Requirements for a loophole-free photonic Bell test using imperfect setting generators

Johannes Kofler,[1] Marissa Giustina,[2,3] Jan-Åke Larsson,[4] and Morgan W. Mitchell[5,6]

[1]*Max Planck Institute of Quantum Optics (MPQ), Hans-Kopfermann-Straße 1, 85748 Garching/Munich, Germany*
[2]*Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences,*
Boltzmanngasse 3, 1090 Vienna, Austria
[3]*Quantum Optics, Quantum Nanophysics, and Quantum Information, Faculty of Physics, University of Vienna,*
Boltzmanngasse 5, 1090 Vienna, Austria
[4]*Institutionen for Systemteknik, Linköpings Universitet, SE-58183 Linköping, Sweden*
[5]*ICFO – Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*
[6]*ICREA – Institució Catalana de Recerca i Estudis Avançats, 08015 Barcelona, Spain*

Experimental violations of Bell inequalities are in general vulnerable to so-called loopholes. In this work, we analyze the characteristics of a loophole-free Bell test with photons, closing simultaneously the locality, freedom-of-choice, fair-sampling (i.e., detection), coincidence-time, and memory loopholes. We pay special attention to the effect of excess predictability in the setting choices due to nonideal random-number generators. We discuss necessary adaptations of the Clauser-Horne and Eberhard inequality when using such imperfect devices and— using Hoeffding's inequality and Doob's optional stopping theorem—the statistical analysis in such Bell tests.

I. INTRODUCTION

Bell's theorem [1] about the incompatibility of a local realist world view with quantum mechanics is one of the most profound discoveries in the foundations of physics. Since the first experimental quantum violation of Bell's inequality [2], countless experimental tests have been performed with various different physical systems, closing all major "loopholes." While it is unlikely that nature exploits these loopholes, let alone different ones for different experiments, there are at least two reasons why a loophole-free test is of great relevance: First, a definitive ruling on local realism is of central importance to our understanding of the physical world. Second, there are quantum information protocols whose security is based on Bell's inequality, and eavesdroppers could actively exploit the loopholes.

This work is structured as follows: We first briefly review Bell's derivation and the five major loopholes—the locality, freedom-of-choice, fair-sampling (detection), coincidence-time, and memory loopholes (Sec. II). Then, we give an analysis of how a photonic Bell test can simultaneously close all of them. We discuss an inequality based on those developed by Clauser-Horne (CH) and Eberhard (Sec. III). The low detection efficiency requirement of this CH-Eberhard inequality is essential given the current status of equipment and technology. We outline the necessary space-time arrangement (Sec. IV) and show how to take into account—by adapting the CH-Eberhard inequality—imperfect random-number generators that sometimes choose settings outside the allowed space-time interval or are for some other reason partially predictable beyond the *a priori* probability (Sec. V). Finally, while allowing both bias and excess predictability of the settings, we demonstrate how to apply Hoeffding's inequality and Doob's optional stopping theorem to achieve high statistical significance of a Bell inequality violation within feasible experimental run time (Sec. VI). Readers who are familiar with loopholes in Bell tests and the CH-Eberhard inequality can skip to Sec. IV.

II. BELL'S THEOREM AND LOOPHOLES

Let us consider the simplest scenario of only two parties called Alice and Bob, who perform measurements on distant physical systems. Alice's and Bob's measurement settings are labeled with a and b, and their outcomes are denoted by A and B, respectively. There are essentially two versions of Bell's theorem:

Deterministic local hidden-variable models. Determinism states that hidden variables determine the outcomes, which are then functions of the form $A = A(a,b,\lambda)$, $B = B(a,b,\lambda)$. *Locality* demands that the local outcomes do not depend on the distant setting,

$$A = A(a,\lambda), \quad B = B(b,\lambda). \tag{1}$$

The original 1964 version of Bell's theorem [1] is based on the assumptions of *perfect anticorrelation* and *locality* which imply determinism. The assumption of perfect anticorrelation was later avoided by Clauser, Horne, Shimony, and Holt (CHSH) in the derivation of their famous inequality [3].

Stochastic local hidden-variable models. Following Refs. [4,5], in the 1976 version of Bell's theorem [6] the assumptions are relaxed to include stochastic models. There, hidden variables only define probabilities for the outcomes, $P(A|a,b,B,\lambda)$, $P(B|a,b,A,\lambda)$, and a joint assumption called *local causality* (or *Bell locality*) demands that the joint probability of Alice's and Bob's outcomes factorizes as follows:

$$P(A,B|a,b,\lambda) = P(A|a,\lambda)\,P(B|b,\lambda). \tag{2}$$

This is equivalent to assuming *outcome independence* $P(A|a,b,B,\lambda) = P(A|a,b,\lambda)$ as well as *setting independence* (or *parameter independence*) $P(A|a,b,\lambda) = P(A|a,\lambda)$, with similar expressions for Bob's outcome probability [7].

The world view in which all physical phenomena can be described by local hidden variables is often referred to as *local realism*. While local causality is implied by the conjunction of determinism and locality, the opposite

implication is not true. Nonetheless, the two classes of local hidden-variable models are mathematically equivalent in the sense that deterministic models are special cases of stochastic ones (where all probabilities are 0 or 1), and that every stochastic model can be viewed as a mixture of deterministic ones [8,9]. Physically, however, the difference is significant. It is conceivable to adhere to a stochastic world view in which the hidden variables only define probabilities, rejecting a hidden determinism, although this determinism might mathematically exist and explain the probabilities.

In addition to local causality (or, stronger, determinism and locality) there is another essential assumption in the derivation of every Bell inequality called *freedom of choice* (or *measurement independence*). It demands that the distribution ρ of the hidden variables λ is statistically independent of the setting values,

$$\rho(\lambda|a,b) = \rho(\lambda). \tag{3}$$

By Bayes' theorem, this assumption can also be written as $\rho(a,b|\lambda) = \rho(a,b)$. The freedom-of-choice assumption was first pointed out in a footnote in Ref. [5] and later discussed in an exchange [6,10,11], which is reprinted in [12].

Bell's theorem states that the joint assumption of local hidden variables and freedom of choice enables the derivation of inequalities that put local realist bounds on combinations of probabilities for Alice's and Bob's measurement results. In Bell experiments, measurements on entangled quantum states can violate Bell's inequality and thus refute the existence of local hidden variables.

The translation from any mathematical expression to a physical experiment employs further physical assumptions, which may render an experimental Bell violation vulnerable to a local realist explanation. In the following, we discuss the five main "loopholes" in Bell tests. For further details on the assumptions in Bell's theorem, the use of entanglement in Bell experiments, and the loopholes that can arise, we refer the reader to the recent reviews [13–16].

A. The locality loophole

The locality loophole refers to the possibility of violating outcome or setting independence via subluminal or luminal influences between the two outcomes or from one setting to the distant outcome. It is generally acknowledged that the best possible way to close the loophole is to invoke special relativity. Spacelike separation of the two outcome events enforces outcome independence, and spacelike separation of each party's independent setting choice event from the opposite party's outcome event enforces setting independence. In this way, the locality loophole is considered to have been closed for photons by the experiments [17–21], and with nitrogen-vacancy (NV) centers by the experiment [22].

This, however, rests on the assumption that there were no prior influences for the setting choice events that could have been communicated to the distant party. Deterministic setting mechanisms as, e.g., the periodic switching used in [23], are predictable into the future and thus, in principle, still allow a local realist explanation [24] unless restrictions are imposed on the information communicated.

B. The freedom-of-choice loophole

The freedom-of-choice loophole refers to the possibility that the freedom-of-choice condition $\rho(\lambda|a,b) = \rho(\lambda)$ fails due to an influence of the hidden variables on the setting choices, an influence of the setting choices on the hidden variables, or, more generally, a common influence on both the setting choices and the hidden variables.

