Chapter 15
How to Avoid the Coincidence Loophole

Jan-Åke Larsson

Bell inequality tests of local realism are notoriously difficult to perform. Physicists have attempted these tests for more than 50 years, and for each attempt, gotten closer and closer to a proper test. So far, every test performed has been riddled by one or more loopholes. While I personally am not overly fond of using the word "loophole", it has become the standard term; the word is usually used in connection to laws, most commonly taxation law, where a "loophole" points to some unintended and/or unexpected circumstance where the law does not apply, so that one can avoid the law without technically breaking it. In connection to Bell tests, the word points to some unexpected circumstance in experiments that makes additional assumptions necessary, we will see some examples of this below.

Here, we will encounter one of the more recently found loopholes, the coincidence loophole, and learn how to avoid it. First, the scene is set by reviewing the relation between the Einstein-Podolsky-Rosen (EPR) paradox and the Bell inequality, and also fix the notation and formal background. Then a few (not all) recent and not so recent experiments will be presented, with some of their qualities and shortcomings. And finally, we have a look at the coincidence loophole, why it seems to be more serious than one could have expected given results from the eighties, and then why the seemingly large impact is in fact not so damaging after all. In short, how to avoid the coincidence loophole.

J.-Å. Larsson (✉)

Institutionen för systemteknik, Linköpings Universitet, 581 83 Linköping, Sweden
e-mail: jan-ake.larsson@liu.se

Local Realism

The concept of local realism is motivated by the question posed by Einstein, Podolsky, and Rosen (EPR) in [12]: "Can [the] quantum-mechanical description of physical reality be considered complete?" EPR argue that the answer to the question is negative, while Bohr in [7] argues that the question is meaningless. Although many physicists agree with Bohr, one should note that EPR were actually correct in a sense: for the system used by EPR, there does exist a more complete description as Bell [5] noted. Bell simply constructed a more complete description of that system. His construction gives the same predictions as quantum mechanics for position and momentum, and also has these two explicit in the description.

Another example was used by Bohm [6] to simplify the experiment, going from the infinite-dimensional quantum description of position-momentum to the much simpler finite-dimensional quantum description of two spin systems. The choice of measuring position or momentum now translates into a choice of axis, or direction, along which a spin component measurement is performed. This choice of direction is denoted ϕ and ψ at the two sites in Fig. 15.1. In a spin-1/2 system, there are only two outcomes: magnetic dipole moment parallel to the direction ("spin up", $+1$) or anti-parallel to the direction ("spin down", -1).

The system is such that the outcomes are (anti-)correlated; if the same direction is chosen on both sides ($\psi = \phi$), the outcomes are always opposite, as indicated in Fig. 15.1. This is called a *singlet* state, which is a special case of a so-called *entangled* state. The terminology is due to Schrödinger who translated the German *Verschräkung* in one of his letters to Einstein into *entanglement*. Obtaining an outcome at one site at direction ϕ allows prediction (with probability 1) of the outcome at the other at direction $\psi = \phi$. If the systems are separated, the choice of direction can be performed so that it (the choice of direction) cannot disturb the remote system, such disturbances being limited by the speed of light. Prediction without disturbance is the reason to believe that the outcome exists, as a physical property, independent of measurement (EPR [12]).

In this system, there are more choices than in the EPR system, since the direction parameter ϕ is has a continuous range (uncountably infinite to a mathematician) rather than the two choices of position and momentum in EPR. Still, if one only con-
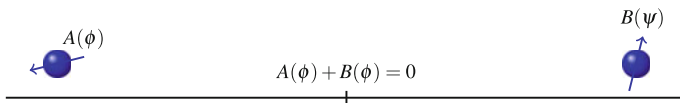

Fig. 15.1 Bohm's modification of the EPR setup. Two spin-1/2 systems are created in a joint state such that measurement of the spin component along the same direction gives opposite results, for all directions. Such a state can be created, it is called a singlet state and has total spin 0. The two systems are then separated, while making sure that the joint state is preserved. Then, a local measurement at one site for one direction can be used to predict, with probability 1, the outcome of a measurement at the remote site, for the same direction.

siders the predictions for equal settings $\psi = \phi$, there exists a more complete description just as for EPR, so that equal settings give opposite outcomes (several authors, but see e.g., [3]). In this sense, EPR are still correct. However, there are many more combinations $\psi \neq \phi$ than in the plain EPR case (position $\neq$ momentum), something John Bell made use of in his famous inequality [4]. The trick is to use three (later four, five, . . .) different settings, and use that anticorrelations are large when the two directions are close but not equal to each other.

To write down the inequality we need some concepts and notions from probability theory. In probability theory, measurement outcomes are given by random variables A and B that can take parameters such as measurement directions, and also depend on a random *sample* λ, here often referred to as a *hidden variable*. The hidden variable takes values in a *sample space* Λ, and subsets of the sample space where, e.g., the random variable takes a certain value is called an *event*. To tell us the probability of events, we need a *probability measure*, so for example the probability that a random variable takes a certain value can be written

$$P\Big(\big\{ \lambda \,:\, A(\phi, \lambda) = -1 \big\} \Big) = P\Big(A(\phi) = -1 \Big). \tag{1}$$

The notation is often simplified by leaving out the set notation and the λ as shown above. The conditional probability of one outcome given another can be calculated as

$$P\Big(A(\phi) = -1 \Big| B(\psi) = -1 \Big) = \frac{P\Big(A(\phi) = -1 \cap B(\psi) = -1 \Big)}{P\Big(B(\psi) = -1 \Big)}. \tag{2}$$

Missing detections (as in lowered efficiency) can be handled through either assigning value 0 to the outcome, or not defining an outcome value at all [18, 19]. Finally, we need the concept of an *expectation value*, so to speak, the average outcome. For brevity I will only give what we need here, the expectation of the product of ± 1 outcomes,

$$E\Big(A(\phi)B(\psi) \Big) = P\Big(A(\phi) = B(\psi) \Big) - P\Big(A(\phi) \neq B(\psi) \Big) \tag{3}$$

