

Challenging local realism with human choices

†The BIG Bell Test Collaboration: C. Abellán,¹ A. Acín,^{1,2} A. Alarcón,^{3,4} O. Alibart,⁵ C. K. Andersen,⁶ F. Andreoli,⁷ A. Beckert,⁶ F. A. Beduini,¹ A. Bendersky,⁸ M. Bentivegna,⁷ P. Bierhorst,⁹ D. Burchardt,¹⁰ A. Cabello,¹¹ J. Cariñe,^{3,4,12} S. Carrasco,¹ G. Carvacho,⁷ D. Cavalcanti,¹ R. Chaves,¹³ J. Cortés-Vega,^{3,12} A. Cuevas,⁷ A. Delgado,^{3,12} H. de Riedmatten,^{1,2} C. Eichler,⁶ P. Farrera,¹ J. Fuenzalida,^{3,12,14} M. García-Matos,¹ R. Garthoff,¹⁰ S. Gasparinetti,⁶ T. Gerrits,⁹ F. Ghafari Jouneghani,^{15,16} S. Glancy,⁹ E. S. Gómez,^{3,12} P. González,^{3,12} J.-Y. Guan,^{17,18} J. Handsteiner,¹⁴ J. Heinsoo,⁶ G. Heinze,¹ A. Hirschmann,¹ O. Jiménez,¹ F. Kaiser,⁵ E. Knill,⁹ L. T. Knoll,^{19,20} S. Krinner,⁶ P. Kurpiers,⁶ M. A. Larotonda,^{19,20} J.-Å. Larsson,²¹ A. Lenhard,¹ H. Li,^{22,23} M.-H. Li,^{17,18} G. Lima,^{3,12} B. Liu,^{24,14} Y. Liu,^{17,18} I. H. López Grande,^{19,20} T. Lunghi,⁵ X. Ma,²⁵ O. S. Magaña-Loaiza,⁹ P. Magnard,⁶ A. Magnoni,²⁰ M. Martí-Prieto,¹ D. Martínez,^{3,12} P. Mataloni,⁷ A. Mattar,¹ M. Mazzeri,¹ R. P. Mirin,⁹ M. W. Mitchell,^{1,2,*} S. Nam,⁹ M. Oppliger,⁶ J.-W. Pan,^{17,18} R. B. Patel,^{15,16} G. J. Pryde,^{15,16} D. Rauch,¹⁴ K. Redeker,¹⁰ D. Rieländer,¹ M. Ringbauer,²⁶ T. Roberson,²⁶ W. Rosenfeld,¹⁰ Y. Salathé,⁶ L. Santodonato,⁷ G. Sauder,⁵ T. Scheidl,¹⁴ C. T. Schmiegelow,²⁰ F. Sciarrino,⁷ A. Seri,¹ L. K. Shalm,⁹ S.-C. Shi,²⁷ S. Slussarenko,^{15,16} M. J. Stevens,⁹ S. Tanzilli,⁵ F. Toledo,^{3,12} J. Tura,^{1,28} R. Ursin,¹⁴ P. Vergyris,⁵ V. B. Verma,⁹ T. Walter,⁶ A. Wallraff,⁶ Z. Wang,^{22,23} H. Weinfurter,^{10,28} M. M. Weston,^{15,16} A. G. White,²⁶ C. Wu,^{17,18} G. B. Xavier,^{3,4,21} L. You,^{22,23} X. Yuan,²⁵ A. Zeilinger,¹⁴ Q. Zhang,^{17,18} W. Zhang,^{22,23} and J. Zhong²⁷

¹ICFO-Institut de Ciències Fòtiques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain

²ICREA – Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

³Center for Optics and Photonics and MSI-Nucleus for Advanced Optics, Universidad de Concepción, Casilla 4016, Concepción, Chile

⁴Departamento de Ingeniería Eléctrica, Universidad de Concepción, 160-C Concepción, Chile

⁵Université Côte d'Azur, CNRS UMR 7010, Institut de Physique de Nice (INPHYNI), Parc Valrose, 06108 Nice Cedex 2, France

⁶Quantum Device Lab, Department of Physics, ETH Zurich, CH-8093 Zurich, Switzerland

⁷Dipartimento di Fisica, Sapienza Università di Roma, I-00185 Roma, Italy

⁸Departamento de Computación, FCEyN, UBA and ICC, CONICET, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina

⁹National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80305, USA

¹⁰Ludwig-Maximilians-Universität, 80799 München, Germany

¹¹Departamento de Física Aplicada II, Universidad de Sevilla, 41012 Sevilla, Spain

¹²Departamento de Física, Universidad de Concepción, 160-C Concepción, Chile

¹³International Institute of Physics, Federal University of Rio Grande do Norte, 59070-405 Natal, Brazil

¹⁴Institute for Quantum Optics and Quantum Information (IQOQI),

Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

¹⁵Centre for Quantum Computation and Communication Technology, Griffith University, Brisbane, Queensland 4111, Australia

¹⁶Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia

¹⁷Shanghai Branch, National Laboratory for Physical Sciences at Microscale and Dept. of Modern Physics, University of Science and Technology of China, Shanghai 230026, China

¹⁸Shanghai Branch, CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 230026, China

¹⁹DEILAP, CITEDEF & CONICET, J.B. de La Salle 4397, 1603 Villa Martelli, Buenos Aires, Argentina

²⁰Departamento de Física, FCEyN, UBA and IFIBA, Conicet, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina

²¹Department of Electrical Engineering, Linköping University, 581 83 Linköping, Sweden

²²State Key Laboratory of Functional Materials for Informatics, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China

²³CAS Center for Excellence in Superconducting Electronics, Shanghai 200050, China

²⁴School of Computer, NUDT, 410073 Changsha, China

²⁵Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

²⁶Centre for Engineered Quantum Systems, Centre for Quantum Computation and Communication Technology, School of Mathematics and Physics, University of Queensland, Brisbane, Queensland 4072, Australia

²⁷Purple Mountain Observatory and Key Laboratory of Radio Astronomy,

Chinese Academy of Sciences, 2 West Beijing Road, Nanjing 210008, China

²⁸Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, Garching 85748, Germany

(Dated: November 12, 2018)

Challenging local realism with human choices

The BIG Bell Test Collaboration[†]

A Bell test is a randomized trial that compares experimental observations against the philosophical worldview of local realism¹, in which the properties of the physical world are independent of our observation of them and no signal travels faster than light. A Bell test requires spatially distributed entanglement, fast and high-efficiency detection and unpredictable measurement settings^{2,3}. Although technology can satisfy the first two of these requirements⁴⁻⁷, the use of physical devices to choose settings in a Bell test involves making assumptions about the physics that one aims to test. Bell himself noted this weakness in using physical setting choices and argued that human ‘free will’ could be used rigorously to ensure unpredictability in Bell tests⁸. Here we report a set of local-realism tests using human choices, which avoids assumptions about predictability in physics. We recruited about 100,000 human participants to play an online video game that incentivizes fast, sustained input of unpredictable selections and illustrates Bell-test methodology⁹. The participants generated 97,347,490 binary choices, which were directed via a scalable web platform to 12 laboratories on five continents, where 13 experiments tested local realism using photons^{5,6}, single atoms⁷, atomic ensembles¹⁰, and superconducting devices¹¹. Over a 12-hour period on 30 November 2016, participants worldwide provided a sustained data flow of over 1,000 bits per second to the experiments, which used different human-generated data to choose each measurement setting. The observed correlations strongly contradict local realism and other realistic positions in bipartite and tripartite¹² scenarios. Project outcomes include closing the ‘freedom-of-choice loophole’ (the possibility that the setting choices are influenced by ‘hidden variables’ to correlate with the particle properties¹³), the utilization of video-game methods¹⁴ for rapid collection of human generated randomness, and the use of networking techniques for global participation in experimental science.

Bell tests, like Darwin’s studies of finches and Galileo’s observations of the moons of Jupiter, bring empirical methods to questions previously accessible only by other means, e.g. by philosophy or theology¹⁵. Local realism, i.e., realism plus relativistic limits on causation, was debated by Einstein and Bohr using metaphysical arguments, and recently has been rejected by Bell tests⁴⁻⁷ that closed all technical ‘loopholes.’ For example, the ‘detection-efficiency loophole’ describes the possibility that the observed statistics are inaccurate due to selection bias, and is closed by high efficiency detection and statistical methods that analyse all trials. Recent work on device-independent quantum information¹⁶ shows how Bell inequality violation (BIV) can also challenge causal determinism¹⁷, a second topic formerly accessible only by metaphysics¹⁸. Central to both applications is the use of free variables to choose measurements: in the words of Aaronson¹⁹ “Assuming no preferred reference frames or closed timelike curves, if Alice and Bob have genuine ‘freedom’ in deciding how to measure entangled par-

ticles, then the particles must also have ‘freedom’ in deciding how to respond to the measurements.”

Prior Bell tests used physical devices^{20,21} to ‘decide’ for Alice and Bob, and thus demonstrated only a relation among physical processes: if some processes are ‘free’ in the required sense (see Methods II), then other processes are similarly ‘free.’ In the language of strong Bell tests, this conditional relation leaves open the freedom-of-choice loophole (FOCL), which describes the possibility that ‘hidden variables’ influence the setting choices. Because we cannot guarantee such freedom within local realism, the tests must assume physical indeterminacy in the hidden-variable theory². Laboratory methods can tighten but never close this loophole²⁻⁶.

Gallicchio, Friedman, and Kaiser²² have proposed choosing settings by observation of cosmic sources at the edge of the visible universe. A Bell inequality violation under such conditions could only be explained within local realism if events across all of history conspire to produce the measured outcomes^{23,24}. Bell himself argued that human choices could be considered ‘free variables’ in a Bell test⁸ (see Methods III), and noted the impracticality of using humans with 1970’s technologies. Here we implement Bell’s idea, using modern crowd-sourcing, networking, and gamification¹⁴ techniques. In this BIG Bell Test (BBT) the Alice and Bob of Aaronson’s formulation are real people. Assuming no faster-than-light communication, such experiments can prove the conditional relation: if human will is free, there are physical events (the measurement outcomes in the Bell tests) that are intrinsically random, i.e., impossible to predict²⁵. We note that this argument in no way uses the theory of quantum mechanics, and yet arrives to one of that theory’s most profound claims. Intrinsic randomness supported by a BIV is central to so-called device-independent quantum technologies^{16,26}.

It is perhaps surprising that human choices, which are known to contain statistical regularities²⁷, are suitably unpredictable for a Bell test. Recent works on the statistical analysis of Bell tests^{3,28,29} clarify this: sequence randomness, i.e., the absence of patterns and correlations in the sequence of choices, is not, *per se*, a requirement for a rejection of local realism. Rather, statistical independence of choices from the hidden variables that describe possible measurement outcomes is required (see Methods II). This independence can fail in different ways, categorized by named loopholes: The FoCL describes the possibility that hidden variables influence the setting choices. The ‘locality loophole’ describes the possibility that a choice at one station could influence a measurement result at the other station. The term ‘locality’ reflects one way of blocking this possibility, by space-like separation of the choice and measurement events (see Methods IV).

Patterns strongly affect statistical strength in experiments that aim to close LL by space-like separation – they allow current choices to be predicted from earlier choices, which have had more time to reach the distant measurement. As described below, the BBT tightens LL using many independent experiments rather than space-like separation. Furthermore, the hu-

man capacity for free choice removes the need for assumptions about physical indeterminism, allowing the FoCL to be closed. Thus, although human choices show imperfect sequence randomness, they nonetheless enable a strong rejection of local realism with the BBT strategy.

A major obstacle to a Bell test with humans has been the difficulty of generating enough choices for a statistically significant test. A person can generate roughly three random bits per second, while a strong test may require millions of setting choices in a time span of minutes to hours, depending on the speed and stability of the experiment. To achieve such rates, we crowd-sourced the basis choices, recruiting in total about 100,000 participants, the *Bellsters*, over the course of the project. Each choice by a participant, encoded as a bit, '0' or '1', was entered in an internet-connected device such as the participant's mobile phone. Servers relayed the entered bits to the 13 experiments, see Fig. 1. The same bits were sent to many experiments, which used them for individual settings without re-use (except experiment ②). To encourage participants to contribute a larger number of more unpredictable bits, the input was collected in the context of a video game, The BIG Bell Quest (available at <https://museum.thebigbelltest.org/quest/>), implemented in javascript to run directly in a device's web browser.

The BIG Bell Quest is designed to reward sustained, high-rate input of unpredictable bits, while also being engaging and informative (see Methods VIII). An interactive explanation first describes quantum nonlocality and the role played by participants and experimenters in the BBT. The player is then tasked with entering a given number of unpredictable bits within a limited time. A machine learning algorithm (MLA) attempts to predict each input bit, modelling the user's input as a Markov process and updating the model parameters using reinforcement learning (see Methods VI). Scoring and level completion reflect the degree to which the MLA predicts the player's input, motivating players to consider their own predictability and take conscious steps to reduce it, but the MLA does not act as a filter: all input is passed to the experiments. Bellster input showed unsurprising deviations from ideal randomness²⁷, e.g., $P(0) \approx 0.5237$ (bias toward '0') while adjacent bits show $P(01) + P(10) \approx 0.6406$ (excess of alternation).

Modern video-game elements were incorporated to boost engagement (animation, sound), to encourage persistent play (progressive levels, power-ups, boss battles, leaderboards) and to recruit new players (group formation, posting to social networks). Different level scenarios illustrate key elements of the BBT: human input, global networking, and measurements on quantum systems, while boss battles against the Oracle (see Methods) convey the conceptual challenge of unpredictability. Level completion is rewarded with 1) a report on how many bits from that level were used in each experiment running at that time, 2) a 'curious fact' about statistics, Bell tests, or the various experiments, and, if the participant is lucky, 3) one of several videos recorded in the participating laboratories, explaining the experiments. The game and BBT website (preserved at <http://museum.thebigbelltest.org>) are available in Chinese, English, Spanish, French, German, Italian and Catalan, making them accessible to roughly three billion first- and second-language speakers.

To synchronize participant activity with experimental operation, the Bell tests were scheduled for a single day, Wednesday

30 November 2016. The date was chosen so that most schools worldwide would be in session, and to avoid competing media events such as the US presidential election. Participants were recruited by a variety of channels, including traditional and social media and school and science museum outreach, with each partner institution handling recruitment in their familiar geographical regions and languages. The media campaign focused on the nature of the experiment and the need for human participants. The press often communicated this with headlines such as "Quantum theory needs your help" (China Daily). A first, small campaign in early October was made to seed "word-of-mouth" diffusion of the story and a second, large campaign 29-30 November was made to attract a wide participant base. The media campaign generated at least 230 headlines in printed and online press, radio and television.

The data networking architecture of the BBT, shown in Fig. 1, includes elements of instant messaging and online gaming, and is designed to efficiently serve a fluctuating number of simultaneous users that is not known in advance and could range from 10 to 100,000. A gaming component handles the BBT website, participant accounts management, delivery of game code (javascript and video), score records and leaderboards. In parallel, a messaging component handles data conditioning, streaming to experiments, and reporting of participant choices generated via the game. Horizontal scaling is used in both components: participants connect not directly to servers but rather to dynamic load balancers that spread the input among a pool of servers dynamically scaled in response to load. The timing of input bits (but not their values) was used to identify robot participants and remove their input from the data stream. Game operation was unchanged, to avoid alerting the robots' masters. A single, laboratory-side server received data from the participant-side servers, concatenated the user input and streamed it to the labs at laboratory-defined rates. See Methods VII for details.

By global time zoning, November 30th defines a 51-hour window, from 0:00 UTC+14h (e.g. Samoa) to 23:59 UTC-12h (e.g. Midway island). Nevertheless, most participants contributed during a 24-hour window centred on 18:00 UTC. Recruitment of participants was geographically uneven, with a notable failure to recruit large numbers of participant from Africa. Despite this, the latitude zones of Asia/Oceania, Europe/Africa, and the Americas had comparable participation, which proved important for the experiment. As shown in Fig. 2, input from any single region dropped to low values during the local early morning, but was compensated by high input from other regions, resulting in a high sustained global bit rate. Over the 12-hour period from 09:00 UTC to 21:00 UTC, 30 November 2016, the input exceeded 10^3 bits per second, allowing a majority of the experiments to run at their full speed. Several experiments posted their results live on social networks. Due to their high speeds, ⑫ and ⑬ accumulated human bits to use in short bursts. As determined by independent measurements, ⑨ and ⑫ were not in condition to observe a BIV on the day of the event; they report later results using stored human bits.

The Earth is only 43 light-ms in diameter, so human choices are too slow to be space-like separated from the measurements. This leaves open the locality loophole regarding the influence of choices on remote detection. Influence of Alice's measurement setting on Bob's detection (and vice-versa) is nonetheless excluded by space-like separation in experiments ③, and ⑬ (see

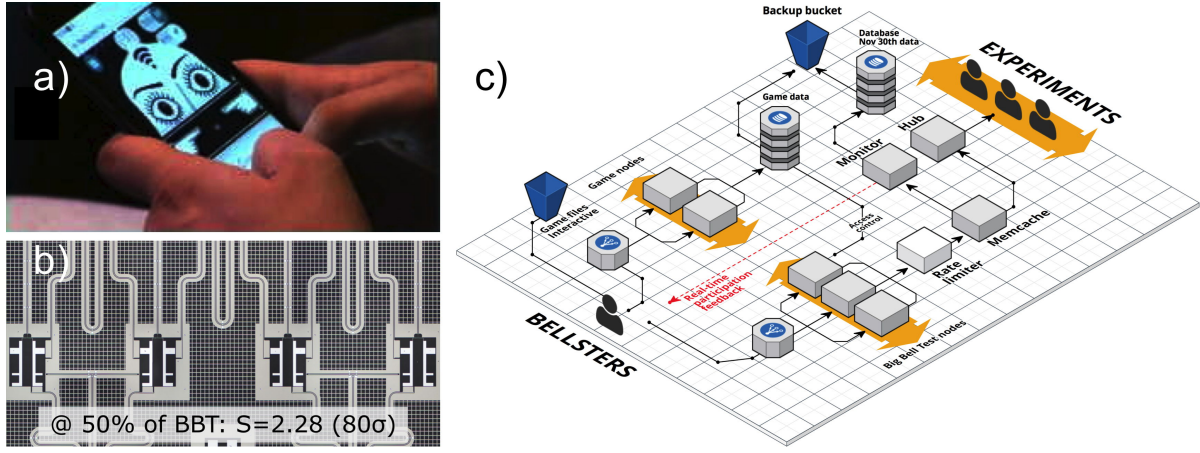


FIG. 1: Structure of The BIG Bell Test. a) human participants, or *Bellsters*, enter ‘0’s and ‘1’s in an online video game that incentivizes sustained generation of unpredictable bits. Image: ICFO/Maria Pascual (Kaitos Games). b) experiments use *Bellster*-generated bits to control measurement-defining elements, such as wave-plates for photons or microwave pulses for matter qubits. Shown is a micrograph of superconducting qubits used in ⑦, with the measured ClauserHorneShimonyHolt Bell parameter S mid-way through the BBT. c) A cloud-based networking system integrates the activities a) and b), serving game elements to *Bellsters*, distributing input bits to connected laboratories, and providing in-game feedback about experimental use of the player’s input. Through this system, *Bellsters* are given direct, if brief, control of the experimental apparatus, so that each measurement setting is determined by a single human choice, traceable to a given user ID and time of entry. See Methods VII.

TABLE I: Experiments carried out as part of the BBT, ordered by longitude, from East to West. Descriptions of the experiments are given in Supplementary Information. Stat. Sig. (statistical significance); indicates number of standard deviations assuming independent and identically distributed trials, unless otherwise indicated. γ signifies photon. Rate indicates the peak rate at which bits were used by the experiments. Due to the limited rate of Bellster input, some experiments had dead times. B , K , S , S_{AB} , S_{BC} , S_{HRN} , S_{QRN} , indicate Bell parameters for the respective experiments, S_{16} is the steering parameter (see Supplementary Information). l_0 indicates the minimum Pütz-Rosset-Barnea-Liang-Gisin measure of settingchoice independence consistent with the observed BIV. USTC, University of Science and Technology of China; EQUUS, Centre for Engineered Quantum Systems; IQOQI, Institute for Quantum Optics and Quantum Information; INFYNI, Institut de Physique de Nice; ICFO, Institut de Ciencies Fotoniques; LMU, Ludwig-Maximilians-Universität; ETHZ, ETH Zurich; CITEDEF, Institute of Scientific and Technical Research for Defence; UdeC, University of Concepción; NIST, National Institute of Standards and Technology.

ID	Lead Institution	Location	Entangled system	Rate	Inequality	Result	Stat. Sig.
①	GRIFFITH	Brisbane, AU	γ polarisation	4 bps	$S_{16} \leq 0.511$	$S_{16} = 0.965 \pm 0.008$	57σ
②	EQUUS	Brisbane, AU	γ polarisation	3 bps	$ S \leq 2$	$S_{AB} = 2.75 \pm 0.05$ $S_{BC} = 2.79 \pm 0.05$	15σ 16σ
③	USTC	Shanghai, CN	γ polarisation	1 kbps	PRBLG ³⁰	$l_0 = 0.10 \pm 0.05$	N/A
④	IQOQI	Vienna, AT	γ polarisation	1.61 kbps	$ S \leq 2$	$S_{HRN} = 2.639 \pm 0.008$ $S_{QRN} = 2.643 \pm 0.006$	81σ 116σ
⑤	SAPIENZA	Rome, IT	γ polarisation	0.62 bps	$B \leq 1$	$B = 1.225 \pm 0.007$	32σ
⑥	LMU	Munich, DE	γ -atom	1.7 bps	$ S \leq 2$	$S_{HRN} = 2.427 \pm 0.0223$ $S_{QRN} = 2.413 \pm 0.0223$	19σ 18.5σ
⑦	ETHZ	Zurich, CH	transmon qubit	3 kbps	$ S \leq 2$	$S = 2.3066 \pm 0.0012$	$p < 10^{-99}$
⑧	INPHYNI	Nice, FR	γ time-bin	2 kbps	$ S \leq 2$	$S = 2.431 \pm 0.003$	140σ
⑨	ICFO	Barcelona, ES	γ -atom ensemble	125 bps	$ S \leq 2$	$S = 2.29 \pm 0.10$	2.9σ
⑩	ICFO	Barcelona, ES	γ multi-frequency-bin	20 bps	$ S \leq 2$	$S = 2.25 \pm 0.08$	3.1σ
⑪	CITEDEF	Buenos Aires, AR	γ polarisation	1.02 bps	$ S \leq 2$	$S = 2.55 \pm 0.07$	7.8σ
⑫	CONCEPCION	Concepcion, CL	γ time-bin	52 kbps	$ S \leq 2$	$S = 2.43 \pm 0.02$	20σ
⑬	NIST	Boulder, US	γ polarisation	100 kbps	$K \leq 0$	$K = (1.65 \pm 0.20) \times 10^{-4}$	8.7σ

Supplementary Information). To tighten LL, we take a strategy we call the BIG test: many simultaneous Bell tests in widely-separated locations using different physical systems, with each experiment’s apparatus constructed and operated by different experimental teams. The only hidden-variable theories that escape this tightening are ones in which choices can simultaneously influence hidden variables in many differently-constructed experiments, to produce in each one a BIV. This strategy is strengthened by using the same bits in many experiments, as

described above.

The suite of 13 BBT experiments, including true Bell tests and other realism tests requiring free choice of measurement, are summarized in Table I and described in the Supplementary Information. Experiments ①, ②, ③, ④, ⑤, ⑧, ⑪, ⑫, and ⑬ used entangled photon pairs, ⑥ used single-photon/single-atom entanglement, ⑨ used single-photon/atomic ensemble entanglement, and ⑦ used entangled superconducting qubits. Experiments ⑦ and ⑬ used high-efficiency detec-

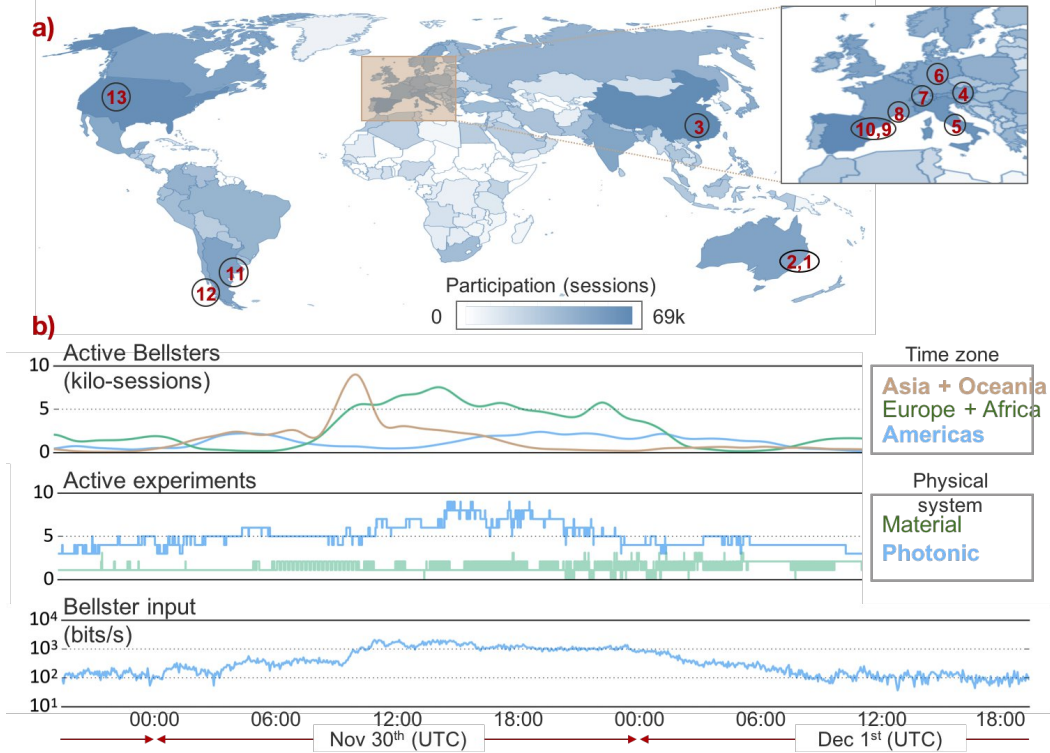


FIG. 2: Geography and timing of the BBT. a) Locations of the 13 BBT experiments, ordered from East to West. Numbers index the experiments, summarized in Table I. Shading shows total sessions by country. Eight sessions from Antarctica are not shown. Map by Giorgio Colangelo using data from OpenStreetMaps, rendered in Wolfram Mathematica. b) Temporal evolution of the project. (top) live sessions versus time for different continent groups, showing a strong drop-off in the local early morning in each region. The spike in Asian participation around 11:00 UTC coincides with a live-streamed event in Barcelona, hosted by D. Jimenez and the CosmoCaixa science museum, re-broadcast live in Chinese by L.-F. Yuan and the University of Science and Technology of China (USTC). (middle) number of connected labs versus time, divided into experiments using only photons and experiments with at least one material component, e.g. atoms or superconductors. (bottom) input bitrate versus time. Data flow remains nearly constant despite regional variations, with Asian *Bellsters* handing off to *Bellsters* from the Americas in the critical period 12:00-00:00 UTC. Session data from Google analytics.

tion to avoid the fair sampling assumption, thus closing simultaneously the detection efficiency and freedom-of-choice loopholes. ⑤ demonstrated a violation of bi-local realism, while ⑩ violated a Bell inequality for multi-mode entanglement. ① demonstrated quantum steering and ② demonstrated temporal quantum correlations with a three-station measurement. ⑫ closed the post-selection loophole typically present in Bell tests based on energy-time entanglement. Analysis of ③ puts bounds on how well a measurement-dependent local model would have to predict *Bellster* behaviour to produce the observed results³⁰. ③, ④, ⑥ and ⑬ tested whether human-generated measurement choices gave different results from machine-generated ones. Most experiments observed statistically strong violations of their respective inequalities, justifying rejection of local realism in a multitude of systems and scenarios.