As with the locality loophole, spacelike separation allows an experiment to exclude certain influences within any local theory. For example, spacelike separation of the pair generation from the setting choices eliminates the pair generation as a possible influence. This has been achieved in the experiments [18–21]. However, again, it is not possible to exclude all possible influences in this way because these could, in principle, extend arbitrarily far into the past.

Note that freedom of choice does not require the factorization $\rho(a,b) = \rho(a)\,\rho(b)$. However, if the setting choices are not spacelike separated with respect to each other, then one of the outcome events will always be in the future light cone of the distant setting event, leaving the locality loophole open.

A second, complementary way to address the freedom-of-choice loophole is to derive the setting choices from events that are plausibly beyond the control of hidden variables, for example spontaneous emission, chaotic evolution, human decision making, or cosmic sources. A Bell inequality violation using one or more of these sources can exclude local realist theories in which the setting events are unpredictable, pushing the unexcluded theories in the direction of a full determinism (cf. Sec. II F).

C. The fair-sampling (detection) loophole

The fair-sampling assumption states that the ensemble detected by Alice and Bob is representative of the total emitted ensemble. This is the case if the detection efficiency depends only on the hidden variable and not on the local setting. Unfair sampling opens the fair-sampling (or detection) loophole [25].

Inequalities that make use of the fair-sampling assumption in their derivation, such as the CHSH inequality [3], can be rendered immune to the fair-sampling loophole only by explicitly demonstrating sufficiently large detection efficiency or by incorporating the undetected events into the inequality [4]. This latter, more elegant approach—not assuming fair sampling in the first place—is used in the derivation of the Clauser-Horne (CH) [5] and the Eberhard inequalities [26]. The fair-sampling loophole has been closed for atoms [27–29], superconducting qubits [30], and NV centers [22]. Using superconducting detectors, it has also been closed for photons [20,21,31,32].

D. The coincidence-time loophole

The fair-coincidence assumption states that the statistics of the identified pairs are sufficiently representative of the statistics of all detected pairs, had they been correctly identified. In experiments where (near-)coincident arrival times are used to identify which detections belong to a pair, the assumption is fulfilled if the local detection time depends only on the hidden variable and not on the local setting. Unfair coincidences open the coincidence-time loophole [33].

This loophole arises in any situation where a (setting-dependent) shift in detection time could alter the number of identified pairs; it is especially applicable to continuous-wave photonic experiments. The loophole can be closed using locally predefined time slots or (for the CH-Eberhard inequality) by employing a window-sum method for coincidence-based identification of pairs [34,35]. Regarding photonic experiments, the loophole was closed in [20,21,31,32,34,36].

E. The memory loophole

One can imagine a situation in which the experimental apparatuses use memory of the previous measurements to skew the apparent significance of a violation. In this case, say, the probability for Alice to find outcome $A^{(m)}$ in the mth measurement can depend not only on her current setting $a^{(m)}$ and hidden variable λ, but also on the $m - 1$ previous settings and outcomes on her side $(a^{(1)}, \ldots, a^{(m-1)}, A^{(1)}, \ldots, A^{(m-1)}$, one-sided memory) and maybe also on Bob's side $(b^{(1)}, \ldots, b^{(m-1)}, B^{(1)}, \ldots, B^{(m-1)}$, two-sided memory), and vice versa for Bob's outcome probability for $B^{(m)}$ [37–40]. Then, the no-memory assumption that successive measurement trials are independent and identically distributed (i.i.d.) is not valid.

The memory loophole does not change a Bell inequality's local realist bound but forbids quantifying the statistical significance of a Bell test by the amount of conventional standard deviations between the observed Bell value and the local realist bound. The loophole could, in principle, be closed by using separate apparatuses and spacelike separation of each of Alice's measurements from all of Bob's measurements. However, this is technologically unfeasible. Thus, a more useful approach is to apply statistical methods, such as hypothesis testing, that can—without the assumption of i.i.d. measurement trials—bound the probability that the data can be explained by a random variation of a local hidden-variable model.

Table I summarizes the assumptions used in derivations of Bell inequalities as well as the corresponding loopholes and the procedures for closing them.

F. Additional assumptions and unclosable loopholes

By attributing significance to spacelike separation, one implicitly assumes that one can localize key events to particular space-time regions. For example, spacelike separation of the setting choices from the detection events closes the locality loophole, but requires that the setting choices are independent of prior conditions. This break between the past and the present means that closure of the locality loophole can only be attempted within nondeterministic (i.e., stochastic) local realism. Within determinism, the settings would also be deterministic and thus predictable arbitrarily far in the past, rendering spacelike separation impossible. Similarly, using spacelike separation to close the freedom-of-choice loophole can only eliminate theories in which the hidden variable is created in a defined space-time region (e.g., at the down-conversion event in a photonic experiment).

Likewise, arguments based on spacelike separation of the detection events from the distant setting choices requires that one knows when the measurement is complete. In all practical scenarios for Bell tests, there is an identifiable time window in which a microscopic observable, such as the polarization of a single photon, becomes amplified into a macroscopic observable, such as a large number of electrons moving in a wire. Usually this conversion to a macroscopic event is taken as the time of the measurement, but there is no logical contradiction in assuming that the measurement happens later ("collapse locality loophole" [41]).

The general feature of all these arguments is that a loophole-free Bell test is possible only when a set of reasonable assumptions about the physical working of the experimental setup is made. Experiments can shift hypothetical effects to more and more absurd scales but can never fully rule them out. In particular, it is, in principle, impossible to rule out "superdeterminism" [42], a world constructed such that Eq. (3) cannot be fulfilled. Therefore, strictly speaking, the locality and freedom-of-choice loopholes can only be addressed (i.e., closed within some assumptions) and cannot be closed in general.

Finally, every Bell test needs to rest on metaprinciples, most notably that the classical rules of logic hold. In 2015,

TABLE I. Summary of the five main loopholes in Bell experiments. The assumptions of outcome and setting independence as well as freedom of choice are minimal in the sense that they enter the derivation of any Bell inequality. The corresponding loopholes are closed by the spatiotemporal construction of the experiment and the means of choosing settings. The other three loopholes are related to auxiliary assumptions and are closed by a suitable choice of Bell inequality (or additional tests) as well as appropriate data analysis. See main text for further details and references.

Minimal assumptions	Auxiliary assumptions	Loopholes	Closed by . . .
Outcome and setting independence		Locality loophole	spacelike separation between the outcome events and between each outcome and the distant setting choice event
Freedom of choice		Freedom-of-choice loophole	(for photonic experiments) spacelike separation between each pair emission event and the setting choice events
	Fair sampling	Fair-sampling (detection) loophole	violation of an inequality free of the fair-sampling assumption (e.g. CH-Eberhard) or explicit demonstration of sufficiently large detection efficiency (e.g., for CHSH)
	Fair coincidences	Coincidence-time loophole	using fixed time slots or (for CH-Eberhard) a window-sum method for identifying coincidences
	No memory	Memory loophole	sufficiently many measurement trials, no i.i.d. assumption

three different groups were able to perform "loophole-free" Bell tests [20–22].

III. THE CH-E INEQUALITY

Eberhard's derivation [26] considered a source that produces photon pairs where the polarization of one photon of every pair is measured by Alice with setting a_1 or a_2, while the other photon's polarization is measured by Bob with setting b_1 or b_2. We label the outcome or "fate" (given by the hidden variable) of every photon by '$+$', '$-$', or '0', which denotes being detected in the first ("ordinary") output beam of the polarizer, being detected in the second ("extraordinary") beam, or remaining undetected, respectively. We denote joint fates for outcomes A (for Alice) and B (for Bob) by AB with $A, B \in \{+, -, 0\}$.

