

High-visibility time-bin entanglement for testing chained Bell inequalitiesMarco Tomasin,^{1,2} Elia Mantoan,^{1,2} Jonathan Jogenfors,³ Giuseppe Vallone,^{1,2} Jan-Åke Larsson,³ and Paolo Villoresi^{1,2}¹*Department of Information Engineering, University of Padova, I-35131 Padova, Italy*²*Istituto di Fotonica e Nanotecnologie, INF-CNR, UOS di Padova, I-35131 Padova, Italy*³*Institutionen för systemteknik, Linköpings Universitet, SE-58183 Linköping, Sweden*

(Received 13 December 2016; published 9 March 2017)

The violation of Bell's inequality requires a well-designed experiment to validate the result. In experiments using energy-time and time-bin entanglement, initially proposed by Franson in 1989, there is an intrinsic loophole due to the high postselection. To obtain a violation in this type of experiment, a chained Bell inequality must be used. However, the local realism bound requires a high visibility in excess of 94.63% in the time-bin entangled state. In this work, we show how such a high visibility can be reached in order to violate a chained Bell inequality with six, eight, and ten terms.

DOI: [10.1103/PhysRevA.95.032107](https://doi.org/10.1103/PhysRevA.95.032107)**I. INTRODUCTION**

In his well-known work of 1989 [1], Franson proposed a Bell inequality for energy-time entanglement to investigate local realism. The proposed experimental configuration consists of a source emitting two correlated photons that are generated by a spontaneous parametric down-conversion (SPDC) process with a continuous-wave (cw) pump laser. The two photons are directed towards two measurement stations, each consisting of identical unbalanced Mach-Zehnder interferometers with path-length difference ΔL .

The measurement stations can either impose or not impose a time delay $\Delta L/c$, depending on the local phase setting. Because of this, photon detection may coincide in time or be delayed on one side with respect to the other. The optical path difference ΔL satisfies $\tau_c \ll \Delta L/c \ll \tau_p$, where τ_c and τ_p are the coherence time of the SPDC photons and the pump, respectively. The lower bound ensures there is no first-order interference of single-photon counts and the upper bound ensures second-order interference for coincident photon counts.

From all two-photon detections, only the 50% coincident detections correspond to entangled photons. Indeed, by postselecting those events where both photons are detected within the time window $\Delta L/c$, two-photon interference occurs and a Bell inequality can be violated, modulo the postselection loophole discussed in Sec. II.

A different way of performing the experiment is by using time-bin entanglement. This method differs from the standard Franson setup in that it uses a pulsed laser pump and an additional unbalanced interferometer in the source. Here, the photon pairs can be generated from the same pump pulse at two different moments in time. In this case, coherence is ensured by using a pump interferometer rather than having to rely on the pump to be coherent, as would have been the case for a cw pump without the interferometer.

As pointed out in several works [2–5], Franson's scheme suffers from an intrinsic loophole because of the aforementioned postselection. Which events that are discarded could, in principle, be influenced by the measurement settings, and, indeed, a local hidden-variable (LHV) model can reproduce the quantum correlations for the second-order interference exploiting the discarded events [2,3,5,6]. The main issue is

that the postselection is “nonlocal” in the sense that it requires communication between the observers to know which events should be discarded.

Note that the postselection loophole is different from the more familiar problem that arises from low experimental efficiency. At low efficiency, detection occurs only on a subset of hidden-variable values that is determined in a completely local fashion. Coincidence, in turn, occurs on a subset that is an intersection of two locally determined sets, so the selection of events is locally separable and therefore has less severe implications; see Ref. [7] for details. The focus of the present paper is the postselection loophole described above, and we will use the fair-sampling assumption [8] in the remainder of the paper.

Different approaches have been proposed to reach a violation of local realism avoiding the postselection loophole. As described by [3–5], there are ways of reestablishing the security of the Franson setup, including modifying the experimental setup [3,9–12]. Additional assumptions on the underlying physics have also been discussed by Franson [13]; these, however, have given rise to new, undesirable loopholes [2,4].

A consequence of using a time-bin entanglement setup instead of energy time is that it introduces a local postselection. The generated photon can be detected only at three possible times ($t_0 - \Delta L/c, t_0, t_0 + \Delta L/c$) due to the pulsed laser pump: the three arrival times will be denoted as early (E), medium (M), and late (L). Each observer can perform a local postselection by discarding all events except those occurring at t_0 .

As will be shown later, such local postselection is not sufficient for a loophole-free violation of the Bell test. Instead, only 50% of all locally postselected events correspond to a coincidence and another, extra postselection step is then required. This extra step requires communication between Alice and Bob and is therefore nonlocal. As we will show below, the LHV model [3,14] previously introduced for energy-time entanglement can be slightly modified to obtain an LHV model also for time-bin entanglement.

