A strong loophole-free test of local realism

Shalm, Meyer-Scott, Christensen, Bierhorst, Wayne, Stevens, Gerrits, Glancy, Hamel, Allman, Coakley, Dyer, Hodge, Lita, Verma, Lambrocco, Tortorici, Migdall, Zhang, Kumor, Farr, Marsili, Shaw, Stern, Abellán, Amaya, Pruneri, Jennewein, Mitchell, Kwiat, Bienfang, Mirin, Knill, and Nam

Nov 10, 2015

Abstract

We present a loophole-free violation of local realism using entangled photon pairs. We ensure that all relevant events in our Bell test are spacelike separated by placing the parties far enough apart and by using fast random number generators and high-speed polarization measurements. A high-quality polarization-entangled source of photons, combined with high-efficiency, low-noise, single-photon detectors, allows us to make measurements without requiring any fair-sampling assumptions. Using a hypothesis test, we compute p-values as small as 5.9×10^{-9} for our Bell violation while maintaining the spacelike separation of our events. We estimate the degree to which a local realistic system could predict our measurement choices. Accounting for this predictability, our smallest adjusted p-value is 2.3×10^{-7}. We therefore reject the hypothesis that local realism governs our experiment.

But if [a hidden variable theory] is local it will not agree with quantum mechanics, and if it agrees with quantum mechanics it will not be local. This is what the theorem says. - John Stewart Bell

Quantum mechanics at its heart is a statistical theory. It cannot with certainty predict the outcome of all single events, but instead it predicts probabilities of outcomes. This probabilistic nature of quantum theory is at odds with the determinism inherent in Newtonian physics and relativity, where outcomes can be exactly predicted given sufficient knowledge of a system. Einstein and others felt that quantum mechanics was incomplete. Perhaps quantum systems are controlled by variables, possibly hidden from us [2], that determine the outcomes of measurements. If we had direct access to these hidden variables then the properties of quantum systems would not need to be treated probabilistically. De Broglie's 1927 pilot-wave theory was a first attempt at formulating a hidden variable theory of quantum physics [3]; it was completed in 1952 by David Bohm [4, 5]. While the pilot-wave theory can reproduce all of the predictions of quantum mechanics, it has the curious feature that hidden variables in one location can instantly change values because of events happening in distant locations. This seemingly violates the locality principle from relativity, which says that objects cannot signal one another faster than the speed of light. In 1935 the nonlocal feature of quantum systems was popularized by Einstein, Podolsky, and Rosen [6], and is something Einstein later referred to as "spooky actions at a distance"[7].

But in 1964 John Bell showed that it is impossible to construct a hidden variable theory that obeys locality and simultaneously reproduces all of the predictions of quantum mechanics [8]. Bell's theorem fundamentally changed our understanding of quantum theory and today stands as a cornerstone of modern quantum information science.

Bell's theorem does not prove the validity of quantum mechanics, but it does allows us to test the hypothesis that nature is governed by local realism. The principle of realism says that any system has pre-existing values for all possible measurements of the system. In local realistic theories, these pre-existing values depend only on events in the past lightcone of the system. Local hidden-variable theories obey this principle of local realism. Local realism places constraints on the behavior of systems of multiple particles—constraints that do not apply to entangled quantum particles. This leads to different predictions that can be tested in an experiment known as a Bell test. In a typical two-party Bell test, a source generates particles and sends them to two distant parties, Alice and Bob. Alice and Bob independently and randomly choose properties of their individual particles to measure. Later, they compare the results of their measurements. Local realism constrains the joint probability distribution of their choices and measurements. The basis of a Bell test is an inequality that is obeyed by local realistic probability distributions but can be violated by the probability distributions of certain entangled quantum particles [8]. A few years after Bell derived his inequality, new forms were introduced by Clauser, Horne, Shimony and Holt [9], and Clauser and Horne [10] that are simpler to experimentally test.

In a series of landmark experiments, Freedman and Clauser [11] and Aspect, Grangier, Dalibard, and Roger [12-14] demonstrated experimental violations of Bell inequalities using pairs of polarization-entangled photons generated by an atomic cascade. However, due to technological constraints, these Bell tests and those that followed (see [15] for a review) were forced to make additional assumptions to show local realism was incompatible with their experimental results. A significant violation of Bell's inequality implies that either local realism is false or that one or more of the assumptions made about the experiment are not true; thus every assumption in an experiment opens a "loophole". No experiment can be absolutely free of all loopholes, but in [16] a minimal set of assumptions is described that an experiment must make to be considered "loophole free". Here we report a significant,

loophole free, experimental violation of local realism using entangled photon pairs. We use the definition of loophole free as defined in [16]. In our experiment the only assumptions that remain are those that can never—even in principle—be removed. We present physical arguments and evidence that these remaining assumptions are either true or untestable.

Bell's proof requires that the measurement choice at Alice cannot influence the outcome at Bob (and vice-versa). If a signal traveling from Alice can not reach Bob in the time between Alice's choice and the completion of Bob's measurement, then there is no way for a local hidden variable constrained by special relativity at Alice to change Bob's outcomes. In this case we say that Alice and Bob are spacelike separated from one another. If an experiment does not have this spacelike separation, then an assumption must be made that local hidden variables cannot signal one another, leading to the "locality" loophole.

