Einstein, Podolsky, and Rosen (EPR) argued that the quantum mechanical wave function is an incomplete description of physical reality [1]. They started their discussion by noting that quantum mechanics predicts perfect correlations between the outcomes of measurements on two distant entangled particles. This is best discussed considering Bohm’s example of two entangled spin-1/2 atoms [2,3], which are emitted from a single spin-0 molecule and distributed to two distant observers, now commonly referred to as Alice and Bob. By angular momentum conservation, the two spins are always found to be opposite. Alice measures the spin of atom 1 in a freely chosen direction. The result obtained allows her to predict with certainty the outcome of Bob should he measure atom 2 along the same direction. Since Alice could have chosen any possible direction and since there is no interaction between Alice and Bob anymore, one may conclude that the results of all possible measurements by Bob must have been predetermined. However, these predeterminate values did not enter the quantum mechanical description via the wave function. This is the essence of the argument by EPR that the quantum state is an incomplete description of physical reality [1].

Bell’s theorem states that quantum mechanics is incompatible with local realism. He showed that if we assume, in line with Einstein’s theory of relativity, that there are no physical influences traveling faster than the speed of light (the assumption of locality) and that objects have physical properties independent of measurement (the assumption of realism), then correlations in measurement outcomes from two distant observers must necessarily obey an inequality [4]. Quantum mechanics, however, predicts a violation of the inequality for the results of certain measurements on entangled particles. Thus, Bell’s inequality is a tool to rule out philosophical standpoints based on experimental results. Indeed, violations have been measured.

Do these experimental violations invalidate local realism? That is not the only logical possibility. The experimental tests of Bell’s inequality thus far required extra assumptions, and therefore left open loopholes that still allow, at least in principle, for a local realist explanation of the measured data. (Note that empirically closing a
loophole might still require the validity of some specific assumptions about the experiment.)

The locality loophole (or communication loophole) is open if the setting choice or the measurement result of one side could be communicated to the other side in time to influence the measurement result there. Spacelike separation of each local measurement from both the distant setting choice and the distant measurement closes the locality loophole. In an experiment, this can be ensured by independently choosing the measurement settings on both sides so quickly that no physical signal (limited by the speed of light) can pass information about the chosen setting or the measurement result to the other side in time to be relevant for the measurement there.

The freedom-of-choice loophole refers to the requirement, formulated by Bell, that the setting choices are “free or random” [5]. For instance, this would prohibit a possible interdependence between the choice of measurement settings and the properties of the system being measured. Following Bell, we describe all properties of the system with the variable $\lambda$, which represents “any number of hypothetical additional complementary variables needed to complete quantum mechanics in the way envisaged by EPR” [5]. This loophole can be closed only under specific assumptions about the origin of $\lambda$. Under the assumption that $\lambda$ is created with the particles to be measured, an experiment in which the settings are generated independently at the measurement stations and spacelike separated from the creation of the particles closes the loophole.

The fair-sampling loophole (or detection loophole) refers to the following issue: It is conceivable under local realism that a subensemble of emitted particles violates a Bell inequality, while the total ensemble does not. The loophole is exploited if an experiment detects only this subensemble and assumes that it represents the entire ensemble [6]. It is possible to close the loophole by detecting the particles with adequate efficiency; the situation can be made even cleaner by using a Bell inequality that does not make such a “fair sampling” assumption in the first place.

There is a long history [7,8] of experimental Bell tests [9–20] addressing individual loopholes, though the long-term goal has always been to close all loopholes in a single experiment. To name a few, Aspect et al.’s 1982 experiment [10] first employed rapid switching of the measurement settings; Weis et al.’s 1998 experiment [11] improved this with fast random switching; Scheidl et al. [16] addressed freedom-of-choice in 2010 while also closing the locality loophole; Rowe et al. [12] were first to close the fair-sampling loophole in 2001 and were followed by several experiments in a variety of systems [13,15,18–20]. It has only recently become possible to address all three aforementioned loopholes in a single experiment [20,21] (where [20] reported a $p$ value of 0.039). Here, we report the violation of a Bell inequality using polarization-entangled photon pairs, high-efficiency detectors, and fast random basis choices spacelike separated from both the photon generation and the remote detection. We simultaneously close all three aforementioned loopholes in a single experiment with high statistical significance and thus provide strong support for the idea that nature cannot be described within the framework of local realism.

