probabilities $p(A = B|a_i b_j)$ as in Eq. (33), then $W = (3+C)N$, with C defined in Eq. (34). As described in Ref. [38], to avoid the ambiguity that would arise by the need to assume a specific prior probability distribution for the various detector settings, we

take the actual number N_{ij} of the occurrences of each settings pair $a_i b_j$ as given. The relevant ensemble with respect to which we calculate probabilities is the ensemble of all possible *orders* in which the settings choices could have occurred. The *p*-value we calculate is then the fraction of orderings for which the local-realist mechanism, using its best strategy, could obtain a value of W greater than or equal to the value obtained in the experiment.

In the absence of noise or errors, the local-realist mechanism could specify which outcomes (A, B) will arise for each of the possible settings (a_i, b_j) . The best plans will win for three of the four possible settings pairs, but will lose for one of the possible settings pairs. Therefore a plan may be fully specified by identifying which settings pair will be the loser for a given trial. In the presence of noise and errors, for each time the settings pair is $a_i b_j$, there is a probability ϵ_{ij} that the trial is corrupt. If the trial is corrupt, it automatically registers as a win. If it is not corrupt, then it has a probability P_{ij}^{win} of registering as a win, where we take $P_{ij}^{\text{win}} = p(A = B | a_i b_j)$ for $(ij) = (22)$, and $p(A \neq B | a_i b_j)$ for the other cases. These considerations motivate the form of W in Eq. (52) [38].

The ensemble average of W takes the form [38]

$$\langle W \rangle = N(3 + \bar{\epsilon}), \quad (53)$$

where we have defined

$$\bar{\epsilon} \equiv \sum_{ij} \frac{\epsilon_{ij}}{1 - \epsilon_{ij}}. \quad (54)$$

To calculate the statistical significance, we also must calculate $\sigma_W^2 \equiv \langle W^2 \rangle - \langle W \rangle^2$. As in Ref. [38], we calculate σ_W^{opt} , subject to the condition that the local-realist mechanism may choose the fractions f_{ij} so as to optimize its strategy, where the f_{ij} are defined as the fraction of trials for which the local-realist mechanism chooses settings pair (ij) to be the losing pair. In Ref. [38] we found

$$f_{ij}^{\text{opt}} = \frac{1}{2} - q_{ij} + \frac{N-1}{2N} \left[\frac{\epsilon_{ij}}{1 - \epsilon_{ij}} - \bar{\epsilon} q_{ij} \right]. \quad (55)$$

Each value of f_{ij} must be non-negative. (If any value is negative, we use a second Lagrange multiplier and re-calculate the f_{ij}^{opt} [38].) For pairs 1 and 2, we find all values $f_{ij}^{\text{opt}} > 0$ when calculated as in Eq. (55), so we may use our expression for σ_W^{opt} from Ref. [38], namely

$$\begin{aligned} (\sigma_W^{\text{opt}})^2 &= \frac{N^2}{4(N-1)} \left[\left(\sum_{ij} \frac{1}{q_{ij}} \right) - 4 \right] - N\bar{\epsilon} \\ &\quad + \frac{N}{4} \sum_{ij} \frac{\epsilon_{ij}}{q_{ij}(1 - \epsilon_{ij})} \\ &\quad - \frac{1}{4}(N-1)\bar{\epsilon}^2 + \frac{1}{4} \sum_{ij} \frac{(N - \epsilon_{ij})\epsilon_{ij}}{q_{ij}(1 - \epsilon_{ij})^2}. \end{aligned} \quad (56)$$

We may then calculate the number of standard deviations $\bar{\nu}$ by which the correlations among the measured coincidence counts exceed those that the local-realist mechanism could have engineered by exploiting excess predictabilities:

$$\bar{\nu} = \frac{W - \langle W \rangle}{\sigma_W^{\text{opt}}}. \quad (57)$$

The quantities W , $\langle W \rangle$, and σ_W^{opt} each depend on ϵ_{a_i} and ϵ_{b_j} as well as on the coincidence counts N_{ij}^{AB} , which we take as given. Whereas Eq. (57) takes into account the excess predictabilities, ϵ_{ij} , however, it does not incorporate the uncertainties on the excess predictabilities, $\sigma_{\epsilon_{a_i}}$ and $\sigma_{\epsilon_{b_j}}$. For the next step, we therefore propagate uncertainties on ϵ_{a_i} and ϵ_{b_j} to compute the uncertainty on $\bar{\nu}$, which we denote Δ_{ν} . The uncertainty Δ_{ν} takes the form [38]