Another requirement in a Bell test is that Alice and Bob must be free to make random measurement choices that are physically independent of one another and of any properties of the particles. If this is not true, then a hidden variable could predict the chosen settings in advance and use that information to produce measurement outcomes that violate a Bell inequality. Not fulfilling this requirement opens the "freedom-of-choice" loophole. While this loophole can never in principle be closed, the set of hidden variable models that are able to predict the choices can be constrained using space-like separation. In particular, in experiments that use processes such as cascade emission or parametric down-conversion to create entangled particles, space-like separation of the measurement choices from the creation event eliminates the possibility that the particles, or any other signal emanating from the creation event, influence the settings. To satisfy this condition, Alice and Bob must choose measurement settings based on fast random events that occur in the short time before a signal traveling at the speed of light from the entangled-photon creation would be able to reach them. But it is fundamentally impossible to conclusively prove that Alice's and Bob's random number generators are independent without making additional assumptions, since their backward lightcones necessarily intersect. Instead, it is possible to justify the assumption of measurement independence through a detailed characterization of the physical properties of the random number generators (such as the examination described in [17, 18]).

3

In any experiment, imperfections could lead to loss, and not all particles will be detected. To violate a Bell inequality in an experiment with two parties, each free to choose between two settings, Eberhard showed that at least 2/3 of the particles must be detected [19] if non-maximally entangled states are used. If the loss exceeds this threshold, then one may observe a violation by discarding events in which at least one party does not detect a particle. This is valid under the assumption that particles were lost in an unbiased manner. However, relying on this assumption opens the "detector" or "fair-sampling" loophole. While the locality and fair-sampling loopholes have been closed individually in different systems [20-24], it has only recently been possible in our system and in two others [25, 26] to close all loopholes simultaneously. These three experiments also address the freedom-of-choice loophole by space-like separation.

Fundamentally a Bell inequality is a constraint on probabilities that are estimated from random data. Determining whether a data set shows violation is a statistical hypothesis-testing problem. It is critical that the statistical analysis does not introduce unnecessary assumptions that create loopholes. A Bell test is divided into a series of trials. In our experiment, during each trial Alice and Bob randomly choose between one of two measurement settings (denoted $\{a, a'\}$ for Alice and $\{b, b'\}$ for Bob) and record either a "+" if they observe any detection events or a "0" otherwise. Alice and Bob must define when a trial is happening using only locally available information, otherwise additional loopholes are introduced. At the end of the experiment Alice and Bob compare the results they obtained on a trial-by-trial basis. Our Bell test uses a version of the Clauser-Horne/Eberhard inequality [10, 19, 27] where, according to local realism,

$$P(+ + |ab) \leq P(+0|ab') + P(0 + |a'b) + P(+ + |a'b')$$ (1)

The terms $P(+ + |ab)$ and $P(+ + |a'b')$ correspond to the probability that both Alice and Bob record detection events (++) when they choose the measurement settings ab or $a'b'$, respectively. Similarly, the terms $P(+0|ab')$ and $P(0 + |a'b)$ are the probabilities that only Alice or Bob record an event for settings ab' and $a'b$, respectively. A local realistic model can saturate this inequality; how- ever, the probability distributions of entangled quantum particles can violate it.

To quantify our Bell violation we construct a hypothesis test based on the

4

inequality in Eq. (1). The null hypothesis we test is that the measured probability distributions in our experiment are constrained by local realism. Our evidence against this null hypothesis of local realism is quantified in a p-value that we compute from our measured data using a test statistic. Our test statistic takes all of the measured data from Alice's and Bob's trials and summarizes them into a single number (see the Supplemental Material for further details). The p-value is then the maximum probability that our experiment, if it is governed by local realism, could have produced a value of the test statistic that is at least as large as the observed value [28]. Smaller p-values can be interpreted as stronger evidence against this hypothesis. These p-values can also be used as certificates for cryptographic applications, such as random number generation, that rely on a Bell test [24, 29]. We use a martingale binomial technique [27] for computing the p-value that makes no assumptions about the distribution of events and does not require that the data be independent and identically distributed [30] as long as appropriate stopping criteria are determined in advance.