In summary, on 30 November 2016, a suite of 13 Bell tests and similar experiments, using photons, single atoms, atomic ensembles and superconducting devices, demonstrated strong disagreement with local realism, using measurement settings chosen by tens of thousands of globally-distributed human participants. The results also show empirically that measurement settings independence, here provided by human agency, is in strong disagreement with causal determinism¹⁷⁻¹⁹, a topic formerly accessible only by metaphysics. The experiments reject

local realism in a wide variety of physical systems and scenarios, set the groundwork for Bell-test based applications in quantum information, introduce gamification of randomness generation, and demonstrate global networking techniques by which hundreds of thousands of individuals can directly participate in experimental science.

Acknowledgements: We are grateful to the many people and organizations who contributed to this project, starting with the *Bellsters* (details at <http://thebigbelltest.org>). We thank the Departament d'Ensenyament de la Generalitat de Catalunya, Ministerio de Educacin, Cultura y Deporte of Spain, CosmoCaixa, Fundacin Bancaria "la Caixa", INTEF, Optical Society of America, European Centers for Outreach in Photonics (ECOP), Muncyct-Museo de Ciencia y Tecnologia de Madrid, Investigacin y Cincia, Big Van, Crea Cincia, Tencent news, Micius Salon, Politecnico di Milano, University of Waterloo Institute for Quantum Computing, Universidad Autnoma de Barcelona, Universidad Complutense de Madrid, Universit degli Studi di Milano, Toptica, EPS Young Minds, Real Sociedad Espaola de Fsica, Ajuntament de Barcelona, Ajuntament de Castelldefels, Universit degli studi di Padova, Universit degli studi del l'Insubria, CNRIFN Istituto di Fotonica e Nanotecnologia, Istituto d'Istruzione Superiore Carlo Livi, Webbile, and Kaitos Games. We are especially thankful for the recruitment efforts of the outreach departments at our institutions.

We acknowledge financial support from CONICET; ANPCyT (Argentina), Australian Research Council Centre for Quantum Computation and Communication Technology (CE110001027, CE170100012); Australian Research Council and the University of Queensland (UQ) Centre for Engineered Quantum Systems (CE110001013, CE170100009). A.G.W. acknowledges the University of Queensland Vice-Chancellor's Research and Teaching Fellowship (Australia); Austrian Academy of Sciences (OEAW), the Austrian Science Fund (FWF) (SFB F40 (FoQuS); CoQuS No. W1210-N16), the Austrian Federal Ministry of Education, Science and Research (BMBWF) and the University of Vienna (project QUESS) (Austria), MEC; MCTIC (Brazil), Barcelona City Hall; Generalitat de Catalunya (SGR 874 and 2014-SGR-1295; CERCA programme) (Catalonia), PIA CONICYT (grants PFB0824 and PAI79160083); FONDECYT (grants 1140635, 1150101, 1160400, 3170596 and 11150325; Becas Chile); the Millennium Institute for Research in Optics (MIRO); Becas CONICYT (Chile), the National Fundamental Research Program (grant 2013CB336800); the National Natural Science Foundation of China (grants 91121022, 61401441 and 61401443); the Chinese Academy of Science (Strategic Priority Research Program (B) XDB04010200); the 1000 Youth Fellowship programme; the National Natural Science Foundation of China (grant 11674193); the Science and Technology Commission of Shanghai Municipality (grant 16JC1400402) (China); ERC (grant agreements AQUMET 280169, 3DQUEST 307783, OSYRIS 339106, ERIDIAN 713682, QITBOX, QUOLAPS, QuLIMA and SuperQuNet 339871); ESA (contract number 4000112591/14/NL/US); the European and Regional Development Fund (FEDER); H2020 (QUIC 641122); the Marie Skłodowska-Curie programme (grant agreement 748549); the FP7-ITN PICQUE project (grant agreement No 608062); the OPTIMAL project granted through FEDER; (European Commission); the Université Côte d'Azur (UCA, France) through its program Quantum@UCA; the Foundation Simone & Cino Del Duca (Institut de France); l'Agence Nationale de la Recherche (ANR, France) for the CONNEQT, SPOCQ and SITQOM projects (grants ANR-EMMA-002-01, ANR-14-CE32-0019, and ANR-15-CE24-0005, respectively); the iX-Core Research Foundation (France); German Federal Ministry of Education and Research (projects QuOReP and Q.com-Q) (Germany); CONACyT graduate fellowship programme (Mexico), MINECO (FIS2014-60843-P, FIS2014-62181-EXP, SEV-2015-0522, FIS2015-68039-P, FIS2015-69535-R and FIS2016-79508-P; Ramon y Cajal fellowship programme; TEC2016-75080-R); the ICFOnest + international postdoctoral fellowship programme (Spain); the Knut and Alice Wallenberg Foundation (project 'Photonic Quantum Information')

(Sweden); NIST (USA), AXA Chair in Quantum Information Science; FQXi Fund; Fundaci Privada CELLEX; Fundaci Privada MIR-PUIG; the CELLEX-ICFO-MPQ programme; Fundaci Catalunya-La Pedrera; and the International PhD-fellowship programme 'la Caixa'-Severo Ochoa.

Author contributions: CA instigator, MWM (contact author) project leader. Coordination, gamification and networking (ICFO): SC general supervision, MM-P project management, MG-M, FAB, CA game design, JT prediction engine, AH, MG-M, FAB Bellster recruitment and engagement strategy, design and execution, CA web infrastructure and networking, MWM (contact author), CA, JT main manuscript with input from all authors. Experiments: ①: GJP, RBP (contact author), FGJ, MMW, SS experiment design and execution. ②: AGW, MR (contact author) experiment design and execution. ③: J-WP (contact author), QZ (contact author) supervision, J-WP, QZ, XM, XY, YL experiment conception and design, ZW, LY, HL, WZ superconducting nanowire single-photon detector (SNSPD) fabrication and characterization, JZ SNSPD maintenance, M-HL, CW, YL photon source design and characterization, J-YG, YL software design and deployment, XY protocol analysis, XY, YL data analysis, J-WP, QZ, CW, XY, YL manuscript, with input from all. ④: TS (contact author), AZ, RU supervision, conception and coordination, BL, JHa, DR experiment execution and analysis. ⑤: GC, LS, FA, MB, FS (contact author) experiment execution and analysis, RC theory support. ⑥: HW, WR (contact author), KR, RG, DB experiment design and execution. ⑦: JHe, PK, YS, CKA, AB, SK, PM, MO, TW, SG, CE, AW (contact author) experiment design and execution. ⑧: ST (contact author), TL, FK, GS, PV, OA, experiment design and execution. ⑨: HdR (contact author), PF, GH experiment design and execution. ⑩: HdR (contact author), AS, AL, MM, DR, OJ, AM experiment design and execution, DC, A. Ac. theory support. ⑪: MAL (contact author) coordination, server communication, LTK, IHLG, AGM experiment design and execution, CTS, AB input data formatting, LTK, IHLG, AGM, CTS, AB, MAL data analysis. ⑫: GX (contact author) coordination, FT optical setup, PG, AAI, JF, A. Cuevas, GC optical setup support, JC electronics design and implementation, JC, FT, experiment execution, DM software, GL, PM, FS experimental support, ACa theory support, JC, ESG, J-ÅL data analysis. ⑬: LKS (contact author), SN, MS, OSM-L, TG, SG, PB, EK, RM experiment design, execution and analysis.

Competing interests statement: The authors declare no competing financial interests.

Correspondence and material requests should be addressed to morgan.mitchell@icfo.eu.

¹ J. S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics* **1**, 195 (1964).

² J.-Å. Larsson, Loopholes in Bell inequality tests of local realism, *Journal of Physics A: Mathematical and Theoretical* **47**, 424003 (2014).

³ J. Kofler, M. Giustina, J.-Å. Larsson, M. W. Mitchell, Requirements for a loophole-free photonic Bell test using imperfect setting generators, *Phys. Rev. A* **93**, 032115 (2016).

⁴ B. Hensen, *et al.*, Loophole-free Bell inequality violation using

electron spins separated by 1.3 kilometres, *Nature* **526**, 682 (2015).

⁵ M. Giustina, *et al.*, Significant-loophole-free test of Bell's theorem with entangled photons, *Phys. Rev. Lett.* **115**, 250401 (2015).

⁶ L. K. Shalm, *et al.*, Strong loophole-free test of local realism, *Phys. Rev. Lett.* **115**, 250402 (2015).

⁷ W. Rosenfeld, *et al.*, Event-ready Bell test using entangled atoms simultaneously closing detection and locality loopholes, *Phys. Rev. Lett.* **119**, 010402 (2017).

- ⁸ J. Bell, *Free variables and local causality*, *Speakable and Un-speakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, 2004), chap. 7.
- ⁹ <http://thebigbelltest.org>.
- ¹⁰ P. Farrera, *et al.*, Generation of single photons with highly tunable wave shape from a cold atomic ensemble, *Nature Communications* **7** (2016).
- ¹¹ A. Wallraff, *et al.*, Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics, *Nature* **431**, 162 (2004).
- ¹² G. Carvacho, *et al.*, Experimental violation of local causality in a quantum network, *Nature Communications* **8**, 14775 (2017).
- ¹³ T. Scheidl, *et al.*, Violation of local realism with freedom of choice, *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19708 (2010).
- ¹⁴ J. Sørensen, *et al.*, Exploring the quantum speed limit with computer games, *Nature* **532**, 210 (2016).
- ¹⁵ A. Shimony, *Bell's Theorem*, *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2005), winter 2016 edn.
- ¹⁶ R. Colbeck, Quantum and relativistic protocols for secure multi-party computation, *Ph.D. Thesis, Cambridge University* (2007).
- ¹⁷ C. Hofer, *Causal Determinism*, *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2005), spring 2016 edn.
- ¹⁸ A. Acín, L. Masanes, Certified randomness in quantum physics, *Nature* **540**, 213 (2016).
- ¹⁹ S. Aaronson, Quantum randomness, *American Scientist* **102**, 266 (2014).
- ²⁰ C. Abellán, W. Amaya, D. Mitrani, V. Pruneri, M. W. Mitchell, Generation of fresh and pure random numbers for loophole-free Bell tests, *Phys. Rev. Lett.* **115**, 250403 (2015).
- ²¹ M. Fürst, *et al.*, High speed optical quantum random number generation, *Optics Express* **18**, 13029 (2010).
- ²² J. Gallicchio, A. S. Friedman, D. I. Kaiser, Testing bell's inequality with cosmic photons: Closing the setting-independence loophole, *Phys. Rev. Lett.* **112**, 110405 (2014).
- ²³ J. Handsteiner, *et al.*, Cosmic Bell test: Measurement settings from milky way stars, *Phys. Rev. Lett.* **118**, 060401 (2017).
- ²⁴ C. Wu, *et al.*, Random number generation with cosmic photons, *Phys. Rev. Lett.* **118**, 140402 (2017).
- ²⁵ M. N. Bera, A. Acín, M. K. M. W. Mitchell, M. Lewenstein, Randomness in quantum mechanics: philosophy, physics and technology, *Reports on Progress in Physics* **80**, 124001 (2017).
- ²⁶ S. Pironio, *et al.*, Random numbers certified by Bell's theorem, *Nature* **464**, 1021 (2010).
- ²⁷ M. Bar-Hillel, W. A. Wagenaar, The perception of randomness, *Advances in Applied Mathematics* **12**, 428 (1991).
- ²⁸ P. Bierhorst, A robust mathematical model for a loophole-free Clauser-Horne experiment, *Journal of Physics A: Mathematical and Theoretical* **48**, 195302 (2015).
- ²⁹ D. Elkouss, S. Wehner, (Nearly) optimal P values for all Bell inequalities, *npj Quantum Information* **2**, 16026 EP (2016).
- ³⁰ G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, N. Gisin, Arbitrarily small amount of measurement independence is sufficient to manifest quantum nonlocality, *Phys. Rev. Lett.* **113**, 190402 (2014).

Methods

I. LOCAL REALISM, BELL PARAMETERS, BELL INEQUALITIES

In their 1935 article³¹, Einstein Podolsky and Rosen employed notions of locality (actions or observations in one location do not have immediate effects at other locations), and realism (observables have values even if we do not observe them) to argue that quantum theory was incomplete and could in principle be supplemented with information about which outcomes actually occur in any given run of an experiment. Bell formalized these notions by defining local hidden variable models, (LHVMs) a class of non-quantum theories that are simultaneously local and realistic. We consider the simplest case, of two systems measured by two observers Alice and Bob. We write $x(y)$ to represent Alice's (Bob's) measurement setting, $a(b)$ to represent their measurement outcomes, and λ to represent the hidden variable, something we cannot measure, but which we include in the model to explain why a and b take on particular values. The predictions of any such bi-partite LHVM are given by

$$P(a, b|x, y) = \sum_{\lambda} P(a|x, \lambda)P(b|y, \lambda)P(\lambda). \quad (1)$$

where $P(\cdot|\cdot)$ indicates a conditional probability. That is, the probability of getting outcome a, b when Alice and Bob measure x and y , respectively, is expressible in terms of the local conditional probabilities $P(a|x, \lambda)$ and $P(b|y, \lambda)$. λ is averaged over, because of our ignorance of the value of this hidden variable. If the probabilities $P(a|x, \lambda)$ and $P(b|y, \lambda)$ are restricted to zero or one we have a deterministic LHVM, in which λ, x , and y fully determine the outcomes a and b . Such LHVMs are explicitly realistic in the EinsteinPodolskyRosen sense. Locality is also explicit in the model: For example, $P(a|x, \lambda)$ depends on neither b nor y , so that the events at Bob's station have no influence on Alice's measurement outcome a . A mathematical notion of 'freedom' is implicit in the LHVM: x and y are included as free parameters, and not, e.g., as functions of λ . If $P(a|x, \lambda)$ and $P(b|y, \lambda)$ are allowed to take intermediate values, we speak of a non-deterministic LHVM. Because the unknown $P(\lambda)$ can take on intermediate values, deterministic and non-deterministic LHVMs are equivalent, and from here on we drop the distinction.

This class of models, which by construction embody the EPR assumptions, was shown by Bell to be incapable of reproducing the predictions of quantum mechanics. For example, if Alice and Bob's local systems are spin-1/2 particles in a singlet state, then their measurements (assumed to be ideal) will agree, i.e., show that both are spin-up or both are spin-down, with probability $P(\text{agree}|\phi_a, \phi_b) \equiv P(\uparrow, \uparrow|\phi_a, \phi_b) + P(\downarrow, \downarrow|\phi_a, \phi_b) = \sin^2(\frac{\phi_b - \phi_a}{2})$, where $\phi_b - \phi_a$ is the angle between Alice's and Bob's analysis directions. Eq. (1) cannot reproduce all the features of this distribution. No choice of $P(a|x, \lambda)$, $P(b|y, \lambda)$, and $P(\lambda)$ can give a probability $P(a, b|x, y)$ that simultaneously depends on the difference $\phi_b - \phi_a$, has high-visibility (ranging from 0 to 1) and sinusoidal.

This difference is efficiently captured by Bell inequalities. A Bell parameter is a linear combination of conditional probabilities $P(a, b|x, y)$, and a Bell inequality indicates the bounds

(within the class of LHVMs) of a Bell parameter. Typically, the Bell inequalities of interest are those that are not obeyed by quantum mechanics, i.e., those for which quantum correlations can be strong enough to violate the Bell inequality. Bell's theorem shows that there are such inequalities, and thus that quantum mechanics cannot be 'completed' with hidden variables.

Bell inequalities also make possible experimental tests of local realism: A Bell test is an experiment that makes many spatially-separated measurements with varied settings to obtain estimates of the $P(a, b|x, y)$ that appear in a Bell parameter. If the observed Bell parameter violates the inequality, one can conclude that the measured systems were not governed by any LHVM. It should be noted that this conclusion is always statistical, and typically takes the form of a hypothesis test, leading to a conclusion of the form 'assuming nature is governed by local realism, the probability to produce the observed Bell inequality violation (or a stronger one) is $P(\text{observed or stronger}|\text{local realism}) \leq p$.' This p -value is a key indicator of statistical significance in Bell tests.

II. 'FREEDOM' IN BELL TESTS

The use of the term 'free' to describe the choices in a Bell test derives more from mathematical usage than from its usage in philosophy, although the two are clearly related. Bell⁸ (see Section III below) states that his use of 'free will,' reflects the notion of 'free variables,' that is, externally-given parameters in physical theories, as opposed to dynamical variables that are determined by the mathematical equations of the theory.

As described above x, y are the settings in a Bell test, a, b are the outcomes, and λ is the hidden variable. Any local realistic hidden variable model is described by Eq. (1). The mathematical requirements for the relevant 'freedom' are made evident by a more general description in which the local realistic model includes also x, y , in which case it specifies the joint probability

$$P(a, b, x, y) = \sum_{\lambda} P(a|x, \lambda)P(b|y, \lambda)P(x, y|\lambda)P(\lambda). \quad (2)$$

Using the Kolmogorov definition of conditional probability $P(A, B) = P(A|B)P(B)$, we find that Eq. (2) reduces to Eq. (1), provided that $P(x, y|\lambda) = P(x, y)$, i.e., provided that the settings are statistically independent of the hidden variables. By Bayes' theorem, this same condition can be written $P(\lambda|x, y) = P(\lambda)$ and $P(x, y, \lambda) = P(x, y)P(\lambda)$. This condition is known in the literature as the freedom-of-choice assumption, although it implies more than just free choices. A more accurate term might be 'measurement setting/hidden-variable independence.' We note that this condition does not require that x be independent of y , nor does it require that $P(x, y)$ be unbiased. Similar observations emerge from the more involved calculations required to assign p -values to observed data in Bell tests³.

The above clarifies the sense in which the basis choices should be 'free.' The desideratum is independence from the hidden variables that describe the particle behaviours, keeping in mind that the choices and measurements could, consistent with relativistic causality, be influenced by any event in their backward light-cones. Because the setting choices

and the measurements will always have overlapping backward light-cones, it is impossible to rule out the possibility of a common past influence through space-time considerations. If human choices are free, however, such influences are excluded. It should also be noted that complete independence is not required, although the tolerance for interdependence can be low^{30,32,33}. The theory that the entire experiment, including choices and outcomes, is pre-determined by initial conditions is known as superdeterminism. Superdeterminism cannot be tested³⁴.

A very similar concept of 'freedom' applies to the entangled systems measured in a Bell test. A Bell inequality violation with free choice and under strict locality conditions implies either indeterminacy of the measurement outcomes or faster-than-light communications and thus closed time-like curves^{16,18}. If Bob's measurement outcome is predictable based on information available to him before the measurement, and if it also satisfies the condition for a Bell inequality violation, namely a strong correlation with Alice's measurement outcome that depends on his measurement choice, then Bob can influence the statistics of Alice's measurement outcome, and in this way communicate to her despite being space-like separated from her. Considering, again, that Bob could in principle have information on any events in his backward light cone, this implies (assuming no closed time-like curves) that Bob's measurement outcome must be statistically independent of all prior events.

In this way, we see that 'freedom,' understood as behaviour statistically independent of prior conditions, appears twice in a Bell test, first as a requirement on the setting choices, and second as a conclusion about the nature of measurement outcomes on entangled systems. These two are linked, in that the second can be demonstrated if the first is present.

Prior tests using physical randomness generators to choose measurement settings thus demonstrate a relationship between physical processes, showing for example^{4,20} that if spontaneous emission is 'free,' then the outcomes of measurements on entangled electrons are also 'free.' By using humans to make the choices, we translate this to the human realm, showing, in the words of Conway and Kochen³⁵, "if indeed there exist any experimenters with a modicum of free will, then elementary particles must have their own share of this valuable commodity." Here 'experimenters' should be understood to refer to those who choose the settings, i.e., the Bellsters. See main text for a discussion of the locality loophole when using humans.

III. JOHN STEWART BELL ON 'FREE VARIABLES'

A brief but informative source for Bell's positions on setting choices is an exchange of opinions with Clauser, Horne and Shimony (CHS)³⁶, in articles titled 'The theory of local beables' and 'Free variables and local causality.' In the first of these articles Bell very briefly considers using humans to choose the measurement settings

It has been assumed [in deriving Bell's theorem] 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.

while the second article defends this choice of method and

compares it against 'mechanical,' i.e. physical, methods of choosing the settings.

Suppose that the instruments are set at the whim, not of experimental physicists, but of mechanical random number generators. Indeed it seems less impractical to envisage experiments of this kind...

Bell proceeds to consider the strengths and weaknesses of physical random number generators in Bell tests, offering arguments why under 'reasonable' assumptions physical random number generators might be trusted, but nonetheless concluding

Of course it might be that these reasonable ideas about physical randomizers are just wrong - for the purpose at hand. A theory might appear in which such conspiracies inevitably occur, and these conspiracies may then seem more digestible than the non-localities of other theories.

In sum, Bell distinguishes different levels of persuasiveness, noting that physical setting generators, while having the required independence in many local realistic theories, cannot be expected to do so in all such theories. In contemporary terminology, what he argues here is that physical setting generators can only tighten, not close the FOCL.

Bell also defends his use of the concept of 'free will' in a physics context, something that had been criticized by CHS. Bell writes

Here I would entertain the hypothesis that experimenters have free will ... it seems to me that in this matter I am just pursuing my profession of theoretical physics.

... A respectable class of theories, including contemporary quantum theory as it is practiced, have 'free' 'external' variables in addition to those internal to and conditioned by the theory. These variables provide a point of leverage for 'free willed experimenters', if reference to such hypothetical metaphysical entities is permitted. I am inclined to pay particular attention to theories of this kind, which seem to me most simply related to our everyday way of looking at the world.

Of course there is an infamous ambiguity here, about just what and where the free elements are. The fields of Stern-Gerlach magnets could be treated as external. Or such fields and magnets could be included in the quantum mechanical system, with external agents acting only on the external knobs and switches. Or the external agents could be located in the brain of the experimenter. In the latter case the setting of the experiment is *not* itself a free variable. It is only more or less correlated with one, depending on how accurately the experimenter effects his intention.

It is clear from the last three sentences that Bell considers human intention - that is, human free will - to be a 'free variable' in the sense he is discussing. That is, he believes human intention fulfils the assumptions of Bell's theorem, as do experimental settings faithfully derived from human intention.

IV. USE OF 'FREEDOM-OF-CHOICE LOOPHOLE' AND 'LOCALITY LOOPHOLE' IN THIS WORK

As noted above, a statistical condition used to derive Bell's theorem is $P(x, y, \lambda) = P(x, y)P(\lambda)$, where x and y are choices and λ describes the hidden variables. This statistical condition, known as the freedom of choice assumption, does not distinguish between three possible scenarios of influence: the condition could fail if the choices influence the hidden variables, if the hidden variables influence the choices, or if a third factor influences both choices and hidden variables^{2,3,13}. According to Bayes' theorem, equivalent forms are $P(x, y|\lambda) = P(x, y)$, which expresses the fact that knowing λ does not give information about (x, y) , and $P(\lambda|x, y) = P(\lambda)$ which expresses the fact that knowing (x, y) does not give information about λ . The latter relationship makes clear that influence (in either direction) is incompatible with the freedom of choice assumption. The name for this condition should not be taken literally; the condition can be false even if the choices are fully free, in the sense of being independent of all prior conditions. This occurs for example if the choices are freely made but then influence the hidden variable.