In this work, we use the so-called chained Bell inequalities [15–21] in order to obtain a violation of local realism in a time-bin experiment. In order to facilitate this, we have designed and realized a source of time-bin entangled photons

with high visibility. It was previously believed that the visibility requirements on such an experiment would be highly demanding, but we can demonstrate visibilities up to 99% that make such a violation possible.

II. THEORY

In a time-bin Bell experiment, a source device generates two time-correlated photons, of which one is sent to Alice and one to Bob. Alice and Bob each randomly select a measurement from the sets $\{A_i\}$ and $\{B_j\}$, respectively. Here, $0 \leq i, j \leq N$ for $N \geq 2$.

The pump interferometer is unbalanced with a difference in path length of ΔL . This leads to two distinct possibilities for when the photons are emitted. We call the short-short (SS) events the photons generated by the pump pulse traveling along the short arm, while long-long (LL) events are photons generated by the pump pulse traveling along the long arm.

The measurement stations consist of unbalanced Mach-Zehnder interferometers with two “paths” whose lengths differ by the same value ΔL . Note that “path” is written in quotation marks since one cannot force particles to have path realism without significantly altering the interpretation of the outcomes [4]. Instead, we view the effect of the measurement station as having either a long or short delay on the particle. Then, due to the pump interferometer and the measurement interferometer, each photon can be detected at three possible times: $(t_0 - \Delta L/c, t_0, t_0 + \Delta L/c)$.

If Alice and Bob perform measurements A_i and B_j , the outcome $(-1, +1)$ and the time of detection are recorded at either end. After a number of trials, Alice and Bob compare their measurement settings and time of arrival, and whenever both photons are detected at t_0 , that trial is used to compute the quantum correlation $\langle A_i B_j \rangle$. These correlations

are then used to compute a statistical measure, for instance the Clauser-Horne-Shimony-Holt (CHSH) value, which uses $N = 2$ measurement settings:

$$S_{\text{CHSH}} = \langle A_1 B_1 \rangle - \langle A_2 B_2 \rangle + \langle A_2 B_1 \rangle + \langle A_1 B_2 \rangle. \quad (1)$$

In a classical system, S_{CHSH} has an upper bound of 2 [8].

As explained above, the *time of arrival* is critical for each trial and is therefore recorded by both Alice and Bob. According to Cabello *et al.* [3] and Aerts *et al.* [14], a trial with differing time of arrival between Alice and Bob has no quantum interference, which means it must be discarded. Many events of this kind can be discarded locally, simply by considering only the events in which the photon arrived at t_0 . However, if Alice and Bob were to only keep the photons arriving at t_0 , then a detection on Alice’s side would have a 50% chance of not corresponding to a detection on Bob’s side, and vice versa.

In addition, a further “nonlocal” postselection is required to weed out noncoincident events. This extra step is an inherent feature of time-bin entanglement experiments and, on average, 50% of the locally selected data will be discarded here. As a consequence, the subensemble of finally selected events becomes dependent on the phase settings, which must be considered when choosing an appropriate Bell test for the experiment. With this dependence on phase settings, it becomes possible to mimic the quantum-mechanical results using a purely classical setup.

As shown in [4], if local realism holds in a time-bin entanglement experiment, the outcomes obey

$$S_{\text{re}} \leq 3, \quad (2)$$

and we note that the value predicted by quantum mechanics, $S_{\text{QM}} = 2\sqrt{2}$, is *not* in violation of inequality (2). In other words, a local hidden-variable (LHV) model such as Fig. 1 can give the same outcome statistics as a truly quantum-

