New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

The binary-outcome detection loophole

To cite this article: Thomas Cope 2021 *New J. Phys.* **23** 073032

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
16 March 2021

REVISED
11 June 2021

ACCEPTED FOR PUBLICATION
1 July 2021

PUBLISHED
22 July 2021

PAPER

The binary-outcome detection loophole

Thomas Cope*

Institut für Theoretische Physik, Leibniz Universität Hannover, Appelstr. 2, 30167 Hannover, Germany
* Author to whom any correspondence should be addressed.

E-mail: thomas.cope@itp.uni-hannover.de

Keywords: quantum information, detection loophole, quantum foundations, nonlocal correlations, hidden variables

Abstract

The detection loophole problem arises when quantum devices fail to provide an output for some runs. If treating these devices in a device-independent manner, failure to include the unsuccessful runs in the output statistics can lead to an adversary falsifying security i.e. Bell inequality violation. If the devices fail with too high frequency, known as the *detection threshold*, then no security is possible, as the full statistics cannot violate a Bell inequality. In this work we provide an intuitive local hidden-variable strategy that the devices may use to falsify any two-party, binary-outcome no-signalling distribution up to a threshold of $2(m_\mathrm{A}+m_\mathrm{B}-8)/(m_\mathrm{A}m_\mathrm{B}-16)$, where m_A, m_B refer to the number of available inputs choices to the two parties. This value is the largest analytically predicted lower bound for no-signalling distributions. We strongly conjecture it gives the true detection threshold for $m_\mathrm{A}=m_\mathrm{B}$, and for computationally tractable scenarios we provide the Bell inequality which verifies this. We also prove that a non-trivial detection threshold remains, even when allowing one party an arbitrary number of input choices.

1. Introduction

Due to the scales on which it operates, quantum technology faces the challenge of single photons or electrons being lost to the environment. This can result in devices failing to give any output. Ignoring these failures leads to the 'detection loophole' [1–3] security flaw. This is where a preprogrammed 'hidden-variable' device can falsely appear to exhibit non-local behaviour. Non-locality is necessary for the security proofs of device-independent quantum cryptography [4–11], therefore understanding and preventing the detection loophole is an extremely relevant problem.

One important question to consider is how low the rate of successful detection events (the efficiency) can be before all observed correlations are describable by a local realistic model. Knowing this threshold allows one to set minimum requirements for commercial devices and benchmark current technology. However, obtaining this bound for quantum states is generally difficult due to the infinite set of extremal quantum correlations, and only a few optimal constructions are known [12, 13].

In this article we present an intuitive local hidden-variable (LHV) construction for two parties, arbitrary inputs, and binary outputs, which will be able to reproduce any no-signalling distribution obtained by the successful runs, up to a detection efficiency dependent on the number of inputs. This provides a lower bound on the threshold for quantum measurements in the same scenario. When both parties have the same number of inputs into their device, this construction achieves numerically known thresholds (for general no-signalling distributions) leading us to conjecture it is optimal for this symmetric case. We furthermore show that in cases with an asymmetric number of measurements, increasing the number of Bob's measurements m_B above $2^{\lceil\log_2 m_\mathrm{A}\rceil}$ provides no additional power in verifying non-local correlations.

Bell's seminal theorem [14] and its subsequent generalisations [15–17] give fundamental constraints on the correlations exhibited by any local realistic model; constraints that quantum theory can violate. These violations have been confirmed experimentally [18, 19]. Due to limitations on technology however, to show Bell violations they relied on a 'fair-sampling' assumption; that the device failures were non-malicious and the successful detections were representative of the underlying system. In cryptographic protocols however, we cannot make that assumption, allowing an adversary (Eve) to pre-program the device to fail. It was not until much later that loophole-free violations, with no fair-sampling assumptions, were experimentally demonstrated [20–22]. The difficulty involved in closing this loophole highlights the importance of obtaining the best theoretical thresholds possible, so that minimal technological developments are required to perform secure protocols.