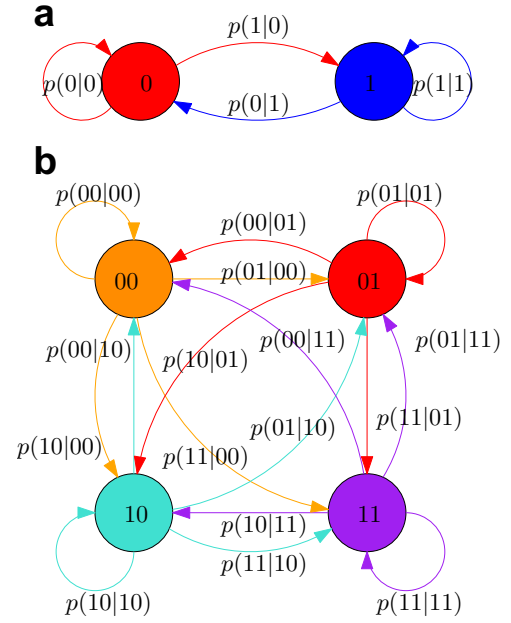
By long tradition, the 'locality loophole' (LL) is the name given to the possibility of influence from Alice's (Bob's) choices or measurements to Bob's (Alice's) measurement outcomes. The term 'freedom-of-choice loophole' was introduced in Scheidl et al.¹³, to describe influence from hidden variables to choices. The text of the definition was "the possibility that the settings are not chosen independently from the properties of the particle pair." We note that this formulation centres on the act of choosing and its independence, which (assuming relativistic causality, an element of local realism) can only be violated by influences from past events, not future events. These loophole definitions employ the concept of influence, which is directional, to explain how the non-directional relation of independence can be broken. Similarly directional definitions have recently been applied to experiments using cosmic sources^{23,24}.

Our use of the term in this paper follows the definition of Scheidl et al.¹³ described above: FOCL refers to the possibility of influences on the choices from any combination of hidden variables and/or other factors within the backward light-cone of the choice, whereas the possibility of choices influencing the hidden variables, which necessarily occur in the forward light-cone of the choice, is included in the locality loophole. Such a division, in addition to fitting the common-sense notion of free choice, avoids counting a single possible channel of influence in both FOCL and the locality loophole.

V. STATUS OF THE FREEDOM-OF-CHOICE LOOPHOLE

The FoCL remains unclosed after recent experiments simultaneous closing locality, detection efficiency, memory, timing, and other loopholes⁴⁻⁷. Space-time considerations can eliminate the possibility of such influence from the particles to the choices^{5,6,13,37}, or from other space-time regions to the choices^{4,7}, but not the possibility of a sufficiently early prior influence on both choices and particles. To motivate freedom of choice in this scenario, well-characterized physical randomizers^{20,21} have been used to choose settings.

In experiments⁴⁻⁶ the physical assumption is that at least



Extended Data Figure 1: Markov chains. **a**, Markov chain for $L = 1$. States are represented by circles of different colours, and transitions between states by arrows coloured as the initial state. The last input bit determines the state of the predictor. The probability $p(a|b)$ of a transition from state b to state a is estimated from the sequence S_k . **b**, Markov chain for $L = 2$, with four states determined by the last two input bits. The transitions model the probability $p(a, b|c, d)$ that the user will input bits ab , given that the last two bits were cd . The final prediction is based on the marginal of the next single bit, which is extracted from these estimated probabilities.

one of: spontaneous emission, thermal fluctuations, or classical chaos²⁰ is uninfluenced by prior events, and thus unpredictable even within local realistic theories. In experiments^{7,13,37,38} the physical assumption is that photodetection is similarly uninfluenced. While still requiring a physical assumption and thus not closing the freedom-of-choice loophole, this strategy tightens the loophole in various ways: First, by using space-like separation to rule out influence from certain events, e.g. entangled pair creation, and from defined space-time regions. Second, by using well-characterized randomness sources, for which the setting choice is known to faithfully derive from a given physical process, it avoids assumptions about the predictability of side-channel processes. Third, in the case of^{4-6,20}, by using a physical variable that can be randomized by each of several processes, the required assumption is reduced from 'x is uninfluenced' to 'at least one of x, y and z is uninfluenced.'

VI. PREDICTION ENGINE

Generation of random sequences by humans has been a subject of study in the field of psychology for decades^{27,39}. Early studies showed that humans perform poorly when asked to produce a random sequence, choosing in a biased manner and deviating from a uniform distribution. It was shown in⁴⁰ that humans playing competitive, zero-sum games that reward uniform random choices tend to produce sequences with fewer

identifiable biases. One such game is matching pennies: Players have to simultaneously choose between heads or tails; one player wins if the results are equal, the other wins if the results are different. This is a standard two-person game used in game theory⁴¹ (see also⁴²) with a mixed-strategy Nash equilibrium: As both players try to outguess the other, by behaving randomly they do not incentive either player to change their strategy.

The BIG Bell Quest reproduces the coin-matching game, with a machine-learning algorithm (MLA) playing the part of the opponent. The MLA operates on simple principles that human players could employ: it maintains a model of the tendencies of the opponent, noting for example “after choosing ‘0,’ ‘0,’ she usually choses ‘1’ as her next bit.” The MLA strategy operates with very little memory, mirroring the limited short-term memory of humans.

Formally, we write $x_i \in \{0, 1\}$ for the i th input bit, $S_k \equiv \{x_1, \dots, x_k\}$ for the sequence of k input bits, and $\mathbf{x}_j^{(L)} \equiv \{x_j, x_{j+1}, \dots, x_{j+L-1}\}$ for the length- L sub-sequence of S_k starting from bit j . Given S_k as input, the algorithm predicts the value of $x_{k+1} \in \{0, 1\}$ that maximizes

$$\max_{L \in [1, L_{\max}]} f_L(x_{k+1} | \mathbf{x}_{k-L+1}^{(L)}) \quad (3)$$

where f estimates the probability of x following \mathbf{x} in S_k :

$$f_L(x | \mathbf{x}) = \frac{\#\{\mathbf{x}_i^{(L+1)} : \mathbf{x}_i^{(L+1)} = \mathbf{x} \parallel x, \quad 1 \leq i \leq k-L\}}{\#\{\mathbf{x}_i^{(L)} : \mathbf{x}_i^{(L)} = \mathbf{x}, \quad 1 \leq i \leq k-L\}}, \quad (4)$$

where \parallel indicates concatenation and $\#A$ indicates the number of elements in set A . Equations (3) and (4) mean that the prediction algorithm identifies the most frequently input sequence, of length $L_{\max} + 1$ or shorter, that the player can form when adding the bit x_{k+1} , and predicts the player will indeed produce the bit needed to complete that sequence. L_{\max} is chosen to be three, reflecting limited memory of a human opponent in the coin-matching game.

In Eq. (4), the estimator $f_L(x | \mathbf{x})$ of the probability that x follows \mathbf{x} in S_k is based on modelling the user’s input as a Markov process⁴³. The MLA keeps a running estimate $T_L(\mathbf{z} | \mathbf{y})$, updated with each new input bit, of the matrix describing probabilities of transitions among length- L words, from word \mathbf{y} to word \mathbf{z} . The estimates are simply the observed frequency of transitions in S_k . The MLA then obtains f_L as a marginal probability distribution: the probability of the first bit of \mathbf{z} being x , conditioned on the tail of S_k being \mathbf{y} (See Extended Data Fig. 1).

VII. NETWORKING STRATEGY AND ARCHITECTURE

The BBT required reliable, robust, and scalable operation of two linked networking tasks: providing the BIG Bell Quest video game experience, and live aggregation and streaming of user input to the running experiments. From a networking perspective, the latter task resembles an instant messaging service, with the important asymmetry that messages from a large pool of senders (the Bellsters) are directed to a much smaller pool of recipients (the labs). The network architecture is shown in Fig. 1c, and was implemented using Amazon Web Services IaaS (Infrastructure as a Service) products.

In the messaging component, we employed a two-layered architecture, shown in Fig. 1c. In the first layer *Big Bell Test nodes* received input bits from the users and performed a real-time health check, described below, to block spamming by robot participants. The data were then sent to the second layer, a single instance *Hub node* that concatenated all the bits from the first stage and distributed them to the labs. The communication between the two layers was implemented using a *memcache* computation node to maximise speed and to simplify the synchronisation between the two layers.

The gaming task was handled by a single layer of *Game nodes* and a database. To protect the critical messaging task from possible attacks on the gaming components, we used separate instances to handle backend gaming tasks, such as user information and rankings, and to handle backend tasks in the messaging chain, such as data logging. Load balancers, networking devices that distribute incoming traffic to a scalable pool of servers, were used in both the gaming and messaging front ends to avoid overloading. This design pattern is known as horizontal scaling, and is a common practice in scalable cloud systems.

This specific architecture was not available as a standard service from web service providers, but was readily constructed from standard component services. The architecture is not specific to the low-bitrate manual input collected for the BBT, and could straightforwardly be adapted to other data collectable by personal devices, e.g. audio or acceleration. The architecture was designed to solve a problem specific to time-limited projects with crowd-sourced input: Due to the single-day nature of the BBT, the unknown number and geography of the participants, and the possibility of hackers/spammers, it was not practical to test the system under full-load conditions prior to the event itself. The two-layer architecture helped maintain all the critical servers isolated and independently operating, and helped us smoothly scale up the system when traffic increased. In the event, the traffic surpassed our initial estimates, and we deployed three additional BIG Bell Test nodes at 09:00 UTC (when Europe was waking up) with no interruption of service. Such scaling-up is expected to be critical for projects, e.g. ref.⁴⁴, that combine laboratory experimentation, which tends to be time-limited due to stability and resource considerations, with crowd-sourcing, which usually entails unknown and fluctuating demand.

We now give more details on each computational resource.

A. Big Bell Test nodes

The first layer of computing resources received data from Bellsters, or more precisely from the BIG Bell Quest running in browsers on their computers and devices. A variable number of servers running the same software functionalities were placed behind a pre-warmed load balancer that was prepared to support up to 10,000 simultaneous connections. Users connected to the load balancer via a public URL end-point, and sent the data from their browsers using *websocket* connections. This first layer of servers aggregated the data from each connection (i.e. from each user) in independent buffers during a $T = 0.5$ s interval.

A simple but important ‘health check’ was performed to identify and block high-speed robotic participants. If a given user

contributed more than ten bits in a single half-second interval, corresponding to a rate of more than 20 keypresses per second, the user account was flagged as being non-human and all subsequent input from that user was removed from the data stream. No feedback was provided to the users in the event their account was flagged, to avoid leaking information on the blocking mechanism. This method could potentially ban honest users due to networking delays and other timing anomalies, but was necessary to prevent the greater risk that the data stream was flooded with robotic input.

B. Hub node

The Hub node aggregated the data from all the Big Bell Test nodes and also handled the connection to the labs. In contrast to the Big Bell Test nodes, which had to service connections from an unknown and rapidly changing number of users, the Hub node was aggregating data from a small and relatively stable number of trusted instances. Overall, the two layer design simplified the networking task of delivering input from a large and variable number of users to end points (the labs) receiving aggregated data streams at variable rates.

Laboratories connected to the Hub instance to receive random bits from the *Bellsters*, which were distributed after aggregating four of the $T = 0.5$ s batches from the Big Bell test nodes, i.e. in intervals of 2 s. At the end of each interval, bits were sent to each running experiment: if an experiment had requested N bits, it was sent bits $\{x_0, x_1, \dots, x_{N-1}\}$, i.e., the earliest bits to arrive in that interval. The same bits were in this way used simultaneously in many experiments. This helps to tighten the locality loophole, since an influence from input bits to measurements would have to operate the same way in several independent experiments and in several locations. With the exception of ⑫ and ⑬, the sent bits were used within the next 2 s interval. Experiments ⑫ and ⑬, in order to run faster than the Bellster input rate, operated in a burst mode, accumulating bits for a time and then rapidly using them. As with the Big Bell Test instances, these connections were established using websocket connections. When connecting to the Hub node, the labs specified their bitrate requirement, which could be dynamically changed. The Hub node then sent a stream of Bellster-generated bits at the requested rate. In the event that insufficient Bellster-generated bits were arriving in real-time, archived bits from BBT participation prior to the day of the experiment were distributed to the labs in advance. In the event, the flux of live bits was sufficient and no experiments used these pre-distributed bits.

C. Memcache node

The interface between The Big Bell Test nodes and the Hub instance was implemented using a memcache node. While adding an extra computing resource slightly increased the complexity of the architecture, it added robustness and simplified operations. The memcache node, in contrast to the Big Bell Test and Hub nodes, had no internet-facing functionality, making its operation less dependent on external conditions. For this reason, both the Big Bell Test nodes and the Hub node were registered and maintained on the memcache node, allowing the

restart of any of these internet-facing instances without loss of records or synchronisation.

In addition, and as detailed in the next section, there was an additional Monitor node in charge of (i) recording all the random bits that were being sent from the *Bellsters* to the labs, and (ii) providing real-time feedback to the *Bellsters*. This functionality was isolated from the operations of the Hub node. Again, by splitting the Monitor and Hub instances, a failure or attack in the public and non-critical real-time feedback functionality had no effect on the main, private, and critical random bit distribution task.

D. Monitor node

For analysis and auditing purposes, all of the bits passing through the first layer of servers were recorded in a database, together with metadata describing their origin (Monitor computing resource in Fig. 1c). In particular, every bit was stored together with the username that created it and the origin timestamp. The random bitstreams sent to the individual labs were similarly recorded bit-by-bit, allowing a full reconstruction of the input to the experiments.

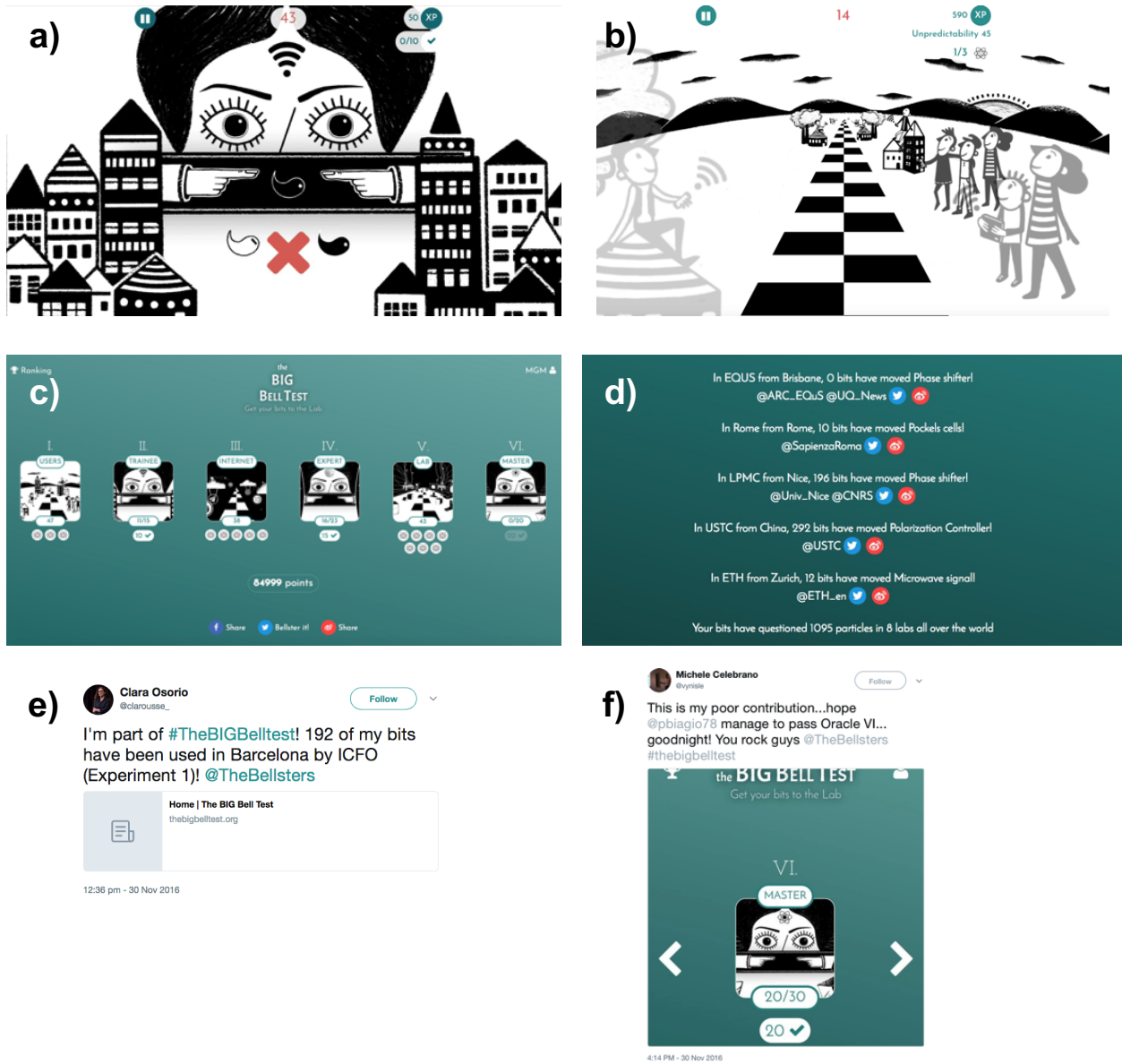
In post-event studies of the input data, we estimated the possible contribution from potentially machine-generated participations that were not blocked by the real-time blocking mechanism. We analysed participants whose contribution were significant, more than 2 kbit in total, and looked for anomalous timing behaviours such as improbably short time spent between missions and improbably large number of bits introduced per mission, both of which are limited by the dynamics of human reactions when playing the game. Flagging participants that contributed such anomalous participations as suspicious, and cross-referencing against the bits sent to the experiments, we find that no experiment received more than 0.1% bits from the eleven suspicious participants.

In the Monitor computing resource, in addition to being used to store in a database all the information that was streamed to the labs, we also implemented a real-time feedback mechanism to improve the *Bellsters'* participation experience. After accomplishing each mission, users were shown a report on the use of their bits in each of the labs running at that moment, as illustrated in Extended Data Fig. 2d. The numbers shown were calculated as a binomial random process $B(n, p_i)$ with parameters $n = N$ and $p_i = R_i/R$, where N is the number of bits introduced by a user in his/her last mission, R_i is the number of bits sent to lab i , and R is the total number of bits entered in the last $T = 0.5$ s interval.

VIII. GAMIFICATION

The BBT required a large number of human-generated random bits in a short time, thus requiring many participants (*Bellsters*), rapid input, and sustained participation. The gamification strategy was designed to maximize all of these factors. The game itself The BIG Bell Quest can be viewed and played at <https://museum.thebigbelltest.org/quest/>. Relevant screenshots are shown in Extended Data Fig. 2.

While still adhering to common conventions of video games (levels, power-ups, boss battles, animations, sound effects,



Extended Data Figure 2: Screenshots from The BIG Bell Quest, illustrating various game elements. **a)** the Oracle opponent uses a machine-learning algorithm to predict user input. **b)** the 'running' component of the game, in which participants are tasked to enter a minimum number of bits with a minimum fraction unpredicted in a limited time. **c)** sequence of levels, of increasing difficulty, interspersed with oracle challenges **d)** in-game feedback on the use of the user's input bits in running experiments. Blue and red buttons allow instant sharing on social networks Twitter and Weibo, respectively **e)** a social media post for direct sharing of participant results **f)** a social media post by a participant who completed the very difficult last Oracle level. BIG Bell Quest artwork by Maria Pascual (Kaitos Games). See also Methods VIII.

etc.), the intended appeal of the BIG Bell Quest is less its entertainment value than the opportunity to contribute to the BBT experiments, and to test one's unpredictability against a computer opponent. The game design incorporated internationalization, connection to social networking, community-building features, and a feedback system to inform users about their contribution to the experiment, all considered essential to attract the necessary tens of thousands of participants.

The game has a classic challenge and reward incentive structure. The challenge is to produce random bits, avoiding being predicted by the Oracle, (see Extended Data Fig. 2a). This reproduces the 'penny-matching' game studied in psychology^{41,42}, and resembles the well-known 'rock-

paper-scissors,' thus requiring little explanation. The Oracle is a machine-learning algorithm that predicts player behaviour based on patterns in past input, described in Section VI. Most player time was spent in a rapid 'speed game' (see Extended Data Fig. 2b), in which the Bellster moves along a road by hitting 0s and 1s. This part of the game requires rapid bit generation (a few bps) in order to complete the level in time. Every 20 bits an indicator shows the player's 'unpredictability,' i.e., the percentage of unpredicted bits entered thus far, and the final score reflects the number of unpredicted bits, with a power-up multiplier for bits entered during a particular time window.

The rewards are multiple: at the individual level, the player

is given a score for each level (to encourage a high fraction of unguessed input), and a cumulative score (to encourage repeated play). At the community level, a sharing platform offers rankings and a tool to create groups, so that *Bellsters* can compare their performance among friends and colleagues. Players can also post their scores to social networks (Facebook, Twitter or Weibo) at the press of a button, see Extended Data Fig. 2e and f. At the scientific level, the game provides a report on which laboratories have used how many of the player's bits, and for what purpose (see Extended Data Fig. 2c). Finally, the player is occasionally rewarded with a short video pre-recorded in one of the laboratories, in which experimentalists explain a part of their experiment. User feedback, see Extended Data Fig. 2e, suggests this approach succeeded in making *Bellsters* feel meaningfully involved in the project, with a positive effect on retention and propagation.

This Speed Game + Oracle structure was repeated through three levels of difficulty, or 'worlds:' 'Users', 'Internet', and 'Laboratory', illustrating the travel of the bits from the fingers of the *Bellsters*, through the Internet, to the labs of the sci-

entists (see Extended Data Fig. 2c). The times, speeds, and required unpredicted fraction in each of the levels were adjusted with the help of beta testers to avoid offering levels of trivial or impossible difficulty. The final Oracle level was objectively difficult even for an experienced player. To pass, it required $n \geq 20$ unguessed bits in a time allowing at most 30 bits to be entered. Even for a sequence of 30 ideal random input bits, the condition $n \geq 20$ occurs less than 5% of the time by binomial statistics. For 30 bits predictable with probability 0.6, the chance of success drops below 0.003. Nevertheless, several players persisted and completed the game (see Extended Data Fig. 2f).

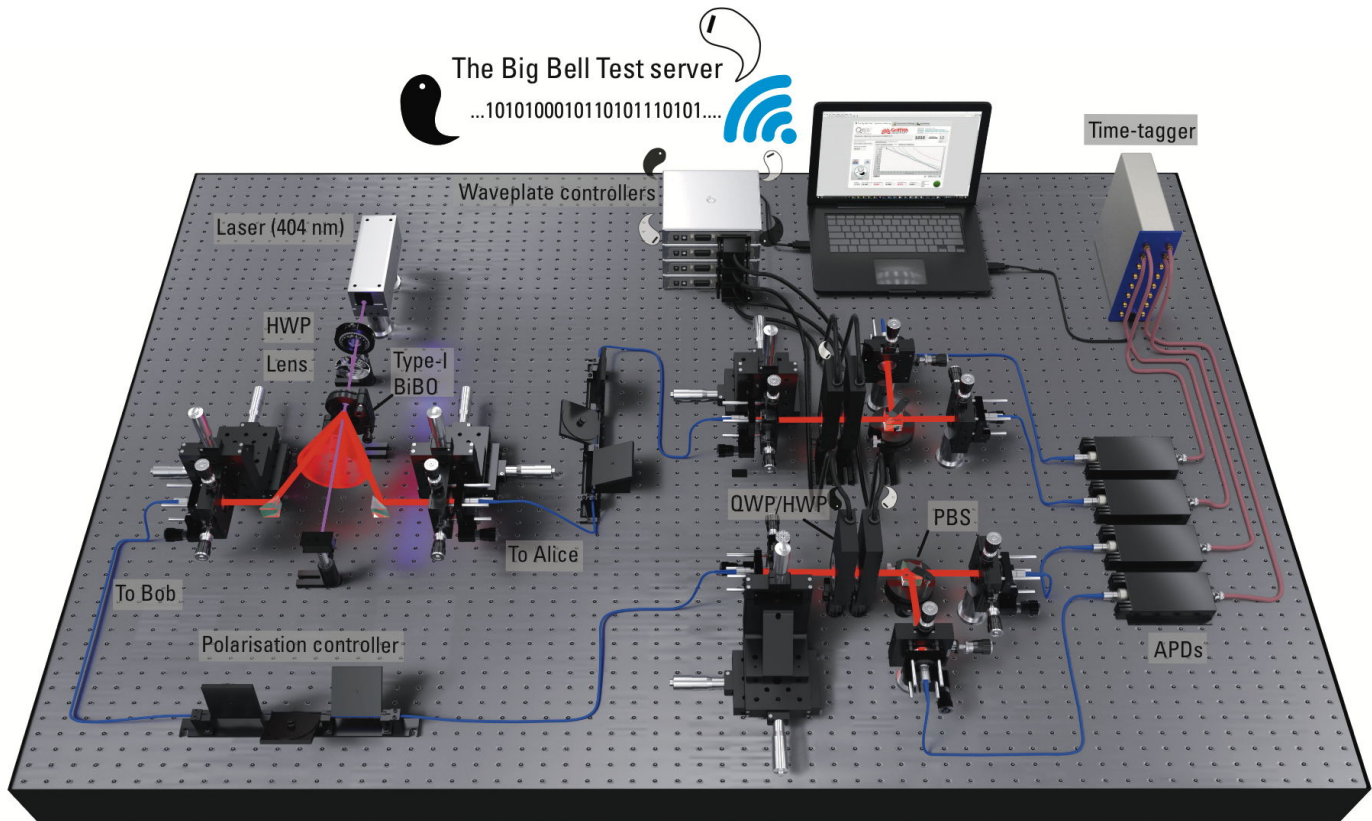
IX. DATA AVAILABILITY

Experimental data are available upon reasonable request from the contact author of each experiment, as indicated in the author contributions. Other project data are available upon reasonable request from the corresponding author.

