

On the equivalence of the Clauser–Horne and Eberhard inequality based tests

Andrei Khrennikov¹, Sven Ramelow^{2,3,4}, Rupert Ursin²,
Bernhard Wittmann^{2,3}, Johannes Kofler⁵ and Irina Basieva⁶

¹International Center for Mathematical Modeling in Physics, Engineering, Economics, and Cognitive Science, Linnaeus University, Växjö-Kalmar, Sweden

²Institute for Quantum Optics and Quantum Information Vienna (IQOQI), Austrian Academy of Sciences, Boltzmannngasse 3, Vienna, Austria

³Quantum Optics, Quantum Nanophysics, Quantum Information, University of Vienna, Faculty of Physics, Boltzmannngasse 5, Vienna, Austria

⁴Cornell University, 271 Clark Hall, 142 Science Dr., Ithaca, 14853 NY, USA

⁵Max Planck Institute of Quantum Optics (MPQ), Hans-Kopfermann-Str. 1, 85748 Garching, Germany

⁶Prokhorov General Physics Institute, Vavilov str. 38D, Moscow, Russia

E-mail: Andrei.Khrennikov@lnu.se

Received 3 December 2014

Accepted for publication 10 April 2014

Published 19 December 2014

Abstract

Recently, the results of the first experimental test for entangled photons closing the detection loophole (also referred to as the fair sampling loophole) were published (Vienna, 2013). From the theoretical viewpoint the main distinguishing feature of this long-aspired to experiment was that the Eberhard inequality was used. Almost simultaneously another experiment closing this loophole was performed (Urbana-Champaign, 2013) and it was based on the Clauser–Horne inequality (for probabilities). The aim of this note is to analyze the mathematical and experimental equivalence of tests based on the Eberhard inequality and various forms of the Clauser–Horne inequality. The structure of the mathematical equivalence is nontrivial. In particular, it is necessary to distinguish between algebraic and statistical equivalence. Although the tests based on these inequalities are algebraically equivalent, they need not be equivalent statistically, i.e., theoretically the level of statistical significance can drop under transition from one test to another (at least for finite samples). Nevertheless, the data collected in the Vienna test implies not only a statistically significant violation of the Eberhard inequality, but also of the Clauser–Horne inequality (in the ratio-rate form): for both a violation $>60\sigma$.

Keywords: loophole free test, Bell inequality, Eberhard inequality, Clauser–Horne inequality, confidence interval, Chebyshev inequality, non-Gaussian distributions

1. Introduction

The experimental realization of a loophole-free test for Bell inequalities [1] will impact both quantum foundations and quantum technologies. In both cases the present situation, e.g., in quantum cryptography [2] and randomness generation [3] (see also [4] for discussion), is unsatisfying from the scientific viewpoint. To experimentally falsify local realism, a so-called loophole-free Bell experiment will have to be

accomplished successfully⁷. This had not been claimed so far in any of the reported experiments. Up to now,

⁷ We remark that *a priori* one still cannot exclude the possibility that in the final loophole-free experiment the Bell inequality would be satisfied. In such a (very improbable) case, since quantum theory predicts that for the state under preparation, Bell's inequality has to be violated, the experiment would imply rejection of the quantum model. Thus the Bell test can also be considered as an attempt to falsify quantum mechanics. (At the initial stage of Bell experimentation the expectation that quantum mechanics would be falsified was common to some extent.)

space-like separation of measurements and basis choices has been accomplished in the pioneering experiments of Aspect *et al* [5, 6] and Weihs *et al* [7], closing the so-called locality loophole⁸. These experimental tests were based on the Clauser–Horne–Shimony–Holt (CHSH) inequality [9]. For this inequality to falsify local realism with maximally entangled states one must either approach very high total detection efficiency (which includes the optical losses of the setup and the efficiency of the detectors) $\eta = 82.8\%$ [10, 11], or proceed under an assumption to circumvent the loss—the so-called fair sampling assumption, see section 2 for discussion. It has been shown that using unfair sampling at sufficiently low detection efficiencies, very simple local models with hidden variables can violate the CHSH-inequality, e.g., [12, 13–21]. In spite of technological progress and the existence of detectors whose efficiency nears unity (e.g., TES detectors with an efficiency of around 95% [22]), it is very challenging to approach the required total detection efficiency.

In 1974, Clauser and Horne [13] proposed a new inequality that is not based on the fair sampling assumption. This inequality is expressed in terms of probabilities and we shall call it *the CH-inequality for probabilities*:

$$p(\alpha_1, \beta_1) + p(\alpha_1, \beta_2) + p(\alpha_2, \beta_1) - p(\alpha_2, \beta_2) \leq p^A(\alpha_1) + p^B(\beta_1), \quad (1)$$

where $p(\alpha, \beta)$ and $p^A(\alpha)$, $p^B(\beta)$ are probabilities for coincidence and single counts⁹, respectively; see section 4 for more detail.

In this paper we shall discuss various forms of the CH-inequality, see section 4 (see also the review of Clauser and Shimony [14] and the Stanford encyclopedia entry of Shimony [23] for details). Therefore we address each form with the corresponding label. However, we restrict our considerations to the class of the CH-inequalities not based on experimentally untestable auxiliary assumptions. Thus we shall not consider the CH-inequality whose derivation is based on the ‘no-enhancement assumption’ [13]: if an analyzer is removed from one of the paths, the resulting probability of detection is at least as great as with an analyzer.