The above would tell us if the outcomes are equal with high probability, which would give a value close to $+1$ or unequal with high probability (close to -1). Using this, Bell could formulate and prove a theorem, later improved by Clauser, Horne, Shimony and Holt [10] into

Theorem 1 (The CHSH inequality, *[10].*) *Under the two assumptions*

Realism: The outcomes can be described by random variables

$$A(\phi_i, \psi_j, \lambda),\ B(\phi_i, \psi_j, \lambda); \quad |A| = |B| = 1$$

Locality: The r.v.s are independent of the remote setting

$$A_i(\lambda) = A(\phi_i, \psi_j, \lambda); \; B_j(\lambda) = B(\phi_i, \psi_j, \lambda)$$

the following inequality holds:

$$\left| E(A_1 B_1) + E(A_1 B_2) \right| + \left| E(A_2 B_2) - E(A_2 B_1) \right| \leq 2.$$

An oversimplified but popular attempt to explain this uses anticorrelated outcomes: one ball in one of two boxes, one box sent to each of two sites, and then opened. If the box at the first site is opened and found to contain a ball, there will be no ball found when the box at the second site is opened. Perfect predictability. The oversimplification is that it uses too few boxes, since the CHSH expression uses four boxes, two sent to one site (A_1 and A_2) and two sent to the second site (B_1 and B_2). The terms $E(A_i B_j)$ tells us how often we can expect to see balls in both or none of the boxes. A negative value tells us that the probability is high that there is a ball in only one of the two boxes. On the other hand, a positive value tells us that the probability is high that there are balls in both or none of them. The CHSH inequality puts bounds on these probabilities: a consequence of the inequality is

$$E(A_2 B_2) \leq 2 + E(A_2 B_1) + E(A_1 B_1) + E(A_1 B_2). \tag{4}$$

This means: if the probability is large that there is one and only one ball in the two boxes A_2, B_1; one in A_1, B_1; and one in A_1, B_2, the terms on the right will be negative. If they are negative enough, the bound will tell us that the left-hand side must be negative, so that the probability must be large that there is only one ball in A_2, B_2 (in essence because there are three pairs on the right-hand side, an odd number). In the extreme case, if all three terms on the right are -1, the term on the left is forced to be -1. For our system, we are already able to predict with probability 1, so there is no news there.

But Bell realized that using directions ϕ_2, ψ_1, ϕ_1, and ψ_2 that are close to each other (in that order) but not equal, the bound tells us something new. As indicated above, the most common example is the singlet state, for which the quantum-mechanical prediction is

$$E_{\mathrm{QM}}(A_i B_j) = -\cos(\phi_i - \psi_j), \tag{5}$$

with $\phi_2 = 0$, $\psi_1 = \pi/4$, $\phi_1 = \pi/2$, and $\psi_2 = 3\pi/4$. Then,

$$2 + E_{\mathrm{QM}}(A_2 B_1) + E_{\mathrm{QM}}(A_1 B_1) + E_{\mathrm{QM}}(A_1 B_2) = 2 - \frac{3}{\sqrt{2}} < 0, \tag{6}$$

while

$$E_{\mathrm{QM}}(A_2 B_2) = \frac{1}{\sqrt{2}} > 0. \tag{7}$$

In other words, for this quantum-mechanical system, there is a large probability of one and only one ball in the box pair A_2, B_1, one in A_1, B_1, and one in A_1, B_2. And the bound tells us that the probability is larger that there is one and only one ball in A_2, B_2, than two or none (the bound is less than 0). But the quantum-mechanical prediction is instead that it is more likely that there are two or no balls in A_2, B_2, than just one (the quantum prediction is larger than 0). The prediction does not obey the bound, there is a *violation*.

If our desired complete description is local realist, then this gives good support for Bohr's argument: there exists no local realist description that gives the quantum-mechanical predictions. The importance of this result cannot be stressed enough: we have learned something about what possible types of mathematical models that can be used to describe quantum-mechanical systems.

Loopholes

At this point, it is important to make one thing clear: even though a local realist description cannot give the quantum-mechanical predictions, it is still possible that it can give the statistics that we see in nature. It could be the case that the underlying assumptions of Theorem 1 does not apply in nature, or more accurately in our experiment. So our task now is to make sure that the theorem really does apply in our experiment, and as it turns out, this may be difficult. There are several loopholes, problems that could occur in experiment that makes the theorem need additional assumptions. These mainly fall into two classes: efficiency and locality (see e.g., [19]).

Lowered efficiency, that not all single systems give outcomes, is a common problem in experiments to test local realism. This may seem like a small problem, can one not just normalize to the subset of systems that do give outcomes? Well, not really: doing this requires an assumption, the *fair sampling assumption*. To be allowed to normalize to that subset, one needs to be sure that the subset is a fair sample of the full ensemble, in a sense that the subset itself does not depend on the local settings. Many modern experiments only violate the inequality under this assumption. And fair sampling is an assumption because in experiment it is impossible to verify that the sample is fair.

If the sample is not fair, the CHSH inequality changes, something first suggested in [27]. The bound increases when the efficiency decreases, making it possible for a local realist description to reach the quantum value. The crossover for CHSH is 82.84 % [13], and below this efficiency, a local realist description can reach the same value as quantum mechanics. Much work has been put into calculating this and other bounds under increasingly general conditions, and also new inequalities for new systems where the effects are smaller.

Many of these modified inequalities need the number of emitted pairs to be known, so that one can get probabilities from counts (by dividing beneficial counts by the total counts). The below version of the inequality gives a bound for the conditional correlations, and the only extra parameter used is the conditional proba-

bility of coincidence, conditioned on single detections. It does not need the number of emissions, but simply the ratio of coincidences to single detections. This data is already available in the experimental record, which simplifies the analysis greatly.