-
- ³¹ A. Einstein, B. Podolsky, N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, *Phys. Rev.* **47**, 777 (1935).
- ³² M. J. W. Hall, Relaxed Bell inequalities and Kochen-Specker theorems, *Phys. Rev. A* **84**, 022102 (2011).
- ³³ J. Barrett, N. Gisin, How much measurement independence is needed to demonstrate nonlocality?, *Phys. Rev. Lett.* **106**, 100406 (2011).
- ³⁴ J. S. Bell, Bell, J., Clauser, M. Horne, and A. Shimony, An exchange on local beables, *Dialectica* **39**, 85–110 (1985).
- ³⁵ J. Conway, S. Kochen, The free will theorem, *Foundations of Physics* **36**, 1441 (2006).
- ³⁶ J. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, Collected papers on quantum philosophy (Cambridge University Press, 2004).
- ³⁷ C. C. Erven, *et al.*, Experimental three-photon quantum nonlocality under strict locality conditions, *Nat Photon* **8**, 292 (2014).
- ³⁸ G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, A. Zeilinger, Violation of Bell's inequality under strict Einstein locality conditions, *Phys. Rev. Lett.* **81**, 5039 (1998).
- ³⁹ W. A. Wagenaar, Generation of random sequences by human subjects: A critical survey of the literature, *Psychological Bulletin* pp. 65–72 (1972).
- ⁴⁰ A. Rapoport, D. V. Budescu, Generation of random series in two-person strictly competitive games., *Journal of Experimental Psychology: General* **121**, 352 (1992).
- ⁴¹ R. Gibbons, *Game theory for applied economists* (Princeton University Press, 1992).
- ⁴² D. Mookherjee, B. Sopher, Learning behavior in an experimental matching pennies game, *Games and Economic Behavior* **7**, 62 (1994).
- ⁴³ R. Serfozo, *Basics of applied stochastic processes* (Springer Science & Business Media, 2009).
- ⁴⁴ R. Heck, *et al.*, Do physicists stop searches too early? A remote-science, optimization landscape investigation (2017).

① Quantum steering using human randomness

Authors: Raj B. Patel, Farzad Ghafari Jouneghani, Morgan M. Weston, Sergei Slussarenko, and Geoff J. Pryde



Supplementary Figure 3: Experimental setup. Alice and Bob's measurements stages are set according to the random stream of bits acquired from the Big Bell Test server. Entangled photon pairs are generated by pumping a sandwiched Type-I BiBO crystal with a continuous wave laser diode at 404 nm. Each photon from the down-converted pair is sent along an optical fiber to polarisation analysis stages, one for Alice and one for Bob. Photons are detected using silicon avalanche photodiode detectors and the electrical signals for each event are time-tagged. A computer is used to calculate the steering parameter S_{16} in real-time. As additional random measurement settings are retrieved from the server, the experiment is iterated for improved statistics.

Schrödinger first coined the term 'steering'⁵¹ as a generalisation of the EPR-paradox. With the advent of quantum technologies, steering has been recognised as being well suited to certain quantum communication tasks. Here we report a demonstration of EPR-steering using polarisation entangled photons, where Alice and Bob's measurement settings are chosen based on data randomly generated by humans. We use a 404 nm UV continuous wave laser diode to pump a pair of sandwiched type-I nonlinear bismuth triborate (BiBO) crystals to generate entangled photon pairs at 808 nm via spontaneous parametric down-conversion. The generated state is the singlet state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H_A V_B\rangle - |V_A H_B\rangle)$. The generated photon pairs are sent to two separate measurement stages consisting of polarisation analysers and single-photon avalanche photodiode detectors. The stages, designated Alice and Bob, were located 50 cm apart from one another. Single photons are measured shot-by-shot. That is, for each random measurement setting, a short burst of detection events are collected and time-tagged. From this set of detections, only the very first joint detection is kept and the others are discarded.

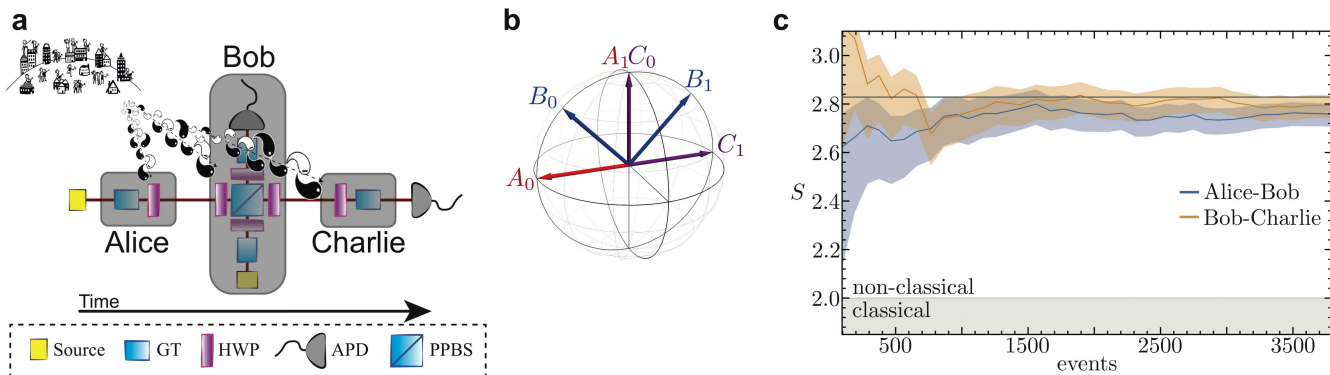
During the Big Bell test event, bits were acquired at a rate of 4 bps for 24 hours. A random four bit sequence represents one of $n = 16$ measurement settings per side. After performing all sixteen measurements, the following steering inequality was calculated, $S_{16} = \frac{1}{n} \sum_{k=1}^n \langle A_k \sigma_k^B \rangle \leq C_n$ (refs^{52,53}). Here S_{16} is referred to as the steering parameter whilst $A_k \in \{-1, 1\}$ and σ_k^B is Alice's measurement outcome and the Pauli operator corresponding to Bob's measurement setting, respectively. The correlation function is bounded by +1 (maximal correlations) and -1 (maximal anti-correlations) with a value of 0 representing no correlation at all. It should be noted that fair sampling of all the detected photons is assumed. Given these parameters we obtain $S_{16} = 0.965 \pm 0.008$ which beats the bound of $C_{16} = 0.511$ by 57 standard deviations. This is first demonstration of

quantum steering with human-derived randomness.

- ⁵¹ E. Schrödinger, Discussion of probability relations between separated systems, *Mathematical Proceedings of the Cambridge Philosophical Society* **31**, 555 (1935).
- ⁵² D. J. Saunders, S. J. Jones, H. M. Wiseman, G. J. Pryde, Experimental EPR-steering using Bell-local states, *Nature Physics* **6**, 845 (2010).
- ⁵³ A. J. Bennet, *et al.*, Arbitrarily loss-tolerant Einstein-Podolsky-Rosen steering allowing a demonstration over 1 km of optical fiber with no detection loophole, *Phys. Rev. X* **2**, 031003 (2012).

② Quantum Correlations in Time

Authors: Martin Ringbauer and Andrew White



Supplementary Figure 4: (a) The experimental setup used to test temporal quantum correlations. A single photon, produced by spontaneous parametric down-conversion (not shown), is first measured by Alice, then by Bob and finally by Charlie. Alice and Charlie use a half-wave plate (HWP) and a Glan-Taylor polarizer (GT) for their measurement, while Bob, being in the middle, measures the quantum system indirectly by entangling it to another photon (using a partially polarizing beam splitter, PPBS) and detecting that photon in an avalanche photo diode (APD). All three observers choose their measurement settings by using the human-generated random bits supplied by the Bellsters to set the angles of their waveplates. **(b) Measurement settings.** Visualization of the measurement settings for Alice (A_0, A_1), Bob (B_0, B_1), and Charlie (C_0, C_1) on the single-qubit Bloch-sphere. **(c) Experimental results.** Shown is the cumulative violation of the CHSH inequality in time between the observers Alice and Bob (blue), and between Bob and Charlie (orange) versus the number of observed detection events. The data shows a strong violation of both inequalities and thus a violation of entanglement monogamy. The shaded regions correspond to 1σ statistical confidence intervals.

In the scenario originally considered by Bell⁵⁴, pairs of entangled particles are shared between spacelike-separated observers, Alice and Bob, who perform local measurements on their particles. Under the conditions of realism, measurement independence, and local causality, the correlations between Alice's and Bob's measurement outcomes must then satisfy a set of Bell inequalities. The correlations of entangled quantum systems, on the other hand, violate these inequalities and are thus said to be in conflict with the above assumptions. These correlations, however, are not limited to spacelike-separated scenarios, which are in fact quite challenging to achieve in practice.

In the present experiment, we thus do not probe traditional Bell-nonlocality, but instead consider correlations between measurements performed on the same quantum system at different points in time. Such experiments are expected to reveal correlations that resemble those observed in entangled quantum systems, and just like in the spatial case, we can test for these correlations using a Bell-inequality⁵⁵

$$S_{AB} = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2, \quad (5)$$

where A_0 and A_1 (B_0 and B_1) are Alice's (Bob's) measurement settings. This inequality is derived under equivalent assumptions as the spatial case, with local causality replaced by an analogous temporal no-fine-tuning assumption, see Ref.^{56,57} for details.

Experimentally we test temporal quantum correlations using the setup in Suppl. Fig. 4a, where a single photon is subject to a sequence of three polarization measurements, first by Alice, then Bob, and finally Charlie. Pairs of single photons at a wavelength of 820 nm are produced via spontaneous parametric downconversion in a β -Barium borate (BBO) crystal, pumped by a femtosecond-pulsed Ti:Sapphire laser at a wavelength of 410 nm and a repetition rate of 76 MHz. One of these photons acts as the system which is subject to a series of projective measurements by Alice, then Bob, and finally Charlie. The second photon is used as an ancilla for Bob's measurement, which is implemented in a non-destructive fashion. Specifically, Bob entangles the system to the ancilla (the meter) using a non-deterministic controlled-NOT gate based on non-classical interference on a partially polarizing beamsplitter, and then measures the meter photon in the computational basis. Operating the experiment at low pump power to suppress higher-order emissions with a $g^{(2)}(0) \sim 10^{-2}$, the gate achieved a relative Hong-Ou-Mandel interference visibility of 1.00 ± 0.01 . Through local rotations of the system before and after the entangling operation, this design thus allowed for arbitrary non-destructive measurements on the system with a fidelity of $\overline{F} = 0.98 \pm 0.02$ with the ideal projective measurement, and an average purity of $\overline{P} = 0.97 \pm 0.03$, as determined using quantum process tomography. The remaining imperfections can be attributed to imperfect mode overlap in the entangling gate, limiting the entanglement generation between Bob and Charlie to a concurrence of $\mathcal{C} = 0.970 \pm 0.006$.

The measurement settings for each party, shown on the Bloch-sphere in Suppl. Fig. 4b, are set through waveplate rotations driven by the human-generated random numbers supplied by the Bellsters. Due to the sequential nature of the experiment a gated detection scheme was chosen, where for each waveplate setting a pair of avalanche photo diodes (APDs) recorded data for 100ms. This resulted in an average event-rate of 1.60 ± 0.07 detected photons per gate window for all combinations of measurement settings. In contrast to the other experiments in the BBT, this sometimes produced multiple events from the same human-generated bits. It should be noted that the experiment in its current form is subject to a detection loophole, and fair

sampling is thus assumed for all observed events. Closing the detection loophole would require not just more efficient detectors, but also an event-ready version of the intermediate non-destructive measurement, instead of the current non-deterministic design. On the other hand, since the experiment is explicitly timelike separated, it is not subject to a locality loophole in the usual sense. There is, however, the related assumption that there is no hidden (i.e. fine-tuned) communication channel between the time steps⁵⁸.

Suppl. Fig. 4c shows the cumulative statistics for inequality (5) between Alice and Bob, and between Bob and Charlie as a function of the number of recorded events. The observed values of $S_{AB} = 2.75 \pm 0.05$ and $S_{BC} = 2.79 \pm 0.05$ not only demonstrate a violation of Eq. (5), but also constitute the first experimental demonstration of a key difference between spatial and temporal quantum correlations. In the spatial case entanglement between Alice and Bob precludes either of them from being entangled with a third party, which is known as monogamy of entanglement⁵⁹. In contrast, our results show that polygamy is possible in the temporal case: Bob can violate a CHSH-inequality with Alice, and, at the same time, with Charlie. This experiment is thus a first step towards exploring the rich structure of multipartite quantum correlations in time, which remains widely unexplored. Understanding the relationship between temporal and spatial correlations might be important for the design of future distributed quantum networks, where both kinds of correlations are expected to play a role. Temporal correlations may also be useful for temporal quantum communication⁵⁵ and computation⁶⁰ schemes. In light of these potential applications it would be of particular interest to explore more complex scenarios involving genuine multipartite temporal correlations and correlations arising from generalized measurements.

⁵⁴ J. S. Bell, On the Einstein Podolsky Rosen Paradox, *Physics* **1**, 195 (1964).

⁵⁵ C. Brukner, S. Taylor, S. Cheung, V. Vedral, Quantum Entanglement in Time, *arXiv:quant-ph/0402127* (2004).

⁵⁶ C. J. Wood, R. W. Spekkens, The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning, *New J. Phys.* **17**, 033002 (2015).

⁵⁷ F. Costa, M. Ringbauer, M. E. Goggin, A. G. White, A. Fedrizzi, A Unifying Framework for spatial and temporal quantum correlations, *arXiv:1710.01776* (2017).

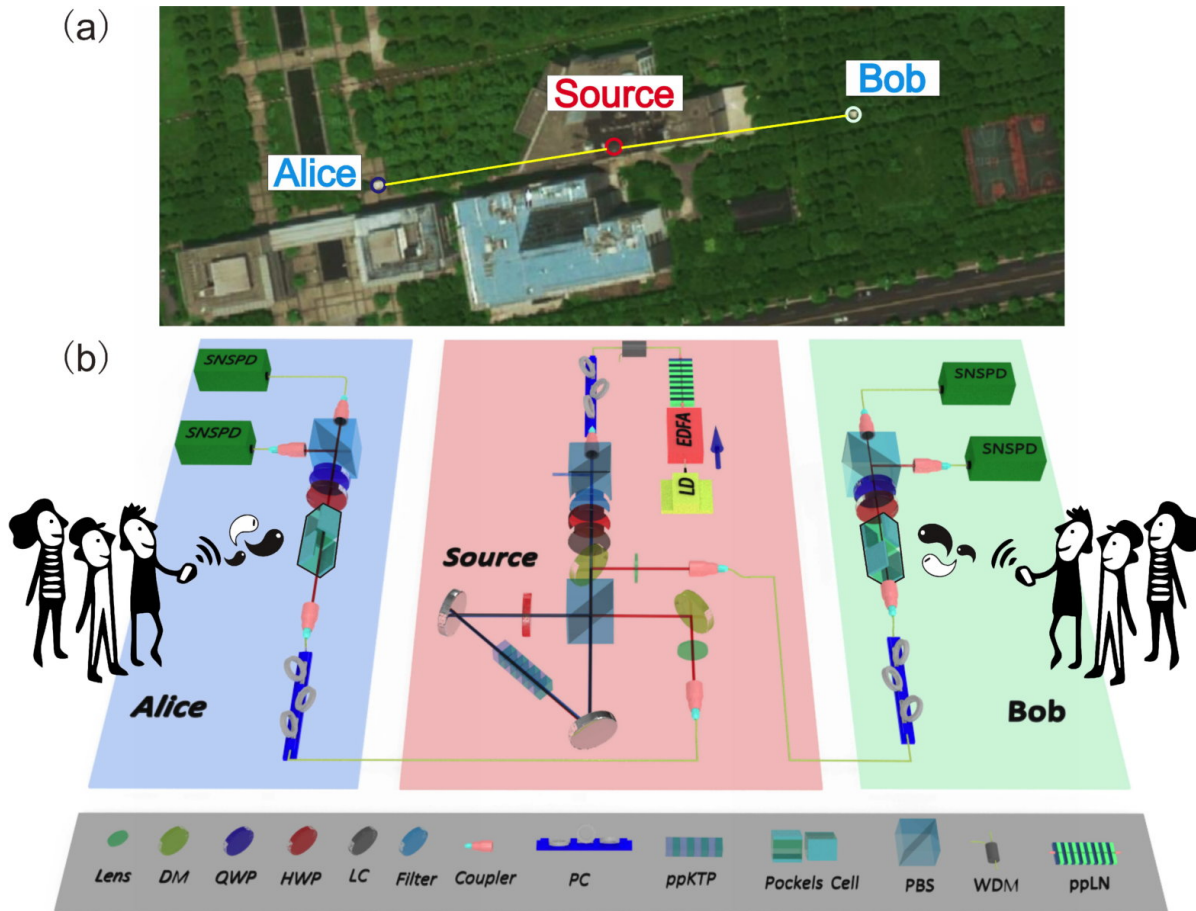
⁵⁸ M. Ringbauer, R. Chaves, Probing the Non-Classicality of Temporal Correlations, *Quantum* **1**, 35 (2017).

⁵⁹ V. Coffman, J. Kundu, W. K. Wootters, Distributed entanglement, *Phys. Rev. A* **61**, 052306 (2000).

⁶⁰ M. Markiewicz, A. Przysiężna, S. Brierley, T. Paterek, Genuinely multipoint temporal quantum correlations and universal measurement-based quantum computing, *Phys. Rev. A* **89**, 062319 (2014).

③ Bell tests with imperfectly random human input

Authors: Yang Liu, Xiao Yuan, Cheng Wu, Weijun Zhang, Jian-Yu Guan, Jiaqiang Zhong, Hao Li, Ming-Han Li, Sheng-Cai Shi, Lixing You, Zhen Wang, Xiongfeng Ma, Qiang Zhang and Jian-Wei Pan



Supplementary Figure 5: Bell test using imperfect input randomness. (a) A bird's-eye view of the entanglement source, Alice's detection and Bob's detection. The distances from the source to Alice and Bob are 87 ± 2 m and 88 ± 2 m, respectively. (b) Schematic setup of the Bell test. A distributed feedback (DFB) laser diode (LD) at $\lambda = 1560$ nm is modulated to produce pulses with a repetition rate of 100 kHz and a pulse width of 10 ns. The pulses are amplified with an erbium-doped fiber amplifier (EDFA) and then are up-converted to 780 nm via second-harmonic generation (SHG) in an in-line periodically poled lithium niobate (PPLN) waveguide. The residual 1560 nm light is filtered with a wavelength-division multiplexer (WDM) and a shortpass filter. After adjusting the polarization using a half-wave plate (HWP) and a liquid crystal (LC), the 780 nm pump light is focused to a periodically poled potassium titanyl phosphate (PPKTP) crystal in a "Sagnac" geometry to generate entangled photon pairs. A series of dichroic mirrors (DMs) are used to remove the residual pump light at 780 nm and fluorescence before the entangled pairs are collected. In Alice's and Bob's detection stations, a polarization controller (PC), a quarter-wave plate (QWP), a HWP and a polarizing beam splitter (PBS) are used to set the measurement angle. Random numbers control the Pockels cells to dynamically select the bases. Superconducting nanowire single-photon detectors (SNSPD) are used to detect the photons after the PBS.

Humans are not perfectly random, and tend to produce patterns that make their choices somewhat predictable. For example, we ran the NIST statistical test suite⁶¹ on the human-generated random numbers from the BBT and of the 14 different tests for uniformity, the human random numbers only passed 2. Nevertheless, human randomness is very attractive for Bell tests because of the element of human free will; if this freedom exists, the human choices are not controlled by hidden variables. Remarkably, it is possible to say how well the hidden variable would have to control the human choices to explain a Bell inequality within local realism. If a, b and x, y denote the binary outputs and inputs, respectively, then the imperfection of the input randomness can be characterized by a bound on the conditional setting probability $P(xy|\lambda)$

$$I = \min_{xy\lambda} P(xy|\lambda). \quad (6)$$

where $I \in [0, 1/4]$ and a smaller value of I indicates more imperfection of the input randomness. We report two Bell tests, one using the well-known Clauser-Horne-Shimony-Holt (CHSH)⁶² inequality and the other using the measurement dependent local (MDL)⁶³ inequalities. The MDL is designed to be more robust against influence on the settings. We observe large violations of the two Bell inequalities and give the I required of the input human randomness to rule out local realism.

Our experimental setup is shown in shown in Suppl. Fig. 5(b). A 1560 nm seed laser with frequency $f = 100$ kHz, width $\Delta t = 10$ ns is amplified and up-converted to 780 nm via second-harmonic generation (SHG). Pumped with this laser, entangled pairs are generated in the Sagnac based setup, and then collected into single mode fibers for detection. The measurement devices in Alice's and Bob's detection station are around 90 meters from the source, as shown in Suppl. Fig. 5(a). The spatial separation makes sure the measurement in Alice's detection station will not affect that in Bob's detection station, and vice versa. A field-programmable gate array (FPGA) board is used to generate synchronizing signals and to distribute the random numbers in real time. ICFO provided us the human generated random numbers, which we re-distribute to modulate the basis immediately. Note that the system only works when there is a batch of random numbers comes in. The random numbers control the Pockels cell by applying a zero or a half-wave voltage, setting the basis to A_0/A_1 for Alice and B_0/B_1 for Bob. Passing through the PBS, the photons are finally detected with a superconducting nanowire single-photon detectors (SNSPD), and recorded by a time-digital convertor (TDC) for off-line data analysis.

First, the maximum entangled state $|\Psi^+\rangle = (|HV\rangle + |VH\rangle)/\sqrt{2}$ is generated and tested. We measure the visibility in the horizontal/vertical basis as 99.2% and in the diagonal/anti-diagonal basis as 98%. State tomography shows the fidelity to the ideal state is around 97.5%. We test the standard CHSH inequality with quantum random numbers, calculate the value $S = 2.804$, with $E(A_1, B_1) = -0.751$, $E(A_1, B_2) = 0.651$, $E(A_2, B_1) = 0.657$, $E(A_2, B_2) = 0.745$ where $E(A_x, B_y) \equiv \sum (-1)^{a+b+xy} p(ab|xy)$. Considering MDL correlation with the randomness measure I , that is $P(xy|\lambda) \geq I, \forall xy\lambda$, the maximal achievable CHSH value with MDL correlation is $4(1 - 2I)^{63,64}$. To achieve the experimentally obtained Bell value with MDL correlation, we thus need $4(1 - 2I) < S = 2.804$ and thus $I < 0.1495$. In other words, if the input human randomness is high enough, i.e., if $I > 0.1495$, the experimentally observed data cannot be explained with any MDL correlation.

Next, the human random numbers are used to test the MDL inequality⁶³ which is defined by $IP(0000) - (1 - 3I)[P(0101) + P(1010) + P(0011)] \leq 0$. Interestingly, as long as $P(xy|\lambda) \geq I$, such an inequality cannot be violated with any MDL correlation. In this case, observing the violation of the MDL inequality indicates the existence of non-classical correlation even in the presence of measurement dependence.

The non-maximal entangled state $\cos(69.1^\circ)|HV\rangle + \sin(69.1^\circ)|VH\rangle$ is generated. State tomography shows the fidelity to the ideal state is 98.7%. Using the setup described above, the locality loophole is closed by considering the spatial distance and the randomness loophole is closed with human free will. The experiment was performed during 30 Nov. and 1 Dec. when the public helps us to generate random numbers. The coincidence values were tested with human-generated bits determining the basis choices. The experimental data are divided into to several one-hour periods for statistical analysis.

As this MDL inequality cannot be violated with MDL correlation satisfying $P(xy|\lambda) \geq I, \forall xy\lambda$, we calculate I_0 , the smallest possible value of I that satisfies the inequality. In experiment, we obtain $I_0 = 0.10 \pm 0.05$ for MDL inequality by using human random numbers. As a comparison, we also use quantum random numbers to choose the basis. With the rest of the setup remaining the same, we obtain $I_0 = 0.106 \pm 0.007$ which gives a similar I value with less fluctuation. This is because the total amount of data using quantum random numbers is much larger than using human random numbers. The input of human random numbers was at most a few thousand per second whereas for quantum random numbers we had a steady 100 kbps for controlling the basis. In total, we accepted and used around 80 Mb of human random numbers for the MDL inequality test in two days. For comparison, we tested the inequality with quantum random numbers in around one and half hours, using more than 500 Mbits.

⁶¹ A. Rukhin, *et al.*, A statistical test suite for random and pseudorandom number generators for cryptographic applications, National Institutes of Standards and Technology, Special Publication (NIST SP) - 800-22 Rev 1a (2010).

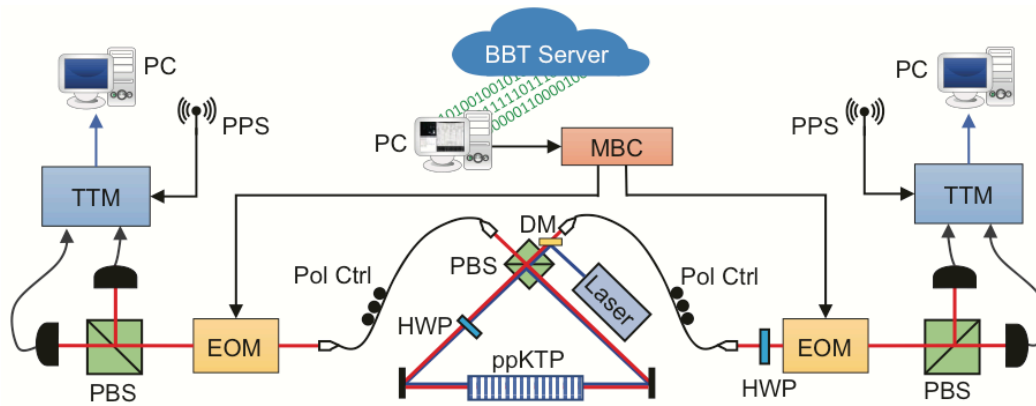
⁶² J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).

⁶³ G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, N. Gisin, Arbitrarily small amount of measurement independence is sufficient to manifest quantum nonlocality, *Phys. Rev. Lett.* **113**, 190402 (2014).