To determine probabilities in (1), one has to know the total number of emitted pairs of photons. As was pointed out by Clauser and Horne [13] (see also the review of Clauser and Shimony [14] for extended discussion), it is practically impossible to determine this number experimentally. To escape this problem, it was proposed [14] to exclude the total number of emitted pairs from consideration by considering a version of the CH-inequality in the form of *ratio of detection*

counts rates¹⁰:

$$T = \frac{R(\alpha_1, \beta_1) + R(\alpha_1, \beta_2) + R(\alpha_2, \beta_1) - R(\alpha_2, \beta_2)}{R^A(\alpha_1) + R^B(\beta_1)} \leq 1, \quad (2)$$

where $R(\alpha, \beta)$ and $R^A(\alpha)$, $R^B(\beta)$ are coincidence and single rates, respectively; see section 4 for details.

For the CH-inequality for probabilities (1), Clauser and Horne [13] formulated the restrictions on the experimental setup in a straightforward way: first presupposing the ‘optimal angles’ (α, β) and then calculating other experimental parameters, namely the degree of entanglement and detection efficiency, to violate the inequality, see equations (5) and (6) in [13]. It can be shown that this procedure leads to very high required detection efficiency. To violate the CH-inequality for probabilities, the detection efficiency has to be at least 82.8%, see [12, 24].

In [25] Eberhard proposed a different approach by *jointly optimizing* all aforementioned parameters. He derived a new Bell inequality which we will abbreviate by ‘E-inequality’:

$$J \equiv n_{oe}(\alpha_1, \beta_2) + n_{ou}(\alpha_1, \beta_2) + n_{eo}(\alpha_2, \beta_1) + n_{uo}(\alpha_2, \beta_1) + n_{oo}(\alpha_2, \beta_2) - n_{oo}(\alpha_1, \beta_1) \geq 0, \quad (3)$$

where $n_{xy}(\alpha_i, \beta_j)$ is the number of pairs detected in a given time period for settings α_i, β_j with outcomes $x, y = o, e, u$ and the outcomes (o) and (e) correspond to detections in the ordinary and extraordinary beams, respectively, and the event that a photon is undetected is denoted by the symbol (u) . We point to the main distinguishing features of the E-inequality:

- (a) derivation without the fair sampling assumption (and without the no-enhancement assumption);
- (b) taking into account undetected photons;
- (c) background events are taken into account;
- (d) the *linear form of presentation* (non-negativity of a linear combination of coincidence and single rates).

The latter feature (which is typically not emphasized in the literature) is crucial to find a simple procedure of optimization of experimental parameters and, hence, it makes the E-inequality the most promising experimental test to close the detection loophole and to reject local realism without the fair sampling assumption. Eberhard’s optimization has two main outputs which play an important role in the experimental design:

- (E1) It is possible to perform an experiment without the fair sampling assumption for detection efficiency less than 82.8%. Nevertheless, the detection efficiency must still be very high, at least 66.6% (in the absence of background).
- (E2) The optimal parameters correspond to non-maximally entangled states.

⁸ Although we discuss only Bell tests for entangled photons, it is relevant that the first closure of the detection loophole was achieved with massive particles [8].

⁹ ‘Single counts’ are defined as all counts registered on one side for a given setting.

¹⁰ We remark that an important implicit assumption of applicability of this inequality is the assumption of a (statistically) constant production rate for pairs of photons, see also section 4 for discussion.

In 2013, the possibility of proceeding with overall efficiencies lower than 82.8% (but larger than 66.6%) was explored for the E-inequality and the first experimental test ('the Vienna test') closing the detection loophole was published [26], for a more detailed presentation of statistical data see also [27, 28].

Almost simultaneously another experiment closing the detection loophole was performed [29, 30] based on the probability version of the CH-inequality [13], see (1).

In this note we analyze the mathematical and experimental equivalence of tests based on the E-inequality (3), and the CH-inequality in the *ratio-rate form* (2)—in fact, its modification for the *ratio of detection counts*, see section 4, the inequality (9). In particular, *one has to distinguish between algebraic and statistical equivalence* (and for the latter, the cases of finite and infinite samples). Although these inequalities are (trivially) algebraically equivalent (see section 3.2), the tests based on them need not be equivalent statistically (for finite samples), i.e., theoretically the level of statistical significance can change essentially under transition from one test to another, see section 3.4. Nevertheless, the data collected in the Vienna test [26] implies not only a statistically significant violation of the E-inequality, but also of the CH-inequality for ratio of detection counts and, hence, for ratio of detection rates: for both a violation $> 60\sigma$.

One of the aims of this note is to determine confidence intervals from the statistics of the data collected in [26]. We remark that if one does not assume that data is Gaussian, then it is impossible to determine the confidence interval exactly with the aid of the standard deviation. However, it is possible to estimate it by using the *Chebyshev inequality*, see, e.g., [31]. This inequality is applicable under the assumption of finite dispersion and independence of the experimental runs. Although the Chebyshev inequality gives only rough estimates of probabilities, in our case (for the data collected in the Vienna test) it is sufficiently powerful to estimate confidence intervals showing that the hypothesis of the local realistic description of the Vienna data must be rejected.

