



HAL
open science

Inadequacy of the CHSH Bell inequality for an application to its associated photon correlation experiments

Gerrit Coddens

► **To cite this version:**

Gerrit Coddens. Inadequacy of the CHSH Bell inequality for an application to its associated photon correlation experiments. 2022. hal-03482574v5

HAL Id: hal-03482574

<https://hal.science/hal-03482574v5>

Preprint submitted on 2 Apr 2022 (v5), last revised 10 Jun 2022 (v6)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inadequacy of the CHSH Bell inequality for an application to its associated photon correlation experiments

Gerrit Coddens (a)

Laboratoire des Solides Irradiés,
Institut Polytechnique de Paris, UMR 7642, CNRS-CEA-Ecole Polytechnique,
28, Route de Saclay, F-91128-Palaiseau CEDEX, France

(a) *retired research physicist of LSI*

2nd April 2022

Abstract. We analyse the proof of a CHSH Bell-type inequality and show that this inequality cannot be applied to the experiments it has been designed for. In the example we study, two polarization values are attributed to each of the two photons of a correlated photon pair, such that four values are assigned simultaneously to a single pair. But in the experiments these four values cannot be measured simultaneously on a single photon pair, such that the assumptions underlying the inequality do not properly translate the experimental situation into mathematics. The CHSH inequality is therefore based on wrong modelling. To derive an inequality that correctly reflects the experimental situation and does contain the experimentally measured probabilities $p(a_j \wedge b_k)$, a combination of four distinct photon pairs (and, where applicable, two single photons to measure $p(a_2)$ and $p(b_2)$ in independent runs) must be considered rather than just a single pair. Using the same methodology as was used for the CHSH inequality one can derive a new inequality, which is now based on correct modelling and properly translates the physics into mathematics. But it has different, less stringent boundaries with the effect that it is no longer violated by the experiments. Larsson has tried to rebut this within the context of the wrong model that has been used to derive the CHSH inequality. Within this wrong context he presented the different boundaries obtained within the correct model as an experimental loophole caused by statistical fluctuations when the counting statistics are not good enough. Individual violations of the inequality would be possible but the averages, i.e. the measured probabilities respect the inequality. This is wrong because what is at stake in the wrong model is a systematic theoretical error caused by the wrong modelling. It has nothing to do with statistical fluctuations and can therefore not be healed by improving the counting statistics. But based on his idea about the averages Larsson is able to prove the inequality for the measured probabilities. However, this is a proof locked up inside the premises of the wrong model, while a comparison between the wrong and the correct model shows that their probability distributions are normalized differently, such that the wrong model is smuggling in a normalization error. The same modelling error occurs in the derivation of other types of Bell inequalities such that none of them can be applied to the experimental data they were designed for. The violation of these Bell inequalities can therefore not be considered as a proof for the existence of entanglement.

PACS. 03.65.Ta, 03.65.Ud, 03.67.-a

1 Introduction - Traditional derivation of a CHSH Bell inequality

The subject matter of the Bell inequalities culminating in the experiments of Aspect *et al.* [1,2] can be supposed to be very well-known. For an introduction we refer to [3]. Further experimental developments have been the recent loophole-free experiments [4,5]. Let us inspect the derivation of the CHSH Bell inequality given in [3]. We consider 4 variables $a_1 \in S$, $a_2 \in S$, $b_1 \in S$, $b_2 \in S$, where $S = \{0, 1\}$. The idea is that 0 corresponds to absorption in a polarizer, and 1 to transmission. Then a_j will correspond to polarizer settings in one arm of the set-up, and b_k to polarizer settings in the other arm. There are thus 16 possible combinations for the values of (a_1, a_2, b_1, b_2) . By making a table of these 16 combinations it is easy to verify that we always have:

$$\forall (a_1, a_2, b_1, b_2) \in S^4 : Q = a_1 b_1 - a_1 b_2 - a_2 b_1 - a_2 b_2 + a_2 + b_2 \in S. \quad (1)$$

Throughout the paper we will use the notation $F(W_1, W_2)$ for the set of functions whose domain is the set W_1 and whose values belong to the set W_2 . Let us consider now functions $a_j \in F(V, S)$ and $b_k \in F(V, S)$. Here V is a set of relevant variables for the experiment. We can call the set V the set of hidden variables. One can imagine that V could be a subset of a vector space \mathbb{R}^n or of a manifold, e.g. a non-abelian Lie group like $SO(3)$ or $SU(3)$. We have then:

$$\begin{aligned} \forall \lambda \in V : \quad 0 \leq Q(\lambda) \leq 1, \quad \text{where:} \\ Q(\lambda) = a_1(\lambda)b_1(\lambda) - a_1(\lambda)b_2(\lambda) - a_2(\lambda)b_1(\lambda) - a_2(\lambda)b_2(\lambda) + a_2(\lambda) + b_2(\lambda). \end{aligned} \quad (2)$$

We can now consider a probability density p over V . The function p belongs then to the set of functions $F(V, [0, \infty[)$ with domain V and values in $[0, \infty[$. We further require that $\int_V p(\lambda)d\lambda = 1$. We can now integrate Equation 2 with p over V . Introducing the notations:

$$\begin{aligned} p(a_j \wedge b_k) &= \int_V a_j(\lambda) b_k(\lambda) p(\lambda) d\lambda, \\ p(a_j) &= \int_V a_j(\lambda) p(\lambda) d\lambda, \quad p(b_k) = \int_V b_k(\lambda) p(\lambda) d\lambda, \end{aligned} \quad (3)$$

we obtain then:

$$0 \leq p(a_1 \wedge b_1) - p(a_1 \wedge b_2) - p(a_2 \wedge b_1) - p(a_2 \wedge b_2) + p(a_2) + p(b_2) \leq 1. \quad (4)$$

Here \wedge is the logical ‘‘and’’ operator. This is the CHSH Bell inequality applied to the photon correlation experiments described in [1] and in its precursor [6]. It is a purely mathematical identity and does not depend on any physical considerations. It is also free of any considerations about statistical correlations and statistical independence, which must be dealt with in the definition of the probability density $p \in F(V, [0, +\infty[)$, i.e. its definition domain V , its values and its normalization to 1. In [1, 6], the quantity $p(a_j \wedge b_k)$ becomes the probability for the event that photon 1 is transmitted by the filter A oriented along a direction defined by the angle α_j while its buddy photon 2 is transmitted by the filter B oriented along β_k .

Rather than using some experimental results with a limited accuracy, we can inject theoretical expressions for the outcomes of these experiments into the inequality and check if they violate it. By adopting theoretical expressions we can avoid the burden of discussing the experimental and statistical uncertainties. Many such uncertainties may occur and lead to the objection that a violation reported could in reality not be genuine and be due to experimental errors. This is then called an experimental loophole. There have therefore been elaborate attempts to make the experiments very precise such as to close all possible experimental loopholes [4, 5]. These experiments and those described in [1, 2, 6] are all extremely meticulous, representing the state of the art of their time. The exact theoretical expressions can be considered to just express what the result of an infinitely precise loophole-free experiment would be. By plugging the exact theoretical expressions rather than experimental results into the inequality we free the argument from all experimental considerations. It becomes immune to them. The probabilities are thus identified with the mathematical expressions for the outcomes of the photon polarization experiments that are considered to be exact:

$$p(a_j \wedge b_k) = \frac{1}{2} \cos^2(\alpha_j - \beta_k), \quad p(a_j) = \frac{1}{2}, \quad p(b_k) = \frac{1}{2}, \quad (5)$$

where α_j and β_k are the angles of the polarizer settings in the two arms of the experiment. These expressions are known to violate the Bell inequality for certain choices of angles $\alpha_1, \alpha_2, \beta_1$ and β_2 . If the theoretical expressions in Equation 5 were not rigorously exact, then we could still test the inequality with the improved rigorous expressions. But the logic of the quest for rigorously exact data to test the Bell inequality falls apart if the Bell inequality in Equation 4 itself is wrong.