In our experiment, the source creates polarization- entangled pairs of photons and distributes them to Alice and Bob, located in distant labs. At the source location a mode-locked Ti:Sapphire laser running at repetition rate of approximately 79.3 MHz produces picosecond pulses centered at a wavelength of 775 nm as shown in figure 1. These laser pulses pump an apodized periodically poled potassium titanyl phosphate (PPKTP) crystal to produce photon pairs at a wavelength of 1550nm via the process of spontaneous parametric downconversion [31]. The downconversion system was designed using the tools available in [32]. The PPKTP crystal is embedded in the middle of a polarization-based Mach-Zehnder interferometer that enables high-quality polarization-entangled states to be generated [33]. Rotating the polarization analyzer angles at Alice and Bob, we measure the visibility of coincidence detections for a maximally entangled state to be 0.999 ± 0.001 in the horizontal/vertical polarization basis and 0.996 ± 0.001 in the diagonal/antidiagonal polarization basis (see [34] for information about the reported uncertainties). The entangled photons are then coupled into separate single-mode optical fibers with one photon sent to Alice and the other to Bob. Alice, Bob, and the source are positioned at the vertices of a nearly right-angle triangle.
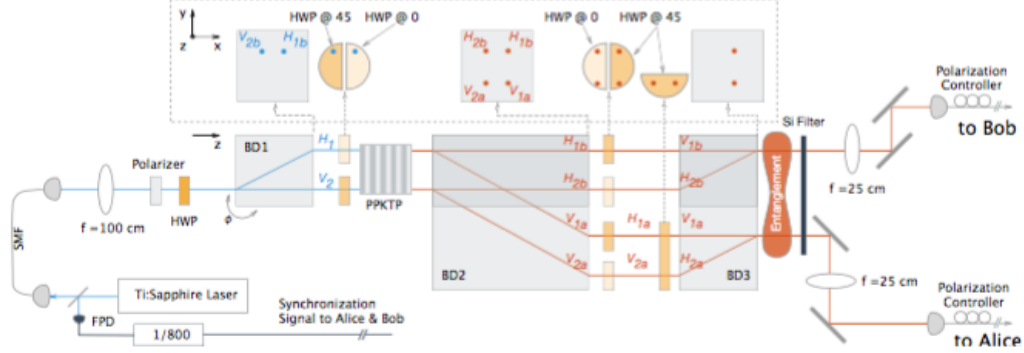
Figure 1: Schematic of the entangled photon source. A pulsed 775 nm-wavelength Ti:Sapphire picosecond mode-locked laser running at 79.3 MHz repetition rate is used as both a clock and a pump in our setup. A fast photodiode (FPD) and divider circuit are used to generate the synchronization signal that is distributed to Alice and Bob. A polarization-maintaining single-mode fiber (SMF) then acts as a spatial filter for the pump. After exiting the SMF, a polarizer and half-wave plate (HWP) set the pump polarization. To generate entanglement, a periodically poled potassium titanyl phosphate (PPKTP) crystal designed for Type-II phasematching is placed in a polarization-based Mach-Zehnder interferometer formed using a series of HWPs and three beam displacers (BD). At BD1 the pump beam is split in two paths (1 and 2): the horizontal (H) component of polarization of the pump translates laterally in the x direction while the vertical (V) component of polarization passes straight through. Tilting BD1 sets the phase, ϕ, of the interferometer to 0. After BD1 the pump state is $(\cos(16°)|H_1\rangle + \sin(16°)|V_2\rangle)$. To address the polarization of the paths individually, semi-circular waveplates are used. A HWP in path 2 rotates the polarization of the pump from vertical (V) to horizontal (H). A second HWP at 0° is inserted into path 1 to keep the path lengths of the interferometer balanced. The pump is focused at two spots in the crystal, and photon pairs at a wavelength of 1550nm are generated in either path 1 or 2 through the process of spontaneous parametric downconversion. After the crystal, BD2 walks the V-polarized signal photons down in the y direction (V_{1a} and V_{2a}) while the H-polarized idler photons pass straight through (H_{1b} and H_{2b}). The $x - y$ view shows the resulting locations of the four beam paths. HWPs at 45° correct the polarization while HWPs at 0° provide temporal compensation. BD3 then completes the interferometer by recombining paths 1 and 2 for the signal and idler photons. The two downconversion processes interfere with one another, creating the entangled state in Eq. (2). A high-purity silicon wafer with an anti-reflection coating is used to filter out the remaining pump light. The idler (signal) photons are coupled into a SMF and sent to Alice (Bob).

Due to constraints in the building layout, the photons travel to Alice and Bob in fiber optic cables that are not positioned along their direct lines of sight. While the photons are in flight toward Alice and Bob, their random num-

6

ber generators each choose a measurement setting. Each choice is completed before information about the entangled state, generated at the PPKTP crystal, could possibly reach the random number generators. When the photons arrive at Alice and Bob, they are launched into free space, and each photon passes through a Pockels cell and polarizer that perform the polarization measurement chosen by the random number generators as shown in Fig. 2.
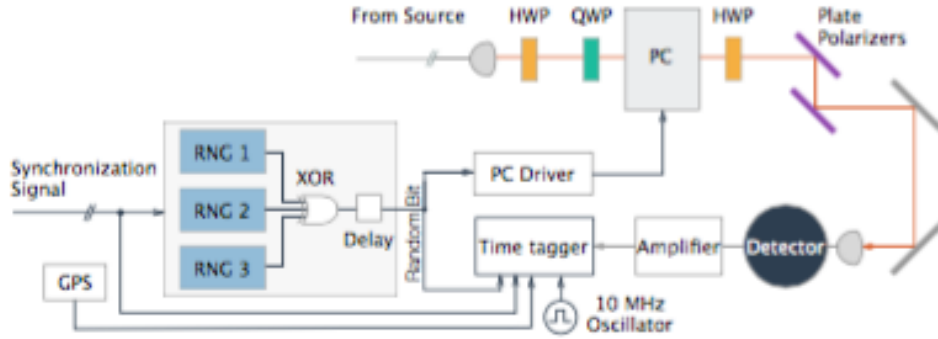