2. Fair sampling assumption

The fair sampling assumption plays a crucial role in justification of tests based on the CHSH-inequality.

Extended discussions on the role of this assumption in the resolution of the classical-quantum dilemma can be found in the papers of Pearle [12] and Clauser and Horne [13] and Aspect's PhD thesis [6], see also Aspect's 'naive experimentalist presentation' of Bell's tests [32]. Later the fair sampling assumption was analyzed in detail in the PhD theses of Larsson [33] and Adenier [34].

The fair sampling assumption is not made in a Bell test based on the E-inequality or any CH-inequality. We also remark that the no-enhancement assumption is not present in the list of assumptions for the derivation of the E-inequality, see [25] (assumptions (i)–(iii)).

Finally, we remark that in this paper we do not discuss two other important loopholes, *the coincidence-time loophole* (and the role of space-time in the Bell argument in general, cf [20, 35–43]) and the *freedom-of-choice loophole* [44] (and its relation to the impossibility of using the conventional model of classical probability theory, the Kolmogorov model, 1933, see [20, 38, 39, 45–50]).

3. On the equivalence of the E-inequality to the CH-inequalities

3.1. E-inequality

We follow Eberhard [25]: photons are emitted in pairs (a, b) . Under each measurement setting combination (α, β) , the events in which the photon a is detected in the ordinary and extraordinary beams are denoted by the symbols (o) and (e) , respectively, and the event that it is undetected is denoted by the symbol (u) . The same symbols are used to denote the corresponding events for the photon b . Therefore, for the pairs of photons there are nine types of events: (o, o) , (o, u) , (o, e) , (u, o) , (u, u) , (u, e) , (e, o) , (e, u) , and (e, e) .

Under the conditions of locality, realism and statistical reproducibility the inequality (3), see section 1, was derived.

3.2. Algebraic equivalence

As was mentioned in section 3.1, originally Eberhard derived his inequality for the four-detectors experiment, one detector at each output of two PBSs. In [26] it was shown that the 'e-outputs' can be eliminated from the E-inequality (3), i.e., the four-detectors experimental design can be transformed into the two-detectors design corresponding to detection of only 'o-outputs'. In this section we demonstrate that the latter version of the E-inequality is *algebraically equivalent* to the CH-inequalities in various forms: for probabilities, ratio of probabilities, ratio of rates, and ratio of detection counts.

As was pointed out [26], the E-inequality can be transformed into the following inequality:

$$\begin{aligned}
 & -n_{oo}(\alpha_1, \beta_1) + S_o^A(\alpha_1) - n_{oo}(\alpha_1, \beta_2) + S_o^B(\beta_1) \\
 & -n_{oo}(\alpha_2, \beta_1) + n_{oo}(\alpha_2, \beta_2) \geq 0,
 \end{aligned}
 \tag{4}$$

where $S_o^A(\alpha_1)$ and $S_o^B(\beta_1)$ are numbers of single counts in the o -channels for 'Alice' and 'Bob', in settings α_1 and β_1 respectively. To match with the CH-inequalities completely, we change the sign and collect singles terms in the right-hand side:

$$\begin{aligned}
 & n_{oo}(\alpha_1, \beta_1) + n_{oo}(\alpha_1, \beta_2) + n_{oo}(\alpha_2, \beta_1) \\
 & -n_{oo}(\alpha_2, \beta_2) \leq S_o^A(\alpha_1) + S_o^B(\beta_1).
 \end{aligned}
 \tag{5}$$

Then by dividing by the number N of emitted pairs per setting combination on both sides (and omitting the index ‘o’) we obtain the *CH-inequality for probabilities*¹¹. (Here we proceed under the assumption of statistically constant production rate for pairs of photons, cf section 4).

$$p(\alpha_1, \beta_1) + p(\alpha_1, \beta_2) + p(\alpha_2, \beta_1) - p(\alpha_2, \beta_2) \leq p^A(\alpha_1) + p^B(\beta_1), \quad (6)$$

where $p(\alpha_i, \beta_j) = n_{oo}(\alpha_i, \beta_j)/N$, $p^A(\alpha_1) = S_o^A(\alpha_1)/N$, $p^B(\beta_1) = S_o^B(\beta_1)/N$.

However, as pointed out by Clauser and Horne [13], this inequality suffers from the problem that the number N and, hence, probabilities, are not well-determined in an experiment. To solve this problem, (6) can be transformed into the inequality:

$$T = \frac{p(\alpha_1, \beta_1) + p(\alpha_1, \beta_2) + p(\alpha_2, \beta_1) - p(\alpha_2, \beta_2)}{p^A(\alpha_1) + p^B(\beta_1)} \leq 1. \quad (7)$$