The experimental setup, located in the sub-basement of the Vienna Hofburg castle, is illustrated in Fig. 1(a). Our pair source [Fig. 1(b)] used spontaneous parametric down-conversion (SPDC) in a periodically poled nonlinear crystal and generated high-quality entanglement using a Sagnac configuration [22]. We optimized the focusing parameters of the pump and down-conversion collection modes for high heralding efficiency [23,24]. The photons were coupled into single-mode fibers and distributed to two distant measurement stations “Alice” and “Bob” [Fig. 1(c)] where polarization measurements were performed in one of two setting angles per side. While the photons were in flight, the choice of measurement setting was made in each station by a random number generator (RNG) [25,26] situated there. The measurement was implemented by a fast electro-optical modulator (EOM) followed by a polarizer and a transition-edge sensor (TES) single-photon detector [27]. The signal from the TES was amplified by a series of superconducting [28] and room-temperature amplifiers, digitized, and recorded locally on a hard drive. In addition, each implemented setting was recorded locally at each measurement station using a time-tagging module. The photon and setting data stored locally in the measurement stations were collected by a separate computer that evaluated the Bell inequality.

To close the freedom-of-choice and locality loopholes, a specific space-time configuration of the experiment was chosen, as depicted in the center plot of Fig. 2. As discussed before, it was necessary to spacelike separate each local setting choice (green bars labeled a and b) from the measurement on the other side (red bars A and B), as well as from the photon emission (blue bar E), which we consider to be the origin of $\lambda$.

To ensure synchronized timing throughout the experiment, we locked the RNG, TTM (time-tagging module), and digitizer to a 10 MHz master oscillator. A 1 MHz clock, phase synchronized to this master oscillator, regulated the laser pulsing and switching of Pockels cells. To confirm the space-time configuration in our experiment, we precisely characterized the delays of all relevant electrical and optical signals relative to this clock using an oscilloscope and a fast photodiode. In particular, we characterized three events in the experiment:

1. Emission.—The origin of the space-time diagram, indicated as a blue dot, represents the earliest possible photon emission. This point corresponds to the leading edge of the pump laser pulse reaching the SPDC crystal in the source. The length of the blue bar E indicates the pulse duration of the pump laser.
Setting choice.—We allow approximately 26 ns for the RNG to produce and deliver a setting choice by generating four raw bits and computing their parity. The randomness in each raw bit is derived from the phase, randomized by spontaneous emission, of an optical pulse. These 26 ns include a creation and throughput time of approximately 11 ns for one raw bit and an additional 15 ns for three additional bits. As described in the Supplemental Material [29], this reduces the chance of predicting the settings to \( \varepsilon_{\text{RNG}} \leq 2.4 \times 10^{-4} \) [25,26]. The solid green horizontal lines in the space-time diagram indicate the latest possible time at which the random phase was sampled inside the respective RNGs for use in a setting choice, while the dashed green lines indicate the earliest possible random phase creation for the first (of the four) contributing raw bit. The configuration ensures conservatively estimated margins of \( \approx 4 \text{ ns} \) for the spacelike separation of each setting from the distant measurement and \( \approx 7 \text{ ns} \) for the spacelike separation of each setting from the emission event (see Fig. 2 for more detail, including error estimates).