We can try to derive the results of Equation 5 by classical reasoning. But when we fail to find such a derivation this does not prove that such a classical derivation does not exist. It could just mean that we are too dull-witted or lacking imagination. An example of such a classical calculation which fails is the following. Let us assume that a source produces photon pairs that are both linearly polarized along an angle φ . Then applying Malus’ law the probabilities for the transmissions of the filters at angles α_j, β_k will be $\cos^2(\alpha_j - \varphi)$ and $\cos^2(\beta_k - \varphi)$ respectively. Under the assumption that the probability of the angle φ is uniformly distributed, we get by integration over φ the following classical expressions for the result of the experiment:

$$\begin{aligned} p(a_j \wedge b_k) &= \int_0^{2\pi} \cos^2(\alpha_j - \varphi) \cos^2(\beta_k - \varphi) d\varphi = \frac{1}{8} + \frac{1}{4} \cos^2(\alpha_j - \beta_k), \\ p(a_j) &= \frac{1}{2}, \quad p(b_k) = \frac{1}{2}. \end{aligned} \quad (6)$$

This is not in agreement with Equation 5 which reproduces the experimental results and is also the result predicted by quantum mechanics. Now we can speculate that our failure could be due to not choosing the right assumptions. This keeps us then wondering if a classical derivation exists or otherwise. In fact, the photons that impinge on a polarization filter have not interacted with another polarization filter beforehand such that using Malus' law might not be in order. The interaction with a filter is also not a mere transmission, else a succession of polarizers at angles $0, \frac{\pi}{2}, \frac{\pi}{4}$ would yield the same result as a succession $0, \frac{\pi}{4}, \frac{\pi}{2}$ which is not the case. The polarizer modifies the state of the photon and the correct description of this modification might be necessary for formulating a correct calculation. The photons are also traversing a medium where they are subject to refraction and their speed becomes smaller than c . But a violation of the Bell inequality puts an end to any such doubts, because it proves that such an alternative derivation just does not exist. It is like Galois' proof that in general we cannot solve a quintic by radicals. It is an impossibility proof which shows that it is pointless to keep searching for another derivation. That is the importance of the inequality. A violation of the inequality also questions realism and/or locality.

2 The logical error and how to correct for it

Let us now explain where this impossibility proof goes wrong. We made the point above that we could rather inject the theoretical values into the inequality because the issue we want to raise is not at all one of some experimental error giving rise to an experimental loophole. This issue we want to raise is a matter of principle, a logical error. It will be very important to keep this distinction in mind in the following. Note that in Equation 1 the number a_1 in $a_1 b_1$ is the same as in $a_1 b_2$, implying that $a_1 b_1$ and $a_1 b_2$ must be measured on a same photon pair. The equation implies that all numbers a_1, a_2, b_1, b_2 must be measured on a same photon pair.

□ Remark 1. This does not mean that the numbers (a_1, a_2, b_1, b_2) are all “inscribed” inside both photons of a pair. As the filter angles α_j and β_k can be given an uncountable infinity of orientations, that would require all possible numbers a_j and b_k to be “registered” inside the photons. It seems unlikely that such a giant amount of information would be stored in a single photon. It is more reasonable to assume that some information is present within the hidden variables of the photon and some information within the hidden variables of the experimental set-up, such that then together these hidden variables determine the outcome of a polarization experiment on a photon. Therefore, when we state that photons have the values (a_1, a_2, b_1, b_2) in the following, it will only be a misuse of language in order to express that the hidden variables of the photons together with those of their context (i.e. the set-up) determine these values.

Presumably, the photon has a few polarization parameters when it leaves the source and the rest is determined by the interaction of the photon with the set-up, i.e. the two polarizers and their orientations. In other words, the outcome of the experiment is produced by the interactions. This raises the question how we can consistently obtain the same response in both arms of the set-up when the two polarizers are parallel. The systematic agreement shows that the interactions that come into play are not the interaction of a photon with an individual electron, atom or molecule of a polarizer. That would give rise to fluctuations with the result that the responses in both arms cannot be consistently the same. The mechanism must therefore be more like the interaction of a photon with a macroscopic field generated by a very large amount of electrons, atoms or molecules of the polarizers. It could e.g. be an electric field. That means that the hidden variables within the polarizer can be replaced by macroscopic quantities that do not allow for fluctuations. All statistical fluctuations must therefore be due to those of the completely correlated pair of photons. We will call these macroscopic quantities, which describe the polarizers, “the fields”.

But as combinations of filter orientations (α_1, β_1) and (α_1, β_2) cannot exist simultaneously, the photon pairs involved in measuring $p(a_1 \wedge b_1)$ and $p(a_1 \wedge b_2)$ must be different (and they are!). Therefore in Equation 4 the probabilities $p(a_1 \wedge b_1)$ and $p(a_1 \wedge b_2)$ are forcedly determined by measuring quantities $a_1 b_1$ and $a_1' b_2'$ on different photon pairs, in contradiction with the assumption underlying Equation 1, where we attribute to a photon simultaneously two different linear polarizations. One of these polarizations in Equation 1 is then revealed by measurement, while the (latent) existence of the other one is not revealed but supposed to be granted by determinism. Even if it might be possible to obtain the value of the latent variable by deterministic reasoning it does not contribute to the statistics (as implied by the algebra) because it is not counted in the set-up wherein it remains latent.

The correct Bell inequality for the experiments, translating them correctly into mathematics, should be based on considering 16 numbers $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$ where ℓ is an index for the set-up used. This will then take into account that the quantities $p(a_j \wedge b_k)$ are all measured on different (sets of) photon pairs. Instead of Q as defined in Equation 1 we can define:

$$\begin{aligned}
q_1 &= a_1^{(1)}b_1^{(1)} - a_1^{(2)}b_2^{(2)} - a_2^{(3)}b_1^{(3)} - a_2^{(4)}b_2^{(4)} + a_2^{(3)} + b_2^{(2)}, \\
q_2 &= a_1^{(1)}b_1^{(1)} - a_1^{(2)}b_2^{(2)} - a_2^{(3)}b_1^{(3)} - a_2^{(4)}b_2^{(4)} + a_2^{(3)} + b_2^{(4)}, \\
q_3 &= a_1^{(1)}b_1^{(1)} - a_1^{(2)}b_2^{(2)} - a_2^{(3)}b_1^{(3)} - a_2^{(4)}b_2^{(4)} + a_2^{(4)} + b_2^{(2)}, \\
q_4 &= a_1^{(1)}b_1^{(1)} - a_1^{(2)}b_2^{(2)} - a_2^{(3)}b_1^{(3)} - a_2^{(4)}b_2^{(4)} + a_2^{(4)} + b_2^{(4)}.
\end{aligned} \tag{7}$$

Here the four alternatives correspond to four different choices for the measurements from which we use a_2 and b_2 to complete the inequality. We could also measure $p(a_2)$ and $p(b_2)$ in independent runs. This will then lead to the definition of yet another quantity \mathcal{Q} (see below). In Equation 7 there are thus $2^{16} = 65536$ combinations of the numbers $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$ to be considered.

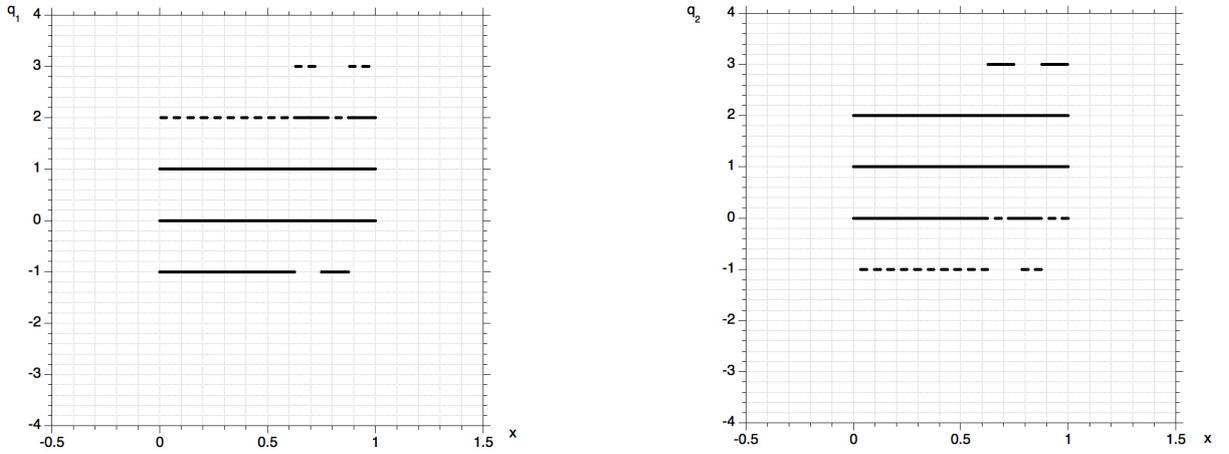


Fig. 1: The quantities q_1 and q_2 as defined in Equation 7 as a function of the variable $x = m/65536$ which labels the 65536 combinations of the 16 variables $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$. It can be clearly seen that these quantities are no longer confined to the set $\{0, 1\}$ but to an interval of length 4.