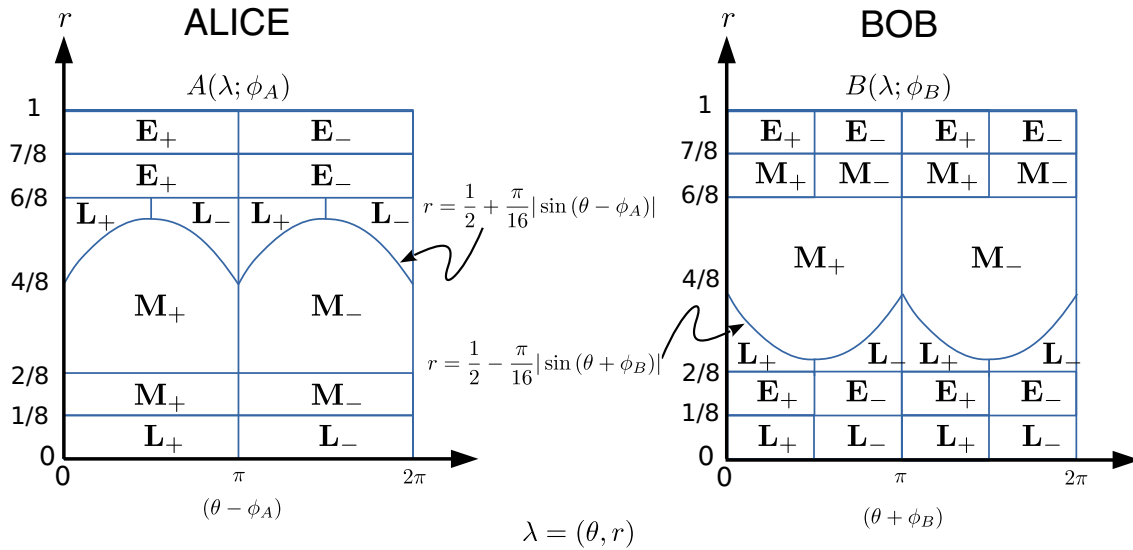


FIG. 1. LHV model for time-bin entanglement. The hidden variable is a pair of numbers $\lambda = (\theta, r)$ uniformly distributed over the rectangle $\Lambda = \{(\theta, r) : 0 \leq \theta < 2\pi, 0 \leq r < 1\}$. The outcomes $\{E_{\pm}, M_{\pm}, L_{\pm}\}$ are determined by the above graphs, where \pm corresponds to the possible outcomes and where early (E), medium (M), and late (L) correspond to the detection time. Indeed, E (M) events correspond to a photon generated from the first (second) pump pulse and traveling through the short (long) arm of the measuring interferometer. M events correspond to a photon generated from the first pump pulse and traveling through the long arm of the measuring interferometer, or vice versa (second pump pulse and short arm). This model fully reproduces the quantum predictions for the time-bin setup.

mechanical experiment, as predicted by [4] and experimentally demonstrated by [5]. This is called the *postselection loophole* [22].

For applications such as quantum key distribution (QKD), it is highly desirable to reestablish a violation of a Bell inequality in the time-bin case. In order to achieve this, we must eliminate the postselection loophole which in turn eliminates any LHV models such as the one in Fig. 1.

A possible way forward is to employ *fast switching* of the phase settings and postselect the coincidence events in which both photons arrived at t_0 . In this case, the Bell inequality holds for the LL events, while a trivial bound is obtained for the SS events (see [4]). Even when taking this into account, the standard Bell inequalities are insufficient as the quantum-mechanical prediction still falls short of the corresponding inequality [4].

However, a violation can be obtained [4] by using a generalized, chained Bell-type inequality. Here, we consider $N \geq 2$ measurements for each observer so that the chained Bell parameter is given by

$$S^N = \langle A_N B_N \rangle + \sum_{k=2}^N [\langle A_k B_{k-1} \rangle + \langle A_{k-1} B_k \rangle] - \langle A_1 B_1 \rangle. \quad (3)$$

If A_i and B_j are phase measurement operators, the set of phases that maximize Eq. (3) are in consecutive measurements separated by an angle of $\theta = \pi/(2N)$. The quantum prediction in this case is

$$S_{\text{QM}}^N = 2N \cos\left(\frac{\pi}{2N}\right), \quad (4)$$

with a classical bound of $2N - 2$ [4,16]. In theory, such a chained inequality can provide a strong violation, but was believed [4,5] to be experimentally demanding to the point of being unfeasible.

However, in a time-bin experiment, the realism bound is weaker than in a standard Bell experiment (see Appendix A for details) and it is equal to

$$S_{\text{LHV}}^N = 2N - 1. \quad (5)$$

The quantum prediction will be in violation of local realism if $S_{\text{QM}}^N > S_{\text{LHV}}^N$, which can occur for $2N \geq 6$. The bound in Eq. (5) is due to the fact that the SS events can be influenced by the phase settings, while the LL subensemble is independent of the phase setting [4].

Finally, there are purely experimental requirements on any Bell experiment. If the experiment has lower than 100% visibility, the bound in Eq. (5) will be weakened so that local realism is not in violation. For $2N \in \{6, 8, 10\}$, the critical visibility $V_{\text{cr}} = S_{\text{LHV}}/S_{\text{QM}}^N$ is at least 94.63% [4], which is experimentally demanding.

As detailed in Appendix A, the inequality can be also written in the CH-Hardy form involving only probabilities as

$$S_{\text{CH}}^N = p(a_N b_N) - \sum_{k=2}^N [p(a_k \bar{b}_{k-1}) + p(\bar{a}_{k-1} b_k)] - p(a_1 b_1) \leq 0, \quad (6)$$

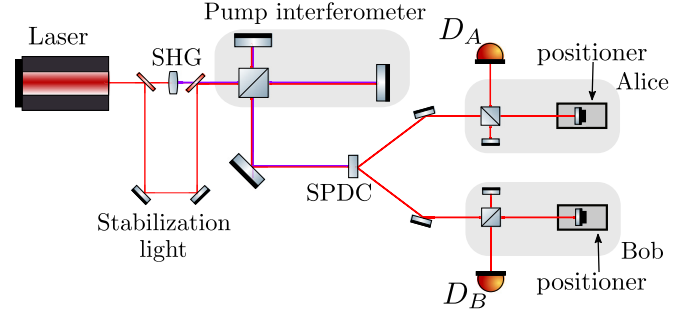


FIG. 2. Setup of the time-bin entanglement experiment.

where $p(a_i b_j)$ is the joint probability of measuring $+1$ on Alice's and Bob's side, respectively, and the bar indicates a -1 outcome. We note that $S^N = 4S_{\text{CH}}^N + 2(N - 1)$. For time-bin entanglement, the correct LHV bound is

$$\frac{3}{4} - N \leq S_{\text{CH}}^N \leq \frac{1}{4}, \quad (7)$$

which is violated by the quantum prediction

$$S_{\text{CH,QM}}^N = \frac{1}{2} - N \sin^2 \frac{\pi}{4N}, \quad (8)$$

when $N \geq 3$.