Measurement.—After a photon pair is emitted by the crystal, the photons are coupled into two single-mode optical fibers that direct one photon each to Alice’s and Bob’s distant locations. At each measurement station, the photons are coupled out of fiber and sent through an electro-optical modulator and a polarizer that transmit a particular polarization based on the setting choice from the RNG. The photons transmitted through the polarizer are coupled back into optical fiber (SMF-28) and sent to the TES. For monitoring purposes, we use an avalanche photodiode to detect the photons that are reflected from the polarizer (black histograms in Fig. 2). Using the arrival time information from this monitoring port, and assuming photons travel at the speed of light in their respective media, we infer that the latest time a photon could arrive at a TES after being emitted from the source is approximately 195 ns. This is represented by the dashed orange and blue lines on the space-time diagram in Fig. 2. After a photon is absorbed by the TES, the resulting detection signal is read out using a SQUID sensor that introduces jitter into the signal. This electrical signal then travels through cables until it reaches a digitizer (the signals take approximately 64.4 and 65.5 ns to travel from the TES to Alice’s and Bob’s digitizers, respectively). Because the shape of the read-out signal depends on the energy of the photons absorbed by the TES, the shape can be used to distinguish both unwanted background light (primarily blackbody photons) and excess noise from the 810 nm...
photon pulses produced by the source. We therefore use the digitizer to record the profiles of these amplified TES pulses. When the amplified signal from the TES crosses a voltage threshold (around 55% of the expected height of an 810 nm photon determined from calibration data), the signal is saved by the data acquisition system for further processing. During the analysis, if the recorded trace crosses a voltage level fixed at around 75% of the expected pulse height, then it is considered a detection event. This level was chosen to eliminate with near certainty lower-energy blackbody photons. The time that the trace crosses a level set at around 20% of the expected pulse height is used to timestamp the detection event. We consider the detection event to be complete and the outcome fixed by this point. Histograms of these detection times relative to the start of each trial are shown in orange and blue in Fig. 2. After accounting for cable delays, all events that fall inside the measurement windows A and B are ensured to be space-like separated from the relevant setting choice at the other party.

Closure of the fair-sampling loophole does not rely on space-time considerations and can be observed in the experimental data. The Clauser-Horne (CH) [36] or the Eberhard [37] inequality can be derived without the fair-sampling assumption. These inequalities can be violated with system heralding efficiencies larger than 2/3. We
employed a CH-Eberhard (CH-E) type inequality, which requires only one detector per side and restricts the probabilities of outcomes—“+” for a detection and “0” for no detection—in the following way [38,39]:

\[
J = p_{++}(a_1b_1) - p_{+0}(a_1b_2) - p_{0+}(a_2b_1) - p_{++}(a_2b_2) \leq 0.
\] (1)

In every trial, Alice chooses setting \(a_1\) or \(a_2\), and Bob chooses \(b_1\) or \(b_2\). They write down their respective outcomes “+” or “0”. Combining their data at the end of the experiment, they estimate the probabilities that appear in the inequality. For example, \(p_{+0}(a_1b_2)\) is the probability that, conditioned on the setting choices \(a_1\) and \(b_2\) for a given trial, Alice observes a detection event and Bob registers no detection. Our experiment employed locally defined time slots and was thus also not vulnerable to the coincidence-time loophole [40,41].

The inequality can be violated using Eberhard states of the form [37]

\[
|\Psi\rangle = \frac{1}{\sqrt{1 + r^2}} (|V\rangle_A|H\rangle_B + r|H\rangle_A|V\rangle_B),
\] (2)

where \(H\) and \(V\) are horizontal and vertical polarizations and the subscripts \(A\) and \(B\) indicate Alice’s and Bob’s photons, respectively.

The optimal values for \(r\) and the setting angles depend on the performance of the setup and can be estimated using a quantum mechanical model [42]. We characterized the system using the product state \((r = 0)\) and the maximally entangled state \((r = -1)\). We found visibilities of over 99% in both the HV and diagonal bases, and system heralding efficiencies of approximately 78.6% in the Alice arm and approximately 76.2% in the Bob arm. These efficiencies represent a ratio of twofold coincidence events divided by singles counts (i.e., total events measured in one detector) directly measured over the entire system and not corrected for any losses. We set a state with \(|\Psi\rangle\) for a natural stopping point—\(7.27 \times 10^{-6}\). For the pure state [Eq. (2)], the above mentioned detection efficiencies, and 3500 down-conversion pairs produced per second (see Supplemental Material [29]), quantum mechanics predicts an optimal \(J\) value of about \(4 \times 10^{-5}\) [42]. That the measured value is smaller can be explained mostly by nonunity state visibility and nonzero background.

We compute the statistical significance of our measured violation under full experimental memory [43–45], without assuming independent and identically distributed (IID) experimental trials [38]. We also account for the excess predictability of the random setting choices and find that under local realism, the probability of observing our measured \(J\) value does not exceed a \(p\) value of \(3.74 \times 10^{-31}\) (see Supplemental Material [29]). Our analysis uses neither Gaussian approximation nor the IID assumption, but for comparison, for a large-sample experiment that allows these two, an 11.5-sigma violation gives this \(p\) value. In light of such an exceedingly small \(p\) value, we remark that the confidence in the experiment as a whole is limited not by the statistical strength of the violation but rather by other more general errors, which might happen in any experiment and could, for example, be systematic, human, or come from other limitations of the apparatus.