⁶⁴ X. Yuan, Q. Zhao, X. Ma, Clauser-horne bell test with imperfect random inputs, *Phys. Rev. A* **92**, 022107 (2015).

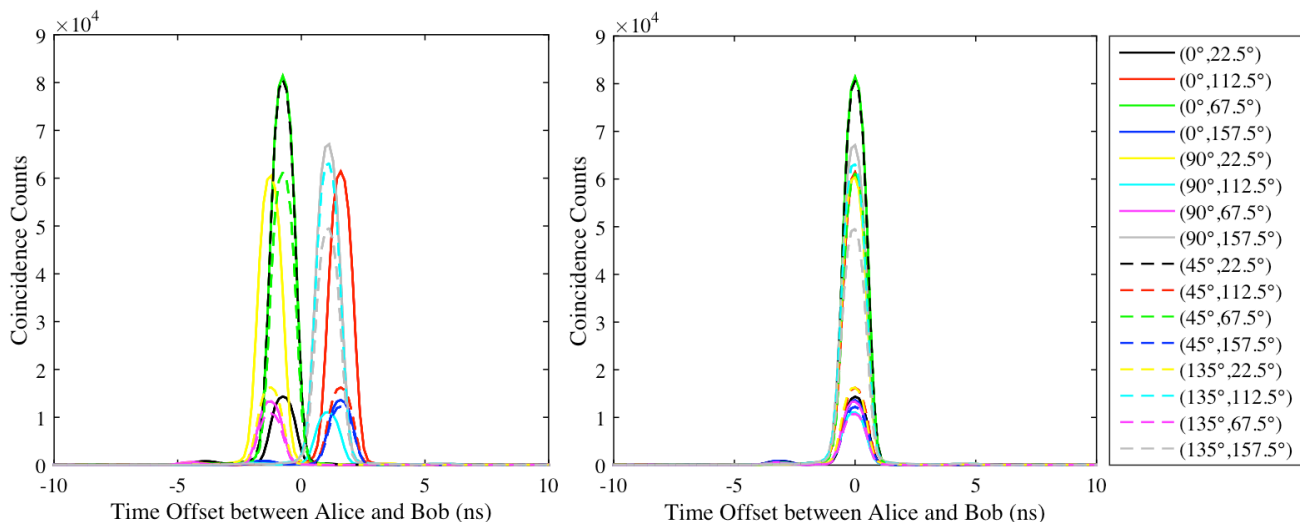
④ Violation of a Bell Inequality using Entangled Photons and Human Random Numbers

Authors: Bo Liu, Johannes Handsteiner, Dominik Rauch, Rupert Ursin, Thomas Scheidl and Anton Zeilinger



Supplementary Figure 6: Experimental setup. The wavelength of the laser is 405nm. MBC: measurement basis controller; Pol Ctrl: polarization controller; DM: dichroic mirror; HWP: half-wave plate; EOM: electro-optical modulator; PBS: polarized beam splitter; TTM: time tagging module; PPS: pulse per second.

In the Big Bell test (BBT) experiment, we tested the CHSH form of the Bell inequality with polarization-entangled photon pairs. As shown in Suppl. Fig. 6, the entangled photon source was based on a Sagnac interferometer generating polarization-entangled photon pairs in the maximally entangled state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|H_A V_B\rangle - |V_A H_B\rangle)$. Each photon pair was guided to two receivers, called Alice and Bob, via single mode fibers. At each receiver, the entangled photons were coupled out of the fiber and guided to an electro-optical modulator (EOM), which allowed for fast switching between complementary measurement bases. The EOM was followed by a polarizing beam splitter (PBS) with a single-photon avalanche diode (SPAD) detector in each output port. The EOM settings were driven by the measurement basis controller (MBC), which read the random bits from the BBT server and converted them in real-time to the corresponding EOM control signals. In this way, real-time measurements in the following linear polarization bases have been implemented for Alice (Bob): $0^\circ/90^\circ$ ($22.5^\circ/112.5^\circ$) if the random number was '0' and $45^\circ/135^\circ$ ($67.5^\circ/157.5^\circ$) if the random number was '1'. Finally, using a time tagging module (TTM) and a computer, Alice and Bob generated and recorded time stamps of all SPAD detection events, the EOM settings and an additional shared pulse per second (PPS) signal.



Supplementary Figure 7: Coincidence counts between Alice and Bob for the experiment with human random numbers (HRN) mixed with quantum random numbers (QRN).

On the BBT day, we performed measurements using all three different types of random numbers provided via the BBT server, i.e. real-time human random numbers (HRN), quantum random numbers (QRN) and human random numbers from the database (DB). In the first experiment, we used a mix of HRN and QRN for a total measurement time of 200 seconds. During post-processing, we separated the results obtained with HRN from the results obtained with QRN. For the HRN (QNR) data we obtained a Bell value of 2.6387 (2.6434) with a 81 (115) sigma violation of the local-realistic bound of 2. The second

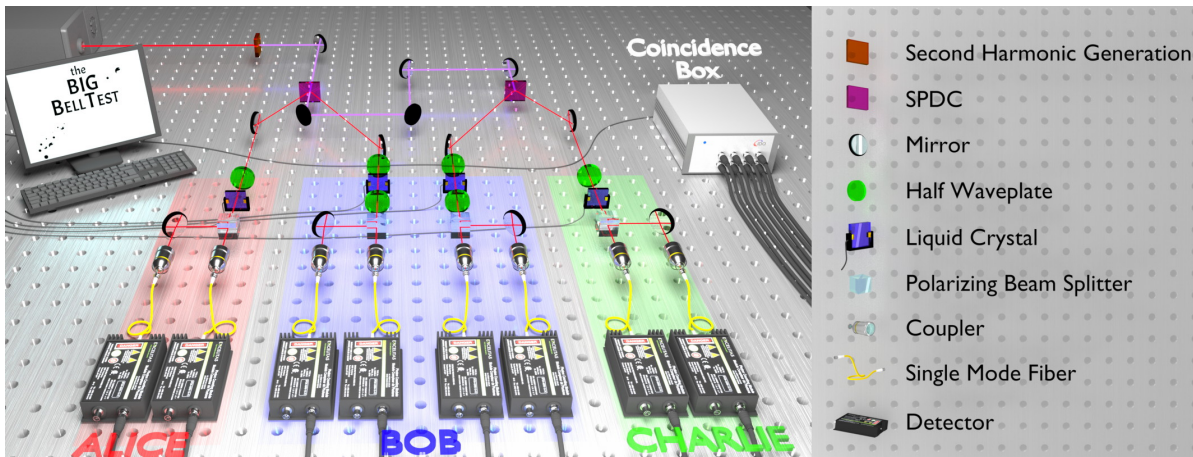
experiment was performed with a mix of HRN and DB for a total measurement time of 210 seconds. For the HRN (DB) data we obtained a Bell value of 2.6403 (2.6371) with a 63 (122) sigma violation of the local-realistic bound of 2. Errors are calculated assuming Poissonian photon statistics. In our experiments, the time delay between Alice's and Bob's time-tags slightly varied for the different detector combinations as can be seen from the coincidence peaks shown in Suppl. Fig. 7. However, this could easily be compensated during data post-processing, such that we could use a coincidence window of 1.25ns. All results are summarized in detail in Table II.

TABLE II: Detailed measurement Results.

Exp #	Time (s)	Source	RN rate (bps)	Total CCs	P1	P2	P3	P4	S Value	Xigma	Visibility
1	200	HRN	1612.88	160178	0.6520	-0.6947	-0.6592	-0.6329	2.6387	81.0265	0.9329
		QRN	2727.68	294465	0.6560	-0.6938	-0.6544	-0.6391	2.6434	115.8174	0.9346
2	210	HRN	910.21	95024	0.6500	-0.7009	-0.6539	-0.6355	2.6403	63.0984	0.9335
		DB	3404.49	337463	0.6498	-0.6984	-0.6589	-0.6300	2.6371	121.7927	0.9324

⑤ Experimental bilocality violation with human randomness

Authors: Luca Santodonato, Gonzalo Carvacho, Francesco Andreoli, Marco Bentivegna, Rafael Chaves and Fabio Sciarrino



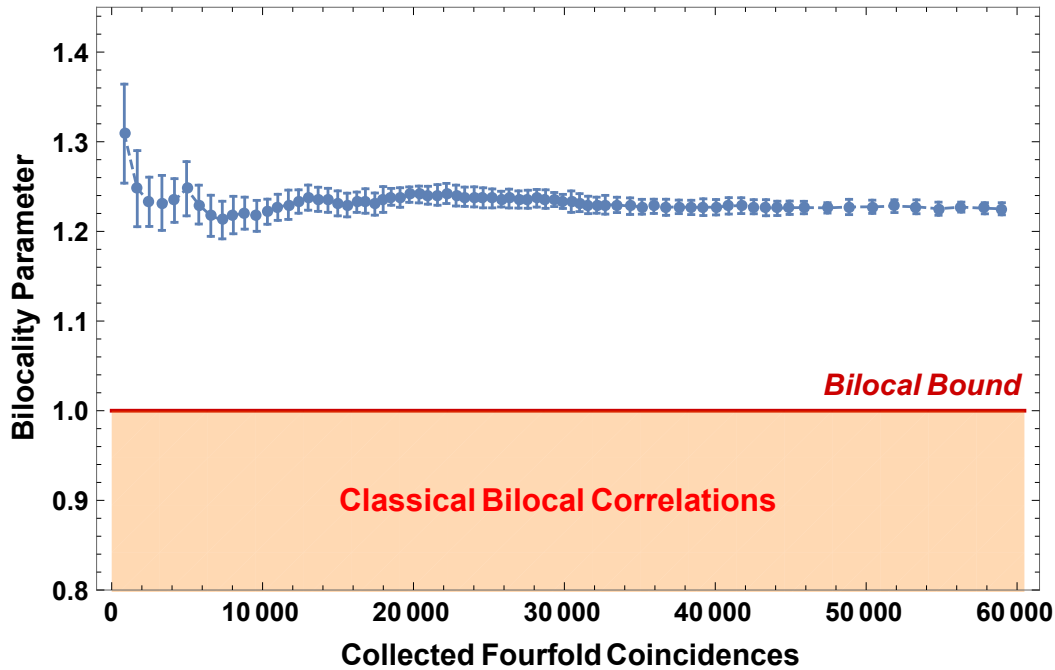
Supplementary Figure 8: Experimental setup for bilocality test exploiting human randomness. Two couples of photons are independently generated in the singlet state by a pulsed pump laser. For each pair, one photon is sent to one of the furthest stations (A or C, respectively), while the other one reaches the central node B. All parties perform single qubit polarization analysis, exploiting Half wave-plates (HWPs) and Liquid Crystals (LC). The measurement choice is controlled by the random bits provided by the ICFO BBT cloud infrastructure, which drives the voltage applied on the LC depending on the value of three random bits.

Nowadays, complex networks are attracting a large interest for the development of future quantum technologies but also for the understanding of generalized Bell-like inequalities. Quantum information processing will certainly rely on quantum networks which until now have been exploited for studies ranging from generalized Bell non-locality⁶⁵⁻⁷³, distribution and control of entanglement⁷⁴⁻⁷⁶, non-Markovian dynamics⁷⁷ to quantum network of clocks with GHZ states⁷⁸. Here by considering a tripartite scenario where two parties (Alice and Charlie) share entangled states (independently generated) with a third one (Bob), we are able to witness a new kind of non-classical behavior called quantum non-bilocality^{65,66}, where the results are incompatible with any possible model described by local realism, two independent hidden variables and the free-will assumption. Alice, Bob and Charlie perform dichotomic measurements respectively labeled by the variables x , y and z , and they return some binary outcomes described by the variables a , b and c . Since we are considering two physically different and independent entangled sources we must consider also two different sets of hidden variables λ_1 and λ_2 . Any correlation allowed by this model can be written as $p(a, b, c|x, y, z) = \sum_{\lambda_1, \lambda_2} p(\lambda_1)p(\lambda_2)p(a|x, \lambda_1)p(b|y, \lambda_1, \lambda_2)p(c|z, \lambda_2)$. The set of correlations compatible with the described model is constrained by the so-called bilocality inequality: $B = \sqrt{|I|} + \sqrt{|J|} \leq 1$ (with $I = \frac{1}{4} \sum_{x,z} \langle A_x B_0 C_z \rangle$, $J = \frac{1}{4} \sum_{x,z} (-1)^{x+z} \langle A_x B_1 C_z \rangle$, (summing on $x, z = 0, 1$) and where: $\langle A_x B_y C_z \rangle = \sum_{a,b,c=0,1} (-1)^{a+b+c} p(a, b, c|x, y, z)$), which

can be violated by quantum mechanics⁷⁹⁻⁸¹. In our experiment we implement a photonic network consisting of three parties and two independent sources of quantum states (see Suppl. Fig. 8) and we evaluate the bilocality parameter B , exploiting random measurements settings driven by the bits provided by the Big Bell Test experiment.

The choice of measurement settings in our stations was physically implemented by the use of liquid crystals, driven in real time by the human random numbers provided by web users during the BBT day. We kept each setting for 5 s, recorded the coincidences in that time and then received fresh random bits for the next setting. This allowed us to witness a strong violation of the bilocality inequality of $B = 1.2251 \pm 0.0066$ by cumulative evaluation of the B parameter for a growing number of collected coincidences (see Suppl. Fig. 9) observing a violation of almost 34 standard deviations. The obtained violation provides clear evidence that, up to the loopholes present in our implementation (low detection efficiencies and the locality loophole), the correlations generated in our experiment cannot be explained by any classical model based on the assumptions of local realism, free will of the participants and the independence of two physically different source of states. That is, the real-time (without post-processing of the data) violation of the bilocality inequality shows that bilocal classical models augmented with free-will are incompatible with our experimental data. This experiment provides a proof-of-principle for generalizations of Bell's theorem to more complex causal networks (involving more than one source of states) and thus can pave the way towards the implementation of new quantum information protocols exploiting non-bilocal quantum correlations as a novel resource.

⁶⁵ C. Branciard, N. Gisin, S. Pironio, Characterizing the nonlocal correlations created via entanglement swapping, *Phys. Rev. Lett.* **104**, 170401 (2010).

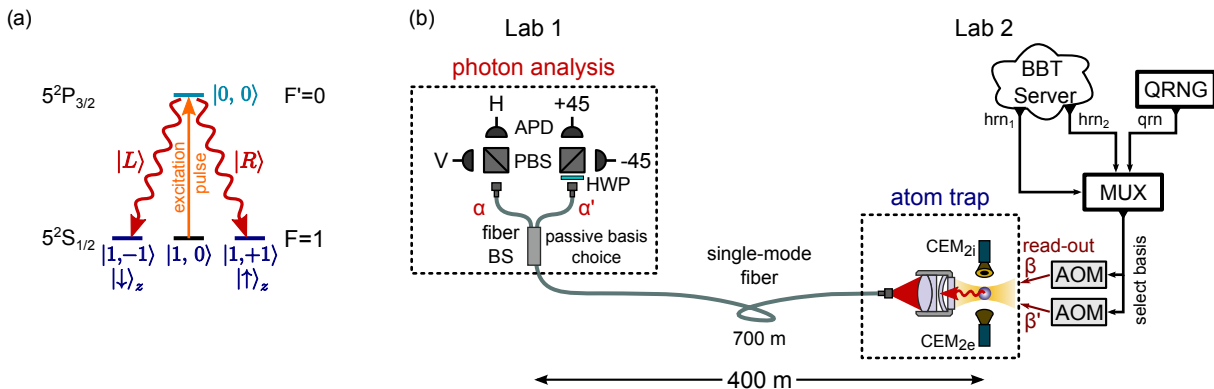


Supplementary Figure 9: Experimental bilocal violation with human randomness. We cumulatively evaluated the bilocal parameter \mathcal{B} and the total number of fourfold coincidences collected. A red line shows the bound of the bilocal inequality, which delimits the classical bilocal region depicted with a shaded red area.

- ⁶⁶ C. Branciard, D. Rosset, N. Gisin, S. Pironio, Bilocal versus nonbilocal correlations in entanglement-swapping experiments, *Phys. Rev. A* **85**, 032119 (2012).
- ⁶⁷ R. Chaves, Polynomial bell inequalities, *Phys. Rev. Lett.* **116**, 010402 (2016).
- ⁶⁸ T. Fritz, Beyond bell's theorem: correlation scenarios, *New Journal of Physics* **14**, 103001 (2012).
- ⁶⁹ A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, A. Acín, Nonlocal correlations in the star-network configuration, *Phys. Rev. A* **90**, 062109 (2014).
- ⁷⁰ D. Rosset, *et al.*, Nonlinear bell inequalities tailored for quantum networks, *Phys. Rev. Lett.* **116**, 010403 (2016).
- ⁷¹ A. Tavakoli, Bell-type inequalities for arbitrary noncyclic networks, *Phys. Rev. A* **93**, 030101 (2016).
- ⁷² N. Gisin, Q. Mei, A. Tavakoli, M. O. Renou, N. Brunner, All entangled pure quantum states violate the bilocal inequality, *Phys. Rev. A* **96**, 020304 (2017).
- ⁷³ F. Andreoli, G. Carvacho, L. Santodonato, R. Chaves, F. Sciarrino, Maximal qubit violation of n-locality inequalities in a star-shaped quantum network, *New Journal of Physics* **19**, 113020 (2017).
- ⁷⁴ S. Perseguers, G. J. L. Jr, D. Cavalcanti, M. Lewenstein, A. Acín, Distribution of entanglement in large-scale quantum networks, *Reports on Progress in Physics* **76**, 096001 (2013).
- ⁷⁵ S. M. Hein, F. Schulze, A. Carmele, A. Knorr, Entanglement control in quantum networks by quantum-coherent time-delayed feedback, *Phys. Rev. A* **91**, 052321 (2015).
- ⁷⁶ D. Cavalcanti, M. L. Almeida, V. Scarani, A. Acín, Quantum networks reveal quantum nonlocality, *Nature comms.* **2**, 184 (2011).
- ⁷⁷ T. Ramos, B. Vermersch, P. Hauke, H. Pichler, P. Zoller, Non-markovian dynamics in chiral quantum networks with spins and photons, *Phys. Rev. A* **93**, 062104 (2016).
- ⁷⁸ P. Komar, *et al.*, A quantum network of clocks, *Nat Phys* **10**, 582 (2014).
- ⁷⁹ G. Carvacho, *et al.*, Experimental violation of local causality in a quantum network, *Nature comms.* **8**, 14775 (2017).
- ⁸⁰ D. J. Saunders, A. J. Bennet, C. Branciard, G. J. Pryde, Experimental demonstration of nonbilocal quantum correlations, *Science Advances* **3** (2017).
- ⁸¹ F. Andreoli, *et al.*, Experimental bilocal violation without shared reference frames, *Phys. Rev. A* **95**, 062315 (2017).

⑥ Violation of Bell's inequality with a single atom and single photon entangled over a distance of 400 m

Authors: Kai Redeker, Robert Garthoff, Daniel Burchardt, Harald Weinfurter, Wenjamin Rosenfeld



Supplementary Figure 10: (a) Scheme of the atomic levels involved in the entanglement between the spin state of the atom and polarization of the photon. The atomic qubit consists of Zeeman spin states $|m_F = \pm 1\rangle$ of $5^2S_{1/2}$, $F = 1$ hyperfine ground level. (b) Overview of the experiment. A single atom, trapped in Lab 2, is entangled with a single photon. The photon undergoes a polarization analysis in Lab 1 where the measurement direction $a \in \{\alpha, \alpha'\}$ is passively selected at a beam-splitter (fiber BS). The measurement direction on the atomic spin $b \in \{\beta, \beta'\}$ is actively selected using input from the BBT server and a local quantum random number generator (QRNG).

We demonstrate a violation of a CHSH Bell inequality with a spin of a single trapped ^{87}Rb atom which is entangled with the polarization state of a single photon. The random numbers generated by human contributors were used here to decide which measurement direction had to be applied to the atomic spin. Our experiment employs 3 different kinds of randomness: randomness in detecting a photon in a certain output of an optical beam-splitter, a quantum random number generator and human-generated random numbers.

In the experiment the generation of entanglement between the spin of the atom and polarization of the photon⁸² is performed by optical excitation as shown in Suppl. Fig. 10(a). The subsequent spontaneous decay leads to the atom-photon state $|\Psi\rangle_{AP} = \frac{1}{\sqrt{2}}(|\downarrow\rangle_z |L\rangle + |\uparrow\rangle_z |R\rangle)$ where $|\uparrow\rangle_z$ and $|\downarrow\rangle_z$ are the atomic spin states and $|L\rangle$, $|R\rangle$ denote the left- and right-circular polarization states of the photon (corresponding to $|\downarrow\rangle_z$, $|\uparrow\rangle_z$ states of the photonic spin, respectively). The emitted photon is coupled into a single-mode fiber and guided to another laboratory which is located at a distance of 400 m, Suppl. Fig. 10(b). There the photon passes a polarization-independent 50 : 50 beam-splitter (BS) where the (passive) choice is made which basis will be used for the photon measurement according to probabilistic rules of quantum mechanics. In one output of the BS the measurement is performed in the H/V basis by means of a polarizing beam-splitter (PBS) and two avalanche photo-diodes (APDs). In the other output of the BS the measurement is performed in the $\pm 45^\circ$ basis with the help of an additional half-wave retarder plate (HWP) oriented at 22.5° .

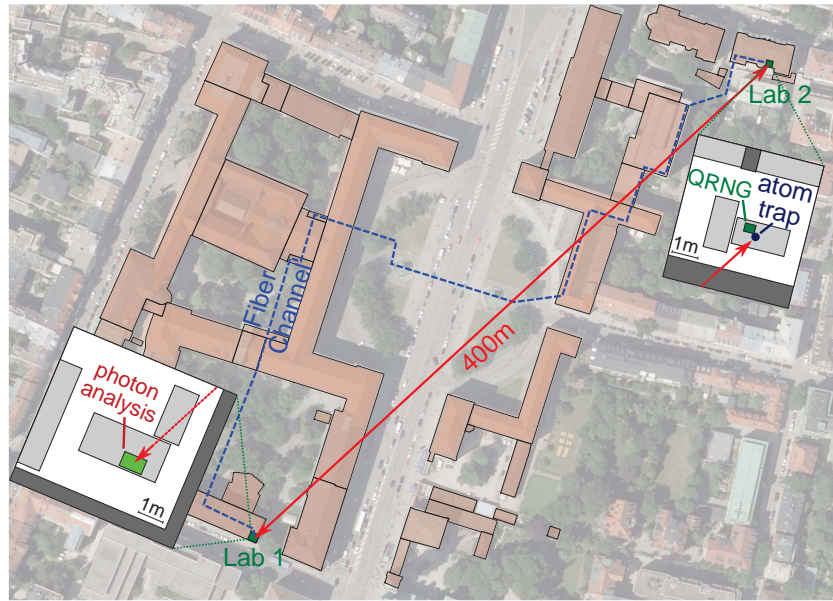
After the state of the photon was measured the read-out of the atomic state is performed. Here, for each measurement round, two random numbers are retrieved from the server. The first one (hrr_1) decides whether in the current round the measurement direction shall be selected by a local quantum random number generator (QRNG)⁸³ or a human contributor. For these means the QRNG output (qrn) and the second random number (hrr_2) from the server are fed into a multiplexer (MUX). Its output activates one of the two acousto-optical modulators (AOMs) thereby determining the polarization of an optical read-out pulse applied to the atom. This polarization defines the basis for the analysis of the atomic spin (some details on the read-out process can be found in⁸⁴). The outcomes of the atom and photon measurements are combined into the CHSH Bell parameter $S = |\langle \sigma_\alpha \sigma_\beta \rangle - \langle \sigma_\alpha \sigma_{\beta'} \rangle| + |\langle \sigma_{\alpha'} \sigma_\beta \rangle + \langle \sigma_{\alpha'} \sigma_{\beta'} \rangle|$ in terms of correlators $\langle \sigma_a \sigma_b \rangle = \frac{1}{N_{a,b}} (N_{a,b}^{\uparrow\uparrow} + N_{a,b}^{\downarrow\downarrow} - N_{a,b}^{\uparrow\downarrow} - N_{a,b}^{\downarrow\uparrow})$, $N_{a,b}^{A,B}$ being the number of events with the respective outcomes A, B for measurement directions $a \in \{\alpha, \alpha'\}$, $b \in \{\beta, \beta'\}$.

During the experimental run of about 13 hours 39614 events were collected, 19716 where atomic measurement direction was chosen by a human random number and 19898 events where this choice was done according to the local QRNG. The results are $S = 2.427 \pm 0.0223$ for the first set and $S = 2.413 \pm 0.0223$ for the second, respectively. Both sets show a strong violation of Bell's inequality by more than 18 standard deviations. No statistically significant difference in the results for the two sets can be observed.

⁸² J. Volz, *et al.*, Observation of Entanglement of a Single Photon with a Trapped Atom, *Phys. Rev. Lett.* **96**, 030404 (2006).

⁸³ M. Fürst, *et al.*, High speed optical quantum random number generation, *Opt. Express* **18**, 13029 (2010).

⁸⁴ W. Rosenfeld, *et al.*, Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes, *Phys. Rev. Lett.* **119**, 010402 (2017).