III. EXPERIMENTAL SETUP

As mentioned in Sec. II, previous works [4,5] described the chained Bell inequality as experimentally demanding as it requires a visibility in excess of 94%. In this section, however, we give details of an experiment, depicted in Fig. 2, that gives us a high enough visibility to have a working implementation of a chained Bell test in a time-bin entangled Franson experiment.

Obviously, this requires a very stable and precise measurement setup. The laser source used in our experiment is a pulsed laser with a wavelength of 808 nm, 76 MHz repetition rate and ≈ 150 fs pulse width coming from a Ti:sapphire mode-locked oscillator, with an average power of 2 W. The Ti:sapphire oscillator beam is used to pump a second harmonic generation (SHG) crystal, which generates pulses at 404 nm.

These pulses pass through a free-space unbalanced Michelson interferometer (pump interferometer) where the light "takes" the short $|S_p\rangle$ or the long $|L_p\rangle$ path and finally pumps a SPDC crystal. As previously described in Sec. II, the length difference between the short and long paths is ΔL . The output of the source is photons with a wavelength of 808 nm that are injected into two measurement stations, one for Alice and one for Bob.

Each measurement station consists of a free-space unbalanced Mach-Zehnder interferometer with a path-length difference of ΔL , matching the path-length difference of the source device. The mirrors on the long paths are mounted on a piezoelectric translator in order to stabilize and change the phase of the interferometers. Finally, the photons are coupled into a single-mode fiber and directed toward two single-photon avalanche photodiodes (SPADs). A time-to-digital converter (TDC) with 81 ps resolution is used to tag each detection event.

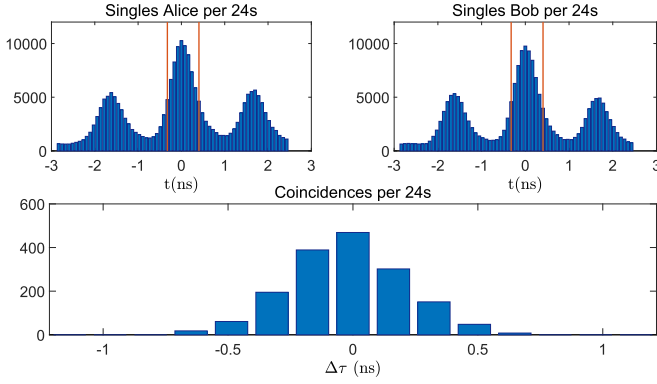


FIG. 3. Time distribution of singles and coincidence count. The red bar on the singles graph represents the time interval within which we consider an event to be coincident. The coincidences graph represents the coincidences between single events within the red bars, in a window of 0.81 ns.

The generated state corresponds to

$$|\Psi_e\rangle = \frac{1}{\sqrt{2}}(|S_A S_B\rangle + |L_A L_B\rangle). \quad (9)$$

Maintaining a constant phase difference between the two events $|S_A S_B\rangle$ and $|L_A L_B\rangle$ is an important step in achieving high visibility in Eq. (9).

In order to achieve a fixed phase difference, Alice's and Bob's phases are actively stabilized to the pump phase. This stabilization uses a small fraction of the original oscillator beam and we call this the *stabilization light*. This stabilization light is injected into the pump interferometer after an appropriate delay in order to prevent a detection overlap between the SPDC photons. The outgoing light is split into two beams and each follows the same path of the two SPDC photons. In this way, the second-order interference generated through the pump and the measurement stations can be used to stabilize and compensate the phase.

The interference pattern depends on the received photons. Therefore, the measurement stations are simultaneously stabilized independently of each other. Piezoelectric positioners with a resolution of 1 nm control the phase by measuring a feedback sensor. A measurement takes 3 s, after which a stabilization process takes over, requiring roughly 1 second before starting over with the measurement procedure. The exact time required by the stabilization depends on the number of single events.

Our stabilization allows us to stabilize not only the phase mismatch introduced in the paths, but also the phase mismatch due to the pump laser. This is important as the wavelength of the pump laser fluctuates due to variations in the environmental room temperature which affects the phase and therefore the visibility of the state.