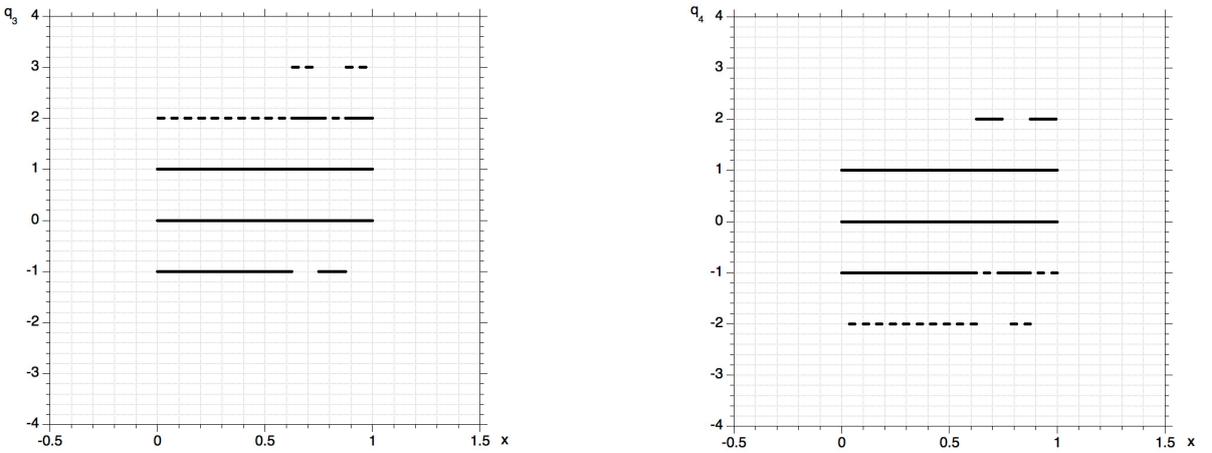


Fig. 2: The quantities q_3 and q_4 as defined in Equation 7 as a function of the variable $x = m/65536$ which labels the 65536 combinations of the 16 variables $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$. It can be clearly seen that these quantities are no longer confined to the set $\{0, 1\}$, but to an interval of length 4.

In Figures 1-2 we represent the graphs of the results of these calculations for q_j . The x -axis corresponds each time to $x = m/65536$ where m is the label $m \in [1, 65536] \cap \mathbb{N}$ of the combination. None of the values q_j is limited to the set $\{0, 1\}$. For comparison, we also define Q_ℓ by:

$$Q_\ell = a_1^{(\ell)}b_1^{(\ell)} - a_1^{(\ell)}b_2^{(\ell)} - a_2^{(\ell)}b_1^{(\ell)} - a_2^{(\ell)}b_2^{(\ell)} + a_2^{(\ell)} + b_2^{(\ell)}, \quad \forall \ell \in \{1, 2, 3, 4\}. \quad (8)$$

This corresponds to the traditional Bell inequalities for the photon pairs in the four set-ups. Figures 3-4 show the graphs of the results of the calculations for Q_ℓ on the same table of 65536 combinations. Now all values Q_ℓ neatly belong to $\{0, 1\}$.

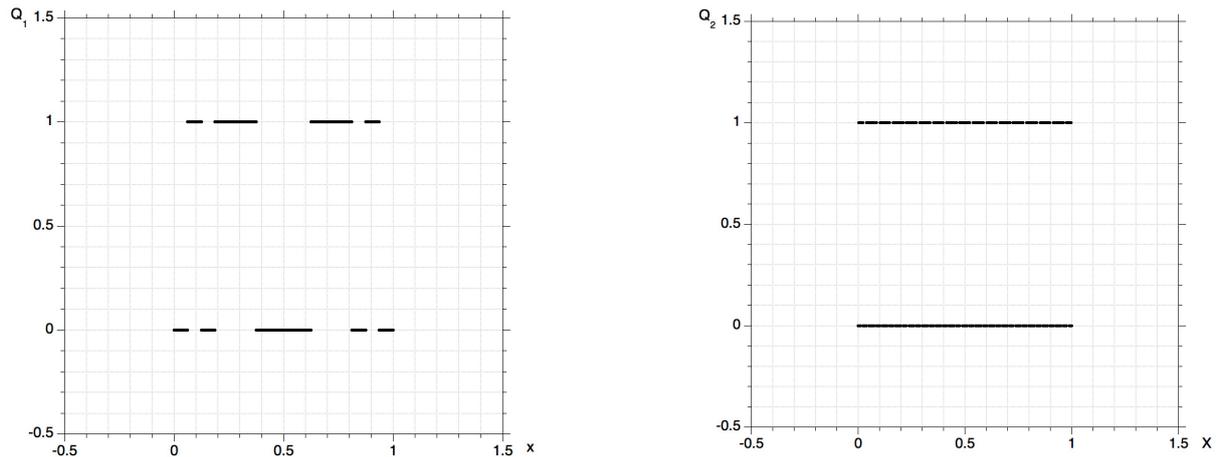


Fig. 3: The quantities Q_1 and Q_2 as defined in Equation 8 as a function of the variable $x = m/65536$ which labels the 65536 combinations of 4 variables $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$, for $\ell = 1$ and $\ell = 2$. Now the quantities remain confined to the set $\{0, 1\}$, which is the basis for the CHSH pristine Bell inequality.

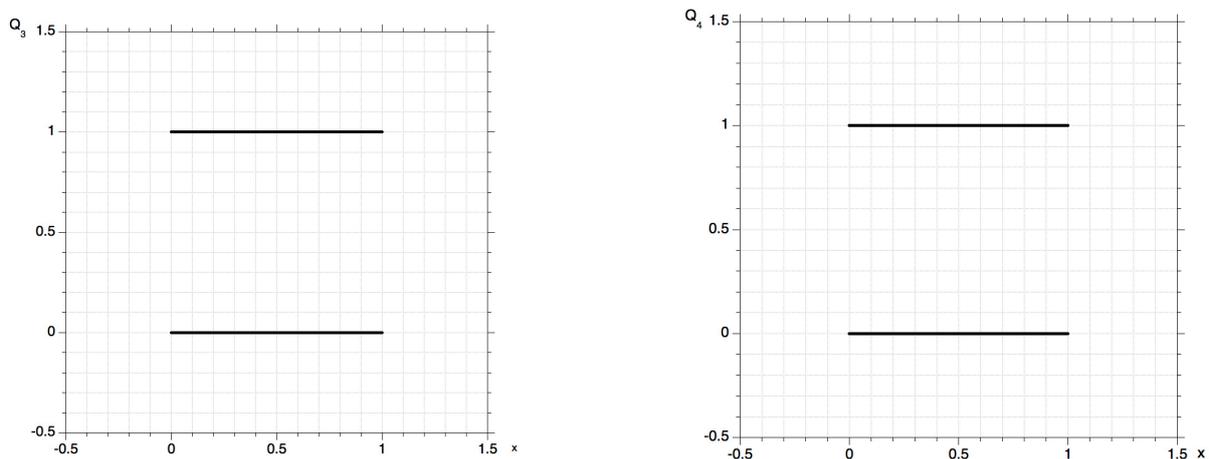


Fig. 4: The quantities Q_3 and Q_4 as defined in Equation 8 as a function of the variable $x = m/65536$ which labels the 65536 combinations of 4 variables $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}, \ell \in \{1, 2, 3, 4\}$, for $\ell = 3$ and $\ell = 4$. Now the quantities remain confined to the set $\{0, 1\}$, which is the basis for the pristine CHSH Bell inequality.

Finally, we have also investigated the case where $p(a_2)$ and $p(b_2)$ are measured in independent runs. We can simplify these calculations by leaving out the variables $a_2^{(1)}, b_2^{(1)}, a_2^{(2)}, b_1^{(2)}, a_1^{(3)}, b_2^{(3)}, a_1^{(4)}, b_1^{(4)}$ from our calculations, because they are not used anyway. We only used them above in order to be able to make a comparison with the data for Q_ℓ . We can define then a set $\mathcal{V} = V_1 \times V_2 \times V_3 \times V_4 \times V_5 \times V_6$ of hidden variables $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ that are necessary

to describe the observables that occur in the six independent runs $\eta \in [1, 6] \cap \mathbb{N}$. The four photon pairs and the two single photons within the six data collection runs have then each their hidden variables. Here λ_1 are hidden variables for $(a_1^{(1)}, b_1^{(1)})$ in set-up $\eta = 1$. We do not consider hidden variables for $(a_2^{(1)}, b_2^{(1)})$ in set-up $\eta = 1$ because they do not intervene in the measurements. In set-up $\eta = 1$, the variables $(a_2^{(1)}, b_2^{(1)})$ are completely free. We consider thus only the measured values $\mathbf{c} = (a_1^{(1)}, b_1^{(1)}, a_1^{(2)}, b_2^{(2)}, a_2^{(3)}, b_1^{(3)}, a_2^{(4)}, b_2^{(4)}, a_2^{(5)}, b_2^{(6)})$. Let us note $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ as $\boldsymbol{\lambda}$ and $d\lambda_1 d\lambda_2 d\lambda_3 d\lambda_4 d\lambda_5 d\lambda_6$ as $d\boldsymbol{\lambda}$. Here λ_η refers to the hidden variables in set-up η , where $\eta \in [1, 6] \cap \mathbb{N}$. The Bell inequality will now be based on the quantity:

$$\mathcal{Q} = a_1^{(1)}b_1^{(1)} - a_1^{(2)}b_2^{(2)} - a_2^{(3)}(\boldsymbol{\lambda})b_1^{(3)} - a_2^{(4)}b_2^{(4)} + a_2^{(5)} + b_2^{(6)}. \quad (9)$$

The result of this calculation for all 1024 combinations \mathbf{c} is shown in Figure 5. We see that the boundaries are now -3 and 3 . The values are now distributed over an interval of length 6. In all our calculations we have thus found an equality between the length of the interval of the values and the number of experiments modelled.