Figure 2: Receiver station setup for Alice and Bob. A photon arrives from the source. Two half-wave plates (HWP), a quarter-wave plate (QWP), a Pockels cell (PC), and two plate polarizers together act to measure the polarization state of the incoming photon. The polarization projection is determined by a random bit from XORing the outputs of two random number generators (RNG1 and RNG2) with pre-determined pseudorandom bits (RNG3). If the random bit is "0", corresponding to measurement setting $a(b)$ for Alice (Bob), the Pockels cell remains off. If the random bit is "1", corresponding to measurement setting $a?(b?)$ for Alice (Bob), then a voltage is applied to the Pockels cell that rotates the polarization of the photons using a fast electro-optic effect. The two plate polarizers have a combined contrast ratio $> 7000{:}1$. The photons are coupled back into a single-mode fiber (SMF) and detected using a superconducting nanowire single-photon detector (SNSPD). The signal is amplified and sent to a time-tagging unit where the arrival time of the event is recorded. The time tagger also records the measurement setting, the synchronization signal, and a one pulse-per-second signal from a global positioning system (GPS). The pulse-per-second signal provides an external time reference that helps align the time tags Alice and Bob record. A 10 MHz oscillator synchronizes the internal clocks on Alice's and Bob's time taggers. The synchronization pulse from the source is used to trigger the measurement basis choice.

After the polarizer, the photons are coupled back into a single-mode fiber and sent to superconducting nanowire single-photon detectors, each with a detection efficiency of $91 \pm 2\%$ [35]. The detector signal is then amplified and sent to a time tagger where the arrival time is recorded. We assume the

measurement outcome is fixed when it is recorded by the time tagger, which happens before information about the other party's setting choice could possibly arrive, as shown in Fig. 3(b).
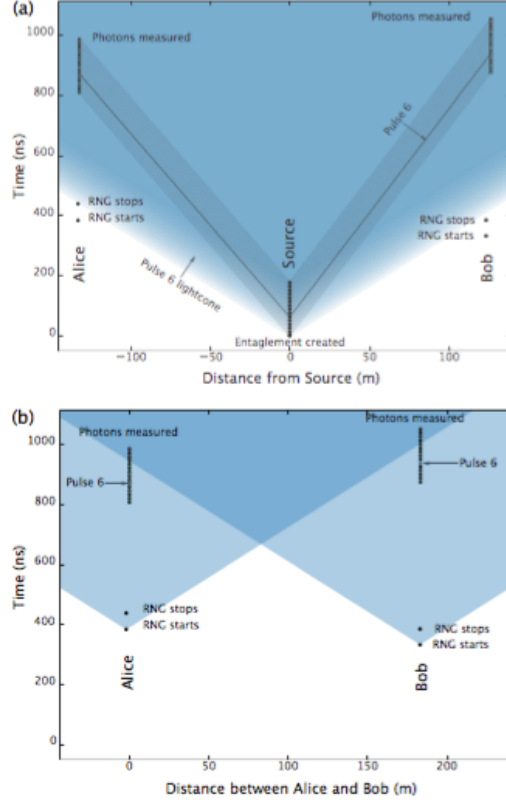

Figure 3: Minkowski diagrams for the spacetime events related to Alice (A) and the source (S) and Bob (B) and the source (a), and Alice and Bob (b). All lightcones are shaded blue. Due to the geometry of Alice, Bob, and the source, more than one spacetime diagram is required. In a) the random number generators (RNGs) at Alice and Bob must finish picking a setting outside the lightcone of the birth of an entangled photon pair. A total of 15 pump pulses have a chance of downconverting into an entangled pair of photons each time the Pockels cells are on. The events related to the first pulse are not spacelike separated, because Alice's RNG does not finish picking a setting before information about the properties of the photon pair can arrive; pulses 2 through 11 are spacelike separated. As shown in (b), pulses 12 through 15 are not spacelike separated as the measurement is finished by Alice and Bob after information about the other party's measurement setting could have arrived. In our experiment the events related to pulse 6 are the furthest outside of all relevant lightcones.

8

Alice and Bob have system detection efficiencies of 74.7±0.3% and 75.6±0.3%, respectively. We measure this system efficiency using the method outlined by Klyshko [36]. Background counts from blackbody radiation and room lights reduce our observed violation of the Bell inequality. Every time a background count is observed it counts as a detection event for only one party. These background counts increase the heralding efficiency required to close the detector loophole above 2/3 [19]. To reduce the number of background counts, the only detection events considered are those that occur within a window of approximately 625ps at Alice and 781 ps at Bob, centered around the expected arrival times of photons from the source. The probability of observing a background count during a single window is 8.9×10^{-7} for Alice and 3.2×10^{-7} for Bob, while the probability that a single pump pulse downconverts into a photon pair is $\approx 5 \times 10^{-4}$. These background counts in our system raise the efficiency needed to violate a Bell inequality from 2/3 to 72.5%. Given our system detection efficiencies, our entangled photon production rates, entanglement visibility, and the number of background counts, we numerically determine the entangled state and measurement settings for Alice and Bob that should give the largest Bell violation for our setup. The optimal state is not maximally entangled [19] and is given by:

$$|\psi\rangle = 0.961 |H_A H_B\rangle + 0.276 |V_A V_B\rangle \qquad (2)$$

where $H(V)$ denotes horizontal (vertical) polarization, and A and B correspond to Alice's and Bob's photons, respectively. From the simulation we also determine that Alice's optimal polarization measurement angles, relative to a vertical polarizer, are $\{a = 4.2°, a' = -25.9°\}$ while Bob's are $\{b = -4.2°, b' = 25.9°\}$.