IV. EXPERIMENTAL RESULTS

Our experimental results are presented in Fig. 3, which shows the distribution of single counts registered by Alice's and Bob's detectors together with the joint coincidence. Thanks to the pulsed pump laser, it is possible to predict

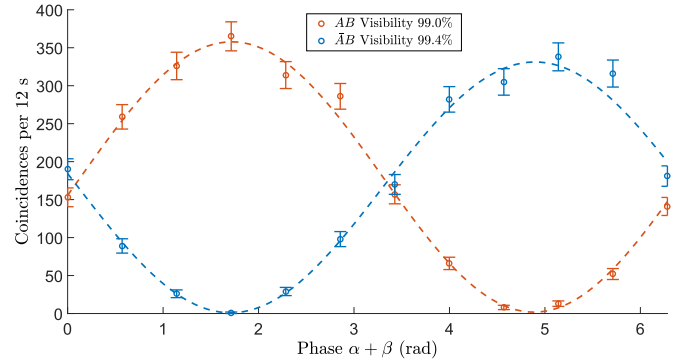


FIG. 4. Raw coincidence counts between Alice's and Bob's detectors for different values of phases (red data). The blue data represents the coincidence counts obtained by shifting Alice's phase by π with respect to the red data.

the arrival time of every generated photon at the detector. As expected, the single counts are concentrated around three distinct peaks, where the central peak corresponds to entangled events.

The coincidence window is indicated by the red bars in Fig. 3 and is determined *a priori* by using the emission time of the pump laser beam as a reference. The horizontal axis in the singles plots of Fig. 3 represents the arrival time of the Alice (or Bob) photon with respect to this reference. Here, $t = 0$ is the detection corresponding to $|L_A L_B\rangle$ and $|S_A S_B\rangle$ events. The coincidence window therefore matches $|t| \leq 0.405$ ns, corresponding to ± 5 bins of the TDC.

An event at Alice's measurement station is said to be coincident with an event at Bob's measurement station if they fall inside this coincidence window. Figure 3 shows the distribution of coincidences as a function of the time difference $\Delta\tau$ between Alice's and Bob's events. Note that, by definition, $|\Delta\tau|$ is smaller than the detection window $|\Delta\tau| \leq 0.810$ ns.

Our experiment uses only one detector each for Alice and Bob. Therefore, four measurements are required to calculate the correlation. This corresponds to a measurement in the standard Franson interferometer if we assume that the correlation obtained with such a measurement is equal to the correlation obtained with two detectors in a single measurement.

The experiment was performed three times, corresponding to a chained Bell inequality using $2N = 6, 8,$ and 10 , respectively. In all the cases, we violated the local hidden-variable bound. In the case of $2N = 6$, by choosing $\theta = \pi/6$ we obtained the experimental value of $S_{\text{CHSH}}^3 = 5.163 \pm 0.033$. For $2N = 8$ and $\theta = \pi/8$, we obtained $S_{\text{CHSH}}^4 = 7.169 \pm 0.034$. Finally, for $2N = 10$ and $\theta = \pi/10$, we measured $S_{\text{CHSH}}^5 = 9.271 \pm 0.031$. Statistical errors were evaluated by propagating the Poissonian errors of the coincidence counts. The measured results from these three experiments are shown in Table I.

As for visibility, our stabilization setup has good performance, as seen in Fig. 4. By measuring the coincidence as a function of the phase $\alpha + \beta$, we obtain a measured visibility up to 99% which is more than suitable for the chained Bell inequality.

TABLE I. Results for three different sets, with $N = 3, 4, 5$. $S_{\text{LHV},i}^N$ correspond to the value predicted by a LHV model. The reported $S_{\text{CH},i}^N$ has been calculated from the raw coincidence by using Eqs. (A3) and (A4), while the S_{CHSH}^N are obtained from Eq. (A8).

i	$S_{\text{LHV},i}^3$	$S_{\text{CH},i}^3$	err_S	Violation
1	0.25	0.289	0.011	3.40σ
2	-2.25	-2.335	0.020	4.15σ
3	-2.25	-2.247	0.020	-0.17σ
4	0.25	0.293	0.012	3.66σ
S_{CHSH}^3	5	5.163	0.033	4.91σ

i	$S_{\text{LHV},i}^4$	$S_{\text{CH},i}^4$	err_S	Violation
1	0.25	0.282	0.011	3.06σ
2	-3.25	-3.299	0.022	2.28σ
3	-3.25	-3.284	0.022	1.59σ
4	0.25	0.302	0.011	4.94σ
S_{CHSH}^4	7	7.169	0.034	4.93σ

i	$S_{\text{LHV},i}^5$	$S_{\text{CH},i}^5$	err_S	Violation
1	0.25	0.307	0.009	6.76σ
2	-4.25	-4.304	0.022	2.64σ
3	-4.25	-4.331	0.021	4.01σ
4	0.25	0.327	0.009	8.89σ
S_{CHSH}^5	9	9.271	0.031	8.67σ

V. CONCLUSIONS

Device independence is a powerful theoretical framework where many of the usual complications in applications such as QKD are reduced to a single statistical Bell test. Instead of a painstakingly complex interpretation where every component, path, defect, etc. impacts the physical interpretation of the experiment (for instance, the idea of ‘‘path realism’’ as discussed by [4]), device independence certifies the entire setup as correct in a single step.