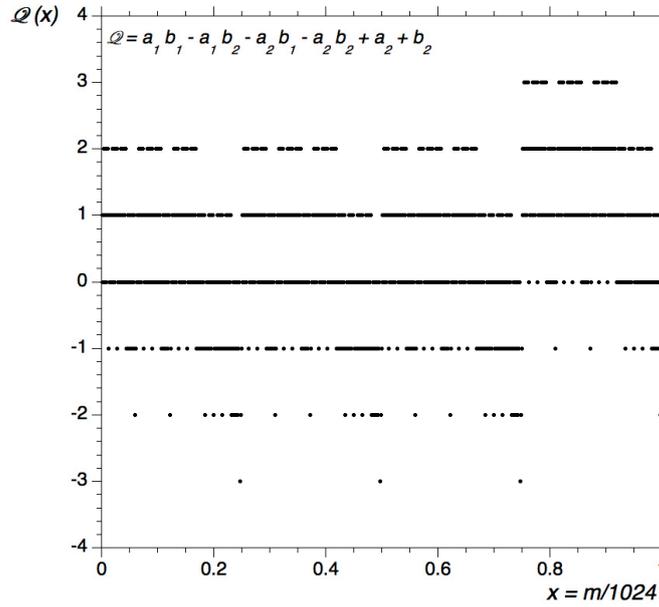


Fig. 5: The quantity \mathcal{Q} as defined in Equation 9 as a function of the variable $x = m/1024$ which labels the $2^{10} = 1024$ combinations \mathbf{c} of possible values for the 10 variables $(a_1^{(1)}, b_1^{(1)}, a_1^{(2)}, b_2^{(2)}, a_2^{(3)}, b_1^{(3)}, a_2^{(4)}, b_2^{(4)}, a_2^{(5)}, b_2^{(6)})$. The quantity \mathcal{Q} is no longer confined to the set $\{0, 1\}$ but rather to the set $[-3, 3] \cap \mathbb{N}$, which is the basis for the corrected Bell inequality in Equation 10. The length of the interval is now 6.

This will now yield really the probabilities measured in 6 runs to be inserted into the inequality:

$$-3 \leq p(a_1 \wedge b_1) - p(a_1 \wedge b_2) - p(a_2 \wedge b_1) - p(a_2 \wedge b_2) + p(a_2) + p(b_2) \leq 3. \quad (10)$$

For the full inequality we must thus integrate the inequality $-3 \leq \mathcal{Q}(\boldsymbol{\lambda}) \leq 3$ with $p(\boldsymbol{\lambda})$ over \mathcal{V} , where the boundaries have been obtained by considering all 1024 possible combinations of $\mathbf{c} = (a_1^{(1)}, b_1^{(1)}, a_1^{(2)}, b_2^{(2)}, a_2^{(3)}, b_1^{(3)}, a_2^{(4)}, b_2^{(4)}, a_2^{(5)}, b_2^{(6)})$. Of course the 6 runs are in reality obtained in one run where one performs ultra-fast switching between the 4 configurations of the two polarizers and two runs where only one parameter is measured.

Hence, the data for q_j and for \mathcal{Q} explore now a larger range than those for Q_ℓ (which correspond to the pristine Bell inequality), because they are less strongly correlated. In fact, the photon pairs in the various runs are completely independent. The observables a_j, b_k measured in different runs can therefore be different. The observables a_j, b_k for a single photon pair do not allow for such a slight of hand in the thought experiment leading to Equation 4, where we imagine that all the information is harvested at once. Hence the traditional Bell inequalities based on Equations 1 or 8 are correct for a single photon pair but it is impossible to test them by real-life experiments. Inequalities that can be tested by experiments must be based on a derivation starting from Equation 7 or Equation 9. By just using the simple argument used in the traditional derivation we can now no longer prove that the boundaries of the inequalities that one can derive from Equation 7 or Equation 9 are as stringent as has been claimed on the basis of Eq 4. That

simple argument stated that we can prove the Bell inequality by just using the results $0 \leq Q_\ell(\lambda) \leq 1$ and integrate it using $\int_V p(\lambda) d\lambda = 1$, where $\forall \lambda : p(\lambda) \geq 0$. We have then $0 \leq \int_V Q_\ell(\lambda) p(\lambda) d\lambda \leq \int_V p(\lambda) d\lambda = 1$. Using the same argument, the boundaries we can establish this way for the inequality are now no longer 0 and 1. We can no longer prove the CHSH inequality. These results indicate that the proof of the Bell inequality which has been applied to the experiments is wrong.

3 Larsson's rebuttal

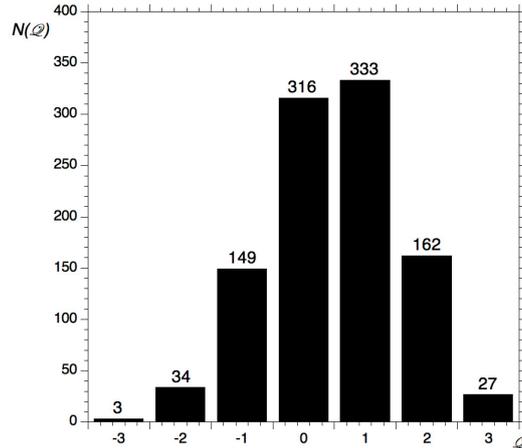


Fig. 6: Histogram of the numbers $N(\mathcal{Q})$ of times each value of the quantity $\mathcal{Q}(\mathbf{c})$ defined in Equation 9 is obtained in the data for all combinations \mathbf{c} reported in Figure 5. The theoretical frequencies are then $\nu(\mathcal{Q}) = N(\mathcal{Q})/1024$. If we determined these frequencies experimentally by coin tossing for all variables $a_j^{(\eta)}$ and $b_k^{(\eta)}$, the measured frequencies may deviate from these theoretical frequencies, due to statistical fluctuations. Certain values of \mathbf{c} may have been under-sampled, while other values might have been over-sampled. But for good enough statistics, the measured frequencies will converge to the theoretical frequencies shown in this histogram. Tossing the coin more times will not change the theoretical interval $[-3, 3] \cap \mathbb{Z}$. The data shown here have no further purpose than illustrating that our argument in the main text has nothing to do with statistical fluctuations. They are of no further use because to calculate the experimental probabilities one must use $p(\boldsymbol{\lambda})$, rather than rather than $p(\mathbf{c})$ or $p(\mathcal{Q})$ as determined by coin tossing and the integral $\int Q(\mathbf{c}(\boldsymbol{\lambda})) p(\boldsymbol{\lambda}) d\boldsymbol{\lambda}$ relies on the use of $p(\boldsymbol{\lambda})$ rather than calculated quantities like $p(\mathcal{Q})$ or $p(\mathbf{c})$.

This problem that in each set-up one has pairs of photons that are different from those in the other set-ups has already been discussed by Larsson [7] who indeed noted that in an experiment $a_1 b_1$ and $a_1 b_2$ are corresponding in general to different values of λ , e.g. $a_1(\lambda_1) b_1(\lambda_1)$ and $a_1(\lambda_2) b_2(\lambda_2)$. His equation 15 contains the same expressions as our Eqs. 7, 9. He then argued that there can be individual violations of the inequality but not in the mean for a large quantity of data. For sufficiently good statistics the individual errors would average out and converge to the mean. This argument would be an experimental loophole that could be solved by improving the statistics. This would then debunk our objection.