Theorem 2 (The CHSH inequality with inefficiency, *[18]*.) *Under the three assumptions*

Realism: *The outcomes can be described by random variables*

$$A(\phi_i, \psi_j, \lambda),\ B(\phi_i, \psi_j, \lambda); \quad |A| \leq 1;\ |B| \leq 1$$

Locality: *The r.v.s are independent of the remote setting*

$$A_i(\lambda) = A(\phi_i, \psi_j, \lambda);\ B_j(\lambda) = B(\phi_i, \psi_j, \lambda)$$

Detection: *Detection is controlled by the local realist model*

$$\eta = \min_{settings} P(remote\ detection | local\ detection)$$

the following inequality holds:

$$\left| E(A_1 B_1 | coinc.) + E(A_1 B_2 | coinc.) \right| + \left| E(A_2 B_2 | coinc.) - E(A_2 B_1 | coinc.) \right| \leq \frac{4}{\eta} - 2$$

So, if the violation is high enough *and* the efficiency is high enough, then a proper test can be performed. It is important to remember that *efficiency* refers to the total efficiency. There is no reference to detectors or other equipment, the focus is on the output data: coincidences compared to local detections.

The requirements for violation are quite high. Also, we need to take experimental noise into account. If we use the singlet state and take noise into account, the amplitude of the cosine correlation is decreased. This is quantified in terms of *visibility*, a multiplicative factor v in front of the cosine. The inequality now translates into

$$2\sqrt{2}v \leq \frac{4}{\eta} - 2 \tag{8}$$

The region of violation is drawn in Fig. 15.2, the small region on the top right.

It is difficult to reach violation when using photons, the system of choice in long-distance experiments. Efficiency and visibility of four example photonic experiments [2, 14, 15, 32] is shown in Fig. 15.3. The most-cited experiment is [2], for which the main intent was to show that the violation remains even with fast-changing settings at the two sites, fast enough that no light-speed signal could carry information on the setting from one site to the other. This is to avoid disturbance in the spirit of EPR, to ensure that the locality assumption of the inequality holds. And the experiment is conclusive, since the violation does remain. The fair sampling assumption is needed because the experiment has an efficiency of approximately 0.1 % which is not visible to the scale of the figure. One reason for the low efficiency of the Aspect experiment
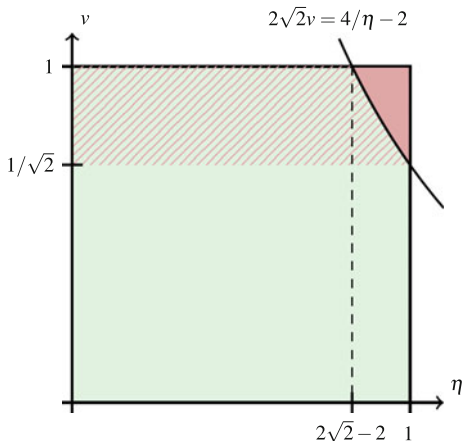
Fig. 15.2 The region of violation of the CHSH inequality for the singlet state, in terms of efficiency η and visibility v. The almost triangular region of violation has corners at visibility 70.71 % and efficiency 82.84 %. The fair-sampling assumption would enable a violation in the entire region above visibility 70.71 %, striped in the figure.

is the wide directional spread of the emissions. More modern sources (for example [17], but there are many improvements after that) give much better collection of photons, and in the experiment of [32] this can be seen, the reported efficiency is 5 %.

That experiment also uses proper random settings rather than the quasiperiodic switching used by Aspect [2]. Another improvement in locality is made in [31], where the random numbers are selected so that no light-speed signal can reach them from the emission event. This requires a large distance resulting in a 35 dB attenuation for the most remote site, or an efficiency of 0.03 %. Impressive for the 144 km distance, but of course lower than our other examples.

At the other end of the scale, we have experiments designed for very high (meaning 100 %) efficiency, these are depicted in the enlarged part of Fig. 15.3. These experiments do not use photons, but instead other more massive systems, or even solid state systems. Our examples are ions in one trap [30], in two separate traps [25], Josephson phase qubits [1], atoms in two separate traps [16], and nuclear spins near a nitrogen-vacancy centre in diamond [28]. In these experiments experimental runs are well defined and do give outcomes, so the efficiency is 100 %, but there is still the locality issue: the systems are not separated enough to guarantee locality. The quotient between the distance you need because of the duration of the relevant measurement, to the distance you have, has improved from 10^9 [30] via 1500 [25], and 300 [1], to 15 [16]. And this is still improving as we speak; the race is really on to achieve and surpass the magic number 1.

Meanwhile, photons are still the system of choice for long-distance quantum experiments, and efficiency numbers have increased steadily over the years. Modern superconducting detectors (e.g., [24] have a very high efficiency, and modern sources (e.g., [29] have very high collection. Coupled with low-loss components for the rest of the setup, an experiment can reach $\eta = 75$ %. There were two such exper-
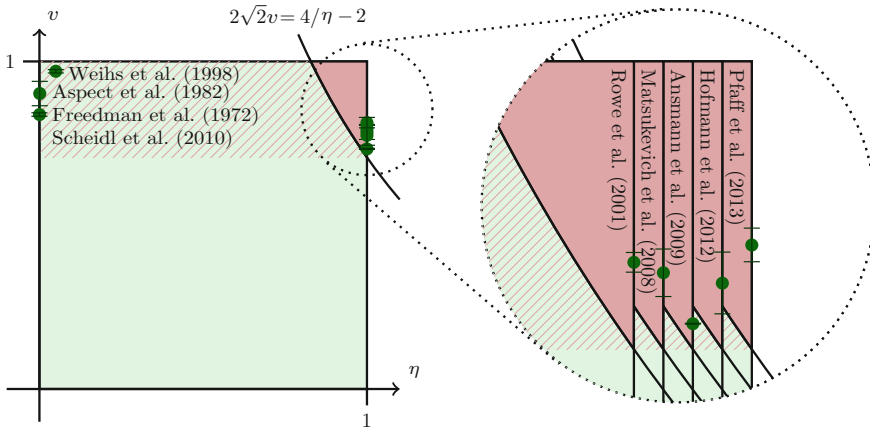
Fig. 15.3 Examples of experimental tests of local realism. The *green dot* shows efficiency and visibility of the experiments. All the papers also report the quality of the visibility estimate in terms of an estimated standard deviation, this is shown as a vertical bar.

iments recently [8, 15], but at this point one may ask: with the bound at 82.84 %, how can an experiment with efficiency 75 % be useful?