Eberhard considered N' pairs emitted for each of the four setting combinations $a_i b_j$ with $i, j \in \{1, 2\}$. For setting combination $a_i b_j$, we denote the number of joint outcomes A and B by $n_{AB}(a_i b_j)$. Note that pairs with joint fate 00 also count as pairs. Hence, $\sum_{A, B \in \{+, -, 0\}} n_{AB}(a_i b_j) = N'$ for each setting combination $a_i b_j$.

In hidden-variable theories, the results for mutually exclusive measurements exist simultaneously. Locality demands that the local fate of a photon must not depend on the distant measurement setting. Freedom of choice assumes that the experimenters' settings are independent of the designated fate. Under these assumptions, Eberhard's inequality bounds the expectation value of a certain combination of outcome numbers [26],

$$\langle +n_{++}(a_1 b_1) - n_{+-}(a_1 b_2) - n_{+0}(a_1 b_2) \\ - n_{-+}(a_2 b_1) - n_{0+}(a_2 b_1) - n_{++}(a_2 b_2) \rangle \leqslant 0. \quad (4)$$

The logical bound of the inequality is N', which can be attained by a model (violating local realism and/or freedom of choice) where all N' pairs for settings $a_1 b_1$ lead to outcome $++$, and no pairs in the other setting combinations ever contribute to the five positive terms. The quantum bound is $(\sqrt{2} - 1) N'/2 \approx 0.207 N'$, which can be attained for perfect detection efficiency (i.e., absence of outcomes 0) on both sides and maximally entangled states. However, for imperfect detection efficiency (i.e., occurrence of outcomes 0), nonmaximally entangled states achieve better violation.

Until now, the derivation has assumed that there was the same number of pairs (N') in each of the four setting combinations. Experiments are not likely to obey this strict constraint, but rather to produce a different number of pairs for every combination. In general, this invalidates the Eberhard inequality (4), as can be seen by considering the case where the setting $a_1 b_1$ is used more often than the others, which will increase the $n_{++}(a_1 b_1)$ contribution (see Refs. [32,43]). A solution is to introduce conditional probabilities $p_{AB}(a_i b_j)$ for outcomes AB given settings $a_i b_j$. As the original Eberhard inequality holds when an equal number of trials is measured in each setting combination, and since under freedom of choice every setting is chosen independently from the source, the same form of inequality holds for the conditional probabilities,

$$+ p_{++}(a_1 b_1) - p_{+-}(a_1 b_2) - p_{+0}(a_1 b_2) \\ - p_{-+}(a_2 b_1) - p_{0+}(a_2 b_1) - p_{++}(a_2 b_2) \leqslant 0. \quad (5)$$

The logical bound of this inequality is 1, and the quantum bound is $(\sqrt{2} - 1)/2 \approx 0.207$. One may drop the distinction between outcomes '$-$' and '0' in the Eberhard inequality (4). Blocking the extraordinary beam such that all '$-$' events become '0' events, the normalized Eberhard inequality (5) is reduced to a one-detector-per-side form with coincidences and exclusive singles (i.e., detections on exactly one side),

$$J \equiv p_{++}(a_1 b_1) - p_{+0}(a_1 b_2) - p_{0+}(a_2 b_1) - p_{++}(a_2 b_2) \leqslant 0. \quad (6)$$

We can define the probabilities of *singles* (photon detections in one particular output beam regardless of the outcome on the other side),

$$p_+^A(a_1)_{b_2} \equiv p_{++}(a_1 b_2) + p_{+0}(a_1 b_2), \quad (7)$$

$$p_+^B(b_1)_{a_2} \equiv p_{++}(a_2 b_1) + p_{0+}(a_2 b_1). \quad (8)$$

Here, the singles probabilities were defined for a particular distant setting, namely b_2 and a_2, respectively. However, due to locality, no-signaling must be fulfilled:

$$p_+^A(a_i)_{b_1} = p_+^A(a_i)_{b_2}, \quad (9)$$

$$p_+^B(b_j)_{a_1} = p_+^B(b_j)_{a_2}, \quad (10)$$

for $i, j \in \{1, 2\}$. Ignoring the conditioning on the distant setting (due to locality) and dropping the index $+$ everywhere, inequality (6) becomes the CH inequality [5],

$$C_H \equiv + p(a_1 b_1) + p(a_1 b_2) + p(a_2 b_1) \\ - p(a_2 b_2) - p^A(a_1) - p^B(b_1) \leqslant 0. \quad (11)$$

Eberhard's main contribution was to realize that nonmaximally entangled states allow a violation of the CH or Eberhard inequality for detection efficiencies as low as 2/3, which is still the lowest known value for qubit systems. In contrast, efficiency of 82.8 % is required for maximally entangled states [44,45]. The use of the CH or Eberhard inequalities and nonmaximally entangled states hence greatly eases the detection efficiency requirements, one of the most challenging aspects of photonic experiments. (We mention that there are also forms of the CH or Eberhard inequality where all terms are divided by the sum of singles probabilities or counts [32,46].)

The inequality (6), which we call the CH-E inequality, will be used in the later sections, as it is the simplest known form, with only four terms that all stem from mutually exclusive setting combinations.

IV. SPACE-TIME ARRANGEMENT AND SETTING PREDICTABILITY

For a photonic Bell test, consider the space-time diagram in Fig. 1, where intervals of space-time events are denoted with (nonitalic) bold letters. A photon pair is emitted by a source at **E**. The photons travel a distance d in fibers (solid blue lines) with refractive index n to Alice and Bob, where they pass the setting devices, indicated by black rectangles. Geometric deviations from a perfectly one-dimensional setup (black dashed lines) and any other additional delays are represented by τ_G. Alice's and Bob's measurement outcomes are restricted to intervals **A** and **B** of duration τ_M. Outcome independence requires spacelike separation of **A** and **B**. Setting independence
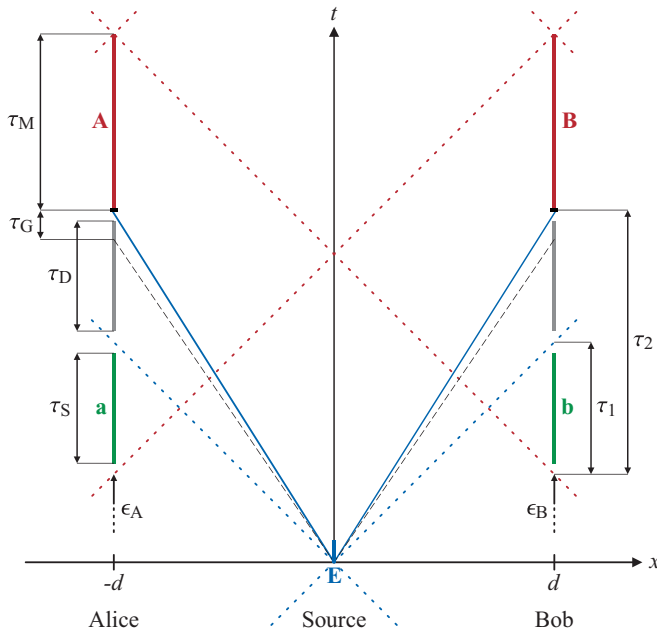
FIG. 1. Space-time diagram of a photonic fiber-based Bell test. **E** represents the emission of a photon pair, **A** and **B** are Alice's and Bob's detection intervals, and **a** and **b** are their setting choice intervals. Relevant light cones are indicated by dotted lines. Knowledge about the distant setting that can be available at Bob's (Alice's) measurement device is quantified by ϵ_A (ϵ_B). See main text for further details.

requires that Alice's setting generation is confined to interval **a**, spacelike separated from Bob's outcome interval **B**, and likewise **b** must be spacelike separated from **A**. Spacelike separation of the setting generations within **a** and **b** from the emission interval **E** closes the freedom-of-choice loophole. (The relevant spacelike separations can only be achieved by using at least three distinct locations. One measurement device may be located at the source [18,19], but then the corresponding setting generator needs to be placed at a distance.) The time duration τ_S for **a** and **b** must be smaller than τ_1, and the time for setting generation as well as deployment of the setting (duration τ_D) must be smaller than τ_2:

$$\tau_S < \tau_1 = \frac{(3-n)\,d}{c_0} - (\tau_G + \tau_M), \qquad (12)$$

$$\tau_S + \tau_D < \tau_2 = \frac{2\,d}{c_0} - \tau_M. \qquad (13)$$

Here, c_0 denotes the speed of light in vacuum.