Synchronization signals enable Alice and Bob to define trials based only on local information. The synchronization signal runs at a frequency of 99.1 kHz, allowing Alice and Bob to perform 99,100 trials/s (79.3 MHz/800). This trial frequency is limited by the rate the Pockels cells can be stably driven. When the Pockels cells are triggered they stay on for ≈ 200 ns. This is more than 15 times longer than the 12.6ns pulse-to-pulse separation of the pump laser. Therefore photons generated by the source can arrive in one of 15 slots while both Alice's and Bob's Pockels cells are on. Since the majority of the photon pulses arriving in these 15 slots satisfy the spacelike separation constraints, it is possible to aggregate multiple adjacent pulses to increase the event rate and statistical significance of the Bell violation.

However, including too many pulses will cause one or more of the spacelike separation constraints to be violated. Because the probability per pulse of generating an entangled photon pair is so low, given that one photon has already arrived, the chance of getting a second event in the same Pockels cell window is negligible ($< 1\%$).

Alice and Bob each have three different sources of random bits that they XOR together to produce their random measurement decisions (for more information see the Supplemental Materials). The first source is based on measuring optical phase diffusion in a gain-switched laser that is driven above and below the lasing threshold. A new bit is produced every 5 ns by comparing adjacent laser pulses [17]. Each bit is then XORed with all past bits that have been produced (for more details see the Supplemental Material). The second source is based on sampling the amplitude of an optical pulse at the single-photon level in a short temporal interval. This source produces a bit on demand and is triggered by the synchronization signal. Finally, Alice and Bob each have a different predetermined pseudorandom source that is composed of various popular culture movies and TV shows, as well as the digits of π, XORed together. Suppose that a local-realistic system with the goal of producing violation of the Bell inequality, was able to manipulate the properties of the photons emitted by the entanglement source before each trial. Provided that the randomness sources correctly extract their bits from the underlying processes of phase diffusion, optical amplitude sampling, and the production of cultural artifacts (such as the movie *Back to the Future*), this powerful local realistic system would be required to predict the outcomes of all of these processes well in advance of the beginning of each trial to achieve its goal. Such a model would have elements of superdeterminism—the fundamentally untestable idea that all events in the universe are preordained.

Over the course of two days we took a total of 6 data runs with differing configurations of the experimental setup [37]. Here we report the results from the final dataset that recorded data for 30 minutes (see the Supplemental Material for descriptions and results from all datasets). This is the dataset where the experiment was most stable and best aligned; small changes in coupling efficiency and the stability of the Pockels cells can lead to large changes in the observed violation. The events corresponding to the sixth pulse out of the 15 possible pulses per trial are the farthest outside all the relevant light-cones. Thus we say these events are the most spacelike separated. To

increase our data rate we aggregate multiple pulses centered around pulse number 6. We consider different Bell tests using a single pulse (number 6), three pulses (pulses 5, 6, and 7), five pulses (pulses 4 through 8), and seven pulses (pulses 3 through 9). The joint measurement outcomes and corresponding p-values for these combinations are shown in Table 1.

Aggregate Pulses	$N(++|ab)$	N_{stop}	Total trials	P-value	Adjusted p-value	Timing Margin (ns)	Minimum distance (m)
1	1257	2376	175,654,992	2.5×10^{-3}	5.9×10^{-3}	63.5 ± 3.7	9.2
3	3800	7211	175,744,824	2.4×10^{-6}	2.3×10^{-5}	50.9 ± 3.7	7.3
5	6378	12127	177,358,351	5.9×10^{-9}	2.3×10^{-7}	38.3 ± 3.7	5.4
7	8820	16979	177,797,650	2.0×10^{-7}	9.2×10^{-6}	25.7 ± 3.7	3.5

Table 1: P-value results for different numbers of aggregate pulses. Here $N(++|ab)$ refers to the number of times Alice and Bob both detect a photon with settings a and b respectively. Before analyzing the data a stopping criteria, N_{stop}, was chosen. This stopping criteria refers to the total number of events considered that have the settings and outcomes specified by the terms in Eq. (1), $N_{stop} = N(++|ab) + N(+0|ab') + N(0+|a'b) + N(++|a'b')$. After this number of trials the p-value is computed and the remaining trials discarded. Such pre-determined stopping criteria are necessary for the hypothesis test we use (see Supplemental Material for more details). The total trials include all trials up to the stopping criteria regardless of whether a photon is detected. The adjusted p-value accounts for the excess predictability we estimate from measurements of one of our random number generators. As discussed in the text, the time difference between Bob finishing his measurement and the earliest time at which information about Alice's measurement choice could arrive at Bob sets the margin of timing error that can be tolerated and still have all events guaranteed to be spacelike separated. We also give the minimum distance between each party and its boundary line (shown in Fig. 4(a)) that guarantees satisfaction of the spacelike separation constraints. In the Supplemental Material the frequencies of each combination of settings choice for 5 aggregate pulses is reported.