But this is a very poor and misleading explanation of the state of affairs. The error at issue is not a statistical error caused by statistical fluctuations but a systematic error produced by a wrong mathematical description of the experiment. It misrepresents the systematic difference between a correct and a wrong model as a statistical error within the framework of the wrong model. We can illustrate this with the aid of Figure 6 which is a histogram for the occurrence of the values $\mathcal{Q}(\mathbf{c}) \in [-3, 3] \cap \mathbb{Z}$ for all possible combinations \mathbf{c} as displayed in Figure 5. The frequencies $\nu(\mathcal{Q}) = N(\mathcal{Q})/1024$ in the histogram of Figure 6 are the exact theoretical frequencies the experimentally measured frequencies will converge to when the statistics are good enough (if we attribute the same probability to each configuration \mathbf{c}). The values of \mathcal{Q} and their frequencies in the histogram cannot be affected by improving the statistics because they are theoretical and exact. The differences we find between \mathcal{Q} and Q_ℓ cannot be belittled as experimental errors due to statistical fluctuations. The only error that exists resides in the wrong choice of the model used to derive Q_ℓ , which just does not describe the physical reality correctly. Both \mathcal{Q} and Q_ℓ are the correct theoretical values for their respective models. Also differences between the boundaries for q_j or \mathcal{Q} and those for Q_ℓ cannot be attributed to statistical fluctuations. The boundaries for q_j or \mathcal{Q} are the exact theoretical values within the framework of a correct

read-out	set-up 1 $+a_1b_1$	set-up 2 $-a_1b_2$	set-up 3 $-a_2b_1$	set-up 4 $-a_2b_2$	set-up 5 $+a_2$	set-up 6 $+b_2$
1	$t_{1,1}$					
2	$t_{2,1}$					
\vdots	\vdots					
j	$t_{j,1}$					
\vdots	\vdots					
u	$t_{u,1}$					
$u+1$		$t_{u+1,2}$				
$u+2$		$t_{u+2,2}$				
\vdots		\vdots				
$u+j$		$t_{u+j,2}$				
\vdots		\vdots				
$2u$		$t_{2u,2}$				
$2u+1$			$t_{2u+1,3}$			
$2u+2$			$t_{2u+2,3}$			
\vdots			\vdots			
$2u+j$			$t_{2u+j,3}$			
\vdots			\vdots			
$3u$			$t_{3u,3}$			
$3u+1$				$t_{3u+1,4}$		
$3u+2$				$t_{3u+2,4}$		
\vdots				\vdots		
$3u+j$				$t_{3u+j,4}$		
\vdots				\vdots		
$4u$				$t_{4u,4}$		
$4u+1$					$t_{4u+1,5}$	
$4u+2$					$t_{4u+2,5}$	
\vdots					\vdots	
$4u+j$					$t_{4u+j,5}$	
\vdots					\vdots	
$5u$					$t_{5u,5}$	
$5u+1$						$t_{5u+1,6}$
$5u+2$						$t_{5u+2,6}$
\vdots						\vdots
$5u+j$						$t_{5u+j,6}$
\vdots						\vdots
$6u$						$t_{6u,6}$

Table 1: Overview of the data collection. The columns show the six set-ups. For each set-up we have listed the terms $a_j b_k$, a_j or b_k for which we want to measure the average value in order to determine the probabilities $p(a_j \wedge b_k)$, $p(a_j)$ or $p(b_k)$. For each set-up there are u registrations, such that in total there are $6u$ registrations.

description of the experiment. The theoretical values in Fig. 6 are like the theoretical values of the Gaussian function $g \in F(\mathbb{R}, \mathbb{R}) : x \rightarrow g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$. The numbers $x \in \mathbb{R}$ and the values $g(x)$ are not empirical statistical fluctuations with respect to the mean μ , that could be eliminated by improving the statistics, but theoretical values.

Statistical fluctuations are a subsidiary problem we do not have to treat in our discussion, because we have adopted the strategy to inject the theoretical values into the inequalities. Behind its confusing formulation, Larsson's argument contains something far more essential than statistical fluctuations. The correct way to explain Larsson's argument is the following. It corresponds to the comparison he makes between his two equations 17 and 18. We consider the experiment and its hidden variables as a purely theoretical thought experiment (*Gedankenexperiment*). We do not perform the experiments, we just do the deterministic calculations that correspond to them. We can e.g. imagine that

we calculate the results of the experiment by a Monte Carlo simulation. For each simulated photon pair we calculate the four parameters a_1, b_1, a_2, b_2 from the hidden variables. We can register them on a first spreadsheet, which we will call the pristine spreadsheet. But we also construct a second spreadsheet. On this second spreadsheet we will simulate the six runs with the six set-ups in the real-life experiment. We do not register during the first run of u measurements in set-up 1 the values obtained for a_2, b_2 in the calculations that we will use to obtain $p(a_1 \wedge b_1)$. We only register the values for $a_1 b_1$. We wipe out all other numbers, because they do not play a rôle in set-up 1.

We can carry out an analogous trick of selective registration for the simulation of the result in each of the six set-ups. For each set-up we register the truncated results for the emission of $u \in \mathbb{N}$ photon pairs. We must select for u a very large number, in order to obtain good statistics. But as we are performing a thought experiment, we can take u as large as we like. It just plays a theoretical rôle in a proof, similar to the one of u in the following definition of a limit:

$$\lim_{x \rightarrow \infty} f(x) = a \quad \Leftrightarrow \quad (\forall \varepsilon > 0)(\exists u \in \mathbb{N})(x > u \Rightarrow |f(x) - a| < \varepsilon), \quad (11)$$

because we have pointed out that we do not bother about statistics in our theoretical approach. This way each of the six measurement is in reality a repetition of a same imaginary experiment where we determine simultaneously all parameters a_1, b_1, a_2, b_2 , but do not register all the results.

In Table 1 we present an overview of the data collection on the second spreadsheet of the Monte-Carlo calculation. We have noted all the data values registered in this data collection under the form $t_{m,n}$ with a line index m for the sequence number of the registration and a column index n for the number of the set-up. We see that e.g. on lines with sequence numbers $m \in [1, u] \cap \mathbb{N}$ the quantities $t_{m,n}$ for $n \neq 1$ have not been registered. In general, only the quantities $t_{(n-1)u+j,n}$ where $j \in [1, u] \cap \mathbb{N}$ have been registered. We can qualify the experiment simulated by the Monte-Carlo method, and whereby we measure all quantities a_1, b_1, a_2, b_2 simultaneously on each photon pair the imaginary experiment, because it is impossible in real life to measure all quantities a_1, b_1, a_2, b_2 simultaneously on a photon pair. If we do not wipe out any of the results of the calculation, such that on each data line m we would register 6 numbers instead of 1, then this imaginary experiment must satisfy the Bell inequality Eq. 4 because the inequality corresponds to the exact translation into mathematics of this imaginary experiment where all the quantities a_1, b_1, a_2, b_2 are supposed to be fully determined by the hidden variables. The imaginary experiment is simulated by the pristine spreadsheet. But after wiping out all the data as specified above and visualized in Table 1 the second spreadsheet of the Monte-Carlo calculation simulates the real-life experiment.

For u sufficiently large the averages of the data $t_{(n-1)u+j,n}$, where n is fixed and $j \in [1, u] \cap \mathbb{N}$ will converge to the same values as the averages of $t_{m,n}$, where n is fixed and $m \in [1, 6u] \cap \mathbb{N}$. Of course only the data for $m = (n-1)u + j$ with $j \in [1, u] \cap \mathbb{N}$ have been registered on the second spreadsheet but all data $t_{m,n}$ are available on the pristine spreadsheet of the imaginary experiment. The same is true for the averages of the data $t_{j,n}$, where n is fixed and $j \in [1, u] \cap \mathbb{N}$. Again, of the latter data only the data $t_{j,1}$ have been registered on the second spreadsheet but the other data are available on the pristine spreadsheet. Combining the two pieces of information, we can conclude that the averages for $(m, n) = ((n-1)u + j, n)$ and for $(m, n) = (j, n)$, where j runs over $[1, u] \cap \mathbb{N}$ are equivalent. This has been called the fair-sampling argument.

Consider now the block structure of the second spreadsheet in Table 1. The equivalence we proved means that we can move all diagonal blocs (n, n) to the positions $(1, n)$, not for the detailed information inside the blocks but for their averages. This means that the six averages we obtain from the six set-ups are the same as we would have obtained in the imaginary experiment by registering all the data for the measurements $m \in [1, u] \cap \mathbb{N}$ (or alternatively the measurements $m \in [1, 6u] \cap \mathbb{N}$ for better statistics) without wiping them out. This proves that the results obtained from the real-life experiment with the six set-ups are theoretically equivalent to those that we would have obtained from the imaginary experiment whereby it would be physically possible to measure always the six values $t_{m,n}$ with $n \in [1, 6] \cap \mathbb{N}$ simultaneously. The real-life experiment is correctly translated into mathematics by Eq. 10. The imaginary experiment is correctly translated into mathematics by Eq. 2. As the averages we obtain from the real-life experiment and from the imaginary experiment have now been shown to be equivalent, the theoretical averages obtained from the six experiments can be plugged into the Bell inequality based on the wrong formulation Eq. 2 and will satisfy it. This description correctly resumes Larsson's rebuttal. We may note that this rebuttal admits that the original proof of the CHSH inequality was wrong and that the real proof is no longer as simple as has been claimed. This should prompt us to be more cautious about the proofs that have been forwarded to us for years as unassailable.

4 The normalization error

Let us now show why this rebuttal is flawed. The motivation for formulating the Bell inequality was to perform a simultaneous measurement of observables for correlated photons. It wanted to test Einstein's original idea that the components L_x and L_y of the angular momentum could be measured simultaneously for a single particle, despite the fact that the corresponding operators \hat{L}_x and \hat{L}_y do not commute. We have discussed this issue in some more detail

in the Appendix 1. Einstein's trick was to make the reputedly impossible simultaneous measurement of L_x and L_y by performing simultaneously a measurement with two different experimental devices on two identical particles. This kept alive the idea that the information determining the outcome in the two experimental devices would be simultaneously inscribed in the particles, e.g. L_x and L_y would be simultaneously defined for a single particle, just like they would be simultaneously defined for a macroscopic object, like e.g. a spinning top.