Closing the locality and freedom-of-choice loopholes requires generation of fast random numbers for the settings (a,b), which must not be able to influence the respective distant outcome (locality) or have a mutual interdependence with the hidden variable λ (freedom of choice). It should be noted, however, that the requirements for (a,b)-λ independence in a Bell test differ in important ways from "randomness" as per the usual definitions. For example, it is common to consider as random a source of independent, identically distributed, unbiased bits x_i, described by the probabilities $p(x_i|x_{j\neq i}) = \frac{1}{2}$. Using such sources to choose (a,b) does not by itself guarantee independence from λ because λ could influence x in such a

way that x is predictable knowing λ, but fully unpredictable absent this knowledge. In contrast, a source that is biased but uninfluenced by λ, e.g., $p(x_i|\lambda) = \frac{3}{4}$, is suitable for generating the required independence, despite being far from random by the usual definitions.

As concerns physical variables and setting choices, we use the term "random" to mean independence from λ. Physically, this independence can be compromised by an influence of λ on x, by an influence of x on λ, or by a common influence. The first two of these can be excluded by spacelike separation of the setting generation from the creation of the hidden variables, while the last one is excluded if λ and/or x is uninfluenced, i.e., stochastic.

An entire Bell experiment, including the setting generation, must be viewed within local realism, and quantum mechanics must not be invoked. Candidate stochastic processes include chaotic dynamics, human decision making [42], and cosmic light sources [47]. Photonic devices use the reflection and transmission at a beam splitter [48], or the emission or detection time [49], population [50], or phase [51,52] of a coherent light source. It bears repeating that a local realist model must contain some stochastic element if it is to be testable.

Any real implementation of a random-number generator will to some extent be influenced by effects prior to the generation, giving nonzero predictive power beyond the *a priori* probability of guessing the eventual setting. We call this the *excess predictability*. This opens the locality loophole or freedom-of-choice loophole to some extent. In general, each setting choice could have a different excess predictability in every trial, such that in trial n the excess predictability takes on values $\epsilon_A^{(i)}$ and $\epsilon_B^{(i)}$ with $\epsilon_A^{(i)}, \epsilon_B^{(i)} \in [0,1]$. Below we model two special cases: We either assume that in a small *fraction* ϵ_A (ϵ_B) of experimental runs, Alice's (Bob's) setting choice is *perfectly* communicable to the distant party Bob (Alice), or we consider that in *every* trial, Bob (Alice) can predict the distant setting with a small certainty ϵ_A (ϵ_B) better than the *a priori* probability. Presumably, the physical situation could be any mixture of these two models.

V. ADAPTATION OF THE CH-E INEQUALITY

To use the CH-E inequality, which employs conditional probabilities, we need the concept of a trial. Without an exact definition of what a trial is, it is unclear how to use normalized counts or the concept of probabilities when employing the CH-E inequality. Normalization with respect to the pair production rate or measurement time for a given fixed setting [43,46] will not be possible for a loophole-free Bell test because the analysis technique for closing the memory loophole relies on the concept of trials. Noting that the particular construction and assumptions involved in a given test might refine the operational definition of a trial in that test, we suggest that the reader consider a trial most basically as a (locally defined) measurement interval, for which each measurement party must record exactly one outcome (possibly including "undetected").

Specifically, we have in mind a pulsed experiment, where every pulse—which might or might not create a down-conversion pair—belongs to exactly one trial. We will not consider anything that happens between the trials. Fixed measurement time windows synchronized with the laser pulses

are also suitable for closing the coincidence-time loophole for the CH inequality [34].

Given that information about a setting will sometimes exist in the backward light cone of the distant outcome event, it is necessary to adapt the CH-E inequality. We now consider two different mathematical models for the communication or excess predictability of the setting values:

Scenario (i): *Communication in some trials.* Here, in a fraction ϵ_A (ϵ_B) of the trials, Alice's (Bob's) setting is perfectly known to Bob (Alice) via communication, while in the rest of the trials the locality condition is perfectly fulfilled. For simplicity, we assume that this fraction is the same for all setting combinations. To be conservative, we shall not assume that the "glitches" of too early settings happen statistically independently on the two sides, but that they may avoid happening in the same trials. We introduce the abbreviation

$$\epsilon_{AB} \equiv \min(\epsilon_A + \epsilon_B, 1) \qquad (14)$$

for the (maximal possible) fraction where one setting is communicable to the distant outcome. Let us consider the subset S_A of trials in which Alice's setting a is communicated to Bob's measurement device, while her measurement device has no information about Bob's setting b. It is conceivable that Alice's devices know when her setting is communicated. Then the strategy is as follows: Alice's measurement device "overrules" whatever fate has been designated and outputs +. Bob also outputs +, unless $a = a_2$ and $b = b_2$, whereupon he outputs 0. For the different setting combinations, their measurement results therefore contribute to $p_{++}(a_1b_1)$, $p_{++}(a_1b_2)$, $p_{++}(a_2b_1)$, and $p_{+0}(a_2b_2)$, and nothing else. The last three terms do not appear in the CH-E inequality (6), and the first is beneficial for its violation. The J value in the subset S_A can therefore reach the logical bound $J = +1$. Importantly, also, those events that would have had fate 00 contributed to the violation.

Straightforwardly, the above arguments can be repeated for the subset S_B of trials where Bob's setting can be communicated but not Alice's and for the subset S_{AB} where both can be communicated. This implies that local hidden variables augmented with setting communication can attain the CH-E value +1 in the total subset $S_{\epsilon_{AB}} = S_A \cup S_B \cup S_{AB}$ whose size is bounded by the fraction ϵ_{AB} of all trials. This means that for the entirety of all trials, such models reach $J = \epsilon_{AB}$. The CH-E inequality $J \leqslant 0$ must therefore be rewritten with an *adapted bound,*

$$J \leqslant \epsilon_{AB}. \qquad (15)$$

In other words, when physical (sub)luminal communication of a setting to a distant outcome is possible in a fraction ϵ_{AB} of trials, the collected results must violate inequality (15) with its adapted bound to rule out a local realist explanation.

An important remark: The above strategy violates the no-signaling condition (10). From subset S_A, one has contributions to the singles probability $p_+^B(b_2)_{a_1}$ but not to $p_+^B(b_2)_{a_2}$. This violation is a general feature of pure strategies with communication. Mixed strategies can hide the communication and obey no-signaling. When the entire setting information is communicated, the predictions of every no-one-way-signaling distribution can be simulated by local hidden variables [53]. The optimal no-signaling strategy is the simulation of a Popescu and Rohrlich (PR) box [54], which works as follows:

For every trial, Alice and Bob share a random variable $r \in \{+, 0\}$ with distribution $p(r = +) = p(r = 0) = \frac{1}{2}$. When Alice transmits her setting a to Bob, she outputs $A = r$. Bob also outputs $B = r$ unless $a = a_2$ and $b = b_2$, in which case he produces the opposite result (+ if $r = 0$, 0 if $r = +$). This strategy obeys no-signaling and, within the subset S_A, reaches $J = \frac{1}{2}$. Note that for the CHSH inequality the logical and the no-signaling bound are identical (equal to 4). This is not the case for the CH-E inequality, where the logical bound is 1 and the no-signaling bound is $\frac{1}{2}$.