For a single pulse we measure a p-value = 2.5×10^{-3}, for three pulses a p-value = 2.4×10^{-6}, for five pulses a p-value = 5.9×10^{-9} and for seven pulses a p-value = 2.0×10^{-7}, corresponding to a strong violation of local realism.

If, trial-by-trial, a conspiratorial hidden variable (or attacker in cryptographic scenarios) has some measure of control over or knowledge about the settings choices at Alice and Bob, then they could manipulate the outcomes to observe a violation of a Bell inequality. Even if we weaken our assumption that Alice's and Bob's setting choices are physically independent from the source, we can still compute valid p-values against the hypothesis of local realism. We characterize the lack of physical independence with predictability of our
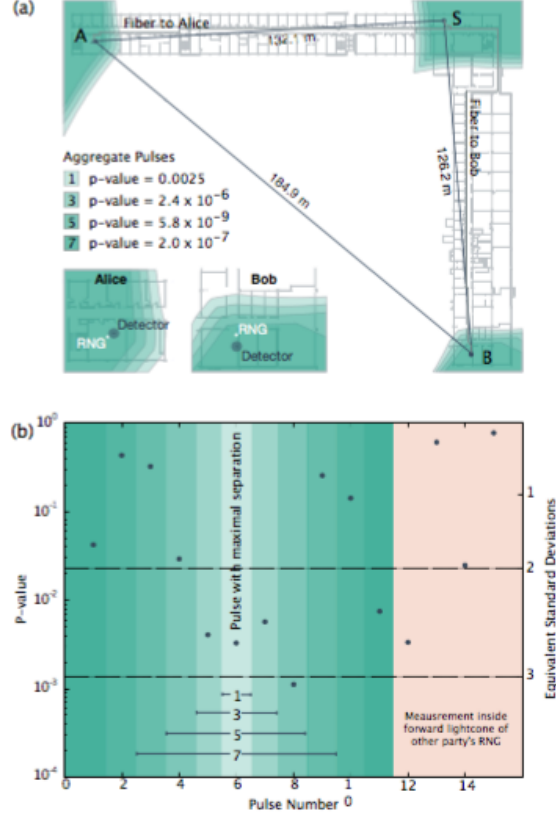
random number generators.

Figure 4: (a) The positions of Alice (A), Bob (B), and the source (S) in the building where the experiment was carried out. The insets show a magnified (×2) view of Alice's and Bob's locations. The white dots are the location of the random number generators (RNGs). The larger circle at each location has a radius of 1m and corresponds to our uncertainty in the spatial position measurements. Alice, Bob, and the source can be located anywhere within the green shaded regions and still have their events be spacelike separated. Boundaries are plotted for aggregates of one, three, five, and seven pulses. Each boundary is computed by keeping the chronology of events fixed, but allowing the distance between the three parties to vary independently. In (b) the p-value of each of the individual 15 pulses is shown. Overlayed on the plot are the aggregate pulse combinations used in the contours in (a). The statistical significance of our Bell violation does not appear to depend on the spacelike separation of events. For reference and comparison purposes only, the corresponding number of standard deviations for a given p-value (for a one-sided normal distribution) are shown.

The "predictability", $\mathcal{P}$, of a random number generator is the probability

with which an adversary or local realistic system could guess a given setting choice. We use the parameter ϵ, the "excess predictability" to place an upper bound on the actual predictability of our random number generators:

$$\mathcal{P} \le \frac{1}{2}(1+\epsilon) \tag{3}$$

In principle, it is impossible to measure predictability through statistical tests of the random numbers, because they can be made to appear random, unbiased, and independent even if the excess predictability during each trial is nonzero. Extreme examples that could cause nonzero excess predictability include superdeterminism or a powerful and devious adversary with access to the devices, but subtle technical issues can never be entirely ruled out. Greater levels of excess predictability lead to lower statistical confidence in a rejection of local realism. In Fig. 5 we show how different levels of excess predictability change the statistical significance of our results [38] (see Supplemental Materials for more details).
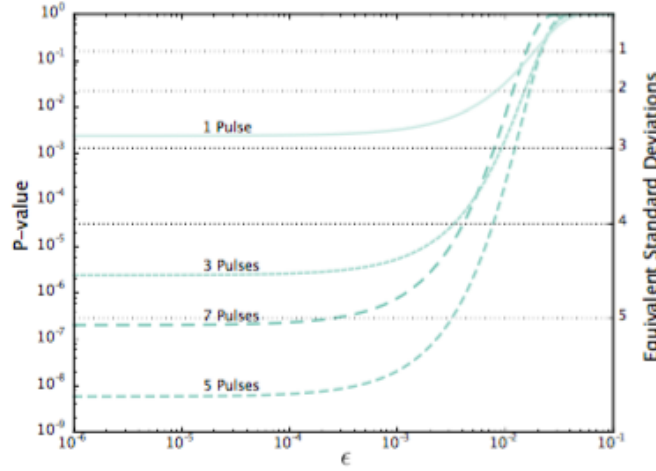
Figure 5: The p-value for different numbers of aggregate pulses as a function of the excess predictability, ϵ, in Alice's and Bob's measurement settings. Larger levels of predictability correspond to a weakening of the assumption that the settings choices are physically independent of the photon properties Alice and Bob measure. As in Fig. 4(b), the p-value equivalent confidence levels corresponding to the number of standard deviations of a one-sided normal distribution are shown for reference.