$$\begin{aligned} \Delta_{\nu}^2 &= \sum_{a_i} \left(\frac{\sigma_{\epsilon_{a_i}}}{\sigma_W^{\text{opt}}} \right)^2 \left[\sum_j \frac{\mathcal{E}_{ij}}{q_{ij}(1 - \epsilon_{ij})^2} \right]^2 \\ &\quad + \sum_{b_j} \left(\frac{\sigma_{\epsilon_{b_j}}}{\sigma_W^{\text{opt}}} \right)^2 \left[\sum_i \frac{\mathcal{E}_{ij}}{q_{ij}(1 - \epsilon_{ij})^2} \right]^2, \end{aligned} \quad (58)$$

where

$$\mathcal{E}_{ij} \equiv N_{ij}^{\text{win}} - N q_{ij} - \left(\frac{\bar{\nu} N}{2\sigma_W^{\text{opt}}} \right) f_{ij}^{\text{opt}}. \quad (59)$$

The naive estimate of the number of standard deviations, $\bar{\nu}$ in Eq. (57), implicitly assumed $\sigma_{\epsilon_{a_i}}, \sigma_{\epsilon_{b_j}} = 0$, and therefore $\Delta_{\nu} = 0$. If we allow for an uncertainty in ν equal to n times the 1- σ uncertainty in $\bar{\nu}$, then we should calculate the *p*-value using

$$\nu_n \equiv \bar{\nu} - n\Delta_{\nu}. \quad (60)$$

If we choose n so that $n = \nu_n$, then we find

$$\nu_n = \frac{\bar{\nu}}{1 + \Delta_{\nu}}. \quad (61)$$

Assuming a Gaussian distribution for large-sample experiments, we conclude that the conditional probability that the local-realist mechanism could achieve a value of W as large as the observed value W_{obs} , assuming that the true value of $\nu \geq \nu_n$, is given by

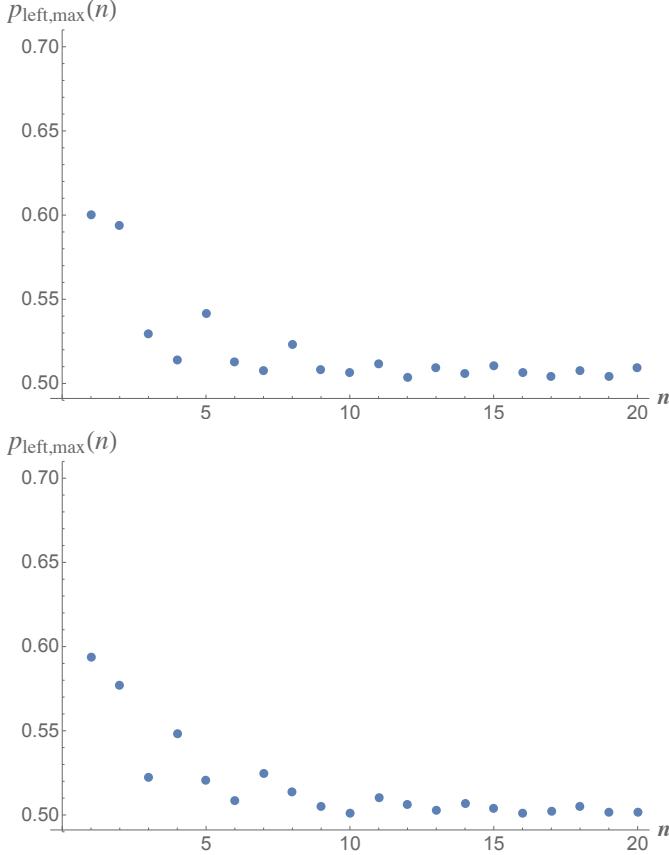
$$p_{\text{cond}} = \frac{1}{2} \text{erfc}(\nu_n/\sqrt{2}). \quad (62)$$