In addition, time-bin entanglement has a higher inherent noise rejection [5] than traditional methods such as polarization encoding. This can lead to simpler devices with less moving parts than required by traditional protocols, opening the door for a wider variety of applications.

In this work, we show that is possible to reach a high visibility in a device-independent experiment using time-bin entanglement in order to violate a chained Bell inequality. By using $2N \geq 6$ settings at each measurement station, no local hidden-variable model can reproduce the predictions of a quantum experiments. The postselection loophole, which is present in the case $2N = 4$ (i.e., the standard CHSH-Bell inequality), is therefore avoided with the generalized, chained Bell inequality with $2N \geq 6$.

Note that in order to fully violate the chained Bell inequality, fast switching must be used so that the phase settings at the measurement stations are randomly chosen at least every $\Delta L/c$. This has not been performed by the current experiment, but could be done by using a fast phase

modulator, synchronized with the pulse laser, to change the phase within the interferometers. If $2N \geq 3$ is combined with fast switching, all requirements set out by [14] are fulfilled.

Generally, chained Bell inequalities demand a high experimental visibility ($\gtrsim 94.63\%$). Our experiment fulfills this condition with a visibility of 99% and shows that, despite the difficulties, it is therefore experimentally possible to implement chained Bell inequalities. This opens the door for applications such as device-independent quantum key distribution based on time-bin entanglement without the weaknesses inherent in the original Franson system.

APPENDIX A: CHAINED BELL INEQUALITIES WITH TIME-BIN ENTANGLEMENT

In this section, we show how to derive the chained Bell inequality for time-bin entanglement in two different forms, CH and CHSH. The chained Bell inequalities are a generalization of the CHSH and Clauser-Horne (CH) inequalities introduced by [15,16]: such inequalities consider the scenario in which Alice and Bob choose among $2N \geq 4$ dichotomic observables (with outputs +1 or -1).

The CH form of the chained Bell inequality is

$$1 - N \leq S_{\text{CH}}^N \leq 0, \quad (\text{A1})$$

where the Bell parameter is given by

$$S_{\text{CH}}^N = p(a_N b_N) + \sum_{k=2}^N [p(a_k b_{k-1}) + p(a_{k-1} b_k)] - p(a_1 b_1) - \sum_{k=2}^N [p(a_k) + p(b_k)]. \quad (\text{A2})$$

In the above expression, $p(a_i b_j)$ is the joint probability of measuring +1 on Alice’s and Bob’s side, respectively. If we assume fair sampling, the single-side probabilities $p(a_k)$ and $p(b_k)$ can be expressed as $p(a_k) = p(a_k b_{k-1}) + p(a_k \bar{b}_{k-1})$ and $p(b_k) = p(a_{k-1} b_k) + p(\bar{a}_{k-1} b_k)$, where \bar{a}_j means the -1 outcome. The above replacement leads to

$$S_{\text{CH},1}^N = p(a_N b_N) - \sum_{k=2}^N [p(a_k \bar{b}_{k-1}) + p(\bar{a}_{k-1} b_k)] - p(a_1 b_1). \quad (\text{A3})$$

Since we can arbitrarily define the measurement outcomes with the +1 and -1 inverted, the inequality also holds when a_k or b_k are replaced by \bar{a}_k and \bar{b}_k , obtaining three other Bell parameters given by

$$S_{\text{CH},2}^N = p(\bar{a}_N \bar{b}_N) - \sum_{k=2}^N [p(\bar{a}_k b_{k-1}) + p(a_{k-1} \bar{b}_k)] - p(\bar{a}_1 \bar{b}_1), \quad (\text{A4})$$

$$S_{\text{CH},3}^N = p(a_N \bar{b}_N) - \sum_{k=2}^N [p(a_k b_{k-1}) + p(\bar{a}_{k-1} \bar{b}_k)] - p(a_1 \bar{b}_1), \quad (\text{A5})$$

$$S_{\text{CH},4}^N = p(\bar{a}_N b_N) - \sum_{k=2}^N [p(\bar{a}_k b_{k-1}) + p(a_{k-1} b_k)] - p(\bar{a}_1 b_1). \quad (\text{A6})$$

By combining Eqs. (A3) to (A6), it is possible to derive the inequality in the CHSH form given by

$$|S_{\text{CHSH}}^N| \leq 2N - 2, \quad (\text{A7})$$

where

$$\begin{aligned} S_{\text{CHSH}}^N &= S_{\text{CH},1}^N + S_{\text{CH},2}^N - S_{\text{CH},3}^N - S_{\text{CH},4}^N \\ &= \langle A_N B_N \rangle - \sum_{k=2}^N [\langle A_k B_{k-1} \rangle + \langle A_{k-1} B_k \rangle] \\ &\quad - \langle A_1 B_1 \rangle. \end{aligned} \quad (\text{A8})$$