2. Preliminaries

In this paper, we are working in the *device-independence framework*. We assume that two parties (named Alice and Bob) have been distributed a joint system, on which they can make measurement choices, also referred to as inputs (labelled by x for Alice, and y for Bob) and receive outcomes (labelled a for Alice and b for Bob). We characterise the joint system only by the conditional probability distribution $p(ab|xy)$, making no assumptions about the underlying state or measurements made. This is known as a *black box* description. However, we do assume that Alice and Bob can isolate their systems, also referred to here as *devices*, from communicating with each other. This imposes the *no-signalling conditions*

$$\sum_b p(ab|xy) = \sum_b p(ab|xy') \quad \forall\, a, x, y, y', \tag{1}$$

$$\sum_a p(ab|xy) = \sum_a p(ab|x'y) \quad \forall\, b, x, x', y. \tag{2}$$

When the number of inputs and outputs are finite, so that $x \in \{0\dots m_{\mathrm{A}}-1\}$, $y \in \{0\dots m_{\mathrm{B}}-1\}$, $a \in \{0\dots n_{\mathrm{A}}-1\}$, $b \in \{0\dots n_{\mathrm{B}}-1\}$, then we may express any no-signalling probability distribution via the vector $\mathbf{p} := [p(00|00)\dots p(n_{\mathrm{A}}-1 n_{\mathrm{B}}-1|m_{\mathrm{A}}-1 m_{\mathrm{B}}-1)]$. The set of such vectors forms a convex set with finitely many extremal points, known as the no-signalling polytope, $\mathcal{NS}$. This restriction is known as the $(m_{\mathrm{A}}, m_{\mathrm{B}}, n_{\mathrm{A}}, n_{\mathrm{B}})$-scenario.

Within this set is a strict subset [23] of *quantumly realisable* distributions, $\mathcal{Q}$. Unlike the full no-signalling space, $\mathcal{Q}$ has an infinite number of extremal points, making it more difficult to deal with computationally. Strictly contained within $\mathcal{Q}$ is the set of local distributions, $\mathcal{L}$. Any distribution $p(ab|xy)$ within $\mathcal{L}$ has a LHV model of the form $p(ab|xy) = \int_\Lambda \mathrm{d}\lambda \rho(\lambda) p(a|x,\lambda) p(b|y,\lambda)$. These distributions may always be expressed as convex combinations of *deterministic* distributions $p(ab|xy) = \delta_{a,a_x}\delta_{b,b_y}$, which are finite in number. Geometrically, this means the structure of $\mathcal{L}$ is also a *polytope*.

$\mathcal{L}$ may be equivalently described by a set of *Bell inequalities*, linear inequalities of the form $\sum_{a,b,x,y} s^{xy}_{ab} p(ab|xy) \leqslant k$, where $p(ab|xy)$ is our input-conditional joint distribution [24]. There is a finite set of *facet* Bell inequalities; if all facets are satisfied by $p(ab|xy)$ it must have a LHV model i.e. it belongs to $\mathcal{L}$. Thus violation of a Bell inequality is used to prove the impossibility of a LHV model. We will also often denote a Bell inequality by a vector $\mathbf{s} = (s^{00}_{00}\dots s^{m_{\mathrm{A}}-1,m_{\mathrm{B}}-1}_{n_{\mathrm{A}}-1,n_{\mathrm{B}}-1})$, though one must also state the sign and magnitude of the inequality.

The typical detection loophole model; and the one considered in this article, is one in which the devices fail to detect with equal probability independently of each other [25]. Whilst not completely general, it is how we would expect the device to behave if the failures were ‘honest’; if we see autocorrelations, or correlations between the joint failures; this is a clear signal of adversarial manipulation. The model considered here adds an extra output to both parties to alter the original distribution $p(ab|xy)$ in the following way:

$$\begin{aligned} p_\eta(ab|xy) &= \eta^2 p(ab|xy),\\ p_\eta(Fb|xy) &= \eta(1-\eta)p(b|y),\\ p_\eta(aF|xy) &= \eta(1-\eta)p(a|x),\\ p_\eta(FF|xy) &= (1-\eta)^2. \end{aligned} \tag{3}$$