As it turns out, Eberhard found that one obtains lower bounds if a different state is used in the test [11]. We need to relax the perfect anticorrelation at equal directions, that we (and Bell, and EPR) thought so important. This is easiest to see if we use a different inequality that relates probabilities directly rather than correlations as in CHSH. The discovery sounds simple but is really profound, so much so that the below inequality sometimes is called *the Eberhard inequality* when used with counts rather than probabilities.

Theorem 3 (The Clauser-Horne (CH) inequality, *[9]*.) *Under the two assumptions*

Realism: *The result can be described by random variables*

$$A(\phi_i, \psi_j, \lambda); \ B(\phi_i, \psi_j, \lambda)$$

Locality: *The r.v.s are independent of the remote setting*

$$A_i(\lambda) = A(\phi_i, \psi_j, \lambda); \ B_j(\lambda) = B(\phi_i, \psi_j, \lambda)$$

the following inequality holds:

$$P(A_1 = B_2 = 1) + P(A_2 = B_1 = 1) - P(A_2 = B_2 = 1)$$
$$\leq P(A_1 = 1) + P(B_1 = 1) - P(A_1 = B_1 = 1).$$

One should be aware that the above is *equivalent* to CHSH, but there is a benefit since the efficiency estimate η does not enter into the calculation. There is no need
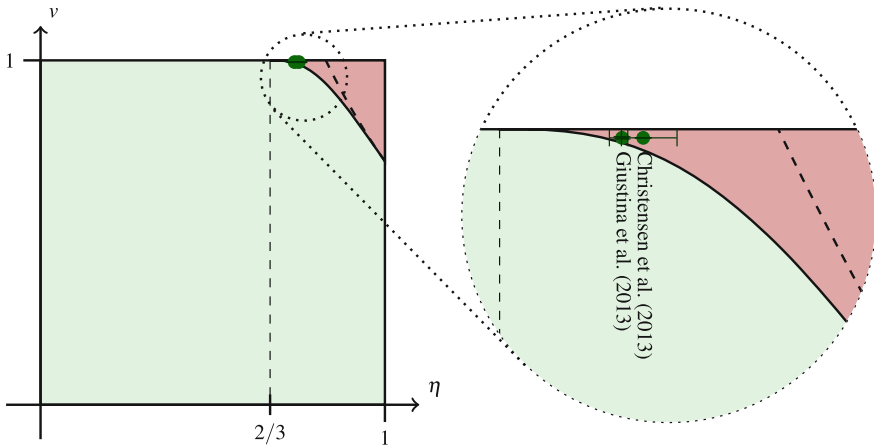
Fig. 15.4 Region of violation for non-maximally entangled quantum states. The region stretches to efficiency 66.67 %, but thins out considerably close to the limit. The two experimental efficiency estimates are qualitatively different and not so easily comparable with the earlier results, or to each other. [15] uses a separate efficiency measurement while [8] uses an estimate from equipment parameters (with a large standard deviation as indicated), and neither of these are necessarily equal to the η of Theorem 2, so the horizontal position of the two dots might not be accurate. On the other hand, the Clauser-Horne inequality does not need an explicit η, and indeed, both datasets violate the bound.

to condition on coincidence here, nor estimate the efficiency of the experiment. We will use this inequality below because of its somewhat simpler structure, since there are fewer outcome combinations that contribute to the expression, but in principle, the same can be done with CHSH.

In any case, for a given efficiency, we can find a non-maximally entangled state (i.e., not our earlier singlet state) and measurement directions that give a violation even if the efficiency is below 82 % but above 66.67 % [11, 21, 23]. The region of violation in terms of visibility and efficiency becomes somewhat larger if this is done, see Fig. 15.4. And the two mentioned experiments are inside the region of violation. This makes photons the only kind of system for which both kinds of loopholes, both efficiency and locality, have been closed, in separate experiments. It remains to close both in the same experiment, but nonetheless, we are gathering evidence that there really exists no local realist description of such a system.

The Coincidence Loophole

There is one problem in photonic experiments that we have not addressed yet: how the experimenter knows which clicks come from the same pair of photons. This is not the same as the efficiency loophole discussed above, which is more the problem of finding photons at all. Typically pairs are identified by *coincidence detection*, by comparing the detection times, either to each other (we'll call this *moving windows*), or to a predetermined reference (which we'll call *fixed time slots*). If the comparison
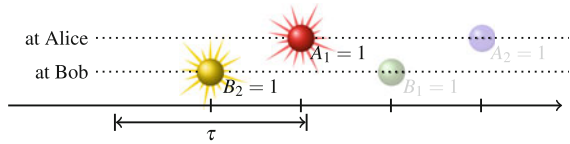
Fig. 15.5 Example local realist model (due to E. Knill). The detection time depends on the hidden variable and the local setting. If Alice measures A_1 and Bob measures B_2, the events are close enough to give a coincidence.
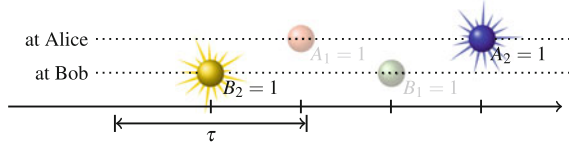

Fig. 15.6 Other choices of setting in the same local realist model will give different detection times. If Alice measures A_2 and Bob measures B_2, the events are *not* close enough to give a coincidence.

is to other events, each event on one side is surrounded with a time interval of length τ, centered on the event, as in Figs. 15.5 and 15.6. If there is a remote event in the time interval, this is considered as belonging to the same pair, or coincident. (In a real experiment there may be several events in the time interval, in which case the closest can be chosen as belonging to the same pair, but we will just consider the simple case here.)