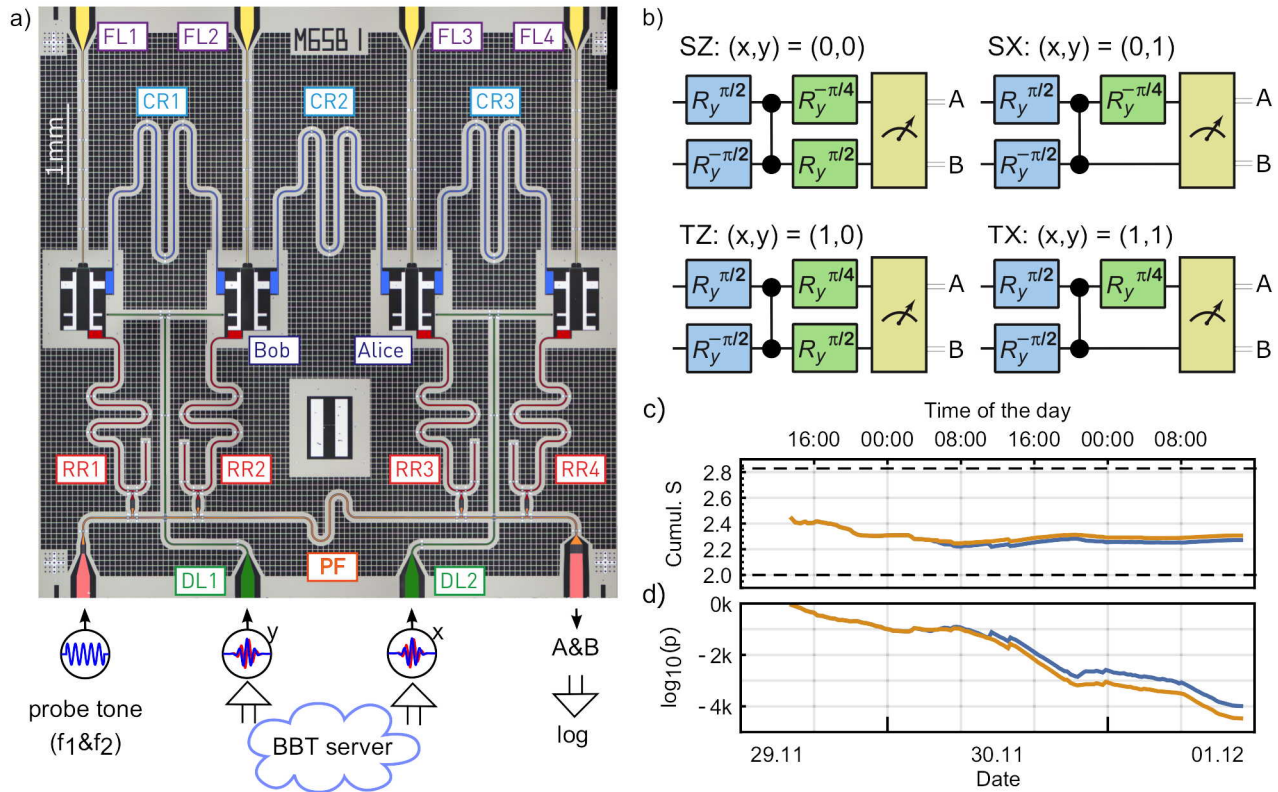


Supplementary Figure 11: Top view of the main campus of LMU. The photon analysis arrangement is located in the of the faculty of physics (Lab 1) while the atomic trap in the basement of the department of economics (Lab 2). Map data was provided by⁸⁵ .

⁸⁵ Bayerisches Landesamt fuer Digitalisierung, Breitband und Vermessung.

⑦ Violation of a Bell inequality using superconducting qubits

Authors: Johannes Heinsoo, Philipp Kurpiers, Yves Salathé, Christian Kraglund Andersen, Adrian Beckert, Sebastian Krinner, Paul Magnard, Markus Oppliger, Theodore Walter, Simone Gasparinetti, Christopher Eichler, Andreas Wallraff



Supplementary Figure 12: (a) False coloured photograph of the superconducting chip used for the Bell test experiment. Each of the four qubits, of which the two central ones were used, has its own flux line (FL) and readout resonator (RR). Neighbouring qubits are coupled through the coupling resonators (CR) with flux controllable effective interaction strength. To perform single qubit gates microwave pulses are applied via drive lines (DL). Probe tones at frequencies f_1 and f_2 are applied to the Purcell filter (PF) to measure the qubit states. (b) Depending on the bit string (x,y) received from the BBT server we execute one of the four quantum gate sequences labeled SZ, SX, TZ and TX. Single qubit gates R_y^ϕ rotate the state vector around the y -axes of the Bloch sphere by an angle ϕ . The state preparation single qubit gates (blue) and the c -PHASE gate (connected black dots) remain unchanged, while the single qubit gates (green) applied prior to the correlation sensitive readout (yellow) depend on the input bits from the server. (c) The cumulative S value and (d) p value as a function of time covering the entire period of the BBT, including all data (blue) and considering only data with successful calibration results (orange). The classical and the non-signalling thresholds are indicated as dashed horizontal lines.

The Bell test in this work is performed using two superconducting qubits^{86–88}, referred to as Alice and Bob, which are located about 1 mm apart from each other as shown in Suppl. Fig. 12a. The qubit transition frequencies are set to $\omega_A/2\pi = 5.963$ GHz and $\omega_B/2\pi = 5.464$ GHz, respectively. Single qubit gates are performed by applying 580 MHz wide DRAG pulses resonant with the respective qubit transition frequencies^{89,90}. The c -PHASE two qubit gate is realized by non-adiabatically tuning the state $|11\rangle$ into resonance with the $|02\rangle$ state for half the oscillation period⁹¹.

For each pair of bits received from the Big Bell test (BBT) server we perform one trial of the experiment by executing one of the four quantum circuits shown in Suppl. Fig. 12b, each of which corresponds to a specific pulse sequence played from an arbitrary waveform generator (AWG). Each of the four sequences comprises the deterministic generation of an entangled state $\frac{1}{\sqrt{2}}(|0+\rangle - |1-\rangle)$ followed by a set of single qubit gates, which rotates the different measurement axes into the computational basis prior to readout. The readout of the qubit state is achieved by measuring the qubit state dependent transmission through the Purcell filter (see Suppl. Fig. 12a) recorded and processed with an analog to digital converter (ADC) and field programmable gate array (FPGA). A single shot fidelity of 96.6% and 93.2% for Alice and Bob respectively is achieved by employing a near-quantum limited parametric amplifier⁹². Between two sequential runs of the experiment, a wait time of $50 \mu\text{s} \approx 5T_1$ allows the qubits to decay back to their ground state. The successful initialization of the qubits using this passive method is verified by reading out the qubit state prior to each experimental run. The data sets with failed state preparation were discarded. About every 20 min both the qubit transition frequencies and two qubit gate parameters were recalibrated. Moreover, Bell state tomography was used to benchmark each round of calibration.

During the 48 hours of continuously running experiments, 16.34 million human generated random numbers were used to perform 8.17 million individual Bell measurements of which 7.69 had successful state initialization and calibration. By counting

the different measured two qubit states for different basis choices we directly evaluate the CHSH inequality⁹³. The resulting S-value is shown in Suppl. Fig. 12c and converges to a final value of 2.271(1) in case all data is included and to 2.307(1) if data sets with failed calibration are ignored. The observed S value is mostly limited by qubit readout fidelity and qubit decay during the 300 ns long pulse sequence.

The confidence of which the experiment violates the CHSH inequality is quantified by estimating the p -values for the data under the hypothesis of a local hidden variable (LHV) model⁹⁴. The estimation based on the Bentkus' inequality⁹⁵ is shown to be tight and accounts for all possible memory effects and input bias⁹⁶. Note that the latter is crucial for reasonable p -value estimation as BBT participants tended to input (1, 0) and (0, 1) pairs twice more often than (1, 1). We observed an exponentially decreasing p -value with increasing number of trials as shown in Suppl. Fig. 12d. The exponentially decreasing p -value is expected as long as the performance of the experiment is not degraded over time, which is ensured by our automated recalibrations. The final p -value of about $\sim 10^{-4000}$ indicates that with high confidence locality is violated either by the incompleteness of a classical description of reality under the additional assumption of free will of the participants or by a locality loophole in our specific experimental setup.

⁸⁶ V. Bouchiat, D. Vion, P. Joyez, D. Esteve, M. H. Devoret, Quantum coherence with a single Cooper pair, *Phys. Scr.* **T76**, 165 (1998).

⁸⁷ J. Koch, *et al.*, Charge-insensitive qubit design derived from the Cooper pair box, *Phys. Rev. A* **76**, 042319 (2007).

⁸⁸ A. A. Houck, *et al.*, Controlling the spontaneous emission of a superconducting transmon qubit, *Phys. Rev. Lett.* **101**, 080502 (2008).

⁸⁹ F. Motzoi, J. M. Gambetta, P. Rebentrost, F. K. Wilhelm, Simple pulses for elimination of leakage in weakly nonlinear qubits, *Phys. Rev. Lett.* **103**, 110501 (2009).

⁹⁰ J. M. Gambetta, F. Motzoi, S. T. Merkel, F. K. Wilhelm, Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator, *Phys. Rev. A* **83**, 012308 (2011).

⁹¹ F. W. Strauch, *et al.*, Quantum logic gates for coupled superconducting phase qubits, *Phys. Rev. Lett.* **91**, 167005 (2003).

⁹² C. Macklin, *et al.*, A near-quantum-limited Josephson traveling-wave parametric amplifier, *Science* **350**, 307 (2015).

⁹³ J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).

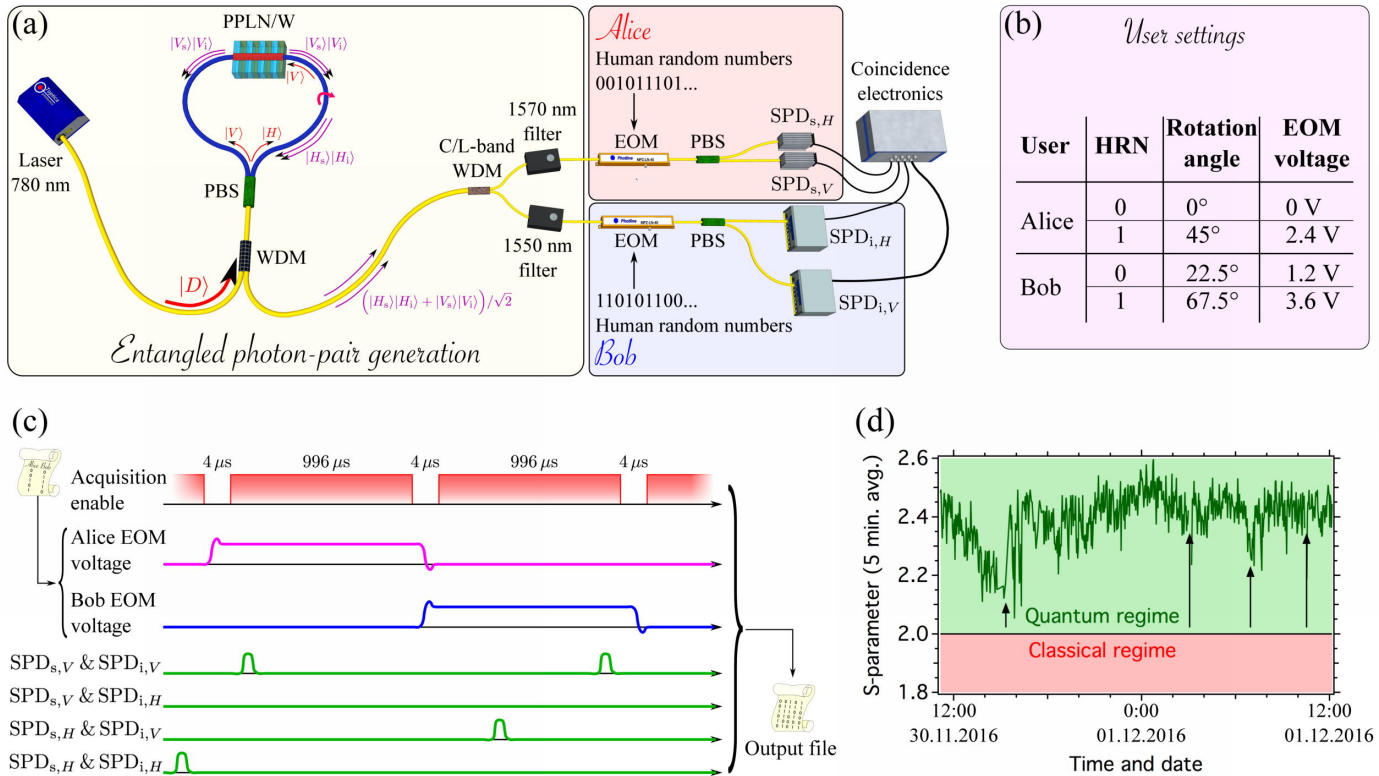
⁹⁴ J.-A. Larsson, Loopholes in Bell inequality tests of local realism, *J. Phys. A: Math. Theor.* **47**, 424003 (2014).

⁹⁵ V. Bentkus, On Hoeffding's inequalities, *Ann. Probab.* **32**, 1650 (2004).

⁹⁶ D. Elkouss, S. Wehner, (Nearly) optimal P values for all Bell inequalities, *npj Quantum Information* **2**, 16026 (2016).

⑧ Telecom-compliant source of polarisation entangled photons for the violation of Bell inequalities driven by human-generated random numbers

Authors: Florian Kaiser, Tommaso Lunghi, Grégory Sauder, Panagiotis Vergeris, Olivier Alibert, and Sébastien Tanzilli



Supplementary Figure 13: (a) Experimental setup. A 780 nm continuous-wave laser pumps a PPLN/W placed in a Sagnac interferometer. This way, polarisation entangled photon pairs are generated in the state $|\psi\rangle = (|H_s\rangle|H_i\rangle + |V_s\rangle|V_i\rangle)/\sqrt{2}$. After being deterministically separated using a C/L-band WDM, the photons are directed to Alice and Bob. Both rotate, accordingly to the HRN, the polarisation state of their respective photon with an EOM. Finally, the polarisation state is analysed using a PBS, followed by two SPDs. (b) User settings that are applied at Alice's and Bob's stations accordingly to the HRN inputs, provided by the Bellsters. (c) Schematic sketch of the data acquisition strategy. (d) Development of the 5-minute averaged S -parameter as a function of time. The four vertical arrows indicate the times at which unexpected incidents occurred to the experiment.

We violate the Bell inequalities by more than 143 standard deviations using polarisation entangled photon pairs^{97–100}. The related experimental setup is shown in Suppl. Fig. 13(a). We developed a photon pair source which is entirely based on guided-wave photonics in order to guarantee the highest possible stability¹⁰¹.

Diagonally polarised light, $|D\rangle$, from a fibre-coupled continuous-wave 780 nm pump laser passes through a wavelength division multiplexer (WDM), which is the fibre optics equivalent of a bulk dichroic mirror. The subsequent fibre polarising beam-splitter (PBS) defines the input and output of a fibre Sagnac loop^{101,102}. Inside the loop, vertically (horizontally) polarised pump light propagates clockwise (anti-clockwise) in polarisation maintaining fibres (indicated by blue lines). By rotating the right-hand side fibre by 90°, the horizontally polarised pump light is effectively turned to be vertical. The pump light is injected in a periodically poled lithium niobate waveguide (PPLN/W) in which vertically polarised pump photons are converted to vertically polarised signal and idler photon pair contributions through type-0 spontaneous parametric downconversion, *i.e.* $|V\rangle \rightarrow |V_s\rangle|V_i\rangle$. Here, the subscripts s and i denote signal and idler photons at wavelengths of $\lambda_s = 1570$ nm and $\lambda_i = 1550$ nm, respectively. The photon pair contribution obtained in the clockwise sense is rotated in the polarisation maintaining fibre, and therefore becomes $|H_s\rangle|H_i\rangle$. Both contributions, *i.e.* $|V_s\rangle|V_i\rangle$ and $|H_s\rangle|H_i\rangle$, are then coherently combined at the PBS and separated from the pump light using the WDM. It is important to note that we choose the pump laser power to be low enough such that the source generates only one photon pair contribution at a time, *i.e.* the probability for generating both contributions $|V_s\rangle|V_i\rangle$ and $|H_s\rangle|H_i\rangle$ simultaneously is negligible. Therefore, the following maximally polarisation entangled photon pair state is generated: $|\psi\rangle = (|H_s\rangle|H_i\rangle + |V_s\rangle|V_i\rangle)/\sqrt{2}$. Signal and idler photons are deterministically separated using a standard telecom C/L-band WDM, filtered down to ~ 1 nm bandwidth and further directed to Alice's and Bob's stations, which are separated by 35 m of optical fibre, corresponding to a physical separation of 5 m. For entanglement analysis, both users are equipped with fibre electro-optic phase modulators (EOM), allowing them to rotate the polarisation state of their respective photon on demand. For entanglement analysis, both users are equipped with fibre electro-optic phase modulators (EOM), allowing them to rotate the polarisation state of their respective photon on demand. EOM phases are set 4 μs prior to the detection window. As such,

there is not space-like separation of the measurement setting from the remote detection. Nevertheless, the 800 m fibre length employed would allow this kind of space-like separation if the fibre were linearly deployed. Suppl. Fig. 13(b) shows the operations that Alice and Bob apply depending on the human random numbers (HRN) that are generated by the Bellsters and provided by the central server at the ICFO¹⁰³. The settings we use correspond to the standard ones for violating the Bell inequality using polarisation entangled photons^{97–100}. Finally, each user analyses the polarisation state of their photon using a PBS, followed by two Indium Gallium Arsenide avalanche single photon detectors (SPD). A coincidence electronic circuit is used to account only for events in which both Alice and Bob receive a photon from the same pair.

A schematic of our data acquisition scheme is shown in Suppl. Fig. 13(c). Every two seconds, a table of 2×1000 random bits (Alice and Bob) is downloaded from the ICFO server, and the following sequence is repeated 1000 times:

1. A $4 \mu\text{s}$ time window is reserved in order to adjust Alice's and Bob's EOM settings accordingly to the HRNs. During this time, no coincidences are recorded. The following information is written into the raw data output file: a time stamp at the end of the $4 \mu\text{s}$ window, and Alice's and Bob's EOM settings.
2. Two-photon coincidences are recorded for $996 \mu\text{s}$. Time stamps and information about the corresponding pairs of detectors are stored in the raw data output file.
3. Thereafter, the experiment is halted for one second to give the computer time to analyse the data and store them on the hard-disk drive.

We note that during the second step in the above-mentioned procedure it might occur that no or multiple coincidence events are recorded. This is why we construct a new data file from the raw data file. Here, no-coincidence events are deleted, and for very rarely occurring multi-coincidence cases only the event belonging to the first time stamp is kept. Single coincidence events are not modified and stored as is.

From the new data file, the S -parameter is constructed in real-time following the standard approach^{97–100}. In Suppl. Fig. 13(d), we show the 25-hour time line of the measured S -parameter with a 5-minute running average. Within the first four hours, we observe a drop due to a strong temperature change in our laboratory which caused a slight misalignment of the PPLN/W. At $\sim 3:20$ pm, the photon pair coupling from the PPLN/W to PMFs was re-optimised, and the resulting S -parameter increased. At $\sim 3:00$ am the experiment was halted for about 10 minutes due to an electric power failure. A second realignment procedure was required at $\sim 7:00$ am due to a fast temperature change in the laboratory induced by the building air conditioning. Finally, the experiment was halted a second time at $\sim 10:30$ am due to a SPD overheating problem.

Note that none of the above-mentioned issues is actually related to the optical design of our fully guided-wave photon pair source, but exclusively due to infrastructural influences that could not be controlled. Nevertheless, the S -parameter is maintained always above 2 which exceeds the classical-quantum boundary, therefore proving the quantumness of our results.

The S -parameter, averaged over the full experiment, which started on 30.11.2016 at 11:09 am and ended on 01.12.2016 at 12:10 pm is calculated to be $S_{\text{full}} = 2.431$ with a standard error of $\sigma_{\text{full}} = \pm 0.003$, corresponding to a violation of the Bell inequalities by more than 143 standard deviations. Throughout the course of the experiment, the Bellsters provided us with $2 \times 19.5 \cdot 10^6$ HRNs, with which we performed $2.9 \cdot 10^6$ successful measurements (coincidences).

⁹⁷ J. S. Bell, On the Einstein Podolsky Rosen paradox, *Physics* **1**, 195 (1964).

⁹⁸ J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).

⁹⁹ G. Greenstein, A. Zajonc, *The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics*, Jones and Bartlett series in physics and astronomy (Jones and Bartlett Publishers, 2006).

¹⁰⁰ M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, New York, NY, USA, 2011), 10th edn.

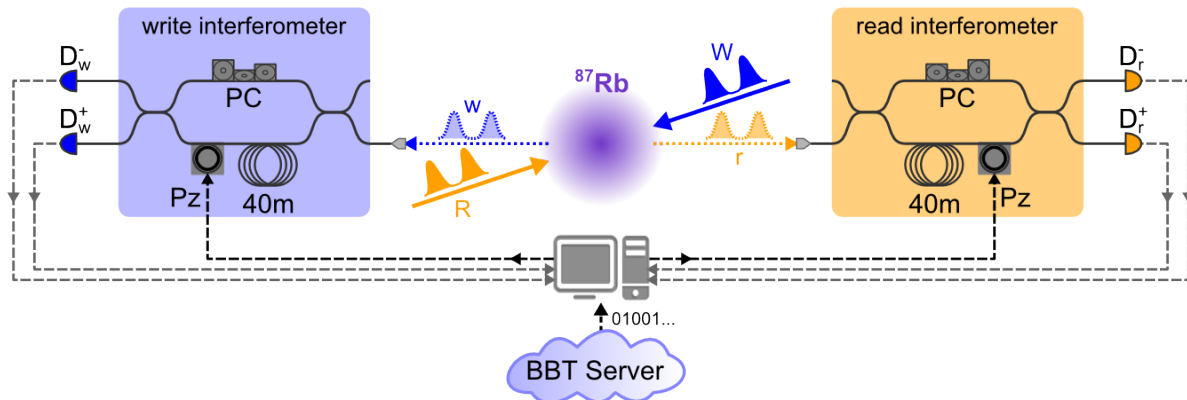
¹⁰¹ P. Vergyris, *et al.*, Fully guided-wave photon pair source for quantum applications, *Quantum Science and Technology* **2**, 024007 (2017).

¹⁰² H. C. Lim, A. Yoshizawa, H. Tsuchida, K. Kikuchi, Wavelength-multiplexed entanglement distribution, *Optical Fiber Technology* **16**, 225 (2010).

¹⁰³ ICFO-Website (2016). www.thebigbelltest.org.

⑨ Bell test using entanglement between a photon and a collective atomic excitation, driven by human randomness

Authors: Pau Farrera, Georg Heinze, Hugues de Riedmatten



Supplementary Figure 14: Schematic drawing of the experiment. Entangled atom-photon time-bin qubits are generated in a cold ^{87}Rb ensemble ($T \approx 100\mu\text{K}$) and analyzed via two imbalanced Mach-Zehnder interferometers. Abbreviations: W write pulse; w write photon; R read pulse; r read photon; PC polarization controller; Pz piezo-electric fiber stretcher; $D_{w(r)}^{+(-)}$ single photon detector.

Our experiment within the Big Bell Test involves the generation and analysis of entanglement between a photon and a collective atomic excitation in a cloud of laser-cooled ^{87}Rb atoms¹⁰⁴. Therefore, we generate photon-atom entangled qubits, which are encoded in the time-bin degree of freedom, as proposed in¹⁰⁵.

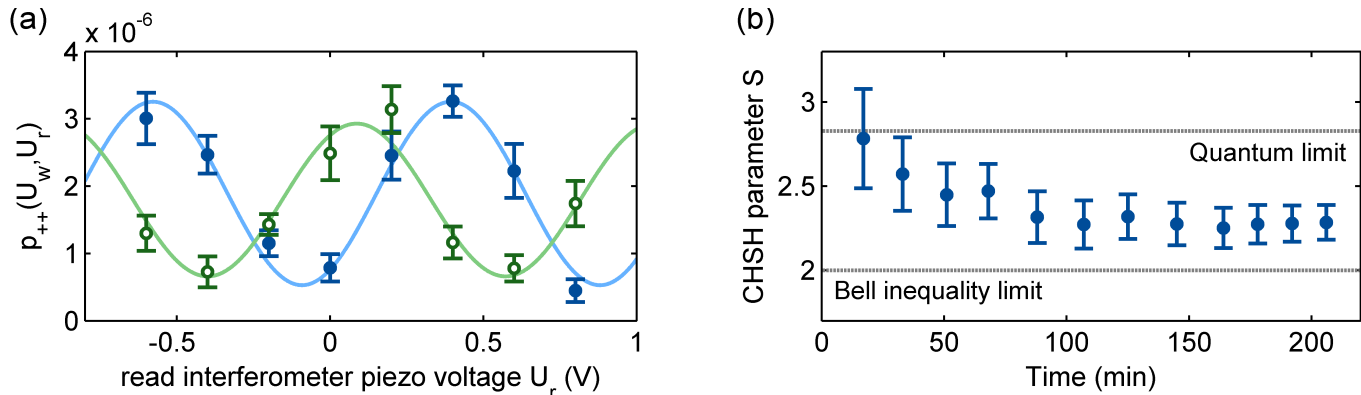
As shown in Suppl. Fig. 14, we follow the DLCZ (Duan-Lukin-Cirac-Zoller) protocol¹⁰⁶ sending a double-peaked write laser pulse (W) to probabilistically generate a write photon (w) which is paired with an excitation in the atomic cloud. The write pulse interacts with many atoms, leading to an excitation that is delocalized over all the atoms, i.e., a collective atomic spin excitation (“spin-wave”). Since this process can happen in any of the two time-bins, the photon and the atomic excitation are entangled in the time-bin degree of freedom. The entangled state can thus be written as $|\Psi_{w,a}\rangle = \frac{1}{\sqrt{2}} (|E_w E_a\rangle + e^{i\varphi} |L_w L_a\rangle)$, where $E_{w(a)}$ denotes a write photon (atomic excitation) generated in the early time-bin, $L_{w(a)}$ denotes a write photon (atomic excitation) generated in the late time-bin, and the phase φ can be controlled by changing the phase between the early and late time-bins of the write and read pulses. In order to temporally distinguish the atomic excitations created in the two different time-bins, we apply an homogeneous magnetic field to induce collective de- and re-phrasings of the stored spin-waves¹⁰⁷. Subsequently, the atomic qubit can be converted into a photonic time-bin qubit (r) using a resonant double-peaked read laser pulse (R). This converts the photon-atom entangled state into a photon-photon entangled state, which is more suited for the entanglement analysis. The analysis is done with two imbalanced Mach-Zehnder interferometers and single photon detectors, which allow us to perform qubit projective measurements in any basis that lies on the equator of the Bloch sphere¹⁰⁸. The bases in which the write and read qubits are projected depend on the phase delays ϕ_w and ϕ_r between the two arms of each interferometer. The quantum correlations of the two photons is then assessed by the CHSH Bell parameter¹⁰⁹

$$S = |E(\phi_w, \phi_r) + E(\phi_w, \phi'_r) + E(\phi'_w, \phi_r) - E(\phi'_w, \phi'_r)| \quad (7)$$

where $E(\phi_w, \phi_r) = [p_{++}(\phi_w, \phi_r) - p_{+-}(\phi_w, \phi_r) - p_{-+}(\phi_w, \phi_r) + p_{--}(\phi_w, \phi_r)] / p(\phi_w, \phi_r)$ are the correlation coefficients, $p_{ij}(\phi_w, \phi_r)$ are the probabilities to detect coincidences between the write and read photons at detectors D_w^i and D_r^j , and $p(\phi_w, \phi_r) = \sum_{i,j=+,-} p_{ij}(\phi_w, \phi_r)$.