One can see this as a linear map $D_\eta : \mathbf{p} \to \mathbf{p}_\eta$, from the set of no-signalling distributions in the $(m_{\mathrm{A}}, m_{\mathrm{B}}, n_{\mathrm{A}}, n_{\mathrm{B}})$-scenario to those in the $(m_{\mathrm{A}}, m_{\mathrm{B}}, n_{\mathrm{A}}+1, n_{\mathrm{B}}+1)$-scenario. The quantity we are interested in is the (quantum) *critical detection efficiency*, $\eta_{\mathrm{c}} := \inf\{\eta|\exists \mathbf{p} \in \mathcal{Q}, \mathbf{p}_\eta \notin \mathcal{L}\}$, where $\mathcal{Q}, \mathcal{L}$ are considered in the $(m_{\mathrm{A}}, m_{\mathrm{B}}, n_{\mathrm{A}}, n_{\mathrm{B}})$-scenario and $(m_{\mathrm{A}}, m_{\mathrm{B}}, n_{\mathrm{A}}+1, n_{\mathrm{B}}+1)$-scenario respectively.

Table 1. Cases for which the no-signalling threshold has been numerically calculated; these provide a lower bound on the corresponding critical detection efficiency for the quantum set. The $*$ indicates numerical evaluation was not attained [26]. Reprinted table with permission from [26], Copyright (2019) by the American Physical Society.

m_{A}	m_{B}				
	2	3	4	5	6
2	2/3	2/3	2/3	2/3	2/3
3		4/7	5/9	5/9	5/9
4			1/2	1/2	1/2
5				4/9	$*$

To check the membership criterion $\mathbf{p}_{\eta} \in \mathcal{L}$, we can calculate the *local weight*. This is defined for an arbitrary distribution $\mathbf{q}$ as:

$$\max_{w\in[0,1]} \mathbf{q} = w\mathbf{q}^{\mathcal{L}} + (1-w)\mathbf{q}', \tag{4}$$

where $\mathbf{q}^{\mathcal{L}}$ is a local distribution and $\mathbf{q}'$ is a general no-signalling distribution. This linear program (see the appendix for details) gives $w = 1$ iff $\mathbf{q}$ is local.

For a given $(m_{\text{A}}, m_{\text{B}}, n_{\text{A}}, n_{\text{B}})$-scenario, we can use the linear weight to lower bound the critical detection threshold η_{c} in the following way. For every extremal no-signalling distribution $\mathbf{p}_j^{\text{NS}}$, we can calculate the local weight of successive distributions $\mathbf{p}_{j,\eta}^{\text{NS}}$—allowing us (e.g. by the binary chop algorithm) to determine the detection threshold of that particular distribution, η_j. By doing this for all extremal points, we find that at $\eta^* = \min_j \eta_j$, the entire $(m_{\text{A}}, m_{\text{B}}, n_{\text{A}}, n_{\text{B}})$ no signalling space is mapped into the $(m_{\text{A}}, m_{\text{B}}, n_{\text{A}}+1, n_{\text{B}}+1)$ local polytope. Thus, η^* is necessarily a lower bound of η_{c}. We will refer to η^* as the *no-signalling threshold*.

This bounding technique was performed in [26] on $m_{\text{A}}, m_{\text{B}} \leqslant 6$ and $n_{\text{A}} = n_{\text{B}} = 2$ for both parties, until the exponential growth in the number of extremal $\mathcal{NS}$ points became too large for numerical calculations.

Reproducing the table of thresholds from [26] in table 1, there are two patterns one observes immediately; that for $m_{\text{A}} = m_{\text{B}} = m$ the bound appears to match $4/(m+4)$, and that, if one fixes m_{A}, the bound for m_{B} decreases with each additional output until $m_{\text{B}} = 2^{\lceil \log m_{\text{A}} \rceil}$. In this article we prove that indeed the threshold for all $m_{\text{A}} = m_{\text{B}} = m$ is bounded below by $4/(m+4)$, and that it remains constant for all $m_{\text{B}} \geqslant 2^{\lceil \log m_{\text{A}} \rceil}$. Instead of doing this via numerical results, we construct an explicit LHV for all $\mathbf{p}_{\eta}$ up to this threshold value.