We are considering local realist models, and it is possible that such a model influences the time of detection, similar to influencing (non-)detection. In this situation, the subset of coincidences could change. It could happen that all of the events are shifted in time so that there always are coincidences if A_2, B_1 are measured, or if A_1, B_1 or A_1, B_2 are measured, but never if A_2, B_2 are measured (see Figs. 15.5 and 15.6). This would give

$$P(A_2 = B_1 = 1 \cap \text{coincidence}) = 1 \tag{9a}$$

$$P(A_1 = B_1 = 1 \cap \text{coincidence}) = 1 \tag{9b}$$

$$P(A_1 = B_2 = 1 \cap \text{coincidence}) = 1 \tag{9c}$$

$$P(A_2 = B_2 = 1 \cap \text{coincidence}) = 0 \tag{9d}$$

$$P(A_1 = 1) = P(B_1 = 1) = 1 \tag{9e}$$

so that

$$P(A_1 = B_2 = 1 \cap \text{coinc.}) + P(A_2 = B_1 = 1 \cap \text{coinc.}) - P(A_2 = B_2 = 1 \cap \text{coinc.})$$
$$= 1 + 1 - 0 = 2$$
$$\nleq P(A_1 = 1) + P(B_1 = 1) - P(A_1 = B_1 = 1 \cap \text{coinc.})$$
$$= 1 + 1 - 1 = 1. \tag{10}$$

The Clauser-Horne inequality does apparently not hold anymore. The apparent violation above is the highest possible, because this reaches the algebraic bound of the CH inequality. But note that the terms containing pairs of outcomes $P(A_i = B_j = 1)$ have been replaced by $P(A_i = B_j = 1 \cap \text{coincidence})$, so it is no longer the same expressions on the right- and left-hand sides. The terms contain an additional restriction in the pair events (that we calculate the probability of), that comes from not knowing beforehand which events belong to the same pair. Clearly, the CH inequality does not apply to the modified expressions, because the left-hand side is larger than the right-hand side.

The effect of this kind of time dependence on the CHSH inequality was studied in [20] (it seems the effect is first discussed in [26]). In [20] a modified CHSH inequality is derived that does apply, and also a local realist model that saturates the inequality (actually containing the above example).

Theorem 4 (The CHSH inequality with coincidence identification, *[20].*) *Under the three assumptions*

Realism: *The outcomes can be described by random variables*

$$A(\phi_i, \psi_j, \lambda),\ B(\phi_i, \psi_j, \lambda);\quad |A| \le 1;\ |B| \le 1$$

Locality: *The r.v.s are independent of the remote setting*

$$A_i(\lambda) = A(\phi_i, \psi_j, \lambda);\ B_j(\lambda) = B(\phi_i, \psi_j, \lambda)$$

Coincidence: *Coincidence is controlled by local realist time delays and*

$$P(coincidence) = \gamma$$

the following inequality holds:

$$\begin{aligned} &\left| E(A_1 B_1 | coinc.) + E(A_1 B_2 | coinc.) \right| \\ &+ \left| E(A_2 B_2 | coinc.) - E(A_2 B_1 | coinc.) \right| \le \frac{6}{\gamma} - 4 \end{aligned}$$

This is similar to Theorem 2, but there are differences. The expression $4/\eta - 2$ is replaced by $6/\gamma - 4$, giving a higher bound (in γ) than the previous (in η). There is no violation unless $\gamma > 87.87\,\%$, which is visible also in Fig. 15.7. Also, γ is not so simple to estimate from experimental data as η, again the emission rate is needed for the estimate. Since the efficiency on the single-photon level is $100\,\%$ in the model example, it has been conjectured that the bound for γ translates directly into an equal bound for η, therefore a comparison with the detection efficiency bound can also be found in Fig. 15.7.

Interestingly, since the model example has single-photon efficiency $100\,\%$, the fair sampling assumption holds. All the photons are registered, so the sample is fair (for the same reason, the *no enhancement* assumption holds, for details see e.g.,
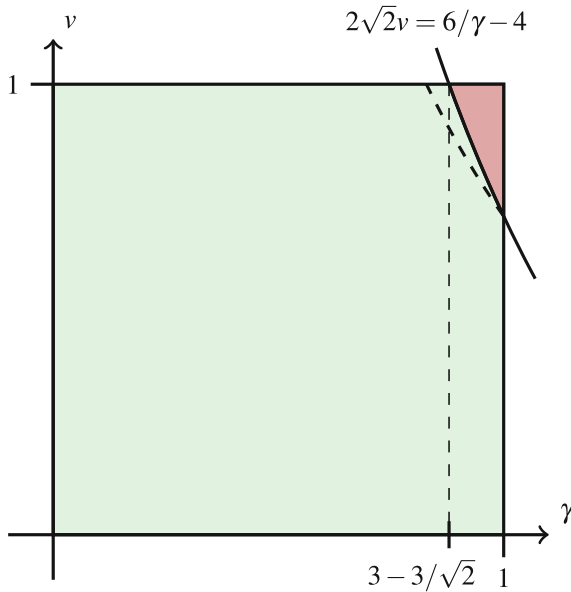
Fig. 15.7 Coincidence probability bound for violation of the modified CHSH inequality, when using a singlet state. The coincidence probability must exceed 87.87 %.