In order to change the phase of the interferometers to the desired values of ϕ_w , ϕ'_w , ϕ_r , and ϕ'_r , a short part of the fiber of the 40 m long interferometer arms is rolled around a piezo-electric ceramic cylinder. Applying a voltage U_w or U_r to the corresponding piezo cylinder stretches the fiber and therefore changes the phase delay between the two arms of the write or the read interferometer. The random numbers generated by the participants of the Big Bell Test were used to decide which voltage is applied to each interferometer, hence controlling the measurement bases. Trapped atom clouds were created at a rate of 59 Hz. For each cloud, we performed 608 entanglement trials. The bases were changed in between each cloud and not between each single trial because of the limited bandwidth of the piezo fiber stretcher. However, the typical write-read photon coincidence detection probability per trial $p(\phi_w, \phi_r)$ was between 10^{-6} and 10^{-5} . We can therefore say that for each detected coincidence event, the bases were chosen randomly.

During the time window of the Big Bell Test we used the live human random numbers to randomly choose between predefined phases in the two analysing interferometers to take the data shown in Suppl. Fig. 15(a). Here, p_{++} is plotted as a function of $U_r \propto \phi_r$ for two different values of $U_w \propto \phi_w$. The two fringes are shifted by 114(9) degrees exhibiting visibilities of $V_1 = 0.72(0.08)$ (blue dots) and $V_2 = 0.63(0.10)$ (green circles). These visibilities are sufficient to prove entanglement between both time-bin qubits, as follows from the Peres separability criterion ($V > 1/3$) under the assumption of equally distributed noise



Supplementary Figure 15: (a) Photon coincidence detection probability between detectors D_w^+ and D_r^+ as a function of the voltage applied to the piezo of the read photon interferometer. The two different curves correspond to two different voltages applied to the piezo of the write photon interferometer: $U_w = 0$ V (blue dots) and $U_w = 0.296$ V (green circles). (b) Accumulated CHSH Bell parameter S as a function of data acquisition time for the measurement with stored human random numbers.

for all possible outcomes¹¹⁰. However, the values didn't surpass the threshold of $V > 1/\sqrt{2} \approx 0.707$ to guarantee Bell-type non-local correlations. We obtained similar visibilities without the use of human random numbers, which confirmed that the measurements were affected by experimental instabilities during the required relatively long integration times of several hours.

After improving the long-term stability of the experiment, we redid a Bell test at a later stage with human random numbers received and stored during the day of the Big Bell Test. Here, the human random numbers were used to switch randomly between the four settings required for the measurement of the CHSH inequality. The experimental results of that Bell Test are shown in Suppl. Fig. 15(b). The data acquisition lasted 3 hours and 26 minutes, during which we performed 36480000 experimental trials. In this time we recorded 1100 photon coincidence events which led to a final CHSH Bell parameter of $S = 2.29 \pm 0.10$. This measurement shows a violation of the Bell inequality $|S| \leq 2$ by approximately three standard deviations.

¹⁰⁴ D. N. Matsukevich, *et al.*, Entanglement of a photon and a collective atomic excitation, *Phys. Rev. Lett.* **95**, 040405 (2005).

¹⁰⁵ P. Farrera, G. Heinze, H. de Riedmatten, Entanglement between a photonic time-bin qubit and a collective atomic spin excitation, *Phys. Rev. Lett.* **120**, 100501 (2018).

¹⁰⁶ L. M. Duan, M. D. Lukin, J. I. Cirac, P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics., *Nature* **414**, 413 (2001).

¹⁰⁷ B. Albrecht, P. Farrera, G. Heinze, M. Cristiani, H. de Riedmatten, Controlled rephasing of single collective spin excitations in a cold atomic quantum memory, *Phys. Rev. Lett.* **115**, 160501 (2015).

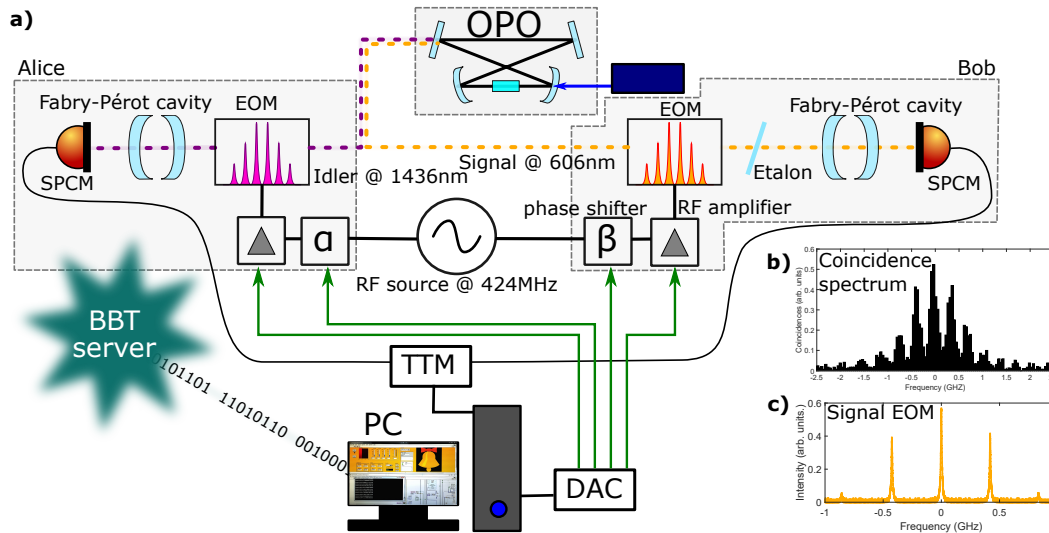
¹⁰⁸ I. Marcikic, *et al.*, Distribution of time-bin entangled qubits over 50 km of optical fiber, *Phys. Rev. Lett.* **93**, 180502 (2004).

¹⁰⁹ J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Physical Review Letters* **23**, 880 (1969).

¹¹⁰ A. Peres, Separability criterion for density matrices, *Phys. Rev. Lett.* **77**, 1413 (1996).

⑩ Violation of a Bell inequality using high-dimensional frequency-bin entangled photons

Authors: Andreas Lenhard, Alessandro Seri, Daniel Rieländer, Osvaldo Jimenez, Alejandro Máttar, Daniel Cavalcanti, Margherita Mazzera, Antonio Acín, and Hugues de Riedmatten



Supplementary Figure 16: a) Experimental setup. Photon pairs are generated via cavity enhanced SPDC in multiple frequency modes. The pairs are split and the Idler photons sent to Alice and Signal photons to Bob. At both ends there are EOMs, controlled by a RF phase shifter and an amplifier while the RF source is shared. Before detecting coincidences, single frequency bins are selected by cavity filters. The basis settings are chosen by HRN and a computer generates analog voltages (DAC) controlling the RF phases and powers accordingly. b) Spectrum of the photons, illustrating frequency-bins. c) Modulation spectrum of Signal EOM.

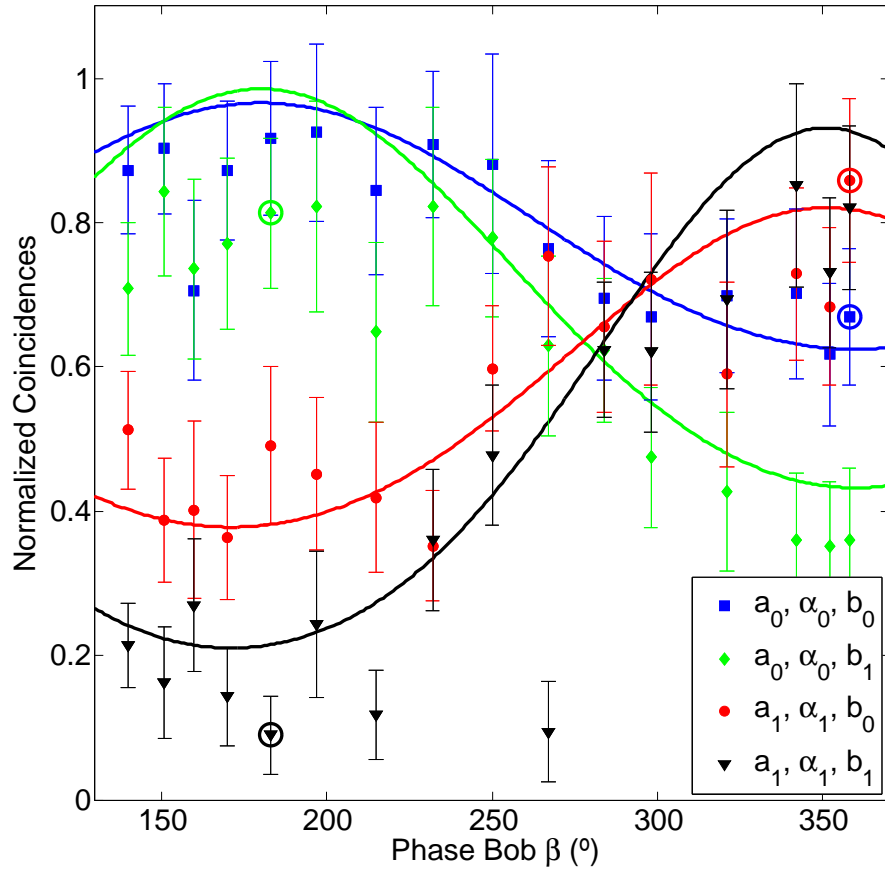
We report a Bell inequality violation with high-dimensional frequency-bin entangled photons. Due to energy conservation, photon pairs generated by spontaneous parametric down-conversion (SPDC) are naturally correlated in frequency, such that the sum of the signal and idler frequencies is always equal to the frequency of the pump beam. One can then define frequency-bins in the spectrum and find correlations between different spectral bins of the two photons, leading to so-called frequency bin entanglement¹¹¹. In our experiment, we generate narrowband photon pairs in discrete frequency modes by SPDC in an optical parametric oscillator (OPO) operated below threshold¹¹². One photon is at visible wavelength (606 nm) while the other is at the telecommunication wavelength (1436 nm). The photon spectrum consists of around 8 equally spaced modes of different amplitudes, separated by the free spectral range of the cavity (FSR=424 MHz, see Suppl. Fig. 16b). The two photons are separated deterministically and sent to Alice and Bob for analysis. The measurement process uses electro-optic phase modulators (EOMs) driven at a radio frequency (RF) corresponding to the FSR of the OPO (see Suppl. Fig. 16c) to mix neighboring modes and introduce a phase shift. Each measurement setting for Alice and Bob corresponds to a given modulation depth (a and b , set by the RF power) and to a given phase (α and β , set by the RF phase). We select single spectral modes of the signal and idler photons via filter cavities and detect the photons with single photon counting modules (SPCM). The full setup is illustrated in Suppl. Fig. 16.

We observe two-photon interference fringes depending on the phase difference between signal and idler EOMs. We develop a theoretical model for our setup from which we estimate the optimal experimental configurations leading to a violation of the CH-inequality¹¹³ $S_{CH} \leq 2$. During the BBT day we used the human random numbers (HRN) to choose in real time one out of four settings for $(a_i \alpha_i b_j; i, j = 0, 1)$ while the second phase setting ($\beta_k; k = 0..15$) was chosen out of 16 possible phases. This allowed us to record the full interference fringes shown in Suppl. Fig. 17. We kept each setting for 10 s, recorded the coincidences in that time and then received fresh HRN for the next setting. Selecting the data points circled in the fringes of Suppl. Fig. 17, our experiment reaches $S = 2.25(8)$, corresponding to a violation of the Bell inequality by three standard deviations. After the BBT, new measurements with stored HRN have been made, considering only the four best settings leading to the maximal violation. In this way data can be collected much faster, resulting in a violation of 8.5 standard deviations¹¹⁴.

¹¹¹ L. Olislager, *et al.*, Implementing two-photon interference in the frequency domain with electro-optic phase modulators, *New Journal of Physics* **14**, 043015 (2012).

¹¹² D. Rieländer, A. Lenhard, M. Mazzera, H. de Riedmatten, Cavity enhanced telecom heralded single photons for spin-wave solid state quantum memories, *New Journal of Physics* **18**, 123013 (2016).

¹¹³ J. F. Clauser, M. A. Horne, Experimental consequences of objective local theories, *Phys. Rev. D* **10**, 526 (1974).

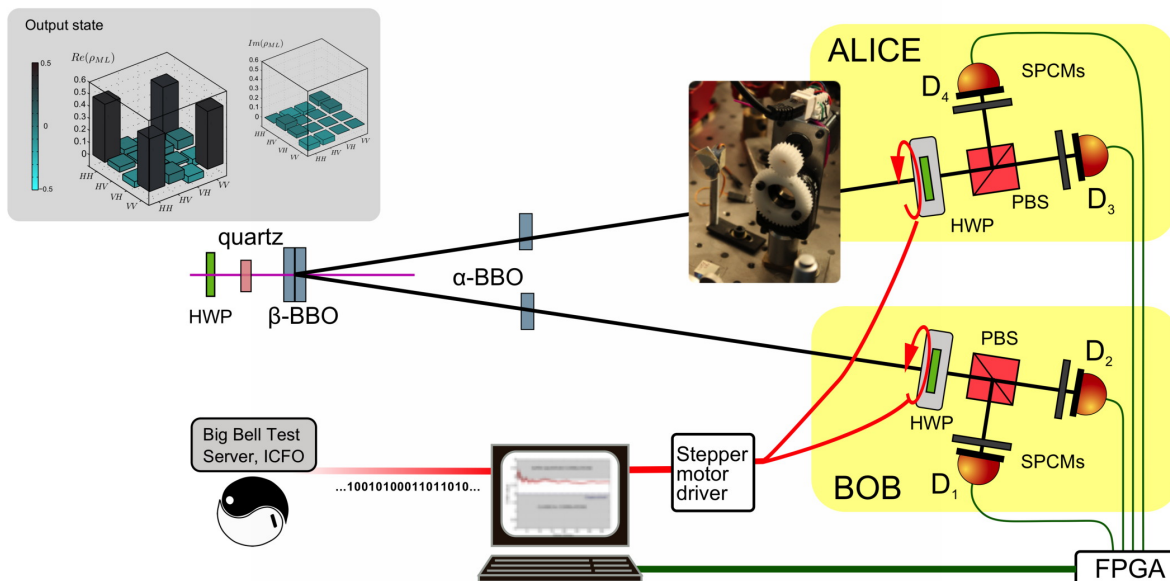


Supplementary Figure 17: Two photon interference fringes obtained by using HRN to choose the setting of modulation amplitude and phase shift for Alice and Bob's EOMs. While the phase of Bob's EOM was chosen among 16 possible values (corresponding to the x-axis), the other values were randomly selected between two possible values: $a_0 = 0.19$, $a_1 = 0.57$, $\alpha_0 = 0^\circ$, $\alpha_1 = 180^\circ$, $b_0 = 0.63$, $b_1 = 1.06$.

¹¹⁴ D. Rieländer, *et al.*, Frequency-bin entanglement of ultra-narrow band non-degenerate photon pairs, *Quantum Science and Technology* **3**, 014007 (2018).

⑪ Violation of the CHSH inequality using polarization-entangled photons

Authors: Laura T. Knoll, Ignacio H. López Grande, Agustina G. Magnoni, Christian T. Schmiegelow, Ariel Bendersky, and Miguel A. Larotonda



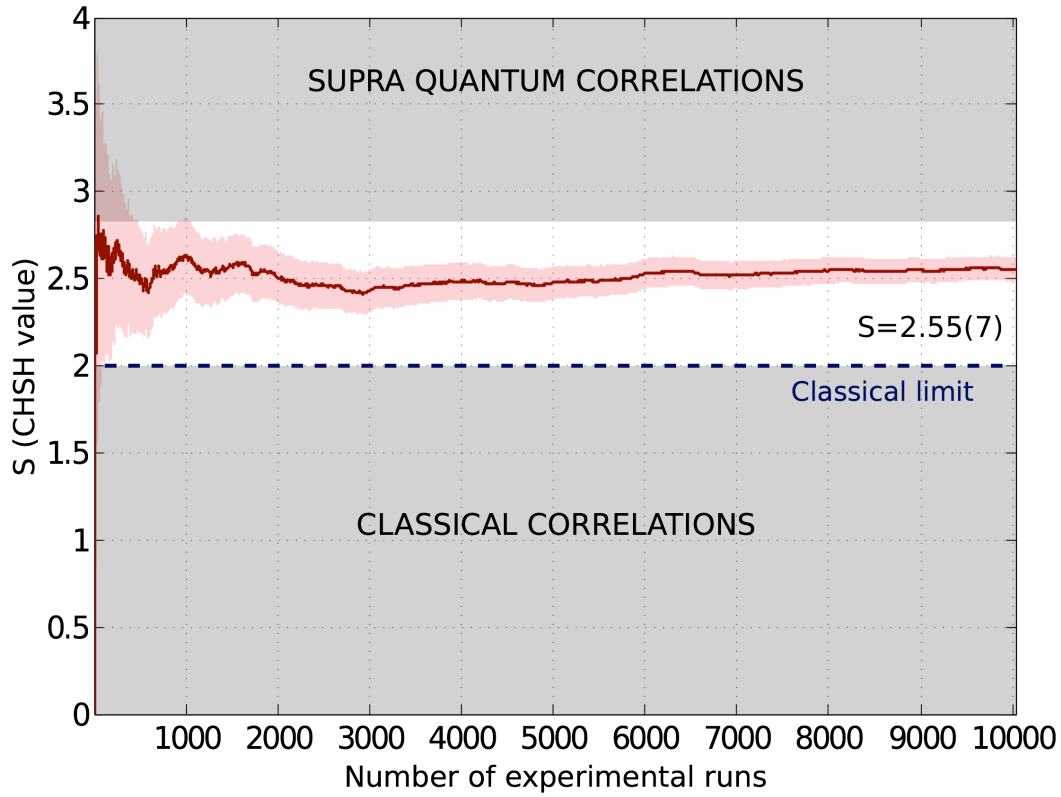
Supplementary Figure 18: Experimental setup. A 405 nm diode laser generates degenerate photon pairs by spontaneous parametric downconversion at 810 nm. A series of birefringent crystals placed at the pump beam and at the downconverted paths compensate temporal and spatial walk-off ensuring a high degree of indistinguishability; state tomography characterization of the produced state is shown in the inset. Random bit strings are retrieved from the BBT server and fed to the stepper motor drivers that control the projective measurements via rotations of the waveplates. The photo shows an actual motorized waveplate mount. Coincidences are registered with an FPGA-based detector and read by the control computer.

We show the results of a two-channel experiment to test a Clauser-Horne-Shimony-Holt (CHSH)-like inequality, using polarization entangled photons. Photon pairs are generated by spontaneous parametric downconversion in a BBO type-I nonlinear crystal arrangement, pumped by a 405 nm cw laser diode polarized 45 degrees with respect to both crystals. Both photons are generated at 810 nm¹¹⁵. A series of birefringence compensating crystals included in the photon pair source guarantee a high level of entanglement: a tomographic reconstruction of the output state of the source is shown at the top left inset on Suppl. Fig. 18)¹¹⁶. Polarization projections are performed with motorized waveplates, which are fed with the random strings generated at the BBT cloud infrastructure. Photons at both sides of the experiment are collected and spatially filtered by means of single mode fibers. Taking advantage of the angular distribution of the parametric fluorescence spectrum, the bandwidth of the correlated photon pairs is filtered down to 10 nm using the same fiber coupling optics. An additional set of longpass filters is used to remove residual scattered light. Two pairs of single-photon counting devices detect the incoming photons and send the detections to an FPGA-based coincidence counter; correlated photons are detected using 8 ns coincidence windows.

The experiment was fully controlled using a single PC. Each experimental run consisted in the following sequence: a pair of random bits were gathered from the BBT server data and fed into the waveplate stepper motor drivers. Once the polarization projections were set, an FPGA-based coincidence counter started a repetition of 20 measurements on the four coincidence detections $D_1 \wedge D_3$, $D_2 \wedge D_3$, $D_1 \wedge D_4$ and $D_2 \wedge D_4$. The full setup is shown in Suppl. Fig. 18. The brightness of our source and a detection window of 200 μ s for each measurement gave a single-channel coincidence probability below 0.1. In particular, the mean $D_1 \wedge D_3$ coincidence number per measurement was $\mu=0.091$ for one particular waveplate setting. This results on a total mean coincidence number per single experimental run (i.e. the expected value of detecting a coincidence on any channel) of $\nu=0.227$. Statistics of occurrence of either of the above situations follow a Poisson distribution. This condition minimizes the occurrence of multiple coincidences. Repetitions of the measurement increased its chance of success: we used the first single-channel coincidence from the list of 20 measurements as the result from each run, therefore increasing the "bit usage efficiency" of the experiment up to 97.5%. This rate also takes into account discarded runs with simultaneous coincidence detections. The total duration for each run was close to 2 s, which accounted for the mechanical rotation of the waveplates and the measuring time. Overall, almost 34 kbit were required from the BBT server at the Buenos Aires Node; some of them were used for initial and mid-term alignment of the optical setup.

The complete experiment carried out during the BBT day built statistics from 10033 polarization settings measured, and resulted in a measurement of the S-parameter value of $S=2.55(7)$, which implies a violation of the CHSH-Bell inequality that bounds local correlations by 7 standard deviations. Since the choice of measurement of both Alice and Bob was obtained through the BBT game, we can be certain about the freedom of choice of both Alice and Bob. We can thus conclude that

the assumption of free will was satisfied, therefore the free will loophole was closed. On the other hand, the locality loophole remains open in our setup. Suppl. Fig. 19 shows the evolution of the S-value obtained throughout the day. The red-shaded area represents the statistical uncertainty derived from the measured CHSH test.



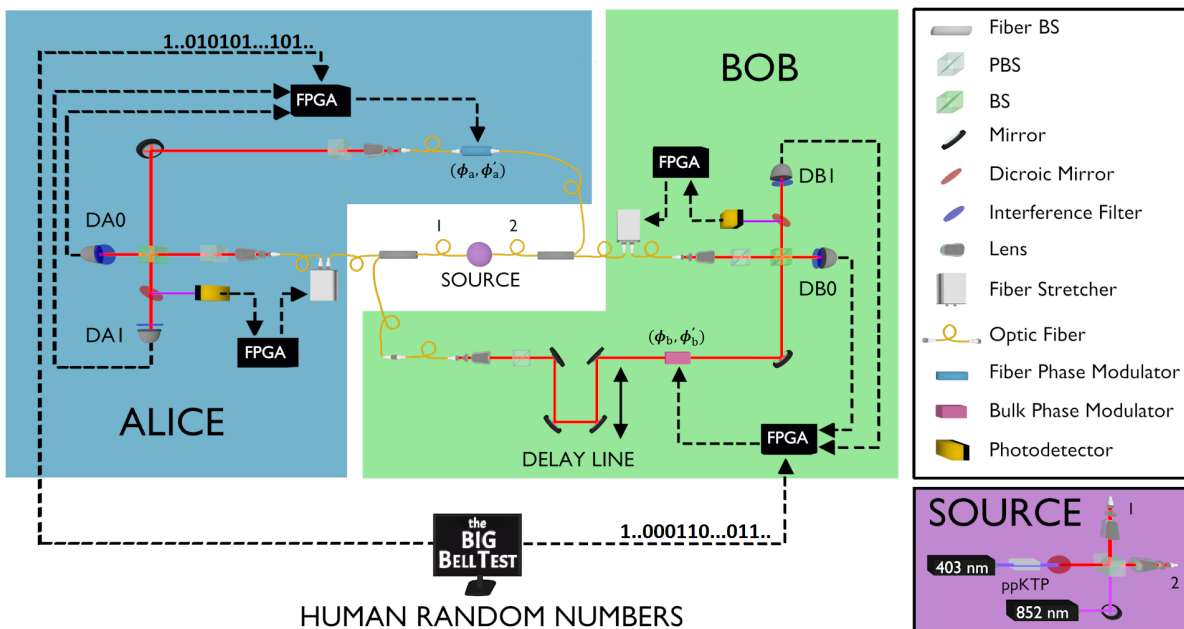
Supplementary Figure 19: Evolution of the CHSH test statistic “S” throughout the BBT day. Approximately 2000 experimental runs are sufficient to obtain a violation of the inequality with 95% confidence level

¹¹⁵ L. T. Knoll, C. T. Schmiegelow, M. A. Larotonda, Remote state preparation of a photonic quantum state via quantum teleportation, *Applied Physics B* **115**, 541 (2014).