We can make estimates of the excess predictability in our system. From additional measurements, we observe a bias of $(1.08 \pm 0.07) \times 10^{-4}$ in the settings reaching the XOR from the laser diffusion random source, which includes synchronization electronics as well as the random number generator. If this bias is the only source of predictability in our system, this level of bias would correspond to an excess predictability of approximately 2×10^{-4}. To be conservative we use an excess predictability bound that is fifteen times larger, $\epsilon_p = 3 \times 10^{-3}$ (see Supplemental Material for more details). If our experiment had excess predictability equal to ϵ_p our p-values would be increased to 5.9×10^{-3}, 2.4×10^{-5}, 2.3×10^{-7}, and 9.2×10^{-6} for one, three, five, and seven pulses, respectively [38]. Combining the output of this random number generator with the others should lead to lower bias levels and a lower excess predictability, but even under the the paranoid situation where a nearly superdeterministic local realistic system has complete knowledge of the bits from the other random number sources, the adjusted p-values still provide a rejection of local realism with high statistical significance.

Satisfying the spacetime separations constraints in Fig. 3 requires precise measurements of the locations of Alice, Bob, and the source as well as the timing of all events. Using a combination of position measurements from a global positioning system (GPS) receiver and site surveying, we determine the locations of Alice, Bob, and the source with an uncertainty of < 1m. This uncertainty is set by the physical size of the cryostat used to house our detectors and the uncertainty in the GPS coordinates. There are four events that must be spacelike separated: Alice's and Bob's measurement choice must be fixed before any signal emanating from the photon creation event could arrive at their locations, and Alice and Bob must finish their measurements before information from the other party's measurement choice could reach them. Due to the slight asymmetry in the locations of Alice, Bob, and the source, the time difference between Bob finishing his measurement and information possibly arriving about Alice's measurement choice is al- ways shorter than the time differences of the other three events as shown in Fig. 3(b). This time difference serves as a kind of margin; our system can tolerate timing errors as large as this margin and still have all events remain spacelike separated. For one, three, five, and seven aggregate pulses this corresponds to a margin of 63.5 ± 3.7 ns, 50.9 ± 3.7 ns, 38.3 ± 3.7 ns, and 25.7 ± 3.7 ns, respectively as shown in Table I. The uncertainty in these timing measurements is dominated by the 1 m positional uncertainty (see Supplemental Material for

14

further details on the timing measurements).

A way to visualize and further quantify the the space- like separation of events is to compute how far Alice, Bob, and source could move from their measured position and still be guaranteed to satisfy the locality constraints, assuming that the chronology of all events remains fixed. In figure 4(a) Alice, Bob, and the source locations are surrounded by shaded green regions. As long as each party remains anywhere inside the boundaries of these regions their events are guaranteed to be spacelike separated. There are specific configurations where all three parties can be outside the boundaries and still be spacelike separated, but here we consider the most conspiratorial case where all parties can collude with one another. The boundaries are over-layed on architectural drawings of the building in which the experiment was performed. Four different boundaries are plotted, corresponding to the Bell test performed with one, three, five, and seven aggregate pulses. Minimizing over the path of each boundary line, the minimum distance that Alice, Bob, and the source are located from their respective boundaries is 9.2 m, 7.3 m, 5.4 m, and 3.5 m for aggregates of one pulse, three pulses, five pulses, and seven pulses, respectively. For these pulse configurations we would have had to place our source and detection systems physically in different rooms (or even move outside of the building) to compromise our spacelike separation. Aggregating more than seven pulses leads to boundaries that are less than three meters away from our measured positions. In these cases we are not able to make strong claims about the spacelike separation of our events.

Finally, as shown in Fig. 4(b), we can compute the 15 p-values for each of the time slots we consider that pho- tons from the source can arrive in every trial. Photons arriving in slots 2 through 11 are spacelike separated while photons in slots 12 through 15 are not. The photons arriving in these later slots are measured after information from the other party's random number generator could arrive as shown in Fig. 3(b). It appears that spacelike separation has no discernible effect on the statistical significance of the violation. However, we do see large slot-to-slot fluctuation in the calculated p-values. We suspect that this is due to instability in the applied voltage when the Pockels cell is turned on. In this case photons receive slightly different polarization rotations depending on which slot they arrive in, leading to non-ideal measurement settings at Alice and Bob. It is because of this slot-to-slot variation that the aggregate of seven pulses has a computed p-value larger than the

five-pulse case. Fixing this instability and using more sophisticated hypothesis test techniques [39-41] will enable us to robustly increase the statistical significance of our violation for the seven pulse case.

The experiment reported here is a commissioning run of the Bell test machine we eventually plan to use to certify randomness. The ability to include multiple pulses in our Bell test highlights the flexibility of our system. Our Bell test machine is capable of high event rates, making it well suited for generating random numbers required by cryptographic applications [29]. Future work will focus on incorporating our Bell test machine as an additional source of real-time randomness into the National Institute of Standards and Technology's public random number beacon (`https://beacon.nist.gov`).