[19]). If anything, it is the *coincidence identification* that is not fair. The subset of coincidences that we register may change with the measurement directions. This happens also when we have lowered efficiency, but there is a difference: detection is a local process, and detection at one site happens on a subset of λs

$$\Lambda_{A_i} = \{\lambda : A_i(\lambda)\text{is detected}\}. \tag{11}$$

A coincidence then occurs on the set

$$\Lambda_{A_i} \cap \Lambda_{B_j}. \tag{12}$$

This is in stark contrast to the coincidence detection we use here, where the detection time T_{A_j} depends on the setting and the hidden variable, and coincidence occurs on the set

$$\Lambda_{A_iB_j} = \left\{\lambda : \left|T_{A_i}(\lambda) - T_{B_j}(\lambda)\right| < \tfrac{1}{2}\tau\right\} \neq \Lambda_{A_i} \cap \Lambda_{B_j} \tag{13}$$

There is even no possibility to rewrite $\Lambda_{A_iB_j}$ so that it can be factorized in this manner. Equivalently, attempting to assign the outcome 0 or ± 1 to "missing coincidences" will lead to a non-local hidden-variable model, because even though the time-dependence is local, the newly assigned missing-coincidence-outcome needs to depend on both settings.

At this point, things look bleak. Even 100 % efficient detectors do not help, and the coincidence probability must exceed 87.87 %, considerably higher than the previ-

ously known efficiency bound 82.84 %. It is not even enough to assume fair sampling, since it already holds in the model example. It seems that the coincidence loophole is much worse than the efficiency loophole ever was. But fear not, there are quite simple modifications that will re-enable the bounds from before.

How to Avoid the Coincidence Loophole

Of course, one could substitute the fair sampling assumption with a *fair coincidence* assumption, that the subset of *pairs* chosen by our coincidence procedure is a fair sample of the whole ensemble. But we want to avoid that kind of assumptions. Better to check why the inequality does not apply, and attempt find conditions under which it does apply.

The timing of the example is such that all the pairs that increase the violation count as coincidences, while the pairs that decrease the violation does not count. One alternative to re-establish the inequality would therefore be to make sure that the latter pairs do count as coincidences. But ensuring that in general would demand that we know which photons belong together in pairs, and as we already have said, this is not possible. There will always be pairs that we cannot identify, so we need to be less ambitious. What we can aim for is to ensure that a pair that would count as a coincidence for all three measurement combinations that increase the violation, also would count as coincidence for the fourth measurement combination. This can be done by having the time window for A_2, B_2 as wide as *all three* of the other windows taken together, we call this the *window sum method* [22]. The reason that this works is that if λ is such that

$$
\begin{cases}
\left|T_{A_2}(\lambda) - T_{B_1}(\lambda)\right| \leq \tau, \text{ and} \\
\left|T_{A_1}(\lambda) - T_{B_1}(\lambda)\right| \leq \tau, \text{ and} \\
\left|T_{A_1}(\lambda) - T_{B_2}(\lambda)\right| \leq \tau,
\end{cases}
\tag{14}
$$

then, for the same λ (see Fig. 15.8), the triangle inequality gives

$$
\begin{aligned}
&\left|T_{A_2}(\lambda) - T_{B_2}(\lambda)\right| \\
&\leq \left|T_{A_2}(\lambda) - T_{B_1}(\lambda)\right| + \left|T_{B_1}(\lambda) - T_{A_1}(\lambda)\right| + \left|T_{A_1}(\lambda) - T_{B_2}(\lambda)\right| \leq 3\tau.
\end{aligned}
\tag{15}
$$

Since this applies to the individual λ values, the subsets (in terms of λ) that give coincidence will obey

$$
\Lambda_{A_1 B_1} \cap \Lambda_{A_1 B_2} \cap \Lambda_{A_2 B_1} \subset \Lambda_{A_2 B_2}.
\tag{16}
$$

That this really fixes the issue will of course need a formal proof, so that there is no possibility for another model to evade our construction. The following theorem will do this for us.
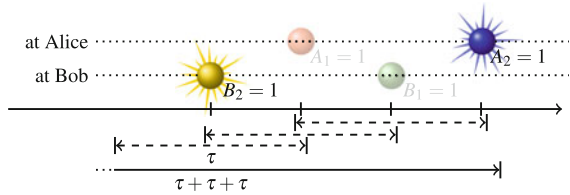
Fig. 15.8 One method to avoid the coincidence loophole is to have longer coincidence windows that decrease the violation (A_2B_2), as long as the sum of the coincidence windows that increase it (A_1B_2, A_1B_1, A_2B_1). This ensures that if there is a chain of possible coincidences that increase the violation ($B_2 \leftrightarrow A_1 \leftrightarrow B_1 \leftrightarrow A_2$), the two endpoints that would decrease the violation ($B_2 \leftrightarrow A_2$) will also be a possible coincidence.

Theorem 5 (The CH inequality, avoiding the coincidence loophole, *[22]*.) *Under the three assumptions*

Realism: *The result can be described by random variables*

$$A(\phi_i, \psi_j, \lambda); \ B(\phi_i, \psi_j, \lambda)$$

Locality: *The r.v.s are independent of the remote setting*

$$A_i(\lambda) = A(\phi_i, \psi_j, \lambda); \ B_j(\lambda) = B(\phi_i, \psi_j, \lambda)$$

Coincidence: *Coincidences are controlled by local realist time delays, are obtained on subsets* $\Lambda_{A_1B_1}; \ \Lambda_{A_1B_2}; \ \Lambda_{A_2B_1};$ *and* $\Lambda_{A_2B_2},$ *of* $\Lambda,$ *and the last coincidence set contains the intersection of the other three,*

$$\Lambda_{A_1B_1} \cap \Lambda_{A_1B_2} \cap \Lambda_{A_2B_1} \subset \Lambda_{A_2B_2}$$

the following inequality holds:

$$P(A_1 = B_2 = 1 \cap \Lambda_{A_1B_2}) + P(A_2 = B_1 = 1 \cap \Lambda_{A_2B_1})$$
$$- P(A_2 = B_2 = 1 \cap \Lambda_{A_2B_2})$$
$$\leq P(A_1 = 1) + P(B_1 = 1) - P(A_1 = B_1 = 1 \cap \Lambda_{A_1B_1}).$$