3. Pre-existing local hidden variable constructions

In order to understand our explicit construction, it is first useful to compare it to a LHV construction for the detection loophole introduced in [25]. Valid for any number of outputs, the construction is simple yet elegant. To emphasise the idea that Alice and Bob's devices are working against them, we introduce Alexa and Boris as the names of their devices, whose goal is to falsify an arbitrary non-local distribution. Beforehand they may agree a strategy (using the LHV λ) but cannot communicate once they have received their input choices. Between themselves, Alexa and Boris first randomly choose a leader, with bias towards Alexa $\alpha \in [0, 1]$; let us suppose for this run Alexa is chosen. They then generate uniformly a prediction for Alexa's input; say $k \in \{0, \ldots m_{\text{A}} - 1\}$. Finally they agree on an output $a \in \{0, \ldots n_{\text{A}} - 1\}$ for Alexa according to her desired marginal probability $p(a|k)$. When separated, once Alexa receives her input, if they have guessed correctly she will return outcome a. If the input received from Alice does not match their prediction, then Alexa outputs a failed detection F. It is clear this occurs with probability $(m_{\text{A}} - 1)/m_{\text{A}}$. Meanwhile, Boris receives his input and returns $b \in \{0, \ldots n_{\text{B}} - 1\}$ according to $p(ab|ky)/p(a|k)$ regardless. Notice that they never jointly output a failure, so in order to fully reproduce inefficient statistics they must with some probability β agree to both output F, regardless of input. This strategy gives rise to the statistics:

$$\begin{aligned}
p^{\text{LHV}}(ab|xy) &= (1-\beta)\left(\frac{\alpha}{m_{\text{A}}} + \frac{1-\alpha}{m_{\text{B}}}\right)p(ab|xy),\\
p^{\text{LHV}}(Fb|xy) &= (1-\beta)\alpha\frac{m_{\text{A}}-1}{m_{\text{A}}}p(b|y),\\
p^{\text{LHV}}(aF|xy) &= (1-\beta)(1-\alpha)\frac{m_{\text{B}}-1}{m_{\text{B}}}p(a|x),\\
p^{\text{LHV}}(FF|xy) &= \beta.
\end{aligned} \tag{5}$$

$$\begin{array}{c|cccccccc|c}
 & & & \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad}^{m_{\mathrm{B}}-2-g}} & \multicolumn{3}{c}{\overbrace{\qquad\qquad}^{g}} & \\
 & S & S & S & \dots & S & L & \dots & L & \\
 & S & A & S/A & \dots & S/A & L & \dots & L & \\
 & S & S/A & S/A & \dots & S/A & L & \dots & L & \\
 & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \left.\right\} m_{\mathrm{A}}-2-h \\
 & S & S/A & S/A & \dots & S/A & L & \dots & L & \\
 & K & K & K & \dots & K & M & \dots & M & \\
 & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \left.\right\} h \\
 & K & K & K & \dots & K & M & \dots & M &
\end{array}$$

$$S=\begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix} \qquad A=\begin{pmatrix}0 & \frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix} \qquad K=\begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ 0 & 0\end{pmatrix} \qquad L=\begin{pmatrix}\frac{1}{2} & 0\\ \frac{1}{2} & 0\end{pmatrix} \qquad M=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$$

Figure 1. The structure of an extremal non-local distribution of $\mathcal{NS}$, where $g \in \{0, 1, \dots m_{\mathrm{B}}-2\}$, $h \in \{0, 1, \dots m_{\mathrm{A}}-2\}$. If g or h is non-zero, then the distribution is a lower input-number extremal point with local deterministic inputs appended to it. As such, its detection threshold cannot be lower than that of the $m_{\mathrm{A}}-h, m_{\mathrm{B}}-g$ case; and is not generally optimal for $m_{\mathrm{A}}, m_{\mathrm{B}}$ inputs.