While the bound ϵ_{AB} in (15) cannot be reached by local hidden-variable models that are augmented by setting communication and obey the no-signaling conditions, the bound is conservative only by a factor of 2 (since, according to the above, the bound for communication strategies obeying no-signaling is $\frac{\epsilon_{AB}}{2}$). Moreover, it has the advantage that one need not additionally check the no-signaling conditions in an experiment. Having quantified ϵ_{AB}, one can solely rely on the inequality (15) itself. Also note that violation of the no-signaling conditions within the subensemble $S_{\epsilon_{AB}}$ could be due to actual (sub)luminal signals and would not be in contradiction with causality.

Scenario (ii): *Excess predictability in all trials.* In this scenario, we assume that in every run, Alice's and Bob's setting choices a and b are partially dependent on external influences that are available also at the distant measurement event. Formally, this corresponds to a violation of the locality and freedom-of-choice assumptions. We can incorporate all of these influences together with the properties λ of the photon pair into a joint set μ of hidden variables. However, similar to Ref. [55], we assume that in every run, the effect of μ cannot alter the probability for a specific setting choice by more than a certain number, quantified by parameters ϵ_A and ϵ_B in the following way:

$$(1 - \epsilon_A)\, p(a) \leqslant p(a|\mu) \leqslant (1 + \epsilon_A)\, p(a), \qquad (16)$$

$$(1 - \epsilon_B)\, p(b) \leqslant p(b|\mu) \leqslant (1 + \epsilon_B)\, p(b). \qquad (17)$$

Using $p(a,b|\mu) = p(a|\mu)\, p(b|\mu)$, which is guaranteed as μ carries all hidden properties, and abbreviating

$$\epsilon_\pm \equiv \epsilon_A + \epsilon_B \pm \epsilon_A \epsilon_B, \qquad (18)$$

we obtain

$$(1 - \epsilon_-)\, p(a)\, p(b) \leqslant p(a,b|\mu) \leqslant (1 + \epsilon_+)\, p(a)\, p(b). \qquad (19)$$

Zero excess predictability implies $p(a,b) = p(a)\, p(b)$, while the converse is not true. Note that the individual setting probabilities $p(a)$ and $p(b)$ can have nonzero biases $\kappa_A, \kappa_B \in \left(-\frac{1}{2}, \frac{1}{2}\right)$,

$$p(a_1) = \tfrac{1}{2} - \kappa_A, \quad p(a_2) = \tfrac{1}{2} + \kappa_A, \qquad (20)$$

$$p(b_1) = \tfrac{1}{2} - \kappa_B, \quad p(b_2) = \tfrac{1}{2} + \kappa_B, \qquad (21)$$

which are neither at variance with the locality or freedom-of-choice assumptions nor problematic in the derivation of the CH-E inequality. The parameters ϵ_A and ϵ_B in (16) and (17) hence quantify *predictability beyond bias.*

Recorded data allows us to estimate total probabilities averaged over μ, that is, $p(A, B, a, b) = E[p(A, B, a, b|\mu)]$,

with E denoting the expectation value, but does not immediately allow us to estimate the conditional probabilities $p_{AB}(ab) \equiv p(A,B|a,b)$. The latter are well-defined for each individual value of μ, which is inaccessible to us. When conditioned on μ, the conditional probabilities obey the CH-E inequality (6):

$$+ p(++|a_1 b_1, \mu) - p(+0|a_1 b_2, \mu)$$
$$- p(0+|a_2 b_1, \mu) - p(++|a_2 b_2, \mu) \leqslant 0. \quad (22)$$

Using (19), we obtain

$$\frac{p(A,B,a,b|\mu)}{p(a)\,p(b)\,(1+\epsilon_+)} \leqslant \frac{p(A,B,a,b|\mu)}{p(a,b|\mu)} \leqslant \frac{p(A,B,a,b|\mu)}{p(a)\,p(b)\,(1-\epsilon_-)}. \quad (23)$$

The inequalities (23) must also hold for expectation values:

$$\frac{p(A,B,a,b)}{p(a)p(b)(1+\epsilon_+)} \leqslant E[p(A,B|a,b,\mu)] \leqslant \frac{p(A,B,a,b)}{p(a)p(b)(1-\epsilon_-)}, \quad (24)$$

where $p(A,B|a,b,\mu) = \frac{p(A,B,a,b|\mu)}{p(a,b|\mu)}$. This allows us to arrive at the following *adapted form* of the CH-E inequality:

$$J_\epsilon \equiv + \frac{p(++,a_1 b_1)}{p(a_1)\,p(b_1)\,(1+\epsilon_+)} - \frac{p(+0,a_1 b_2)}{p(a_1)\,p(b_2)\,(1-\epsilon_-)}$$
$$- \frac{p(0+,a_2 b_1)}{p(a_2)\,p(b_1)\,(1-\epsilon_-)} - \frac{p(++,a_2 b_2)}{p(a_2)\,p(b_2)\,(1-\epsilon_-)} \leqslant 0. \quad (25)$$

The inequality holds because, due to (24), the left-hand side is bounded by $E[p(++|a_1 b_1, \mu) - p(+0|a_1 b_2, \mu) - p(0+|a_2 b_1, \mu) - p(++|a_2 b_2, \mu)]$, which, due to (22), is bounded by 0.

It is important to note that the adaptation in scenario (i) is "absolute," while the one in scenario (ii) is "fractional." The adapted bound violation from a given measured J can withstand a larger value of $\epsilon_\pm$ in scenario (ii) than ϵ_{AB} in scenario (i).

We conclude that the concrete adapted form of the CH-E inequality depends on the physical scenario of how setting choices are communicable to or predictable at the remote side.

VI. STATISTICAL SIGNIFICANCE AND RUN TIME

In published experimental tests of Bell's inequality, it is common to report a violation as the number of standard deviations separating the measured value from the local realist bound, assuming Poissonian statistics. This quantifies the chance that a value consistent with local realism is still in agreement with the experimental data. In fact, we are interested in a different question: What is the chance that the violation observed in the experiment could have been produced under local realism? Moreover, to close the memory loophole, we may no longer assume that the trials are i.i.d. Employing the concept of hypothesis testing, for instance using the Hoeffding inequality [56], one can put a bound on the probability that local realism produced the data in a given experiment, even when allowing memory.

Based on the works [40,57], we present a statistical analysis with the following three key features, all of which are essential for a photonic Bell state with current technology:

(1) We allow for a bias in the setting choices.

(2) We take into account a communication or excess predictability (beyond bias) of the setting choices, using adapted versions of the CH-E inequality.

(3) We apply Doob's theorem to get rid of noncontributing trials and reduce the experimental run time to an acceptable level.

While all points are well understood individually, point 3 becomes nontrivial when combined with 1 and 2.

A *supermartingale* is a stochastic process for which, at any time in the sequence, the expectation value of the next value in the sequence does not exceed the expectation value of the current value in the sequence, given knowledge of all of the measurements in the history of the process. (One can think of it as a random walk with memory and strictly nonpositive drift.)

We consider an experiment with N trials. In each trial $n = 1, \ldots, N$, a measurement involves choosing a pair of settings and recording an outcome for each party, leading to an experimental value $\Delta^{(n)}$ for that trial, according to the inequality. Consider the random process Z_Δ: $Z^{(0)} = 0, Z^{(1)}, \ldots, Z^{(N)}$ with $Z^{(l)} = \sum_{n=1}^{l} \Delta^{(n)}$, whose increments $\Delta^{(n)}$ fall within range r_Δ. Then, Hoeffding's inequality

$$p(Z^{(N)} - E[Z^{(N)}] \geqslant c\sqrt{N}) \leqslant e^{-\frac{2}{r_\Delta^2}c^2} \quad (26)$$

bounds the probability that after N trials, $Z^{(N)}$ can exceed the value $E[Z^{(N)}] + c\sqrt{N}$, where c is a positive number and $E[Z^{(N)}]$ is the expectation value of $Z^{(N)}$. Inequality (26) holds for i.i.d. trials but also (in the weaker case) when $Z^{(N)}$ is a supermartingale, i.e., $E[Z^{(l+1)}|Z^{(1)}, \ldots, Z^{(l)}] \leqslant Z^{(l)}$ for all l or, equivalently, $E[\Delta^{(n)}|\Delta^{(1)}, \ldots, \Delta^{(n-1)}] \leqslant 0$ for all n.