Moreover, if we assume Gaussian statistics for the uncertainty in ν , then there is an equal probability that the true value of ν is less than ν_n , in which case we must conservatively assume that W might exceed W_{obs} . Therefore the *p*-value corresponding to the total probability that $W \geq W_{\text{obs}}$ is bounded by

$$p_{\text{no-m}} = 2p_{\text{cond}}. \quad (63)$$

Pair	W	$\langle W \rangle$	σ_W^{opt}	$\bar{\nu}$	Δ_ν	ν_n	p_{cond}	$p_{\text{no-m}}$	$\nu_{\text{no-m}}$	B	p	ν
1	72 224.1	69 319.1	290.222	10.01	0.0576	9.46	1.48×10^{-21}	2.96×10^{-21}	9.39	0.6001	7.41×10^{-21}	9.29
2	51 110.3	49 268.0	242.745	7.59	0.0395	7.30	1.43×10^{-13}	2.86×10^{-13}	7.21	0.5937	7.03×10^{-13}	7.08

TABLE XIII. Quantities relevant to calculating the statistical significance.

FIG. 8. The quantity $p_{\text{left},\text{max}}(n)$ is the maximum probability that \tilde{W} moves to the left in n trials. Shown here is $p_{\text{left},\text{max}}(n)$ for pair 1 (top) and pair 2 (bottom).

Again assuming Gaussian statistics, we may then calculate

$$\nu_{\text{no-m}} = \sqrt{2} \operatorname{erfc}^{-1}(2p_{\text{no-m}}). \quad (64)$$

The subscript “no-m” stands for “no-memory,” and indicates that these quantities have been calculated without taking into account possible memory effects, which the local-realist mechanism might have been able to exploit. Using the coincidence counts in Eqs. (37) and (38) and the values of various quantities in Tables VII, X, and XI, we find the numerical values relevant to the calculation of $p_{\text{no-m}}$ and $\nu_{\text{no-m}}$ as shown in Table XIII.

The expressions for $\langle W \rangle$ and σ_W^{opt} in Eqs. (53) and (57)

neglect any advantages that the local-realist model could gain by exploiting memory of previous trials. In the last step of our analysis, we incorporate possible effects from such memory, again following closely the discussion in Ref. [38]. We consider the quantity

$$\tilde{W} \equiv W - (3 + \bar{\epsilon})N, \quad (65)$$

with W given in Eq. (52). From Eq. (53), we see that $\langle \tilde{W} \rangle = 0$ when averaged over all N trials. Hence the local-realist mechanism cannot change $\langle \tilde{W} \rangle$, but presumably it *could* affect the standard deviation of \tilde{W} . If we denote by \tilde{W}_0 the value of \tilde{W} obtained in the experiment after all N trials, then the p -value we seek is the probability that the local-realist mechanism could have achieved $\tilde{W} \geq \tilde{W}_0$ by chance.

We again assume Gaussian statistics for experiments with sufficiently large numbers of trials, $N \gg 1$. Then we expect that as long as $\tilde{W}_n \leq \tilde{W}_0$, the best strategy for the local-realist mechanism is to maximize $\sigma_{\tilde{W}}$, where \tilde{W}_n is the value of \tilde{W} after $n < N$ trials. In this way, the local-realist mechanism would require the smallest number of standard deviations to reach its goal. When and if \tilde{W}_n exceeds \tilde{W}_0 , on the other hand, then the best strategy for the local-realist mechanism is to minimize $\sigma_{\tilde{W}}$, so as to minimize the probability that \tilde{W} might backslide to $\tilde{W} < \tilde{W}_0$. The quantity we aim to calculate is therefore p , which is bounded by

$$p \leq \frac{p_{\text{no-m}}}{1 - B}, \quad (66)$$

where $p_{\text{no-m}}$ is given in Eq. (63). The quantity B is defined as

$$B \equiv \max_n \{p_{\text{left},\text{max}}(n)\}, \quad (67)$$

where $p_{\text{left},\text{max}}(n)$ is the maximum probability that the quantity \tilde{W} moves to the left after $n < N$ trials. As in Ref. [38], we compute B numerically using the values of q_{ij} and ϵ_{ij} for pairs 1 and 2. As we had found in Ref. [38], the maximum value of $p_{\text{left},\text{max}}(n)$ occurs for $n = 1$. (See Fig. 8.) For pair 1, we find $B = 0.6001$, which yields $p \leq 7.41 \times 10^{-21}$, corresponding to $\nu = 9.29$. For pair 2, we find $B = 0.5937$, which yields $p \leq 7.03 \times 10^{-13}$, corresponding to $\nu = 7.08$.