One can equate equations (3) and (5) to find this LHV strategy can reproduce statistics up to $\eta \leqslant \frac{m_{\mathrm{A}}+m_{\mathrm{B}}-2}{m_{\mathrm{A}}m_{\mathrm{B}}-1}$. By comparison to results in table 1, one can easily check for e.g. $m_{\mathrm{A}} = m_{\mathrm{B}} = 3$ this is not optimal.

4. A new local hidden variable construction

4.1. The model

We will look to improve this strategy on extremal binary-output $\mathcal{NS}$ points, thereby bounding the threshold for the entire space. To do this, we need to understand better the extremal points themselves. Fortunately, for binary outputs a complete characterisation has been provided in [27]. One can see their general form in figure 1. They may also be expressed in the simple form

$$p(ab|xy) = \begin{cases} 1/2 & \text{if} \quad a \oplus b = G(x,y) = \sum\limits_{i=1}^{2^{n_y}} Q_i(x)R_i(y) \equiv \sum\limits_{j=1}^{2^{n_x}} S_j(y)T_j(x) \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

where $Q_i(x)$ are polynomials in the binary digits[1] of x, which we label $\mathbf{x}_2$, and $R_i(y)$ are monomials in the binary digits of y (labelled $\mathbf{y}_2$). Similarly, $S_j(y)$ are polynomials of $\mathbf{y}_2$ and $T_j(x)$ monomials. $n_x = \lceil \log_2 m_{\mathrm{A}} \rceil$ is the length of $\mathbf{x}_2$ and similarly for n_y. The most famous example of this is the (generalised) PR box [23], which has the form

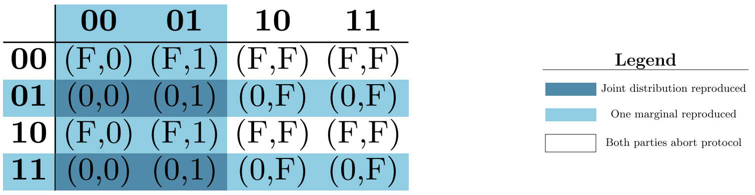
$$p(ab|xy) = \begin{cases} 1/2 & \text{if} \quad a \oplus b = \mathbf{x}_2 \cdot \mathbf{y}_2 \mod 2 \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

For all the numerically evaluated cases presented in table 1, the generalised PR box achieves the no-signalling threshold η^*.

In particular, given any extremal $\mathcal{NS}$ point, the conditional output distribution for two input pairs either match exactly or are exactly anti-matching. This allows the following strategy: Alexa and Boris with probability α randomly choose a leader; suppose it is Alexa. They generate uniformly a prediction for Alexa's input; say $k_1 \in \{0 \dots m_{\mathrm{A}}-1\}$; then another from the remaining $m_{\mathrm{A}}-1$ choices a second prediction, $k_2 \in \{0 \dots m_{\mathrm{A}}-1\} \backslash \{k_1\}$. They also with probability 1/2 decide whether they will use a matching or unmatching strategy. Finally, they decide uniformly on a value for a, $a_{\mathrm{L}} \in \{0, 1\}$. Once Alexa receives her input, if it matches k_1 she returns outcome a_{L}. If she receives k_2, for the matching strategy she returns a_{L}, and if they are following the unmatching strategy $a_{\mathrm{L}} \oplus 1$. If her input does not match k_1 or k_2, then she outputs a failed detection F. It is clear this occurs with probability $(m_{\mathrm{A}}-2)/m_{\mathrm{A}}$. Meanwhile, Boris receives his input $z \in \{0 \dots m_{\mathrm{B}}-1\}$ and checks if $G(k_1, z) = G(k_2, z)$. If these values match *and* they chose the matching strategy he outputs $a_{\mathrm{L}} \oplus G(k_1, z)$, otherwise he outputs F. If the two values are unequal and they chose the unmatching strategy he outputs $a_{\mathrm{L}} \oplus G(k_1, z)$, and F otherwise. They still with some

[1] $\oplus$ denotes addition modulo 2.