(We call this inequality [14] *CH-inequality for ratio of probabilities*, see remark 1 later). And finally T can be represented as the ratio of detection count rates:

$$T = \frac{R(\alpha_1, \beta_1) + R(\alpha_1, \beta_2) + R(\alpha_2, \beta_1) - R(\alpha_2, \beta_2)}{R^A(\alpha_1) + R^B(\beta_1)} \leq 1, \quad (8)$$

where $R(\alpha, \beta)$ and $R^A(\alpha_1)$, $R^B(\beta_1)$ are coincidence and single rates, respectively. (Following Clauser and Shimony [14] we call this inequality *CH-inequality for ratio of detection count rates* or simply the *ratio-rates CH-inequality*, see again Remark 1). This inequality is evidently equivalent to the following inequality in Eberhard’s notation, i.e., with the total numbers of coincidences and single counts, instead of the rates:

$$T = \frac{n(\alpha_1, \beta_1) + n(\alpha_1, \beta_2) + n(\alpha_2, \beta_1) - n(\alpha_2, \beta_2)}{S^A(\alpha_1) + S^B(\beta_1)} \leq 1. \quad (9)$$

(We call this inequality *CH-inequality for ratio of detection counts* or simply the *ratio-counts CH-inequality*, see again Remark 1.)

This inequality can be directly derived from (5) without relying on making any statements or assumptions about N . We can also proceed another way around and derive the E-inequality in the form (5) from the CH-inequality in the form (9). Thus *these two inequalities are algebraically equivalent*. The problem of their statistical equivalence will be studied in section 3.4. And it is more complicated.

Remark 1 (On terminology). The CH-inequality for probabilities, see (6), is sometimes called simply the CH-inequality. On the other hand, in the paper of Clauser and Shimony [14] this inequality was considered as just an

¹¹ This inequality is sometimes referred to simply as ‘CH-inequality’; see remark 1 for a short discussion on the terminology related to the papers of Clauser and Horne [13] and Clauser and Shimony [14].

intermediate step towards the ratio-rates CH-inequality, see (8). Thus it would be natural to refer to the latter as the CH-inequality. This terminology is used, e.g., by Shimony [23], who called the ratio-rates inequality, the BCH-inequality, see also the experimental proposal of Fry and Walther [52]. The main source of referring to (6) as the CH-inequality is that the material in the original CH-paper [13] was presented in a very compact form; in particular, the inequalities for ratios of probabilities and rates, see (7) and (8), were not written anywhere. The authors just remarked that the upper limit in (6) can be experimentally testable without N being known. This statement can be interpreted as simply the (algebraic) equivalence of the inequality (6) to the E-inequality (5) achieved with multiplication of the right- and left-hand sides of (6) by N (again under the assumption of statistically constant production rate).

3.3. Application of Chebyshev’s inequality to data from the Vienna test

Consider statistical data that is normally distributed. The information about the mean value μ and the standard deviation σ is sufficient to find the spread of these data relative to the number of standard deviations from the mean value. Denote the mean value and the standard deviation by the symbols μ and σ , respectively. It is known that 68% of these data are within 1σ -deviation from μ , 95% of the data are within 2σ -deviations from μ , and approximately 99% of the data are within 3σ -deviation from μ .

However, if the statistical data set is not normally distributed, i.e., its density deviates from the bell shape, then a different amount could be within $k\sigma$ -deviation, $k = 1, 2, 3, \dots$. In this case one can apply *Chebyshev’s inequality* [51]—a powerful tool to get to know what fraction of the statistical data falls within a few standard deviations from the mean value. We recall that the Chebyshev theorem states that, for any random variable ξ with finite second moment¹², i.e., $E(|\xi|^2) < \infty$, where E denotes the expectation value, and any positive number c ,

$$P(|\xi| \geq c) \leq \frac{E(|\xi|^2)}{c^2}. \quad (10)$$

Typically in applications one starts with a random variable η and in (10) selects $\xi = \eta - \mu$, where $\mu = E(\eta)$ is the mean value of η . Thus

$$P(|\eta - \mu| \geq c) \leq \frac{E(|\eta - \mu|^2)}{c^2} = \left(\frac{\sigma}{c}\right)^2, \quad (11)$$

where σ is standard deviation of η . We remark that a violation with $c = k\sigma$ results in confidence probability for a violation of $(1 - 1/k^2)$.

By using the values for the mean value and standard deviation calculated in [26] and by applying Chebyshev’s

¹² In particular, the Chebyshev inequality is applicable to any bounded random variable.

inequality, we can find (without knowing the probability distribution exactly, but only assuming that the dispersion is finite, see footnote 13) that the E-inequality is violated statistically significantly.

Remark 2. It is important to present the procedure of calculation of the empirical values m_j and s_j , which was used by Giustina *et al* [26]: ‘After recording for a total of 300 s per setting we divided our data into 10-s blocks and calculated the standard deviation of the resulting 30 different J values.’

We consider random variables corresponding to the left-hand and right-hand sides of the inequality (5):

$$J_{\text{pair}} = n_{oo}(\alpha_1, \beta_1) + n_{oo}(\alpha_1, \beta_2) + n_{oo}(\alpha_2, \beta_1) - n_{oo}(\alpha_2, \beta_2) \quad (12)$$

and

$$J_{\text{single}} = S_o^A(\alpha_1) + S_o^B(\beta_1). \quad (13)$$

Thus

$$J = J_{\text{single}} - J_{\text{pair}}. \quad (14)$$

Denote the mean value and standard deviation of the random variable J by the symbols μ_j and σ_j , respectively¹³.