In Eq. (A8), the correlations are defined as

$$\begin{aligned} \langle A_k B_j \rangle &= p(a_k b_j) + p(\bar{a}_k \bar{b}_j) - p(\bar{a}_k b_j) - p(a_k \bar{b}_j) \\ &= 4p(a_k b_j) - 2p(a_k) - 2p(b_j) + 1. \end{aligned} \quad (\text{A9})$$

As discussed in Sec. II, Eq. (A7) does not apply to time-bin entanglement due to the postselection of events. Instead, we must distinguish between the early-early (EE) and the late-late (LL) events: Eq. (A1) holds only for the LL events, while the early-early (EE) events are only governed by the trivial inequality

$$|S_{\text{CHSH,EE}}^N| \leq 2N \quad (\text{A10})$$

corresponding to the following bound in the CH form:

$$\frac{1}{2} - N \leq S_{\text{CH,EE}}^N \leq \frac{1}{2}. \quad (\text{A11})$$

The bound is trivial since $S_{\text{CHSH,EE}}^N$ contains $2N$ terms and each term has an absolute value ≤ 1 . Since EE and LL events equally contribute to the coincidences, the right-hand side of the inequality is therefore given by the average of the EE and LL right-hand sides, namely, $\tilde{S}^N = \frac{1}{2}S_{\text{EE}}^N + \frac{1}{2}S_{\text{LL}}^N$. We then get the following correct inequalities for time-bin entanglement:

$$|\tilde{S}_{\text{CHSH,EE}}^N| \leq 2N - 1 \quad (\text{A12})$$

and

$$\frac{3}{4} - N \leq \tilde{S}_{\text{CH}}^N \leq \frac{1}{4}. \quad (\text{A13})$$

We note that since S_{CH} involves only joint probabilities and not correlation, the inequality is valid when we use one detector at each side to measure the probabilities $p(a_i b_j)$. Since we have demonstrated that the CHSH form can be derived from the CH form, it implies that the CHSH inequality also holds when one detector at each side is used.

APPENDIX B: LHV MODEL FOR TIME-BIN ENTANGLEMENT

In this section, we give the detail of the LHV model for time-bin entanglement. Consider the general time-bin scheme,

where each observer detects the photon at three possible different arrival times: $t_0 - \Delta T$ (here we call this event early, E), t_0 (medium, M), and $t_0 + \Delta T$ (late, L), where $\Delta L = c\Delta T$ is the length difference between the short and long paths. For each detection time, E , M , and L , a measurement indicated by a and b for Alice and Bob can have two outcomes, $+$ and $-$.

Now consider the probability

$$P(a, b | \phi_A, \phi_B), \quad (\text{B1})$$

where ϕ_A and ϕ_B are the measurement settings for Alice and Bob. For the detection $M - M$, quantum mechanics predicts the following probabilities:

$$\begin{aligned} P(M_+, M_+ | \phi_A, \phi_B) &= P(M_-, M_- | \phi_A, \phi_B) \\ &= \frac{1}{16} [1 + \cos(\phi_A + \phi_B)], \\ P(M_+, M_- | \phi_A, \phi_B) &= P(M_-, M_+ | \phi_A, \phi_B) \\ &= \frac{1}{16} [1 - \cos(\phi_A + \phi_B)]. \end{aligned}$$

In the other cases, we have

$$\begin{aligned} P(M_{\pm}, E_{\pm} | \phi_A, \phi_B) &= P(E_{\pm}, M_{\pm} | \phi_A, \phi_B) = \frac{1}{32}, \\ P(M_{\pm}, L_{\pm} | \phi_A, \phi_B) &= P(L_{\pm}, M_{\pm} | \phi_A, \phi_B) = \frac{1}{32}, \\ P(E_{\pm}, E_{\pm} | \phi_A, \phi_B) &= P(L_{\pm}, L_{\pm} | \phi_A, \phi_B) = \frac{1}{32}, \\ P(E_{\pm}, L_{\pm} | \phi_A, \phi_B) &= P(L_{\pm}, E_{\pm} | \phi_A, \phi_B) = 0. \end{aligned}$$

A local hidden-variable (LHV) model can mimic the above correlations in the following way. Consider a hidden variable $\lambda = (\theta, r)$, with θ uniformly distributed between 0 and 2π , and r uniformly distributed between 0 and 1. The outcome $A(\lambda; \phi_A)$ and $B(\lambda; \phi_B)$ can then be determined from Fig. 1 in the following way: Given a fixed value (θ, r) , the LHV model deterministically establishes the outcomes of Alice and Bob as functions of ϕ_A and ϕ_B given by $A(\lambda; \phi_A)$ and $B(\lambda; \phi_B)$.