Random Variable	Abort Protocol	Leader (L)	k_1	k_2	Match	a
Distribution	$P(\text{yes}) = \beta$	$P(\text{Alexa}) = \alpha$	$\mathbb{U}(\{0 \dots M_{\mathrm{L}} - 1\})$	$\mathbb{U}(\{0 \dots M_{\mathrm{L}} - 1\} \backslash \{k_1\})$	$P(\text{Match}) = 1/2$	$\mathbb{U}(\{0, 1\})$
Outcome of λ	no	Alexa	2	4	Match	0



	00	01	10	11
00	(F,0)	(F,1)	(F,F)	(F,F)
01	(0,0)	(0,1)	(0,F)	(0,F)
10	(F,0)	(F,1)	(F,F)	(F,F)
11	(0,0)	(0,1)	(0,F)	(0,F)

Figure 2. An illustration of the improved LHV strategy. The table gives the distribution for each component variable of the underlying randomness λ. The matrix highlights for which input pairs given to Alexa and Boris they can reproduce the output correlations successfully, and when they must abort, for an example λ. The detection threshold for which one can reproduce all correlations is found by averaging over all possible λ outcomes.

probability β agree to both output F, regardless of input. This gives statistics:

$$\begin{aligned}
p^{\mathrm{LHV}}(ab|xy) &= (1-\beta)\left(\alpha\frac{1}{m_{\mathrm{A}}} + (1-\alpha)\frac{1}{m_{\mathrm{B}}}\right)p(ab|xy),\\
p^{\mathrm{LHV}}(Fb|xy) &= (1-\beta)\left(\alpha\frac{m_{\mathrm{A}}-2}{2m_{\mathrm{A}}} + (1-\alpha)\frac{1}{m_{\mathrm{B}}}\right)p(b|y),\\
p^{\mathrm{LHV}}(aF|xy) &= (1-\beta)\left(\alpha\frac{1}{m_{\mathrm{A}}} + (1-\alpha)\frac{m_{\mathrm{B}}-2}{2m_{\mathrm{B}}}\right)p(a|x),\\
p^{\mathrm{LHV}}(FF|xy) &= \beta + (1-\beta)\left(\alpha\frac{m_{\mathrm{A}}-2}{2m_{\mathrm{A}}} + (1-\alpha)\frac{m_{\mathrm{B}}-2}{2m_{\mathrm{B}}}\right).
\end{aligned} \tag{8}$$

The advantage of such a strategy becomes apparent in the final term; to achieve the joint failure rate $(1-\eta)^2$, they can devote fewer runs to deterministically outputting FF, since their guessing strategy will also output a joint failure some of the time; unlike the single input guessing strategy. Equating equations (3) and (8) one finds one can replicate $\eta \leqslant 2(m_{\mathrm{A}} + m_{\mathrm{B}} - 8)/(m_{\mathrm{A}}m_{\mathrm{B}} - 16)$. In the case where $m_{\mathrm{A}} = m_{\mathrm{B}} = m$, one can see this simplifies[2] to $4/(m+4)$, which matches the known no-signalling threshold in numerically evaluated cases.

4.2. Asymptotic power of the model

We now prove that the no-signalling detection threshold cannot be improved by increasing asymmetrically one party's possible measurements beyond the limit $m_{\mathrm{B}} = 2^{\lceil \log_2 m_{\mathrm{A}} \rceil}$. One may express any extremal point as having $p(ab|xy) = 1/2$ when $a \oplus b = G(x, y) = \sum_{j=1}^{2^{nx}} S_j(y)T_j(x)$, with $n_x = \lceil \log_2 m_{\mathrm{A}} \rceil$. In particular this implies there are at *most* 2^{n_x} functions of x defined by the inputs of Bob. Equivalently, it implies that for any extremal point of a scenario with $m_{\mathrm{B}} > 2^{n_x}$, then for any input choice $y > 2^{n_x}$ the joint distribution $p(ab|xy)$ is identical to the joint distribution $p(ab|xy')$ of some $y' \leqslant 2^{n_x}$, $\forall\, a, b, x$. Therefore, if one has a valid LHV strategy for $m_{\mathrm{A}}, m_{\mathrm{B}} = 2^{\lceil \log_2 m_{\mathrm{A}} \rceil}$ inputs up to efficiency η; one also has a valid strategy for all $m_{\mathrm{A}}, m_{\mathrm{B}} > 2^{\lceil \log_2 m_{\mathrm{A}} \rceil}$ which will also achieve efficiency η. This strategy simply treats $y > 2^{\lceil \log_2 m_{\mathrm{A}} \rceil}$ identically to the corresponding $y' \leqslant 2^{\lceil \log_2 m_{\mathrm{A}} \rceil}$.