Thus $\mu_j = E(J)$ and $\sigma_j^2 = E(J - \mu_j)^2$.

Our aim is to estimate the confidence interval for the mean. As always we shall use the statistical estimates of the mean and dispersion:

$$\bar{J} = \frac{1}{L} \sum_{i=1}^L J_i, \quad s_j^2 = \frac{1}{L-1} \sum_{i=1}^L (J_i - \bar{J})^2 \quad (15)$$

(in the experiment [26] $L = 30$, this is the number of the 10-s blocks, see Remark 2, each block is used to calculate J_i , the number of pairs in each block is very large and in such a framework its exact number is not so important).

We shall also use *the standard error of the mean* (the standard deviation of the sample-mean’s estimate of a population mean):

$$SE_j = \frac{s_j}{\sqrt{L}}. \quad (16)$$

and *the standard deviation of the mean*:

$$SD_j = \frac{\sigma_j}{\sqrt{L}}. \quad (17)$$

We remark that SE_j decreases as \sqrt{L} with increasing size of sample L ¹⁴.

¹³ We remark that, since, for experimental runs of fixed duration, the number of emitted photon pairs is a bounded random variable, the random variable J is bounded and, hence, $E|J|^2 < \infty$. Therefore dispersion is well defined and the Chebyshev inequality is applicable.

¹⁴ This intuitively makes a lot of sense also: the larger the sample one has, the smaller the confidence interval for the mean value. At the same time, one must not overestimate the role of getting a very small standard error of the mean. It mainly means that one was able to perform measurements for very long runs of the experiment and, in particular, to guarantee the stability of the functioning of the source and the measurement devices.

By using the Chebyshev inequality we obtain:

$$P(|\bar{J} - \mu_j| \geq c) \leq \frac{SD_j^2}{c^2}. \quad (18)$$

We proceed by using the standard error of the mean, instead of the standard deviation of the sample mean (in the formal mathematical presentation one has to use the correction related to the finite L , see [53, 54] for details):

$$P(|\bar{J} - \mu_j| \geq c) \leq \frac{SE_j^2}{c^2}. \quad (19)$$

From [26] we take the values $\bar{J} \approx -4224$ and $SE_j \approx 61.23$. This is a $>60\sigma$ violation, where $\sigma \equiv SE_j$. To avoid assuming normally distributed data further analysis is needed. By using the the Chebyshev inequality we estimate the confidence interval corresponding to the confidence level 99.95%. (In applied statistics often the level 95% is considered as sufficiently high.) Here $c \approx 2738$. Thus:

$$\begin{aligned} P(m_j \in [\bar{J} - 2738, \bar{J} + 2738]) \\ = P(m_j \in [-6962, -1486]) \geq 0.9995. \end{aligned} \quad (20)$$

Thus the confidence that can be placed in the result of the Vienna test is very high. The demonstrated violation of the E-inequality is very unlikely a matter of chance.

3.4. Statistical (non-)equivalence

Generally (i.e., without additional assumptions on probability distributions) the algebraic equivalence of two tests does not imply their statistical equivalence, at least for finite samples. In particular, the tests based on the E-inequality and the CH-inequality in the form (9) are not statistically equivalent. The latter means that violation of one of them with $k\sigma$, where k is sufficiently large, need not imply that another will be violated with the same k . It may happen that the significance of the violation changes essentially. This is a general statistical feature, i.e., it is not coupled rigidly with the two statistical tests under consideration, see appendix.

We remark that if the conditions of the central limit theorem are satisfied (in particular, for identically distributed independent random variables), then by using the δ -method [55] (error propagation method) one can prove the statistical equivalence of the E-test and the CH-test for $L \rightarrow \infty$. However, for finite L , in general it is true that by looking at different functions of statistics of interest and using the delta method, one can get any answer one likes. All of these answers are just approximations, and some approximations are better than others. As was pointed out, for $L \rightarrow \infty$, they give the same answer, but for fixed L they all give different answers.

3.5. Statistically significant violation of the CH-inequality (in the ratio form) for the Vienna test

Thus on the basis of purely theoretical arguments one cannot derive a statistically significant violation of the CH-inequality

in the form (9) from statistically significant violation of the E-inequality. One has to use again the experimental data.

In this section we apply Chebyshev's inequality to show that the statistical data collected in the Vienna test also implies statistically significant violation of the (ratio-counts) CH-inequality. Set $\mu_T = E(T)$ (the mean value of the random variable T) and $\sigma_T^2 = E(T - \mu_T)^2$ (its dispersion). We shall use the statistical estimates of the mean and dispersion: $\bar{T} = \frac{1}{L} \sum_{i=1}^L T_i$, $s_T^2 = \frac{1}{L-1} \sum_{i=1}^L (T_i - \bar{T})^2$ (in the experiment [26] $L = 30$, this is the number of the 10-s blocks, each block is used to calculate T_i). We also consider the standard deviation of the mean $SD_{\bar{T}} = \frac{\sigma_T}{\sqrt{L}}$ and and the standard error of the mean $SE_{\bar{T}} = \frac{s_T}{\sqrt{L}}$.