¹¹⁶ P. G. Kwiat, E. Waks, A. G. White, I. Appelbaum, P. H. Eberhard, Ultrabright source of polarization-entangled photons, *Physical Review A* **60**, R773 (1999).

⑫ Post-selection loophole-free energy-time Bell test fed with human-generated inputs

Authors: Jaime Cariñe, Felipe Toledo, Pablo González, Álvaro Alarcón, Daniel Martínez, Jorge Fuenzalida, Jean Cortés, Álvaro Cuevas, Gonzalo Carvacho, Aldo Delgado, Fabio Sciarrino, Paolo Mataloni, Jan-Åke Larsson, Adán Cabello, Esteban S. Gómez, Gustavo Lima and Guilherme B. Xavier



Supplementary Figure 20: Experimental setup. Dashed and red lines represent electrical connections and free-space beams, respectively. The propagation distance between the source to Alice is 15 m and from the source to Bob is 45 m, determined mainly by the length of the piezoelectric spooled fiber-optical stretchers (10 m long for Alice's interferometer and 40 m for Bob's). Please see text for further details.

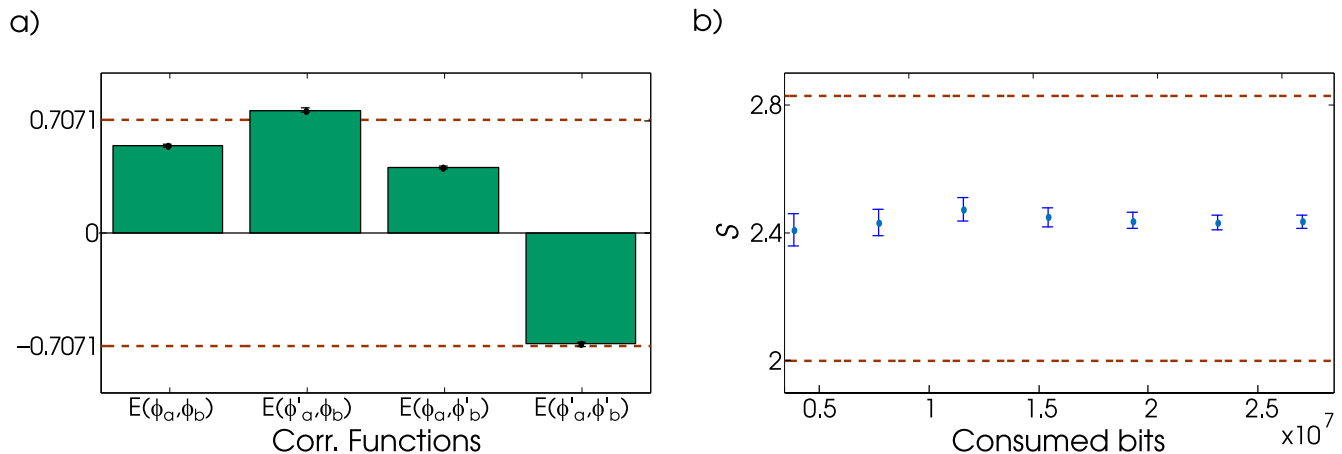
Energy-time Bell tests based on the standard and widely employed Franson configuration¹¹⁷ are affected by the post-selection loophole¹¹⁸. The only experimentally demonstrated solution to this problem is the so-called “hug” configuration^{119–121}. Here we perform the first post-selection loophole-free energy-time Bell test where the measurement settings are randomly chosen by humans. The human random numbers were provided by the Big Bell Test (BBT) server and were used to choose the settings for each detected event. Several optoelectronic upgrades compared to our previous efforts were implemented to support the randomly and dynamically changing measurement settings. The experimental setup is shown in Suppl. Fig. 20, where the source generates energy-time entangled photon pairs through the process of spontaneous parametric down-conversion. A periodically poled potassium titanyl phosphate (ppKTP) 20 mm long waveguide crystal emits degenerate type-II (orthogonally polarized) photon pairs at 806 nm, when pumped with a single-mode longitudinal continuous wave laser at 403 nm. The single-photons are deterministically split with a polarizing beam splitter (PBS), with each photon being coupled to a single-mode fiber and then guided to one of the two bidirectional 50:50 fiber couplers whose outputs are then cross-sent to Alice and Bob. The main characteristic of the “hug” configuration is the presence of two crossed (or “hugging”) Mach-Zehnder interferometers (MZI) serving as the communication channel between the source and the communicating parties. Alice and Bob place single-photon detectors at the outputs of each interferometer. The single-photon detectors are free-running silicon-based single-photon avalanche detectors, D_{A1} and D_{A2} for Alice and D_{B1} and D_{B2} for Bob. The interferometers are asymmetrical, as in the standard Franson scheme, with each consisting of a short (S) and a long (L) arm.

Each interferometer needs to be actively stabilized against environmental phase drifts. In order to do so, a reference laser beam (852 nm) is superposed to the path of the single-photons in the source, before coupling to the single-mode fibers. The intensity of this reference signal is detected by p-i-n photodiodes after both interferometers (split from the single-photons with dichroic mirrors) and are used to close individual electronic control loops based on field programmable gate array (FPGA) electronics. Each control loop employs a piezoelectric fiber-optical stretcher, capable of dynamically changing the length of the fiber¹²². With each interferometer independently stabilized, electro-optical phase modulators (PMs) are used to apply a relative phase difference in a short time window (5 μ s), compared to the control response time. With such a technique the control in itself is not disturbed, and thus independent settings may be applied while the two interferometers remain stable¹²².

A motorized delay line is used to adjust the long-short path difference in both interferometers to be within the coherence length of the photon pair. Once this is achieved the following state is generated without the need for temporal post-selection: $|\Phi\rangle = 1/\sqrt{2}(|SS\rangle + e^{i(\phi_a + \phi_b)}|LL\rangle)$, where ϕ_a and ϕ_b are phase shifts within the long arm of Alice and Bob's respective MZIs. We test the well-known Clauser-Horne-Shimony-Holt (CHSH) inequality $S = E(\phi_a, \phi_b) + E(\phi'_a, \phi_b) + E(\phi_a, \phi'_b) - E(\phi'_a, \phi'_b)$ ¹²³,

where $E(\phi_a, \phi_b) = P_{11}(\phi_a, \phi_b) + P_{22}(\phi_a, \phi_b) - P_{12}(\phi_a, \phi_b) - P_{21}(\phi_a, \phi_b)$, with $P_{ij}(\phi_a, \phi_b)$ corresponding to the probability of a coincident detection at Alice and Bob's detectors i and j respectively, while the relative phases ϕ_a and ϕ_b are applied to Alice and Bob's interferometers. For the maximum violation of the Bell CHSH inequality the phase settings are $\phi_a = \pi/4$, $\phi_b = 0$, $\phi'_a = -\pi/4$ and $\phi'_b = \pi/2$.

For each measurement round two bits are used from the BBT server to choose one out of the four possible setting combinations (two settings for Alice and two for Bob). The detection electronics (also FPGA-based) register if a coincident detection occurred, and record the values for the employed phases. This is done at a rate of 52 kbit/s while consuming approximately 27 million bits from the BBT database. The average expected values for the entire run are shown in Suppl. Fig. 21a), with the cumulative S parameter in Suppl. Fig. 21b), whose final value is 2.4331 ± 0.0218 , corresponding to a violation of 19.87 standard deviations for a total of 20676 detection events.



Supplementary Figure 21: Experimental results. a) Correlation functions E as a function of the different settings for the CHSH inequality. Dashed lines correspond to the value of E for a maximal violation of the CHSH inequality. b) Cumulative CHSH violation (S parameter) as a function of the amount of consumed bits. Dashed red lines represent the local bound (2) and the quantum bound ($2\sqrt{2}$). Error bars in both subfigures represent the standard deviation.

¹¹⁷ J. D. Franson, Bell inequality for position and time, *Phys. Rev. Lett.* **62**, 2205 (1989).

¹¹⁸ S. Aerts, P. Kwiat, J.-Å. Larsson, M. Żukowski, Two-photon Franson-type experiments and local realism, *Phys. Rev. Lett.* **83**, 2872 (1999).

¹¹⁹ A. Cabello, A. Rossi, G. Vallone, F. De Martini, P. Mataloni, Proposed Bell experiment with genuine energy-time entanglement, *Phys. Rev. Lett.* **102**, 040401 (2009).

¹²⁰ A. Cuevas, *et al.*, Long-distance distribution of genuine energy-time entanglement, *Nature Communications* **4**, 2871 (2013).

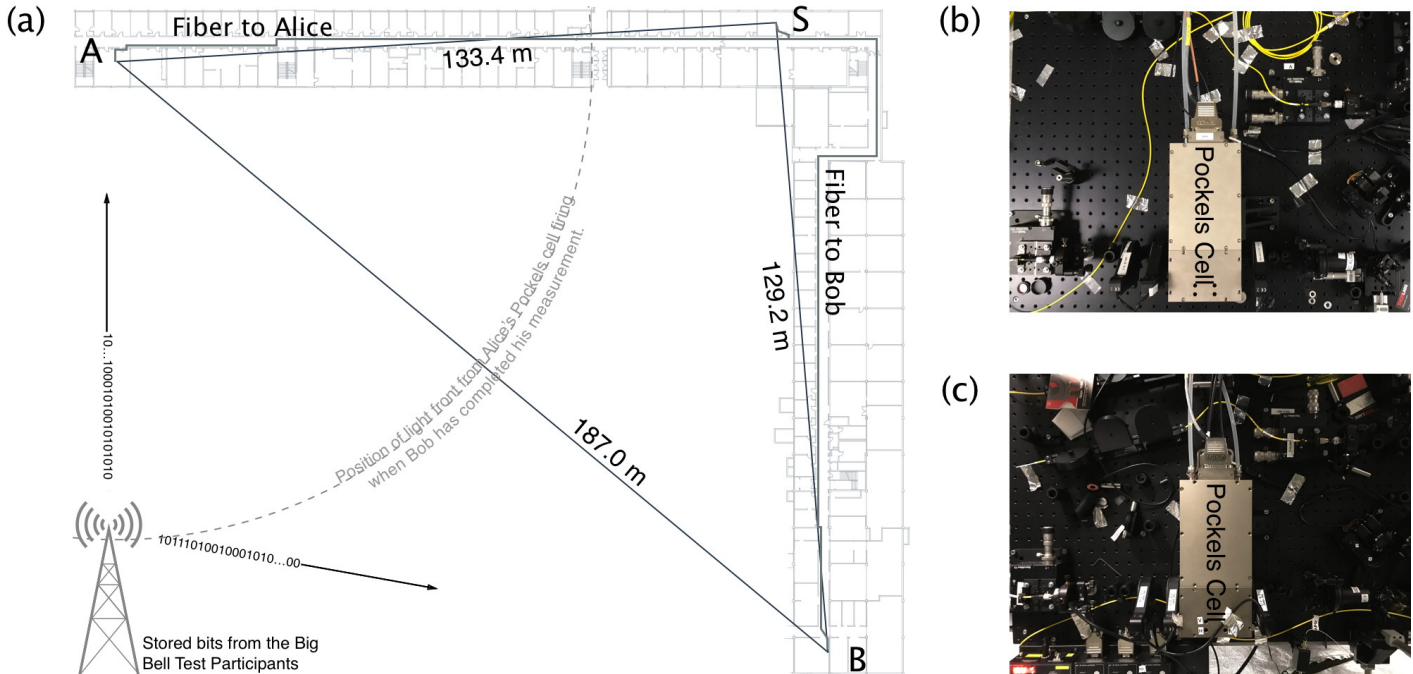
¹²¹ G. Carvacho, *et al.*, Postselection-loophole-free Bell test over an installed optical fiber network, *Phys. Rev. Lett.* **115**, 030503 (2015).

¹²² G. B. Xavier, J. P. von der Weid, Stable single-photon interference in a 1 km fiber-optic Mach-Zehnder interferometer with continuous phase adjustment, *Optics Letters* **36**, 1764 (2011).

¹²³ J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).

⑬ Using human generated randomness to violate a Bell inequality without detection or locality loopholes

Authors: Lynden Krister Shalm, Martin Stevens, Omar S. Magaña-Loaiza, Thomas Gerrits, Scott Glancy, Peter Bierhorst, Emanuel Knill, Richard Mirin, Sae Woo Nam



Supplementary Figure 22: Experimental setup. (a) The locations of the Source (S), Alice (A), and Bob (B). Each trial, the source lab produces a pair of photons in the non-maximally polarization-entangled state $|\psi\rangle \approx 0.982 |HH\rangle + 0.191 |VV\rangle$, where H (V) denotes horizontal (vertical) polarization. One photon is sent to Alice's lab while the other is sent to Bob's lab to be measured. Alice's computed optimal polarization measurement angles, relative to a vertical polarizer, are $\{a = -3.7^\circ, a' = 23.6^\circ\}$ while Bob's are $\{b = 3.7^\circ, b' = -23.6^\circ\}$. Each trial, Alice and Bob each use a bit from the Bellsters to choose which polarization measurement to make. Panel (b) shows the Pockels cell that Alice uses, and panel (c) shows Bob's Pockels cell. Alice and Bob are (187 ± 1) m apart from one another. At this distance, in any given trial when Bob has completed his measurement, information traveling at the speed of light from Alice's Pockels cell firing is still (88.25 ± 1.1) m away from his measurement setup (represented by the light front). When Alice completes her measurement, Bob's light front from his Pockels cell firing is (127.2 ± 1.1) m away from her measurement apparatus.

We report a violation of a Bell inequality free of fair-sampling and locality assumptions. We use polarization-entangled photons generated by a nonlinear crystal pumped by a pulsed, picosecond laser. The laser repetition rate is 79.3 MHz, and each pulse that enters the crystal has a probability of ≈ 0.002 of creating an entangled photon pair. The two entangled photons from each pair are then separated, with each one being sent to one of two measurement stations (187 ± 1) m apart, named Alice and Bob. At Alice and Bob, a Pockels cell and polarizer combine to allow the rapid measurement of the polarization state of the incoming photons. Alice's and Bob's decisions on how to set their Pockels cells are dictated by the bits supplied by the Big Bell Test participants (Bellsters). The photons are then detected using fiber-coupled superconducting single-photon nanowire detectors, each operating at over 90% efficiency¹²⁴.

Our experimental setup, shown in Suppl. Fig. 22, is nearly identical to the experimental setup used in the fully loophole-free Bell Test conducted in 2015¹²⁵. The primary difference is that random numbers used to set Alice's and Bob's polarization settings are provided by the Bellsters instead of random number generators. Like the 2015 experiment, we consider the measurement of Alice's or Bob's photon detection to be complete when the amplified electrical signal from Alice's or Bob's detector reaches the input of the time tagger that records the outcomes. The polarization measurement is carried out using a high-voltage Pockels cell. When this device switches states, it is possible for an electro-magnetic pulse to be emitted. If these electro-magnetic signals travel at the speed of light, then Bob completes his measurement (294.4 ± 3.7) ns before any signal from Alice's Pockels cell could arrive at his station. Similarly, Alice completes her measurement (424.2 ± 3.7) ns before a switching signal from Bob's Pockels cell could arrive at her location. Under these conditions the implementation of Alice's setting and Bob's measurement outcome are space-like separated from one another and vice-versa, so it is impossible for one station's setting to influence the other station's measurement.

Testing the hypothesis of local realism requires repeating the experiment many times, each of which is called a trial. We then compare the trials' statistics to those that local realistic theories predict. In our setup we perform 100,000 trials per second. This is limited by the speed of our Pockels cells at Alice and Bob. Because the probability of generating a photon pair with each pump pulse is small, Alice and Bob look for photons generated by 15 successive pulses during each trial. This increases our probability of creating a photon pair per trial by more than an order of magnitude, allowing us to achieve a stronger violation

using fewer trials. Each trial then consists of 15 different time bins in which Alice and Bob may register photon detections. In our experiment we collected data for just under 7 minutes and performed 40,559,990 trials. Each trial, Alice and Bob each used one bit provided by the Bellsters to set their polarization measurement choice (consuming a total of 81,119,980 bits over the course of the experiment). Immediately after collecting data using bits provided by the Bellsters, we repeated the experiment using random bits provided solely by computer-based random number generators located at both Alice and Bob. In this second data run we performed a total of 74,400,000 trials.

Each trial, Alice uses a bit from the Bellsters to choose between two settings a, a' while Bob uses a bit from the Bellsters to choose between two different settings b, b' . Because we include 15 different time bins each trial, when certain settings are chosen there is a significant probability in a given trial of Alice and Bob both detecting photons, but in different time bins. Using conventional Bell inequalities, these detections would normally be considered a coincidence event even though they involve unentangled photons from different photon pairs. To help limit the effect of these unwanted coincidences, we use a modified Clauser-Horne (CH) inequality¹²⁶ that accounts for the 15 individual time bins that a detection can take place in at either Alice or Bob. During a trial if Alice (Bob) detects a photon in time bin i , we label the event i_A (i_B). If no detection is observed, then the outcome is labelled with a 0_A for Alice or 0_B for Bob. If two separate photon events are detected by either Alice or Bob in the same trial, only the first bin is counted. The modified inequality is:

$$K = \sum_{i=1}^n P(i_A, i_B|ab) - \sum_{i=1}^n P(i_A, i_B|a'b') - \sum_{i=1}^n P(0_A, i_B|a'b) - \sum_{i=1}^n \sum_{j \neq i} P(j_A, i_B|a'b) - \sum_{i=1}^n P(i_A, 0_B|ab') - \sum_{i=1}^n \sum_{j \neq i} P(i_A, j_B|ab') \leq 0 \quad (8)$$

Here $n = 15$ for our experiment. The term $P(i_A, i_B|ab)$ represents the probability that Alice and Bob both detect photons in time bin i , given that Alice chooses setting a and Bob chooses setting b . The term $P(i_A, i_B|a'b')$ is the same, but for settings a' and b' . This eliminates from the inequality the high probability of accidental coincidences where Alice and Bob, with joint settings a' and b' , detect photons in different time bins ($i \neq j$) during the same trial. The terms $P(0_A, i_B|a'b)$ and $P(i_A, 0_B|ab')$ correspond to cases where only one party detects a photon in a given trial. Finally, the $P(j_A, i_B|a'b)$ and $P(i_A, j_B|ab')$ terms correspond to trials where Alice and Bob both detect photons, but not in the same time bins.

To see that equation 8 is a valid Bell inequality, we note that if local realism holds then each term can be written as an integral over hidden variables, λ . Then, because Alice and Bob are far enough apart, Alice's measurement setting (S_A) and outcome (O_A) can be factored from Bob's measurement setting (S_B) and outcome (O_B):

$$P(O_a, O_b|S_a S_b) = \int P(O_a, O_b|S_a S_b \lambda) \rho(\lambda) d\lambda = \int P(O_a|S_a \lambda) P(O_b|S_b \lambda) \rho(\lambda) d\lambda. \quad (9)$$

Since the probability of some event, E , occurring is $P(E) = 1 - P(\bar{E})$, where $P(\bar{E})$ is the probability the event does not occur, the two terms with double sums can be re-written, as

$$\begin{aligned} \sum_{i=1}^n \sum_{j \neq i} P(j_A, i_B|a'b) &= \sum_{i=1}^n \int P(i_B|b\lambda) [1 - P(i_A|a'\lambda) - P(0_A|a'\lambda)] \rho(\lambda) d\lambda, \\ \sum_{i=1}^n \sum_{j \neq i} P(i_A, j_B|ab') &= \sum_{i=1}^n \int P(i_A|a\lambda) [1 - P(i_B|b'\lambda) - P(0_B|b'\lambda)] \rho(\lambda) d\lambda. \end{aligned} \quad (10)$$

Now, since everything is just a sum over i , we can first pull the sum and then the integral out of all terms. A number of terms cancel, leaving us with:

$$K = \sum_{i=1}^n \int [P(i_A|a\lambda)P(i_B|b\lambda) - P(i_A|a'\lambda)P(i_B|b'\lambda) - P(i_B|b\lambda) + P(i_B|b\lambda)P(i_A|a'\lambda) - P(i_A|a\lambda) + P(i_A|a\lambda)P(i_B|b'\lambda)] \rho(\lambda) d\lambda. \quad (11)$$

The fact that the four remaining probabilities, $P(i_A|a\lambda)$, $P(i_B|b\lambda)$, $P(i_A|a'\lambda)$, and $P(i_B|b'\lambda)$ are all bounded between 0 and 1 implies that the maximum value of each integrand is 0, following the arguments of¹²⁶.

Using this inequality, we measure a Bell parameter of $K = (1.70 \pm 0.20) \times 10^{-4}$, corresponding to an $8.7\text{-}\sigma$ violation of a Bell inequality (under the assumptions of a normal distribution for the Bell parameter, and of each trial being independent and identical). Table III contains the number of photons detected and the number of trials for each term in Equation 8. We also analyzed the data from the experiment that used computer-based randomness, and obtained a Bell parameter of $K = (1.55 \pm 0.14) \times 10^{-4}$ corresponding to an $11.2\text{-}\sigma$ violation a Bell inequality. The violation is larger because we performed nearly twice as many trials with the computer-based random number generator, allowing us to obtain better statistics. Both of these results constitute a strong rejection of local realism. A full record of the data, along with the code used to analyze it, is available on request.

It is also possible to carry out a hypothesis test to see how well a local-realistic theory could reproduce the correlations we observed in our experiment. Such a test does not require assumptions about memory effects and the statistical distributions of each trial's outcomes. We devised the following analysis protocol before running the experiment. The first 5% of the data was used for training purposes to estimate the rate of relevant events, given by terms in equation 9. From this, an estimate of

TABLE III: The number of photons detected for each of the terms in equation 8. Due to the large bias towards zeros in the bits received from the Bellsters, it is necessary to normalize the counts by the number of trials conducted at each setting. A total of 40,559,990 trials were carried out over the course of the experiment. A second experiment 74,400,000 trials long that used only unbiased random bits from a computer was run immediately afterwards.

	$\sum_{i=1}^{15} P(i_A, i_B ab)$	$\sum_{i=1}^{15} P(i_A, i_B a'b')$	$\sum_{i=1}^{15} P(0_A, i_B a'b)$	$\sum_{i=1}^{15} \sum_{j \neq i} P(j_A, i_B a'b)$	$\sum_{i=1}^{15} P(i_A, 0_B ab')$	$\sum_{i=1}^{15} \sum_{j \neq i} P(i_A, j_B ab')$
Random bits from Bellsters						
Events	20265	712	7475	143	8181	123
Trials	11126350	9202147	10125716	10125716	10105777	10105777
Random bits from computers						
Events	33543	1677	13911	249	14589	228
Trials	18600599	18598891	18599346	18599346	18601164	18601164

the total number of remaining events was made. A stopping criteria, n_{cut} , defined as 90% of the estimated remaining relevant events was defined. Once n_{cut} relevant events were processed the analysis was stopped, and the p-value was computed using a binomial analysis of equation 8 (see¹²⁵ for more details).

Our hypothesis test protocol was flawed, and failed to observe a violation. When we initially ran the protocol on our data we obtained a p-value 3.3×10^{-70} , with $n_{cut} = 30,916$ events. This is a much smaller p-value than expected. Looking more closely at the data, we realized that there is an approximately 5% bias of excess zeros compared to ones submitted by the Bellsters. Given that over 100,000 individuals from around the world independently submitted zeros and ones, such a large bias was surprising. We do not have a definitive explanation for this, but we suspect it might be related to the majority of humans being right-hand dominant. Most keyboards use the right hand to input a zero. On most specialized number pads, the zero key is also substantially larger. For our setup the ab setting, which contributes to events that violate equation 8, occurs more frequently than the other settings due to the bias. Since our hypothesis test protocol was devised under the assumption that all the settings are equiprobable, we must correct our p-value to account for this large excess bias. After performing this correction, we obtain a p-value of ≈ 1 , and we are unable to reject the null hypothesis that local realism governs our experiment whose random inputs were generated by humans. The data from the computer-based random number generator, however, had the unbiased settings distribution required by our hypothesis test protocol. From this data set we obtained a p-value of 2.6×10^{-27} with $n_{cut} = 54,720$, corresponding to a strong rejection of local realism.

The null result from the hypothesis test with human-based randomness does not mean that local realism governed the experiment. The hypothesis test protocol we followed assumed that there was no bias in both Alice's and Bob's settings choices. An improved protocol, for future experiments, would include a step that first estimates the bias from the training data in order to determine the optimal Bell inequality to test. As long as there are only small deviations from this estimated bias (on the order of $< 0.2\%$) over the remainder of the experiment, it would be possible to compute a rigorous p-value significantly smaller than 1. Additionally, our measurement of the Bell parameter value is good evidence that local realism is not consistent with the experimental correlations we observed. The Bell parameter is normalized by the total trials for each settings choice, and is therefore insensitive to bias in the settings distribution. Under the assumptions that each trial was independent and identically distributed and that the Bell parameter was normally distributed, the measured Bell parameter exceeds the maximum bound predicted by local realism by a statistically significant amount. Even if these assumption are not true, it would take a local realistic theory with a contrived distribution to appreciably lower this significance. For some applications, like using a loophole-free Bell test to extract randomness, it is important to use a rigorous hypothesis testing framework to guard against a malicious adversary who might be able to produce such contrived distributions to attack the system. However, one hopes that nature is not so malicious when testing local realism.

We wanted to see if there was a difference between the results of a Bell test where humans chose the settings at Alice and Bob versus a Bell test that used non-human random number generators. After accounting for the bias in the human inputs, data from both the human-based input and the computer-based input were able to violate a Bell inequality.

¹²⁴ F. Marsili, *et al.*, Detecting single infrared photons with 93% system efficiency, *Nature Photonics* **7**, 210 (2013).

¹²⁵ L. K. Shalm, *et al.*, Strong loophole-free test of local realism, *Phys. Rev. Lett.* **115**, 250402 (2015).

¹²⁶ J. F. Clauser, M. A. Horne, Experimental consequences of objective local theories, *Phys. Rev. D* **10**, 526 (1974).