It has been 51 years since John Bell formulated his test of local realism. In that time his inequality has shaped our understanding of entanglement and quantum correlations, led to the quantum information revolution, and transformed the study of quantum foundations. Until recently it has not been possible to carry out a complete and statistically significant loophole-free Bell test. Using advances in random number generation, photon source development, and high-efficiency single-photon detectors, we are able to observe a strong violation of a Bell inequality that is loophole free, meaning that we only need to make a minimal set of assumptions. These assumptions are that our measurements of locations and times of events are reliable, that Alice's and Bob's measurement outcomes are fixed at the time taggers, and that during any given trial the random number generators at Alice and Bob are physically independent of each other and the properties of the photons being measured. It is impossible, even in principle, to eliminate a form of these assumptions in any Bell test. Under these assumptions, if a hidden variable theory is local it does not agree with our results, and if it agrees with our results then it is not local.

16

[1] J. S. Bell, Epistemological Letters , 2 (1975).

[2] P. Holland, Found. Phys. 35, 177 (2005), arXiv:quant-ph/0401017.

[3] L. de Broglie, J. Phys. Radium 8, 225 (1927).

[4] D. Bohm, Phys. Rev. 85, 166 (1952).

[5] D. Bohm, Phys. Rev. 85, 180 (1952).

[6] A. Einstein, B. Podolosky, and N. Rosen, Phys. Rev. 47, 777 (1935).

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

[8] J. S. Bell, Physics 1, 195 (1964).

[9] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).

[10] J. F. Clauser and M. A. Horne, Phys. Rev. D 10, 526 (1974).

[11] S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. 28, 938 (1972).

[12] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. 47, 460 (1981).

[13] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. 49, 91 (1982).

[14] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. 49, 1804 (1982).

[15] M. Genovese, Phys. Rep. 413, 319 (2005), arXiv:quant-ph/0701071.

[16] J.-Â. Larsson, J. Phys. A 47, 424003 (2014), arXiv:1407.0363.

[17] C. Abella?n, W. Amaya, D. Mitrani, V. Pruneri, and M. W. Mitchell, (2015), arXiv:1506.02712.

[18] M. W. Mitchell, C. Abellan, and W. Amaya, Phys. Rev. A 91, 012314 (2015).

[19] P. H. Eberhard, Phys. Rev. A 47, R747 (1993).

[20] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. 81, 5039 (1998), arXiv:9810080.

[21] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, Nature 409, 791 (2001).

[22] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. K. Langford, T. Jennewein, and A. Zeilinger, Proc. Nat. Acad. Sci. USA 107, 19708 (2010), arXiv:0811.3129.

[23] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, Nature 497, 227 (2013), arXiv:1212.0533.

[24] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat, Phys. Rev. Lett. 111, 130406 (2013), arXiv:1306.5772.

[25] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Nature 526, 682 (2015).

[26] M. Guistina, Private communication.

[27] P. Bierhorst, J. Phys. A 48, 195302 (2015).

[28] J. Shao, Mathematical Statistics, Springer Texts in Statistics (Springer, New York, 1998) See 2nd edition pages 126-127.

[29] S. Pironio, A. Ac??n, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature 464, 1021 (2010), arXiv:0911.3427.

[30] R. D. Gill, in Mathematical Statistics and Applications: Festschrift for Constance van Eeden, Vol. 42, edited by M. Moore, S. Froda, and C. Léger (Institute of Mathematical Statistics. Beachwood, Ohio, 2003) pp. 133-154, arXiv:quant-ph/0110137.

[31] P. B. Dixon, D. Rosenberg, V. Stelmakh, M. E. Grein, R. S. Bennink, E. A. Dauler, A. J. Kerman, R. J. Molnar, and F. N. C. Wong, Phys. Rev. A 90, 043804 (2014).

[32] L. K. Shalm, K. Garay, J. Palfree, A. L. Migdall, A. U?Ren, and S. W. Nam, "Spontaneous parametric downcoversion calculator", `http://www.spdcalc.org`.

[33] P. G. Evans, R. S. Bennink, W. P. Grice, T. S. Humble, and J. Schaake, Phys. Rev. Lett. 105, 253601 (2010).

[34] All uncertainties U and error bars correspond to an estimated standard deviation, σ, and a coverage factor $k = 1$ as $U = k\sigma$.

[35] F. Marsili, V. B. Verma, J. A. Stern, S. Harrington, A. E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. D. Shaw, R. P. Mirin, and S. W. Nam, Nature Photonics 7, 210 (2013), arXiv:1209.5774.

[36] D. N. Klyshko, Sov. J. Quantum Electron. 10, 1112 (1980).

[37] Raw data from each experimental run will be made available online.

[38] P. Bierhorst, (2013), arXiv:1312.2999.

[39] Y. Zhang, S. Glancy, and E. Knill, Phys. Rev. A 84, 062118 (2011), arXiv:1108.2468.

[40] Y. Zhang, S. Glancy, and E. Knill, Phys. Rev. A 88, 052119 (2013), arXiv:1303.7464.

[41] E. Knill, S. Glancy, S. W. Nam, K. Coakley, and Y. Zhang, Phys. Rev. A 91, 032105 (2015).