In this discussion, all the information was locked up in the particles. But in the photon correlation experiment the full information is no longer confined to the correlated pair of identical photons. A part of the information is also stored in each of the polarizers. The full information is spread out over different locations in the set-up that can be very remote one from another. We are therefore not at all testing Einstein's original idea anymore. And in our attempts to make sense of the experimental results we have even ended up with the condition that all the macroscopic fields must be simultaneously inscribed in the hidden variables for a set-up (which corresponds to a single choice for the orientation of each of the filters). This can definitely not be the case, because we are now no longer discussing quantum properties of particles, but macroscopic properties of set-ups. There is no simultaneous information about several set-ups available in a single set-up.

The problem is that in the thought experiment we must consider for both polarizers two possible orientations simultaneously. Theoretically, there is absolutely no obstacle to the simultaneous definition of all the information about the hidden variables involved in the two possible orientations of both filters in the imaginary thought experiment, but experimentally there is.

Larsson's argument considers the statistical averages obtained from the exact model description in Eq. 9 and puts them into correspondence with the statistical averages that occur in the wrong model description of Eq. 2. The discussion is kept confined inside the context of the wrong model.

The Bell inequality implies that the hidden variables for the four combinations of polarizer fields must all be contained in the complete set V of all relevant hidden variables. We must construct this set V . Let us note the sets of hidden variables when the polarizer orientations are α_j as A_j , those when the polarizer orientations are β_k as B_k , and those for the photon pairs C . Then the set of hidden variables for the set-up 1 will be $V_1 = A_1 \times B_1 \times C$. This will be the set of hidden variables for the real-life measurement of $p(a_1 \wedge b_1)$. The hidden variables for a_2 and b_2 do not play a rôle in the outcome of this experiment. We must now define the probability distribution on the set V_1 . This description will not contain a single clue about the fields that will be present in the measurements involving a_2 and b_2 . Nature can absolutely not know that we intend to perform experiments involving a_2 and b_2 later on. The joint probability distribution over the set V_1 of hidden variables for the photon pairs and the two fields involved in measuring a_1 and b_1 (*and nothing more*) will yield an accurate description of what really happens in nature. This joint distribution will be the exact probability distribution that determines the outcome for $p(a_1 \wedge b_1)$.

If we now want to draw also the hidden variables involved in the measurement of a_2 and b_2 into the description such that we can calculate all the probabilities that occur in the Bell inequality for the imaginary experiment simultaneously, we will have to extend the set of all possible hidden variables such as to include also those for the fields involved in the measurements of a_2 and b_2 . This will now be the set we already called V . We will describe this construction in Appendix 2. The result is that the set V must be larger than the set $V_1 \subsetneq V$.

The extension from V_1 to V will have to be accompanied by a redefinition of the probability distribution which will now have a definition domain V that is larger than the set V_1 . And this redefinition will require a change of the normalization of the probability distribution, because the integral of the probability distribution will now have to be normalized to 1 over V while it was previously normalized to 1 over V_1 . This redefinition will change the integral of the probability distribution over $V_1 \subsetneq V$, which will now acquire a value smaller than one. This implies that the normalizations of the probability distributions for the calculations of $p(a_j \wedge b_k)$ in real life and of those for the calculations in the imaginary experiment are different, because in real life the probability distributions in each set-up are normalized to 1. We can therefore not plug the real-life values into the Bell-inequality for the imaginary experiment. The values we would have to plug into the inequality must be smaller. That is the sting inherent to the need for a common description that does not correspond to real life.

Note that this remark would not have surfaced if we had defined the set of hidden variables as $A_1 \times A_2 \times B_1 \times B_2 \times C$ because the probability distributions over all the sets that occur in the Cartesian product could be normalized to 1, such that the integrals over $A_1 \times A_2 \times B_1 \times B_2 \times C$ and over $A_1 \times B_1 \times C$ would be both 1. But as explained in Appendix 2, this does not take into account the correlations and the mutually exclusive character of A_1 and A_2 and of B_1 and B_2 . The formulation of the definition domain is too coarse. To put it simple, in the real-life experiments of $p(a_j \wedge b_k)$ we just define a probability distribution on a subset of $A_j \times B_k$. While to define a common distribution function whereby a_1, b_1, a_2, b_2 are all defined for a same photon pair, we must combine the four definition domains of the four set-ups in an all-encompassing common definition domain. For didactical purposes one often represents probabilities on probability trees. Each set-up corresponds to a tree. But if we want to consider the probabilities from four different set-ups simultaneously we must unite the four trees within a single tree by giving the four trees a common root. And the probabilities must then be normalized on the larger tree instead of on one of the four original trees. One may now conceive the idea that one can avoid this by presenting the whole situation as a single tree with a definition domain

$A_1 \times A_2 \times B_1 \times B_2 \times C$. But the definition domain then still remains a disjoint union of four sets. The need to define a common probability distribution and to translate the physics correctly into mathematics is the whole essence of what has been forgotten in the wrong modelling of the CHSH inequality.

5 Epilogue

In a first step Larsson tacitly replaced the correct model for the real-life experiments by the wrong model of the imaginary experiment. We can see that in a second step his argument about the averages holds as long as one stays inside the context of the imaginary experiment (as we did in the Monte Carlo simulation). However, his argument does not survive the confrontation of the imaginary experiment with the real-life experiments, which also has to be performed in order to validate his first tacit step. Hence, when we claimed that the experiment visualized in Table 1 simulated the real-life experiment, this was actually a falsehood because the probabilities in the six real-life set-ups (with their individual probability distributions) are normalized differently than the six probabilities in the imaginary experiment (which is governed by a common probability distribution). The contexts are different and the context of the imaginary experiment does not translate the real-life contexts correctly into mathematics.

Of course Larsson did not pass his first step deliberately under silence. His first step consisted in misinterpreting the systematic errors induced by the wrong modelling as statistical fluctuations of the wrong model (from which the CHSH inequality is derived), which he continued to consider as correct. This way he became trapped within the wrong model. And then there were just more nasty issues involved in sticking to the wrong model than he could treat after being trapped. Therefore Larsson's argument is a fallacy that manifests itself by the fact that his imaginary experiment has a hidden normalization problem. The reason for this is that the CHSH inequality requires a global context that allows to consider all probabilities simultaneously such that we can insert them into the inequality, while the real-life experiments are taking place in a restricted context, which ignores our intentions to insert the results into a Bell inequality later on.

Note that when we make the individual experiments in real life, we normalize the data from each run to a same amount of u emitted photon pairs. This normalization problem is different from the normalization problem we are describing here to counter Larsson's rebuttal, because in his proof the data we did not register are in principle representative of those we did register and vice versa, such that his proof implies carrying out an identical normalization to a same number u of emitted pairs of photons, as is clearly visible from Table 1. That is why his proof looked correct.

Hence, contrary to what has been claimed by Larsson [7] his work cannot save the comparison between the theoretically and experimentally obtained values in the inequality. Our objection does not address an experimental loophole due to poor statistical accuracy. Remember why we proposed to inject the correct theoretical values for the probabilities into the inequalities. We adopted this strategy because when we proceed this way we do not have to bother about statistical-accuracy loopholes of the kind evoked by Larsson. It warrants us a rock-solid immunity against becoming bamboozled by arguments about statistical fluctuations. We are then avoiding such issues all together and making sure that we focus our attention only onto possible logical flaws in the theory.

6 Conclusion

For a given value of ℓ , the quantities $a_1^{(\ell)}, a_2^{(\ell)}, b_1^{(\ell)}, b_2^{(\ell)}$ are just the data for a single photon pair, which according to classical thinking could be obtained by deterministic reasoning from the hidden variables and/or revealed by measurement. Hence, the traditional Bell inequalities based on Equations 1 or 8 are (in theory) correct for a single photon pair but it is impossible to test them by experiments, let alone that they could be violated by experiments. The inequality that can be tested by experiments is e.g. given in Equation 10. Our results show clearly that a wrong Bell inequality has been applied to the experiments. Larsson tried to repair for this by averaging, but he missed the point that the wrong formulation of the physics leads also to a wrong normalization of the probabilities.

We may note that in some photon correlation experiments the inequalities used are different from the CHSH inequality because they are based on a different type of polarizers (beam-splitting polarizers leading to experimental outcomes $+1$ and -1 rather than 0 and 1). One uses e.g. so-called CH and CH-Eberhard Bell inequalities. There exists a whole plethora of Bell-type inequalities, but in the derivation of all these inequalities the same logical error persists. The violation of such Bell inequalities is thus not an experimental proof that classical thinking would have been defeated and that locality or realism would have to be abandoned. There is no evidence based on Bell-type experiments in favour of the existence of entanglement.