Another alternative to re-establish the inequality would be to avoid counting some pairs that increase the violation. Here, instead, we ensure that if a pair does not count as a coincidence for the measurement combination that decreases the violation, it is also not counted for at least one of the three measurement combinations that increase the violation. This can be ensured, in general, by introducing fixed points on the time-line that break pairs if the detection times are on different sides of such a point (see Fig. 15.9). And this is what happens when using fixed time slots. Formally, we choose a number of fixed time slots $I_k = \{t : t_k \leq t \leq t_k + \tau\}$ (these must be disjoint), and count a coincidence in slot k when the detection time of both events are in the time
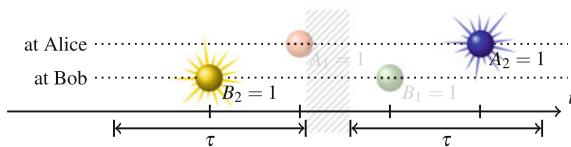
Fig. 15.9 Another method to avoid the coincidence loophole is to have fixed time slots, so that if two events are far enough apart to not give a coincidence ($B_2 \not\leftrightarrow A_2$), there cannot be a chain of possible coincidences from one event to another, that connects the two (here $B_2 \leftrightarrow A_1 \not\leftrightarrow B_1 \leftrightarrow A_2$).
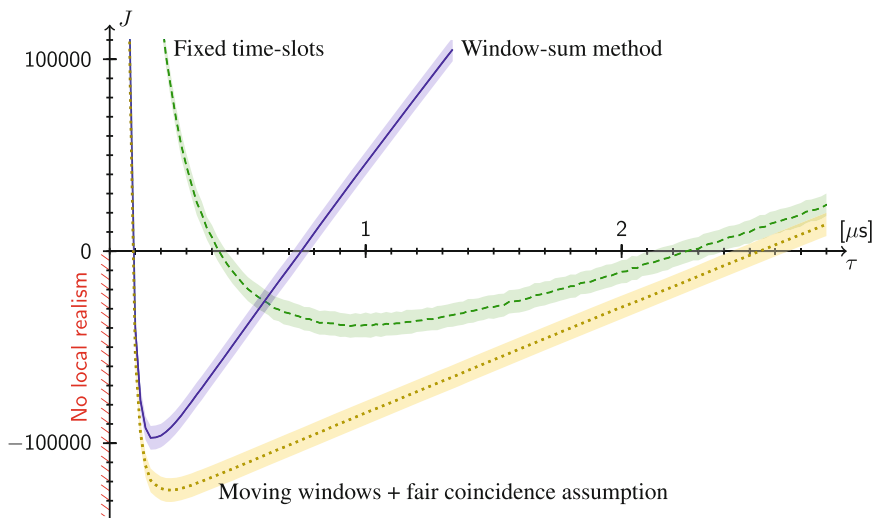

Fig. 15.10 Violation of the Clauser-Horne bound under local realism, using three different methods of handling the coincidence loophole [data from [14]]. The vertical axis scale is in counts rather than probability, a negative value is a violation. The *dotted yellow line* uses standard moving windows and the fair-coincidence assumption. The *green dashed line* uses fixed time slots. The *blue solid line* uses the window-sum method with length τ of the three short windows. The latter two are not vulnerable to the coincidence loophole, and the first one can avoid it by using the fair coincidence assumption. The *shading* corresponds to plus or minus three estimated standard deviations.

slot,

$$\Lambda_{A_i B_j}^{(k)} = \left\{ \lambda : T_{A_i}(\lambda) \in I_k, \ T_{B_j}(\lambda) \in I_k \right\}. \tag{17}$$

Using these time slots, if λ is such that the events A_2, B_1 are in the same time slot, and also B_1, A_1, and A_1, B_2, then of course A_2, B_2 are in the same time slot, so that

$$\Lambda_{A_1B_1}^{(k)} \cap \Lambda_{A_1B_2}^{(k)} \cap \Lambda_{A_2B_1}^{(k)} \subset \Lambda_{A_2B_2}^{(k)}. \tag{18}$$

Since $\Lambda_{A_iB_j}$ is the disjoint union over k of $\Lambda_{A_iB_j}^{(k)}$, Eq. (16) immediately follows. As does the modified CH inequality.

Conclusion

Thus, there are (at least) two ways to avoid the coincidence loophole, either by using the window-sum method, or by using fixed time slots. This will re-enable the CH inequality, so that the bound for γ (coincidence probability) will be replaced by a bound for η (efficiency, as defined in Theorem 2). Essentially, we have restored the violation region from the small region of Fig. 15.7 to the larger region of Fig. 15.4

This means that both experiments from 2013 do violate the modified CH inequality. The experiment in [8] has a natural time slot structure since the experiment uses a pulsed-pump source, so Theorem 5 applies directly. The experiment in [15] does not have this natural structure because it uses a continuously pumped source, but time slots can still be used. We can also use the window-sum method, and interestingly the best time slot and time window lengths are different for the two methods, see Fig. 15.10.

Even though the coincidence loophole did seem to be very serious, e.g., that 100 % efficiency did not help, we have found two methods that remove it. And both available photonic experiments can use one or both methods, so that there is a violation. In conclusion, we are approaching a proper violation; locality remains to be enforced, but we are getting closer and closer. The importance of performing a conclusive experiment cannot be stressed enough: then we will have learned something fundamental about what possible types of mathematical models that can be used to describe physical reality.