Our experiment showed a strong violation of local realism using exacting experimental technique and rigorous statistical analysis. Employing state of the art random number generators, we spacelike separated the setting choices, measurements, and emission event to close the locality and freedom-of-choice loopholes simultaneously. We achieved high system heralding efficiencies and closed the fair-sampling loophole as well. In addition, we closed the coincidence-time loophole in our experiment by using locally defined time slots. We closed the memory loophole by computing the statistical significance of the violation without assuming independently and identically distributed experimental trials. Our experiment provides the strongest support to date for the viewpoint that local realism is untenable.

By closing the freedom-of-choice loophole to one natural stopping point—the first moment at which the particles come into existence—we reduce the possible local-realist explanations to truly exotic hypotheses. Any theory seeking to explain our result by exploiting this loophole would require \(\lambda\) to originate before the emission event and to influence setting choices derived from spontaneous emission. It has been suggested that setting choices determined by events from distant cosmological sources could push this limit back by billions of years [46].

![Bar chart of the four joint probabilities entering the Bell inequality (1). Since the green bar representing \(p_{++}(a_1b_1)\) outweighs the sum of the other three red bars, the \(J\) value is positive and the CH-Eberhard inequality is violated.](250401-5)
We thank Reinhold Sahl and the Vienna Hofburg for the use of their sub-basement. We acknowledge Scott Glancy, Sven Ramelow, Johannes Steurer, Max Tillmann, and Witleif Wieczorek for technical assistance and helpful discussions. M. G. acknowledges support by the program CoQuS of the FWF (Austrian Science Fund). J. K. thanks Lucas Clemente for help with implementing data analysis code and acknowledges support from the EU Integrated Project SIQS. C. A., W. A., V. P., and M. W. M. acknowledge the European Research Council Project AQUOMET, European Union Project QUIC (Grant Agreement No. 641122), Spanish MINECO under the Severo Ochoa programme (Grant No. SEV-2015-0522) and Projects MAGO (Grant No. FIS2011-23520) and EPEC (Grant No. FIS2014-62181-EXP), Catalan AGAUR 2014 SGR 1295 and No. 1623, the European Regional Development Fund (FEDER) Grant No. TEC2013-46168-R, and by Fundació Privada CELLEX. This work was also supported by the Austrian Science Fund (FWF) with SFB F40 (FOQUS). J. K. thanks discussions. M. G. acknowledges support by the program Witlef Wieczorek for technical assistance and helpful use of their sub-basement. We acknowledge Scott Glancy, Sven Ramelow, Johannes Steurer, Max Tillmann, and Witleif Wieczorek for technical assistance and helpful discussions. M. G. acknowledges support by the program CoQuS of the FWF (Austrian Science Fund). J. K. thanks Lucas Clemente for help with implementing data analysis code and acknowledges support from the EU Integrated Project SIQS. C. A., W. A., V. P., and M. W. M. acknowledge the European Research Council Project AQUOMET, European Union Project QUIC (Grant Agreement No. 641122), Spanish MINECO under the Severo Ochoa programme (Grant No. SEV-2015-0522) and Projects MAGO (Grant No. FIS2011-23520) and EPEC (Grant No. FIS2014-62181-EXP), Catalan AGAUR 2014 SGR 1295 and No. 1623, the European Regional Development Fund (FEDER) Grant No. TEC2013-46168-R, and by Fundació Privada CELLEX. This work was also supported by the Austrian Academy of Sciences (ÖAW), the European Research Initiative. This project was supported by the Austrian Development Fund (FEDER) Grant No. TEC2013-46168-R, and by Fundació Privada CELLEX. This work was also supported by the NIST Quantum Information Science Initiative. This project was supported by the Austrian Academy of Sciences (ÖAW), the European Research Council (SIQS Grant No. 600645 EU-FP7-ICT), and the Austrian Science Fund (FWF) with SFB F40 (FOQUS).

†marissa.giustina@univie.ac.at
anton.zeilinger@univie.ac.at