We will examine separately the case where local realism (LR) holds fully, the case where local realism fails in the way described in scenario (i) in Sec. V, and the case where it fails in the way described in scenario (ii) in Sec. V. We call the latter two situations "ϵ local realism" (ϵLR).

Under local realism. Consider the random process Z_J: $Z_J^{(0)} = 0, Z_J^{(1)}, \ldots, Z_J^{(N)}$ with $Z_J^{(l)} = \sum_{n=1}^{l} J^{(n)}$, where the measured value (i.e., the increment of the process) in run n is denoted by $J^{(n)}$. We abbreviate $p_{ij} \equiv p(a_i)\,p(b_j)$ which, under freedom of choice, equals $p(a_i b_j)$, i.e., the probability that Alice chooses setting a_i and Bob chooses b_j. Due to the setting biases, these four values need not be $\frac{1}{4}$. Furthermore, we label with X_{ij}^{AB} those trials where Alice chooses setting a_i and observes outcome $A \in \{+,0\}$, and Bob chooses b_j and observes outcome $B \in \{+,0\}$. The increments $J^{(n)}$ are defined as

$$J^{(n)} \equiv \begin{cases} +\frac{1}{p_{11}} & \text{for } X_{11}^{++} \\ -\frac{1}{p_{12}} & \text{for } X_{12}^{+0} \\ -\frac{1}{p_{21}} & \text{for } X_{21}^{0+} \\ -\frac{1}{p_{22}} & \text{for } X_{22}^{++} \\ 0 & \text{else.} \end{cases} \quad (27)$$

The probability for a trial X_{ij}^{AB} is given by the probability p_{ij} that the setting combination $a_i b_j$ is chosen, multiplied
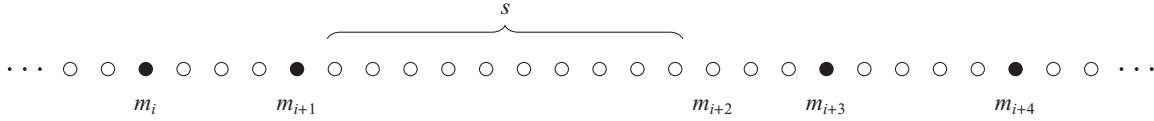
FIG. 2. Illustration of the dilution scheme. The circles represent the experimental trials $n = 1, \ldots, N$, where white and black fillings correspond to the values $K^{(n)} = -\epsilon_{AB}$ and $K^{(n)} \neq -\epsilon_{AB}$, respectively. The concentrated process has "stopping times" m_i, which encompass all black trials as well as all those, including white, which are preceded by a streak of s subsequent occurrences of white trials.

with the conditional probability to observe the outcomes A and B given this setting choice: $p(X_{ij}^{AB}) = p_{ij} \, p_{AB}(a_i, b_j)$. The definition (27) thus assures that the expectation value of $J^{(n)}$ is precisely given by J from (6) and hence, under local realism, is bounded by zero. (Note that, unlike J, the process Z_J scales with N unboundedly.) Even allowing memory, the expected value of every increment is still bounded by zero: $E[J^{(n)}|J^{(1)}, \ldots, J^{(n-1)}] \leqslant 0$, making the process Z_J a supermartingale. The increments fall within the range

$$r_J = \frac{1}{p_{11}} + \max\left(\frac{1}{p_{12}}, \frac{1}{p_{21}}, \frac{1}{p_{22}}\right), \tag{28}$$

which is close to 8 for small biases. The Hoeffding inequality for Z_J reads

$$p_{LR}\left(Z_J^{(N)} \geqslant c\sqrt{N}\right) \leqslant e^{-\frac{2}{r_J^2}c^2}. \tag{29}$$

Scenario (i). Communication in some trials. Now we consider the case of ϵLR in scenario (i) of Sec. V, i.e., the adapted inequality (15). If LR fails altogether, the expectation value of $J^{(n)}$ can reach 1. If LR fails only due to rare communication events and if we assume these failures are independent of the history of the experiment, then the expectation value of $J^{(n)}$ can reach ϵ_{AB}. This means that under ϵLR, Z_J is no longer a supermartingale. We define the process Z_K with increments

$$K^{(n)} \equiv J^{(n)} - \epsilon_{AB}. \tag{30}$$

Due to $E[K^{(n)}|K^{(1)}, \ldots, K^{(n-1)}] \leqslant 0$, the process Z_K is a supermartingale also in scenario (i), ϵLR. The trial values $K^{(n)}$ still have range r_J. The Hoeffding inequality then reads

$$p_{\epsilon LR}\left(Z_K^{(N)} \geqslant c\sqrt{N}\right) \leqslant e^{-\frac{2}{r_J^2}c^2}, \tag{31}$$

where, using Eq. (30), one can replace $Z_K^{(N)}$ by $Z_J^{(N)} - N\epsilon_{AB}$.

If we denote by R the frequency of trials and by J the experimentally expected value, then, assuming small bias, the condition $Z_J^{(N)} \geqslant N\epsilon_{AB} + c\sqrt{N}$ in (31) is likely to be reached after a run time of $\frac{c^2}{R(J - \epsilon_{AB})^2}$. In a photonic Bell experiment with total collection efficiency $\eta \approx 75\%$ [31,32], one down-conversion pair in 10^3 pulses, and reasonable state visibility and rate of dark and background counts, the CH-E value would be of the order of $J \sim 10^{-6}$. (The low probability for a pair production dominates, but the state and measurement angles used at this detection efficiency also contribute to the smallness of this number.) Assuming a pulse rate of $R \approx 1$ MHz, $\epsilon_{AB} \approx 10^{-7}$, and $r_J \approx 8$, particle-physics "gold standard" significance of $p \sim 10^{-6}$ (i.e., $c \approx 20$) would only be reached after a run time of approximately 16 years, which exceeds the typical PhD student project duration.

Fortunately, however, this result can be improved using Doob's optional stopping theorem. Following Ref. [40], we

first estimate the fraction of all trials n for which the $J^{(n)}$ value is nonzero:

$$f = \frac{\sharp\{n : J^{(n)} \neq 0\}}{N}. \tag{32}$$

By inspection of (4), these are the trials $X_{11}^{++}, X_{12}^{+0}, X_{21}^{0+}, X_{22}^{++}$. All other combinations of settings and outcomes do not contribute to the CH-E value, i.e., have $J^{(n)} = 0$ and hence $K^{(n)} = -\epsilon_{AB}$. With the experimental parameters from above, we estimate that the fraction of contributing trials is $f \approx 2 \times 10^{-5}$.

With Doob's optional stopping theorem, it is possible to increase the statistical significance of a given data set by looking at a "concentrated process." If Z_J were a supermartingale, which it is only in LR and not in scenario (i) ϵLR, then the procedure would be rather straightforward as one could simply skip all noncontributing trials with $J^{(n)} = 0$. Our case is more complicated, as those noncontributing trials have (negative) value $K^{(n)} = -\epsilon_{AB}$ and hence do in fact contribute to Z_K.