References

1. M. Ansmann, H. Wang, R.C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A.D. O'Connell, D. Sank, M. Weides, J. Wenner, A.N. Cleland, J.M. Martinis, Violation of Bell's inequality in Josephson phase qubits. Nature **461**, 504–506 (2009). doi:10.1038/nature08363
2. A. Aspect, J. Dalibard, G. Roger, Experimental test of Bell's inequalities using time-varying analyzers. Phys. Rev. Lett. **49**, 1804–1807 (1982). doi:10.1103/PhysRevLett.49.1804
3. J.E. Baggott, *The Meaning of Quantum Theory* (Oxford University, Press, 1992)
4. J.S. Bell, On the Einstein-Podolsky-Rosen paradox. Physics (Long Island City, N. Y.) **1**, 195–200 (1964). http://philoscience.unibe.ch/documents/TexteHS10/bell1964epr.pdf

5. J.S. Bell, EPR correlations and EPW distributions. Ann. N. Y. Acad. Sci. **480**, 263–266 (1986). doi:10.1111/j.1749-6632.1986.tb12429.x

6. D. Bohm, Quantum Theory. Reprinted by Dover, New York, 1989 (Prentice Hall, New York, 1951). ISBN: 0-486-65969-0

7. N. Bohr, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **48**, 696–702 (1935). doi:10.1103/PhysRev.48.696

8. B.G. Christensen, K.T. McCusker, J.B. Altepeter, B. Calkins, T. Gerrits, A.E. Lita, A. Miller, L.K. Shalm, Y. Zhang, S.W. Nam, N. Brunner, C.C.W. Lim, N. Gisin, P.G. Kwiat, Detection-Loophole-free test of quantum nonlocality, and applications. Phys. Rev. Lett. **111**, 130406 (2013). doi:10.1103/PhysRevLett.111.130406

9. J.F. Clauser, M.A. Horne, Experimental consequences of objective local theories. Phys. Rev. D **10**, 526–535 (1974). doi:10.1103/PhysRevD.10.526

10. J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, Proposed experiment to test local hidden-variable theories. Phys. Rev. Lett. **23**, 880–884 (1969). doi:10.1103/PhysRevLett.23.880

11. P.H. Eberhard, Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. Phys. Rev. A **47**, R747–R750 (1993). doi:10.1103/PhysRevA.47.R747

12. A. Einstein, B. Podolsky, N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777–780 (1935). doi:10.1103/PhysRev.47.777

13. A. Garg, N.D. Mermin, Detector inefficiencies in the Einstein-Podolsky-Rosen experiment. Phys. Rev. D **35**, 3831–3835 (1987). doi:10.1103/PhysRevD.35.3831

14. S.J. Freedman, J.F. Clauser, Experimental Test of Local Hidden-Variable Theories. Phys. Rev. Lett. **28**, 938–941 (1972). doi:10.1103/PhysRevLett.28.938

15. M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S.W. Nam, R. Ursin, A. Zeilinger, Bell violation using entangled photons without the fair-sampling assumption. Nature **497**, 227–230 (2013). doi:10.1038/nature12012

16. J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, H. Weinfurter, Heralded entanglement between widely separated atoms. Science **337**, 72–75 (2012). doi:10.1126/science.1221856

17. P.G. Kwiat, P.H. Eberhard, A.M. Steinberg, R.Y. Chiao, Proposal for a loophole-free Bell inequality experiment. Phys. Rev. A **49**, 3209–3220 (1994). doi:10.1103/PhysRevA.49.3209

18. J.-Å. Larsson, Bell's inequality and detector inefficiency. Phys. Rev. A **57**, 3304–3308 (1998). doi:10.1103/PhysRevA.57.3304

19. J.-Å. Larsson, Loopholes in Bell inequality tests of local realism. J. Phys. A **47**, 424003 (2014). doi:10.1088/1751-8113/47/42/424003

20. J.-Å. Larsson, R.D. Gill, Bell's inequality and the coincidence-time loophole. Europhys. Lett. **67**, 707–713 (2004). doi:10.1209/epl/i2004-10124-7

21. J.-Å. Larsson, J. Semitecolos, Strict detector-efficiency bounds for nsite Clauser-Horne inequalities. Phys. Rev. A **63**, 022117 (2001). doi:10.1103/PhysRevA.63.022117

22. J.-Å. Larsson, M. Giustina, J. Kofler, B. Wittmann, R. Ursin, S. Ramelow, Bell-inequality violation with entangled photons, free of the coincidencetime loophole. Phys. Rev. A **90**, 032107 (2014). doi:10.1103/PhysRevA.90.032107

23. G. Lima, E.B. Inostroza, R.O. Vianna, J.-Å. Larsson, C. Saavedra, Optimal measurement bases for Bell tests based on the Clauser-Horne inequality. Phys. Rev. A **85**, 012105 (2012). doi:10.1103/PhysRevA.85.012105

24. A.E. Lita, A.J. Miller, S.W. Nam, Counting near-infrared singlephotons with 95 (2008)

25. D.N. Matsukevich, P. Maunz, D.L. Moehring, S. Olmschenk, C. Monroe, Bell inequality violation with two remote atomic qubits. Phys. Rev. Lett. **100**, 150404 (2008). doi:10.1103/PhysRevLett.100.150404

26. S. Pascazio, Time and Bell-type inequalities. Phys. Lett. A **118**, 47–53 (1986). doi:10.1016/0375-9601(86)90645-6

27. P. Pearle, Hidden-variable example based upon data rejection. Phys. Rev. D **2**, 1418–1425 (1970). doi:10.1103/PhysRevD.2.1418

28. W. Pfaff, T.H. Taminiau, L. Robledo, H. Bernien, M. Markham, D.J. Twitchen, R. Hanson, Demonstration of entanglement-by-measurement of solidstate qubits. Nat. Phys. **9**, 29–33 (2013). doi:10.1038/nphys2444

29. S. Ramelow, A. Mech, M. Giustina, S. Gröblacher, W. Wieczorek, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S.W. Nam, A. Zeilinger, R. Ursin, Highly efficient heralding of entangled single photons. Opt. Express **21**, 6707–6717 (2013). doi:10.1364/OE.21.006707

30. M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, D.J. Wineland, Experimental violation of a Bell's inequality with efficient detection. Nature **409**, 791–794 (2001). doi:10.1038/35057215

31. T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N.K. Langford, T. Jennewein, A. Zeilinger, Violation of local realism with freedom of choice. Proc. Nat. Acad. Sci. **107**, 19708–19713 (2010). doi:10.1073/pnas.1002780107

32. G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, A. Zeilinger, Violation of Bell's inequality under strict Einstein locality conditions. Phys. Rev. Lett. **81**, 5039–5043 (1998). doi:10.1103/PhysRevLett.81.5039