That the application of the CHSH Bell inequality to the photon correlation experiments was not correct had already been pointed out by many authors, starting with Kupczynski back in 1987 (see [8,9] and the many references therein). Many authors have stressed the importance of contextuality and evoked the lack of proof for the existence of a common probability distribution, as stipulated by integrating Eq. 2 over $p(\lambda)$. The latter objection may have

looked to be mere mathematical fault-finding. Our calculations are actually introducing constructions of common probability distributions. But the common probability distribution that underlies the CHSH inequality is different from the real-life probability distributions which define the measured probabilities which we want to inject into the CHSH inequality. Of course everything is rendered complicated by the multidimensional character of the sets of hidden variables, which we even do not know to specify in full detail, because we have no inkling of a physical mechanism that supposedly would explain the experiment classically, as illustrated by the wrong calculation that led to Eq. 6.

It is well-known that probability calculus is teeming with pitfalls and paradoxes such that questioning locality by relying on probability calculus was always a kind of reckless. Locality is nothing less than a cornerstone of the theory of relativity and questioning it should preferably be based on methods that are far less treacherous and error-prone.

Appendix 1: commutation relations

In quantum mechanics (QM) we learn that the operators

$$\hat{L}_z = \frac{\hbar}{i} \left[x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right] (\text{cycl.}), \quad (12)$$

are the operators for the components of the angular momentum $\mathbf{L} = (L_x, L_y, L_z)$. They obey the commutation relations:

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z (\text{cycl.}). \quad (13)$$

From this result, Bohr drew the conclusion that two components L_x and L_y of the angular momentum \mathbf{L} could not *exist* simultaneously. This is in the least highly counterintuitive. Heisenberg was less radical and in his discussions with Bohr, he defended the viewpoint that the commutation relation would rather imply that L_x and L_y cannot be *measured* simultaneously. Einstein did not accept these conclusions and in order to refute them he considered two identical particles that would travel in opposite directions. On one of them one would measure L_x and on the other L_y . Some clarification is needed here. First of all, there is also a commutation relation for \hat{p}_x and \hat{x} :

$$[\hat{x}, \hat{p}_x] = i\hbar \mathbb{1}. \quad (14)$$

But in quantum field theory Dirac replaced this by:

$$[\phi(t, \mathbf{r}_1), \chi(t, \mathbf{r}_2)] = i\hbar \delta(\mathbf{r}_1 - \mathbf{r}_2). \quad (15)$$

Here ϕ and χ are operators for conjugated variables. This implies that the position operator in one place \mathbf{r}_1 and the momentum operator in a different place \mathbf{r}_2 commute, but that position and momentum operators in the same place $\mathbf{r}_1 = \mathbf{r}_2$ do not commute. This remark actually undermines Einstein's strategy to rebut Bohr's conclusion based on the EPR argument, because it makes him miss the target (of addressing non-commuting variables), as we can actually expect similarly:

$$[\hat{L}_x(\mathbf{r}_1), \hat{L}_y(\mathbf{r}_2)] = i\hbar \hat{L}_z \delta(\mathbf{r}_1 - \mathbf{r}_2) (\text{cycl.}), \quad (16)$$

such that it is indeed possible to measure the two quantities simultaneously on two correlated particles. But there is another confusion here, because the real meaning of \hat{L}_x is that it is the operator for the angular momentum \mathbf{L} when it is aligned with the x -axis [10]. The angular-momentum commutation relations are then no longer mysterious. In fact, such commutation relations occur trivially in the Lie algebra of every non-abelian Lie group and are entirely classical. The operators \hat{L}_x and \hat{L}_y simply do not commute because the angular momentum cannot be simultaneously aligned with the y -axis, when it is already aligned with the x -axis [10]. The situation is completely analogous for electron spin operators.

Let us consider only polarizers whose planes are parallel to a given plane. We can then characterize their orientation by an angle in this plane. We have noted these orientations by the angles α_j for the filter A and β_k for filter B . We do not know what kind of commutation relations one should define for polarization operators. But conceptually we must take into account that we cannot orient simultaneously a filter A along α_2 in \mathbf{r}_1 when it is already aligned along α_1 in \mathbf{r}_1 . However, it is possible to align a filter B in a position \mathbf{r}_2 along β_k when A is already aligned along α_j in \mathbf{r}_1 . In other words we cannot measure simultaneously the polarizations along α_1 and α_2 in \mathbf{r}_1 or along β_1 and β_2 in \mathbf{r}_2 but we can measure simultaneously any combination of α_j in \mathbf{r}_1 and β_k in \mathbf{r}_2 . That is not only true in QM, it is also true in classical physics.

Appendix 2 - Common probability distribution for the CHSH inequality

We will only describe the part of the construction that allows to consider the four probabilities $p(a_j \wedge b_k)$. We will make the construction in several steps. We will treat the sets of hidden variables in terms of positions within the filters, despite the considerations we formulated in terms of fields in Remark 1. The idea behind this is that we want to depend as little as possible on assumptions. E.g. the magnitude of a field might vary while only its direction has to remain fixed. All relevant hidden variables will be functions of the positions. This turns the sets of hidden variables into geometrical sets. We cannot have the product $A_1 \times A_2$ within the expression of the definition domain as this would imply that we can also measure $p(a_1 \wedge a_2)$, while the two sets never acquire physical reality simultaneously. Also $A_1 \cup A_2$ does not make sense because $A_1 \cap A_2$ contains position coordinates $\mathbf{r}_1 \in A_1$ and A_2 position coordinates $\mathbf{r}_2 \in A_2$ whereby $\mathbf{r}_1 = \mathbf{r}_2$. The sets A_1 and A_2 should have no real intersection because the two filter orientations they correspond to cannot exist simultaneously. We therefore must perform the simultaneous description of A_1 and A_2 in \mathbb{R}^6 .

This repeats in a sense Einstein's trick of identical duplication. We replace the pair of identical photons by two identical pairs of identical photons, whereby one pair travels in a first universe $E_1 = \mathbb{R}^3$ which contains the polarizers with orientations (α_1, β_1) and the other pair in a second universe $E_2 = \mathbb{R}^3$ which contains the polarizers with orientations (α_2, β_2) . We can consider E_1 and E_2 as simultaneously embedded in \mathbb{R}^6 under the form $\overline{E}_1 \cup \overline{E}_2 \subset \mathbb{R}^6$, where

$$\begin{aligned}\overline{E}_1 &= \{(\mathbf{r}_1, \mathbf{0}) \in \mathbb{R}^6 \mid \mathbf{r}_1 \in E_1\}, \\ \overline{E}_2 &= \{(\mathbf{0}, \mathbf{r}_2) \in \mathbb{R}^6 \mid \mathbf{r}_2 \in E_2\}.\end{aligned}\tag{17}$$

This way $E_1 \sim \overline{E}_1$ and $E_2 \sim \overline{E}_2$ become equivalent to orthogonal subspaces of \mathbb{R}^6 . We define the isomorphism: $f_1 \in F(E_1, \overline{E}_1)$ by $\forall \mathbf{r} \in E_1 : f_1(\mathbf{r}) = (\mathbf{r}, \mathbf{0})$. We note this isomorphism as $\mathbf{r} \sim (\mathbf{r}, \mathbf{0})$. We also define the isomorphism: $f_2 \in F(E_2, \overline{E}_2)$ by $\forall \mathbf{r} \in E_2 : f_2(\mathbf{r}) = (\mathbf{0}, \mathbf{r})$. We note this isomorphism also as $\mathbf{r} \sim (\mathbf{0}, \mathbf{r})$. Finally we define an isomorphism $f \in F(\overline{E}_1, \overline{E}_2)$ by $\forall (\mathbf{r}, \mathbf{0}) \in \overline{E}_1 : f(\mathbf{r}, \mathbf{0}) = (\mathbf{0}, \mathbf{r}) \in \overline{E}_2$. We also write this isomorphism as $(\mathbf{r}, \mathbf{0}) \sim (\mathbf{0}, \mathbf{r})$. Under this isomorphism the trajectories of the first pair of photons are always mapped on the trajectories of the second pair. We have then an isomorphism f between $\overline{E}_1 \sim E_1$ and between $\overline{E}_2 \sim E_2$. The sets \overline{E}_1 and \overline{E}_2 are disjoint. We see that by construction $\overline{E}_1 \perp \overline{E}_2$, while we could have qualified E_1 and E_2 colloquially as "parallel" universes. The experiment has now four arms, with the second pair of arms perpendicular to the first one. Despite the fact that E_1 and E_2 are both equivalent to \mathbb{R}^3 , we have $E_1 \sim E_2$ and $E_1 \perp E_2$ rather than $E_1 = E_2$.

We can then describe the probabilities $p(a_1 \wedge b_1)$ and $p(a_2 \wedge b_2)$ simultaneously on $((A_1 \times B_1) \times (A'_2 \times B'_2)) \times C$, where we use the accents to note subsets of E_2 , such that $A_1 \cap A'_2 = \emptyset$ and $B_1 \cap B'_2 = \emptyset$. But this is not yet a completely accurate definition of the definition domain of the probability distribution, because the photons are impinging onto the polarizers in positions that are symmetrical with respect to the origin of \mathbb{R}^3 . We must take into account this correlation.