We propose the following solution to this problem (see Fig. 2): Let us consider the aggregated value $Z_K^{(m)} = \sum_{n=1}^{m} K^{(n)}$ at M specific "stopping times" $m \in \{m_1, m_2, \ldots, m_M\}$, namely those where (a) $K^{(m)} \neq -\epsilon_{AB}$ or (b) $K^{(m)} = -\epsilon_{AB}$ when preceded by a "streak" of s contiguous occurrences of $K^{(n=m-s, \ldots, m-1)} = -\epsilon_{AB}$. Every stop starts a fresh streak. Let us abbreviate $\frac{1}{w} = \max(\frac{1}{p_{12}}, \frac{1}{p_{21}}, \frac{1}{p_{22}})$. This choice of m ensures, without looking into the future, that the increment from any $Z_K^{(m_i)}$ to $Z_K^{(m_{i+1})}$ is between $-\frac{1}{w} - (s+1)\epsilon_{AB}$ (which one gets for $m_{i+1} = m_i + s + 1$ when there is a streak of s occurrences of $K^{(n=m_i+1, \ldots, m_{i+1}-1)} = -\epsilon_{AB}$ and then a final $K^{(n=m_{i+1})} = -\frac{1}{w} - \epsilon_{AB}$) and $\frac{1}{p_{11}} - \epsilon_{AB}$ (which one gets for $m_{i+1} = m_i + 1$, and $K^{(n=m_{i+1})} = \frac{1}{p_{11}} - \epsilon_{AB}$). This implies that the concentrated process $Z_K^{(m_1)}, Z_K^{(m_2)}, \ldots, Z_K^{(m_M)}$ is a supermartingale with range

$$r_{J,s} = r_J + s\,\epsilon_{AB}. \tag{33}$$

The Hoeffding inequality (31) is now altered in two ways. First, the range increases from r_J to $r_{J,s}$. Second, in the concentrated process, "time is now running faster" [40], which means that N gets replaced by the concentrated process length M, which is the number of stopping times above. Hence,

$$p_{\epsilon LR, M}\left(Z_K^{(m_M)} \geqslant c\sqrt{M}\right) \leqslant e^{-\frac{2}{r_{J,s}^2}c^2}. \tag{34}$$

Note that with $s = 0$, one recovers the original process, i.e., $r_{J,s=0} = r_J$ and $M = N$, and thus inequality (31). Using Eq. (30), $Z_K^{(m_M)}$ can be replaced by $Z_J^{(m_M)} - m_M\epsilon_{AB}$. When most of the trials are noncontributing, one can choose s such that $M \ll N$ while $m_M \approx N$.

Now we focus our attention again to an estimation of the experimental run time. For our purposes, it is not necessary to

find the optimal value for s, which will in general depend on f and ϵ_{AB}. We will see the remarkable power of Doob's theorem already by choosing $s = \lfloor \epsilon_{AB}^{-1} \rfloor$ which, for small biases and small ϵ_{AB}, leads to range $r_{J,s} \approx 9$. Because of $f \gg \epsilon_{AB}$, it (almost) never happens that there are full streaks of $\lfloor \epsilon_{AB}^{-1} \rfloor + 1$ subsequent occurrences of noncontributing $J^{(n)}$ trials. Hence we can take $M \approx fN$, meaning that the concentrated process stops at (almost) exactly the contributing trials. The condition in (34) is likely to be reached after a run time of $\frac{c^2 f}{R(J - \epsilon_{AB})^2}$. To obtain the same statistical significance as before ($p \sim 10^{-6}$), we now need to increase $c \approx 20$ by a factor of $\frac{r_{J,s}}{r_J} \approx \frac{9}{8}$ to $c \approx 22.5$. In total, Doob's theorem leads to a remarkable reduction of the run time by a factor of $\frac{9^2}{8^2} f$ from 16 years to 3 hours, right into the range of experimental feasibility. (We note that although the specific experimental values in a future optical Bell test may differ substantially from our estimates, it is very likely that within the near future, the application of Doob's theorem as just outlined is essential to achieve good statistical significance within a feasible run time.)

Scenario (ii). Excess predictability in all trials. We now consider the case of ϵLR in scenario (ii) of Sec. V, i.e., the adapted inequality (25). This situation is simpler than scenario (i). We can define the increments of a process Z_{J_ϵ} as

$$
J_\epsilon^{(n)} \equiv
\begin{cases}
+\dfrac{1}{p_{11}(1+\epsilon_+)} & \text{for } X_{11}^{++} \\
-\dfrac{1}{p_{12}(1-\epsilon_-)} & \text{for } X_{12}^{+0} \\
-\dfrac{1}{p_{21}(1-\epsilon_-)} & \text{for } X_{21}^{0+} \\
-\dfrac{1}{p_{22}(1-\epsilon_-)} & \text{for } X_{22}^{++} \\
0 & \text{else,}
\end{cases}
\tag{35}
$$

with $p_{ij} \equiv p(a_i)\, p(b_j)$, which need not equal $p(a_i b_j)$. This process has the range

$$
r_{J_\epsilon} = \frac{1}{p_{11}(1+\epsilon_+)} + \max\left(\frac{1}{p_{12}}, \frac{1}{p_{21}}, \frac{1}{p_{22}} \right)\frac{1}{1-\epsilon_-}, \tag{36}
$$

which, for small biases and small $\epsilon_\pm$, is close to 8. Even allowing memory, the expectation value of $J^{(n)}$ is precisely given by J_ϵ from (25), making the process Z_{J_ϵ} a supermartingale. Doob's theorem can be applied right away and all noncontributing trials can be discarded. With M contributing trials, the Hoeffding inequality reads

$$
p_{\epsilon LR, M}\left(Z_{J_\epsilon}^{(N)} \geqslant c\sqrt{M} \right) \leqslant e^{-\frac{2}{r_{J_\epsilon}^2}c^2}. \tag{37}
$$

If the bound (19) fails sometimes, say with probability q_f, then the algebraic bound of J_ϵ, which is $\frac{1}{1-\epsilon_-}$, can be reached in these trials. The above formulas have to be adapted in the following way, using the logic from scenario (i): A process Z_{K_ϵ} is defined such that $Z_{K_\epsilon}^{(N)} = Z_{J_\epsilon}^{(N)} - N\frac{q_f}{1-\epsilon_-}$ with range $r_{K_\epsilon} + s\frac{q_f}{1-\epsilon_-}$, using streak length s.

We note that ϵ_A, ϵ_B—and thereby ϵ_{AB} in scenario (i), Eq. (14), and $\epsilon_\pm$ in scenario (ii), Eq. (18)—as well as the setting probabilities p_{ij} must be estimated in some way, presumably from experimental characterization of the setting choice generation process. To preserve the statistical conclusions, the used values of ϵ_A, ϵ_B, and p_{11} should be conservative overestimates, while p_{12}, p_{21}, p_{22} should be conservative underestimates. Estimates of this kind, including p values for ϵ_A, ϵ_B, have

recently been reported for phase-diffusion random-number generators [52,58]. The p value for ϵ_A, ϵ_B can be taken into account by including the failure probability q_f into the process counting procedure, explained in the previous paragraph. In general, then, an experiment can thus lead to two p values: one for the process value and one for the p_{ij} estimates. These p values can be used in a single test, for example, using the Bonferroni method: to reach significance α, perform two separate hypothesis tests of the two hypotheses (bounded p_{ij} and local realism) with significance $\alpha/2$.

For random-number generators with small bias [52,58], it might be more efficient to quantify with ϵ_A, ϵ_B the excess predictability beyond probability $\frac{1}{2}$, despite the presence of the bias. Expression (16) then becomes $\frac{1}{2}(1-\epsilon_A) \leqslant p(a|\mu) \leqslant \frac{1}{2}(1+\epsilon_A)$, and similar for Bob. This has the advantage that it suffices to estimate ϵ_A, ϵ_B—which in this definition now include both bias itself and excess predictability beyond bias—and their failure probability q_f, so that estimates of the p_{ij} are not required. The expressions (35)–(37) still hold, with all four $p_{ij} = \frac{1}{4}$; in this case, an experiment leads to only one p value (for the process value). This procedure was used in Ref. [20].