By using the data from [26] and more detailed presentation in [28] we obtain $\bar{T} \approx 1.0394$ and $SE_{\bar{T}} \approx 0.0006$. This yields a $>60\sigma$ violation, where $\sigma = SE_{\bar{T}}$. From the Chebyshev inequality

$$P(|\bar{T} - \mu_T| \geq c) \leq \frac{SE_{\bar{T}}^2}{c^2} \quad (21)$$

we estimate the confidence interval corresponding to the confidence level 99.95%. We have $c \approx 0.027$.

$$\begin{aligned} P(m_T \in [\bar{T} - 0.027, \bar{T} + 0.027]) \\ = P(m_T \in [1.0124, 1.0664]) \geq 0.9995. \end{aligned} \quad (22)$$

Thus the confidence that can be placed in the result of the Vienna test is very high. The demonstrated violation of the (ratio-counts) CH-inequality is very unlikely a matter of chance.

4. The Vienna test for the CH-inequality (in the ratio form): taking into account intensity drift

In all previous considerations we assumed, as Eberhard did originally [25], that the number of emitted pairs N is constant during the experiment and does not depend on angles (α_i, β_j) . In the real experiment, the intensity drift was very small [26]. In [28], a data analysis procedure was proposed, based on the following assumption:

$$N(\alpha_1, \beta_1)/N(\alpha_2, \beta_1) = S_o^B(\alpha_1, \beta_1)/S_o^B(\alpha_2, \beta_1), \dots$$

Then one proceeds not simply with coincidence and single counts $n_{oo}, n_{oe}, \dots, S_o^A, S_o^B$, but with their normalized values based on the above mentioned proportion of intensities. We denote normalized quantities as $\tilde{n}_{oo}, \tilde{n}_{oe}, \dots, \tilde{S}_o^A, \tilde{S}_o^B$. For such normalized numbers of coincidences and singles, we can use the E-inequality:

$$\begin{aligned} \tilde{n}_{oo}(\alpha_1, \beta_1) + \tilde{n}_{oo}(\alpha_1, \beta_2) + \tilde{n}_{oo}(\alpha_2, \beta_1) \\ - \tilde{n}_{oo}(\alpha_2, \beta_2) \leq \tilde{S}_o^A(\alpha_1) + \tilde{S}_o^B(\beta_1). \end{aligned} \quad (23)$$

And we jump directly to the CH-inequality for the total numbers of coincidences and singles (which is equivalent to the CH-inequality for the rates):

$$T' = \frac{\tilde{n}_{oo}(\alpha_1, \beta_1) + \tilde{n}_{oo}(\alpha_1, \beta_2) + \tilde{n}_{oo}(\alpha_2, \beta_1) - \tilde{n}_{oo}(\alpha_2, \beta_2)}{\tilde{S}_o^A(\alpha_1) + \tilde{S}_o^B(\beta_1)}. \quad (24)$$

Using data collected in [26] and [28], we get $\bar{T}' \approx 1.0384$ with a violation $>60\sigma$ which shows that also the experimental data taking into account intensity drift lead to the same amount of non-classical correlations.

5. Concluding remark

The statistical data collected in the Vienna test [26, 28] violated statistically significantly not only the E-inequality, but equivalently also the CH-inequality, for ratio of detection counts and, hence, the ratio-rate inequality, cf [29, 30]. Thus, we can consider this experimental test as closing the detection loophole also with regard to the CH-inequality in the ratio-form.

6. Appendix: Statistical non-equivalence of (algebraically equivalent) linear and ratio test

Let X be a set of data sampled from realizations of some random variable x . Take two (for a moment arbitrary positive valued) functions, $J_1(x)$ and $J_2(x)$; set $J(x) = J_1(x) - J_2(x)$ and $T(x) = \frac{J_2(x)}{J_1(x)}$. Consider two tests for statistical data:

$$J(x) \geq 0; \quad (25)$$

$$T(x) \leq 1. \quad (26)$$

Suppose that the data X showed $k\sigma_J$ violation of the first inequality, where k is large. Thus violation of this inequality is significant. Our aim is to show that the same data can in principle show insignificant violation of the second inequality, say $\gamma\sigma_T$, where γ is very small.

Suppose, for example, that the data X was obtained as the result of measurements of a discrete random variable $x = x_1, x_2$, where x_1, x_2 are two arbitrary real numbers. It takes these values with probabilities p_1 and $p_2 = 1 - p_1$. Here we need to set the values of all functions only in the two points, x_1 and x_2 :

$$J_1(x_i) = A_i, J_2(x_i) = B_i, A_i, B_i > 0, i = 1, 2.$$

We take $B_i = (1 + \epsilon_i)A_i, \epsilon_i > 0$. We have:

$$\mu_J = -(\epsilon_1 A_1 p_1 + \epsilon_2 A_2 p_2), \quad \sigma_J^2 = p_1 p_2 (\epsilon_1 A_1 - \epsilon_2 A_2)^2;$$

$$\mu_T = 1 + \epsilon_1 p_1 + \epsilon_2 p_2, \quad \sigma_T^2 = p_1 p_2 (\epsilon_1 - \epsilon_2)^2;$$

$$R_j = \frac{-\mu_j}{\sigma_j} = \frac{\epsilon_1 A_1 p_1 + \epsilon_2 A_2 p_2}{\sqrt{p_1 p_2} |\epsilon_1 A_1 - \epsilon_2 A_2|};$$

$$R_T = \frac{\mu_T - 1}{\sigma_T} = \frac{\epsilon_1 p_1 + \epsilon_2 p_2}{\sqrt{p_1 p_2} |\epsilon_1 - \epsilon_2|}.$$