We could take these correlations also into account by making them part of the definition of the probability distribution, by writing it sloppily, in a notation where the accents are tacitly implied, as:

$$p(\mathbf{r}_{A,1})p(\mathbf{r}_{A,2})p(\mathbf{r}_{B,1})p(\mathbf{r}_{B,2}) \left[\sum_{j=1}^2 \delta(\mathbf{r}_{A,j} + \mathbf{r}_{B,1}) \delta(\mathbf{r}_{A,3-j} + \mathbf{r}_{B,2}) \right],\tag{18}$$

where $\mathbf{r}_{A,j}$ and $\mathbf{r}_{B,k}$ are noting position vectors in A_j and B_k . Note that the Dirac measures can be represented as limits of test functions which are valid probability distribution functions. But all this is of course not rigorous, because without the accents it represents the sets A_1 and A_2 as coexisting simultaneously in \mathbb{R}^3 . The same is true for the sets B_1 and B_2 . Furthermore, the correlations are reducing the true sizes of the definition domains. Such reductions are hidden by identifying the definition domains with $A_j \times B_k$ and relegating the correlations to the definition of the probability distribution. It impedes us to appreciate the true sizes of the definition domains. It becomes then less obvious to show that the size of definition domain V of the common probability distribution is four times larger than the size of the definition domain for a single set-up $V_1 \subsetneq V$, as we will do. It is therefore better to treat the correlations by incorporating them in the definition of the sets. Therefore, instead of $A_1 \times B_1$ and $A'_2 \times B'_2$, we define the sets:

$$\begin{aligned}E_{11} &= \{(\mathbf{r}, \mathbf{r}') \in A_1 \times B_1 \mid \mathbf{r} + \mathbf{r}' = \mathbf{0}\} \\ E_{22} &= \{(\mathbf{r}, \mathbf{r}') \in A'_2 \times B'_2 \mid \mathbf{r} + \mathbf{r}' = \mathbf{0}\} \\ E_{11,22} &= \{(\boldsymbol{\rho}, \boldsymbol{\rho}') \in E_{11} \times E_{22} \mid \boldsymbol{\rho} \sim \boldsymbol{\rho}'\}.\end{aligned}\tag{19}$$

This will allow us to define simultaneously $p(a_1 \wedge b_1)$ and $p(a_2 \wedge b_2)$. We must also define simultaneously $p(a_1 \wedge b_2)$ and $p(a_2 \wedge b_1)$. We must therefore make a similar construction in a different space \mathbb{R}^6 :

$$\begin{aligned} E_{12} &= \{(\mathbf{r}, \mathbf{r}') \in A_1'' \times B_2'' \mid \mathbf{r} + \mathbf{r}' = \mathbf{0}\} \\ E_{21} &= \{(\mathbf{r}, \mathbf{r}') \in A_2''' \times B_1''' \mid \mathbf{r} + \mathbf{r}' = \mathbf{0}\} \\ E_{12,21} &= \{(\boldsymbol{\rho}, \boldsymbol{\rho}') \in E_{12} \times E_{21} \mid \boldsymbol{\rho} \sim \boldsymbol{\rho}'\}. \end{aligned} \quad (20)$$

To define simultaneously the four probabilities $p(a_j \wedge b_k)$ we must then combine the two \mathbb{R}^6 representations as orthogonal spaces in \mathbb{R}^{12} , the same way as we combined two \mathbb{R}^3 representations as orthogonal spaces in \mathbb{R}^6 . The set of hidden variables is then:

$$E \times C = \{(\mathbf{R}, \mathbf{R}') \in E_{11,22} \times E_{12,21} \mid \mathbf{R} \sim \mathbf{R}'\} \times C. \quad (21)$$

We end up in \mathbb{R}^{12} because we must reconcile four mutually incompatible set-ups in \mathbb{R}^3 wherein we measure the same photon pair, as requested by the definition of the CHSH Bell inequality, where all quantities are simultaneously defined for the same photon pair. There is also an ambiguity in the notation (a_1, b_1, a_2, b_2) because it does not specify if in the real world this is obtained as $((a_1, b_1), (a_2, b_2))$ or as $((a_1, b_2), (a_2, b_1))$. We can always use C , because when we know one photon, we know all of them as the four photons are identical. Moreover C does not contain position coordinates but information about the polarizations of the identical photons. Up to isomorphisms the definition domain is a subset of $(A_1 \cup A_2') \times (B_1 \cup B_2') \times C$. This subset can serve as a set of hidden variables because the other variables, like the fields, are defined on A_1, A_2', B_1, B_2' .

Of course, when the fields are constant, a much more simple construction of the set of hidden variables is in principle possible, whereby we no longer specify the position coordinates but the fields. But we then still must define the set of hidden variables as the union of four disjoint sets, determined by the four possible combinations $(\mathcal{F}(A_j), \mathcal{F}(B_k))$ of the fields \mathcal{F} . The set V_1 is then a subset of $A_1 \times B_1 \times C$ and one of these four disjoint sets.

We can write V as a subset of $\cup_{jk} A_j \times B_k \times C$, with the accents implicitly understood. Obviously $V_1 \subsetneq V$. On the set $E \times C$ we can describe as a “real” experiment what we are forced to consider as an “imaginary” experiment in \mathbb{R}^3 , where the use of the qualifier “imaginary” is used to pinpoint the fact that two different orientations of the same filter cannot simultaneously exist. On $E \times C$ this contradiction does no longer exist, because the different filter orientations have been relegated to different orthogonal spaces. It is in this sense that we can call the thought experiment then “real”: it is contradiction-free. The measure of the set of $(A_1 \cup A_2') \times (B_1 \cup B_2')$ will in principle be four times the measure of any set of $A_j \times B_k$ if the polarizers are strictly identical. The same will in principle be true for the subsets which define the hidden variables [11].

This research received no external funding.

The author declares no conflict of interest.

References

1. Aspect, A.; Dalibard, J.; and Roger, G.; Experimental Test of Bell’s Inequalities Using Time-Varying Analyzers. *Phys. Rev. Lett.* **49**, 1804 (1982).
2. Aspect, A.; Grangier, P.; Roger, G.; Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell’s Inequalities, *Phys. Rev. Lett.* **49**, 91 (1982).
3. Shimony, A.; Conceptual foundations of quantum mechanics; in *The New Physics*, ed. by P. Davies, Cambridge University Press, Cambridge, (1989), pp. 373-395.
4. Giustina, M.; Versteegh, M.A.M.; Wengerowsky, S.; Handsteiner, J.; Hochtner, A.; Phelan, K.; Steinlechner, F.; Kofler, J.; Larsson, J.-Å.; Abellán, C.; Amaya, W.; Pruneri, V.; Mitchell, M.W.; Beyer, J.; Gerrits, T.; Lita, A.E.; Shalm, L.K.; Nam, S.W.; Scheidl, T.; Ursin, R.; Wittmann, B.; and Zeilinger, A.; Significant-Loophole-Free Test of Bell’s Theorem with Entangled Photons. *Phys. Rev. Lett.* **115**, 250401 (2015).
5. Shalm, L.K.; Meyer-Scott, E.; Christensen, B.G.; Bierhorst, P.; Wayne, M.A.; Stevens, M.J.; Gerrits, T.; Glancy, S.; Hamel, D.R.; Allman, M.S.; Coakley, K.J.; Dyer, S.D.; Hodge, C.; Lita, A.E.; Verma, V.B.; Lambrocco, C.; Tortorici, E.; Migdall, A.L.; Zhang, Y.; Kumor, D.R.; Farr, W.H.; Marsili, F.; Shaw, M.D.; Stern, J.A.; Abellán, Amaya, C.W.; Pruneri, V.; Jennewein, T.; Mitchell, M.W.; Kwiat, P.G.; Bienfang, J.C.; Mirin, R.P.; Knill, E.; and Nam S.W.; Strong Loophole-Free Test of Local Realism. *Phys. Rev. Lett.* **115**, 250402 (2015).
6. Freedman, S.A.; and Clauser, J.F.; Experimental test of local hidden variables. *Phys. Rev. Lett.* **28**, 938 (1972).
7. Larsson, J.-Å.; Loopholes in Bell inequality test of local realism. *J. Phys. A* **47**, 424003 (2014).

8. Kupczynski, M.; Is the Moon there if Nobody Looks: Bell Inequalities and Physical Reality. *Frontiers in Physics* **8**, 273 (2020).
9. Kupczynski, M.; Closing the Door on Quantum Nonlocality. *Entropy* **20**, 877 (2018).
10. Coddens, G. The exact meaning of the angular-momentum and spin operators in quantum mechanics. <https://hal.archives-ouvertes.fr/hal-03323780> (2021).
11. Khrennikov, A. Classical probability model for Bell inequality. *Journal of Physics: Conference Series* **504**, 012019 (2014).