We finally remark that the Hoeffding bounds used above are not optimal and better bounds are known [57], and that there are elegant methods of testing local realism even without assuming any specific form of a Bell inequality. They use the Kullback-Leibler divergence [59], which measures the mathematical difference of the probability distribution obtained from experimental data and that of any given local realist model. We refer the reader to Refs. [35,60–63].

VII. CONCLUSION

A Bell test claiming violation of the CH or Eberhard inequality bound by some few standard deviations could suffer from an incomplete consideration of the task at hand. Even disregarding world views such as superdeterminism that are inaccessible to the scientific method, it is possible to enforce spacelike separation only up to a limit due to imperfections in even state-of-the-art setting generators. In turn, to truly violate local realism in photonic Bell tests, it is necessary to modify the CH-Eberhard inequality based on the known imperfections of the setting generator in use. We showed how to derive such modifications in two different physical scenarios. Moreover, in the statistical analysis, we applied Doob's optional stopping theorem which dramatically reduces the run time for reasonable experimental parameters.

ACKNOWLEDGMENTS

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

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

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

[4] J. S. Bell, in *Proceedings of the International School of Physics 'Enrico Fermi'*, Course IL: *Foundations of Quantum Mechanics* (Academic, New York, 1971).

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

[6] J. S. Bell, Epist. Lett. **9**, 11 (1976).

[7] J. P. Jarrett, Nous **18**, 569 (1984).

[8] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[9] M. J. W. Hall, arXiv:0909.0015.

[10] A. Shimony, M. A. Horne, and J. F. Clauser, Epist. Lett. **13**, 1 (1976).

[11] J. S. Bell, Epist. Lett. **15**, 79 (1977).

[12] J. S. Bell, J. F. Clauser, M. A. Horne, and A. Shimony, Dialectica **39**, 85 (1985).

[13] J.-W. Pan, Z.-B. Chen, C.-Y. Lu, H. Weinfurter, A. Zeilinger, and M. Zukowski, Rev. Mod. Phys. **84**, 777 (2012).

[14] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).

[15] H. M. Wiseman, J. Phys. A: Math. Theor. **47**, 424001 (2014).

[16] J.-Å. Larsson, J. Phys. A: Math. Theor. **47**, 424003 (2014).

[17] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **81**, 5039 (1998).

[18] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. Langford, T. Jennewein, and A. Zeilinger, Proc. Natl. Acad. Sci. USA **107**, 19708 (2010).

[19] C. Erven, E. Meyer-Scott, K. Fisher, J. Lavoie, B. L. Higgins, Z. Yan, C. J. Pugh, J.-P. Bourgoin, R. Prevedel, L. K. Shalm, L. Richards, N. Gigov, R. Laflamme, G. Weihs, T. Jennewein, and K. J. Resch, Nat. Photon. **8**, 292 (2014).

[20] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, Phys. Rev. Lett. **115**, 250401 (2015).

[21] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, Phys. Rev. Lett. **115**, 250402 (2015).

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

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

[24] A. Zeilinger, Phys. Lett. A **118**, 1 (1986).

[25] P. M. Pearle, Phys. Rev. D **2**, 1418 (1970).

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

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

[28] D. N. Matsukevich, P. Maunz, D. L. Moehring, S. Olmschenk, and C. Monroe, Phys. Rev. Lett. **100**, 150404 (2008).

[29] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, Science **337**, 72 (2012).

[30] M. Ansmann, H. Wang, R. C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, A. N. Cleland, and J. M. Martinis, Nature (London) **461**, 504 (2009).

[31] 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 (London) **497**, 227 (2013).

[32] 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).

[33] J.-Å. Larsson and R. D. Gill, Europhys. Lett. **67**, 707 (2004).

[34] J.-Å. Larsson, M. Giustina, J. Kofler, B. Wittmann, R. Ursin, and S. Ramelow, Phys. Rev. A **90**, 032107 (2014).

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

[36] M. B. Agüero, A. A. Hnilo, and M. G. Kovalsky, Phys. Rev. A **86**, 052121 (2012).

[37] J. Barrett, D. Collins, L. Hardy, A. Kent, and S. Popescu, Phys. Rev. A **66**, 042111 (2002).

[38] L. Accardi and M. Regoli, arXiv:quant-ph/0007005; arXiv:quant-ph/0007019; arXiv:quant-ph/0110086.

[39] R. Gill, in *Mathematical Statistics and Applications: Festschrift for Constance van Eeden*, edited by M. Moore, S. Froda, and C. Léger, Vol. 42 (Institute of Mathematical Statistics, Beachwood, OH, 2003), p. 233; R. Gill, arXiv:quant-ph/0110137.

[40] R. Gill, in *Proceedings of Foundations of Probability and Physics 2*, Vol. 5, Mathematical Modeling in Physics, Engineering, and Cognitive Science (Växjö University Press, Växjö, Sweden, 2003), p. 179; arXiv:quant-ph/0301059.

[41] A. Kent, Phys. Rev. A **72**, 012107 (2005).

[42] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Rev. ed. (Cambridge University Press, Cambridge, 2004), pp. 243–244.

[43] J. Kofler, S. Ramelow, M. Giustina, and A. Zeilinger, arXiv:1307.6475.

[44] A. Garg and N. D. Mermin, Phys. Rev. D **35**, 3831 (1987).

[45] J.-Å. Larsson, Phys. Rev. A **57**, R3145(R) (1998).

[46] A. Khrennikov, S. Ramelow, R. Ursin, B. Wittmann, J. Kofler, and I. Basieva, Phys. Scr. **T163**, 014019 (2014).

[47] J. Gallicchio, A. S. Friedman, and D. I. Kaiser, Phys. Rev. Lett. **112**, 110405 (2014).

[48] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and Anton Zeilinger, Rev. Sci. Instrum. **71**, 1675 (2000).

[49] M. Fürst, H. Weier, S. Nauerth, D. G. Marangon, C. Kurtsiefer, and H. Weinfurter, Opt. Express **18**, 13029 (2010).

[50] M. Stipčević and R. Ursin, Sci. Rep. **5**, 10214 (2015).

[51] C. Abellán, W. Amaya, M. Jofre, M. Curty, A. Acin, J. Capmany, V. Pruneri, and M. W. Mitchell, Opt. Express **22**, 1645 (2014).

[52] C. Abellán, W. Amaya, D. Mitrani, V. Pruneri, and M. W. Mitchell, Phys. Rev. Lett. **115**, 250403 (2015).

[53] D. Bacon and B. F. Toner, Phys. Rev. Lett. **90**, 157904 (2003).

[54] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[55] G. Pütz, D. Rosset, T. J. Barnea, Y.-C. Liang, and N. Gisin, Phys. Rev. Lett. **113**, 190402 (2014).

[56] W. Hoeffding, J. Amer. Stat. Assoc. **58**, 13 (1963).

[57] P. Bierhorst, J. Phys. A: Math. Theor. **48**, 195302 (2015).

[58] M. W. Mitchell, C. Abellán, and W. Amaya, Phys. Rev. A **91**, 012314 (2015).

[59] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).

[60] W. van Dam, R. D. Gill, and P. D. Grünwald, IEEE Trans. Inf. Theory **51**, 2812 (2005).

[61] Y. Zhang, E. Knill, and S. Glancy, Phys. Rev. A **81**, 032117 (2010).

[62] Y. Zhang, S. Glancy, and E. Knill, Phys. Rev. A **84**, 062118 (2011).

[63] Y. Zhang, S. Glancy, and E. Knill, Phys. Rev. A **88**, 052119 (2013).