This LHV model is local because the outcomes of Alice do not depend on ϕ_B , and vice versa. The probabilities predicted by the LHV model are

$$P(a, b | \phi_A, \phi_B)_{\text{LHV}} = \frac{1}{2\pi} \int_{S(a, b, \phi_A, \phi_B)} d\theta dr,$$

where $S(a, b, \phi_A, \phi_B)$ is the subset of the possible hidden-variable values such that $A(\theta, r; \phi_A) = a$ and $B(\theta, r; \phi_B) = b$. In the above equation, the possible values of a and b are $\{E_{\pm}, M_{\pm}, L_{\pm}\}$ for Alice and Bob. It is easy to check that the above probabilities coincide with the predictions of quantum mechanics.

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

[2] J.-Å. Larsson, Bell's inequality and detector inefficiency, *Phys. Rev. A* **57**, 3304 (1998).

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

- [4] J. Jogenfors and J.-Å. Larsson, Energy-time entanglement, elements of reality, and local realism, *J. Phys. A: Math. Theor.* **47**, 424032 (2014).
- [5] J. Jogenfors, A. M. Elhassan, J. Ahrens, M. Bourennane, and J.-Å. Larsson, Hacking the Bell test using classical light in energy-time entanglement-based quantum key distribution, *Sci. Adv.* **1**, e1500793 (2015).
- [6] L. De Caro and A. Garuccio, Reliability of Bell-inequality measurements using polarization correlations in parametric-down-conversion photon sources, *Phys. Rev. A* **50**, R2803(R) (1994).
- [7] J.-Å. Larsson, M. Giustina, J. Kofler, B. Wittmann, R. Ursin, and S. Ramelow, Bell-inequality violation with entangled photons, free of the coincidence-time loophole, *Phys. Rev. A* **90**, 032107 (2014).
- [8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).
- [9] G. Lima, G. Vallone, A. Chiuri, A. Cabello, and P. Mataloni, Experimental Bell-inequality violation without the postselection loophole, *Phys. Rev. A* **81**, 040101 (2010).
- [10] G. Vallone, I. Gianani, E. B. Inostroza, C. Saavedra, G. Lima, A. Cabello, and P. Mataloni, Testing Hardy's nonlocality proof with genuine energy-time entanglement, *Phys. Rev. A* **83**, 042105 (2011).
- [11] A. Cuevas, G. Carvacho, G. Saavedra, J. Cariñe, W. a. T. Nogueira, M. Figueroa, A. Cabello, P. Mataloni, G. Lima, and G. B. Xavier, Long-distance distribution of genuine energy-time entanglement, *Nat. Commun.* **4**, 2871 (2013).
- [12] G. Carvacho, J. Cariñe, G. Saavedra, Á. Cuevas, J. Fuenzalida, F. Toledo, M. Figueroa, A. Cabello, J.-Å. Larsson, P. Mataloni, G. Lima, and G. B. Xavier, Postselection-loophole-free Bell test over an installed optical fiber network, *Phys. Rev. Lett.* **115**, 030503 (2015).
- [13] J. D. Franson, Inconsistency of local realistic descriptions of two-photon interferometer experiments, *Phys. Rev. A* **61**, 012105 (1999).
- [14] S. Aerts, P. Kwiat, J.-Å. Larsson, and M. Żukowski, Two-photon Franson-type experiments and local realism, *Phys. Rev. Lett.* **83**, 2872 (1999).
- [15] P. M. Pearle, Hidden-variable example based upon data rejection, *Phys. Rev. D* **2**, 1418 (1970).
- [16] S. Braunstein and C. Caves, Wringing out better Bell inequalities, *Ann. Phys.* **202**, 22 (1990).
- [17] L. Hardy, Nonlocality for two particles without inequalities for almost all entangled states, *Phys. Rev. Lett.* **71**, 1665 (1993).
- [18] D. Boschi, S. Branca, F. De Martini, and L. Hardy, Ladder proof of nonlocality without inequalities: Theoretical and experimental results, *Phys. Rev. Lett.* **79**, 2755 (1997).
- [19] T. R. Tan, Y. Wan, S. Erickson, P. Bierhorst, D. Kienzler, S. Glancy, E. Knill, D. Leibfried, and D. J. Wineland, Chained Bell inequality experiment with high-efficiency measurements, [arXiv:1612.01618](https://arxiv.org/abs/1612.01618).
- [20] B. G. Christensen, Y.-C. Liang, N. Brunner, N. Gisin, and P. G. Kwiat, Exploring the limits of quantum nonlocality with entangled photons, *Phys. Rev. X* **5**, 041052 (2015).
- [21] T. E. Stuart, J. A. Slater, R. Colbeck, R. Renner, and W. Tittel, Experimental bound on the maximum predictive power of physical theories, *Phys. Rev. Lett.* **109**, 020402 (2012).
- [22] J.-Å. Larsson, Loopholes in Bell inequality tests of local realism, *J. Phys. A* **47**, 424003 (2014).