By playing with parameters we want to make $R_j \gg 1$ and at the same time $R_T \ll 1$. First we set $\epsilon_1 = \lambda \epsilon_2$, where $\lambda > 1$. We have

$$R_T = \sqrt{\frac{p_1}{p_2}} + \frac{1}{\sqrt{p_1 p_2} (\lambda - 1)}.$$

To make the first term very small, we select $p_1 = \delta^2 \ll 1$, so $p_2 \approx 1$; to make the second term very small, we select λ in such a way that its denominator is very large, i.e., $\sqrt{p_1 p_2} (\lambda - 1) \gg 1$. Thus, for the model parameters satisfying conditions

$$\delta \ll 1, \lambda \gg 1/\delta, \tag{27}$$

the inequality (26) is violated insignificantly. We remark that, since the parameter ϵ_2 has not yet been constrained, it is possible to make the absolute value of expectation very large, $\mu_T \gg 1$. We now want to make $R_j \gg 1$. We represent $A_1 = a A_2$, and we have: $R_j = \frac{a \lambda p_1 + p_2}{\sqrt{p_1 p_2} |\lambda a - 1|}$. Now we select a as $a \lambda - 1 > 0$. \tag{28}

Thus

$$R_j = \frac{a \lambda p_1 + p_2}{\sqrt{p_1 p_2} (a \lambda - 1)} = \sqrt{\frac{p_1}{p_2}} + \frac{1}{\sqrt{p_1 p_2} (a \lambda - 1)}.$$

The first summand is negligibly small, see (27), it does not play any role in our considerations. The parameter a has to be selected in such a way that the second summand will be very large. Thus $\sqrt{p_1 p_2} (a \lambda - 1) \ll 1$, or by taking into account (28) we obtain that $0 < a \lambda - 1 \ll 1/\sqrt{p_1 p_2} \approx 1/\delta$. Take very large natural number k . Suppose that

$$a = 1/\lambda + 1/k\delta\lambda. \tag{29}$$

Then $R_j \approx k^{15}$.

Acknowledgments

We would like to thank all authors of the original paper [26] as they contributed to the data we used in this paper; in particular, M Giustina whose numerous comments and suggestions substantia improved our paper. We would like to thank R Gill, and R Pettersson for discussions on statistical aspects of the paper, and J-Å Larsson and G Adenier for discussions on the role of the fair sampling assumption in the

¹⁵ For example, set $\delta = 0.1$, i.e., $p_1 = 0.01$, $p_2 = 0.99$. We remark that to get very small R_T and at the same time very large R_j the probability distribution of our model has to be strongly asymmetric. Then $\lambda = 102$ guarantees that $R_T < 0.2$. Finally, by setting $k = 69$ we obtain that it is sufficient to take $a = 0.0112$.

derivation of various versions of Bell’s inequality. This work was partially supported by MPNS COST Action MP1006, Fundamental Problems in Quantum Physics (I Basieva), a visiting fellowship (A Khrennikov) to Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, and EU Marie-Curie Fellowship (S Ramelow), PIOF-GA-2012-32985.

References

- [1] Bell J S 1987 *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press)
- [2] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145
- [3] Pironio S et al 2010 *Nature* **464** 1021–4
- [4] Fürst H, Weier H, Nauerth S, Marangon D G, Kurtsiefer Ch and Weinfurter H 2010 *Opt. Express* **18** 13029–37
- [5] Aspect A, Dalibard J and Roger G 1982 *Phys. Rev. Lett.* **49** 1804
- [6] Aspect A 1983 *Three Experimental Tests of Bell Inequalities by the Measurement of Polarization Correlations Between Photons* (Orsay: Orsay University Publications)
- [7] Weihs G, Jennewein T, Simon C, Weinfurter R and Zeilinger A 1998 *Phys. Rev. Lett.* **81** 5039–43
- [8] Rowe M A, Kielinski D, Meyer V, Sackett C A, Itano W M, Monroe C and Wineland D J 2001 *Nature* **409** 791–4
- [9] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880–4
- [10] Garg A and Mermin N D 1987 *Phys. Rev. D* **35** 3831
- [11] Larsson J-Å 1998 *Phys. Rev. A* **57** R3145
- [12] Pearle P M 1970 *Phys. Rev. D* **2** 1418–25
- [13] Clauser J F and Horne M A 1974 *Phys. Rev. D* **10** 526–35
- [14] Clauser J F and Shimony A 1978 *Rep. Prog. Phys.* **41** 1881–927
- [15] Fine A 1982 *Synthese* **50** 279–94
- [16] Selleri F and Zeilinger A 1988 *Found. Phys.* **18** 1141
- [17] Thompson C H 1996 *Found. Phys. Lett.* **9** 357
- [18] Larsson J-A 1998 *Phys. Rev. A* **57** 3304–8
- [19] Adenier G and Khrennikov A 2007 *J. Phys. B: Atom. Mol. Opt. Phys.* **40** 131–41
- [20] Khrennikov A 2009 *Contextual Approach to Quantum Formalism* (Berlin: Springer)
- [21] Khrennikov A 2012 *J. Phys. A: Math. Theor.* **45** 215301
- [22] Lita A E, Miller A J and Woo Nam S 2008 *Opt. Express* **16** 3032–40
- [23] Shimony A 2012 Bell’s Theorem *The Stanford Encyclopedia of Philosophy* ed E N Zalta (Stanford: Stanford University Press) <http://plato.stanford.edu/archives/win2012/entries/belltheorem>
- [24] Mermin N D 1986 ed D M Greenberger *Techniques and Ideas in Quantum Measurement Theory* (New York: New York Academy of Science) pp 422–8
- [25] Eberhard Ph H 1993 *Phys. Rev. A* 477–750
- [26] Giustina M et al 2013 *Nature* **497** 227–30
- [27] Ramelow S 2013 *Proc. Int. Conf. on Quantum Theory: Advances and Problems, A Marcus Wallenberg Symp. (Växjö, Sweden)*
- [28] Kofler J, Ramelow S, Giustina M and Zeilinger A 2013 On ‘Bell violation using entangled photons without the fair-sampling assumption’ arXiv:1307.6475 [quant-ph]
- [29] Kwiat P et al 2013 *Proc. Int. Conf. on Quantum Information and Measurement, Session* (New York: Rochester)
- [30] Christensen B G et al 2013 *Phys. Rev. Lett.* **111** 1304–06

- [31] Gibbons J D and Chakraborti S 2003 *Nonparametric Statistical Inference* 4th edn (New York: Dekker)
- [32] Aspect A 2004 Bell's Theorem: the naive view of an experimentalist arXiv:[quant-ph/0402001](https://arxiv.org/abs/quant-ph/0402001)
- [33] Larsson J-A 2000 *Quantum Paradoxes, Probability Theory, and Change of Ensemble* (Linköping: Linköping University Press)
- [34] Adenier G 2008 *Local Realist Approach and Numerical Simulation of Nonclassical Experiments in Quantum Mechanics* (Växjö: Växjö University Press)
- [35] Larsson J-A, Giustina M, Kofler J, Wittmann B, Ursin R and Ramelow S 2013 Bell violation with entangled photons, free of the coincidence-time loophole arXiv:[1309.0712](https://arxiv.org/abs/1309.0712)
- [36] Khrennikov A Yu and Volovich I V 2002 Local realism, contextualism and loopholes in Bell's experiments *Proc. Int. Conf. on Foundations of Probability and Physics-2 (Ser. Math. Model vol 5)* ed A Yu Khrennikov (Växjö: Växjö University Press) pp 325–44
- [37] Hess K and Philipp W 2005 Critique of proofs, with and without inequalities *Proc. Int. Conf. on Foundations of Probability and Physics-3 (AIP Conf. Proc. vol 750)* ed A Khrennikov (Melville: AIP) pp 516–23
- [38] Hess K and Philipp W 2002 *Europhys. Lett.* **57** 775
- [39] Hess K, Michielsen K and De Raedt H 2009 *Europhys. Lett.* **87** 60007
- [40] Gill R D 2003 Time, finite statistics, and Bell's fifth position *Proc. Int. Conf. on Foundations of Probability and Physics* — 2 vol 5 (Växjö: Växjö Univ. Press) pp 179–206 (arXiv:[quant-ph/0301059](https://arxiv.org/abs/quant-ph/0301059))
- [41] Gill R D, Weihs G, Zeilinger A and Zukowski M 2003 *Europhys. Lett.* **61** 282
- [42] Gill R D, Weihs G, Zeilinger A and Zukowski M 2002 *Proc. Natl. Acad. Sci. USA* **99** 14632–5
- [43] Gill R D Statistics, causality and Bell's theorem www.math.leidenuniv.nl/~gill/causality.pdf
- [44] Scheidl T et al 2010 *Proc. Natl. Acad. Sci. USA* **107** 19708–13
- [45] Kupczynski M 1987 *Phys. Lett. A* **121** 205
- [46] Fine A 1989 *Found. Phys.* **19** 453–62
- [47] Khrennikov A 2008 *Entropy* **10** 19–32
- [48] Boole G 1862 *Phil. Trans. R. Soc.* **152** 225–42
- [49] Boole G 1958 *An Investigation of the Laws of Thought* (New York: Dover)
- [50] Vorob'ev N N 1962 *Theor. Prob. Appl.* **7** 147–62
- [51] Tchebichef P 1867 *J. Math. Pures Appliquées* **2** 177–84
- [52] Fry E S and Walther T 1997 *A Bell Inequality Experiment Based on Molecular Dissociation Extension of the Lo-Shimony Proposal to 199Hg (nuclear spin 1/2) dimers in Experimental Metaphysics* ed R S Cohen, M Horne and J Stachel (Dordrecht: Kluwer) pp 61–71
- [53] Saw J G, Yang M C K and Mo T Ch 1984 *Am. Statistician* **38** 130–2
- [54] Kabán A 2012 *Stat. Comp.* **22** 375–85
- [55] Cramer H 1946 *Mathematical Models of Statistics*