4.3. Comparison to numerically known no-signalling thresholds

Although the bound derived in the previous section holds for all pairs $(m_{\mathrm{A}}, m_{\mathrm{B}})$, we see from the numerical evidence in table 1 it is not generally tight. In the case where $m_{\mathrm{A}} = 3$, $m_{\mathrm{B}} = 4$, we know the no-signalling detection threshold to be $\eta^* = 5/9$; however, the hidden variable strategy we have proposed only simulates arbitrary distributions up to $\eta = 1/2$. To reproduce correlations up to η^*, one can *mix* our strategy with the pre-existing one [25] presented earlier in this paper. By choosing the pre-existing strategy, which guesses a single input, 20% of the time and our strategy, predicting two inputs, 80% of the time, and by choosing Alexa solely as the leader for both strategies one can achieve $\eta \leqslant 5/9$. This mixing of strategies does not extend to higher dimensional asymmetric scenarios though; for $m_{\mathrm{A}} = 5$, $m_{\mathrm{B}} = 6$ no combination of the two strategies beats the bound given by equation (8).

[2] The simplified bound also holds for $m_{\mathrm{A}} = m_{\mathrm{B}} = 4$, since cancellation prevents the denominator vanishing.

As the number of input choices increases, one could propose a more general variation; in which the leader (say Alexa) chooses many input predictions $k_1 \ldots k_n \in \{0 \ldots m_{\mathrm{A}} - 1\}$, $n \leqslant m_{\mathrm{A}}$. With this strategy, they must beforehand predict whether $G(z_i k_i)$ will coincide with $G(z, k_1)$, for each $i = 2 \ldots n$. This is analogous to the 'matching/unmatching' choice seen earlier. The probability of guessing this correctly scales as 2^{n-1}. However, the benefit of predicting additional inputs only scales as n/m_{A}. This implies the probability of a correct output will scale as $\frac{n}{m_{\mathrm{A}}} \frac{1}{2^{n-1}}$, which takes its maximal value at $n = 1, 2$ only. Trying to incorporate this strategy to simulate $m_{\mathrm{A}} = 5, m_{\mathrm{B}} = 6$ distributions, our optimisation never chose strategies with $n > 2$. This suggests for the asymmetric case a more nuanced joint strategy is required. However, we stress that when $m_{\mathrm{A}} = m_{\mathrm{B}}$, the bound predicted by this model matches all numerically obtained bounds.

In order to prove that our conjecture of $\eta^* = 4/(m+4)$ for $m_{\mathrm{A}} = m_{\mathrm{B}} = m$ is correct, one would need to provide an extremal $\mathcal{NS}$ distribution p^{NS}, and corresponding Bell inequality $s^{xy}_{a'b'}$, such that $\sum_{a',b',x,y} s^{xy}_{a'b'} p^{\mathrm{NS}}_{\eta}(a'b'|xy) \nleq k$, $\forall\, \eta > \eta^*$. Here we have used a', b' to explicitly remind the reader that a' ranges both in the original values of a and F; that is, it is a three-outcome inequality. From numerical results, the generalised PR box is the best candidate for the extremal point, but we found no obvious generalisation of the witnessing Bell inequalities, which are provided for evaluated cases in the appendix.

4.4. Comparison to quantumly realisable thresholds

As stated above, to prove the no-signalling threshold η^* for a given scenario requires a Bell inequality violation $\sum_{a',b',x,y} s^{xy}_{a'b'} p^{\mathrm{NS}}_{\eta}(a'b'|xy) \nleq k$, $\forall \eta > \eta^*$. Therefore $s^{xy}_{a'b'}$ is the 'optimal' Bell inequality, in that it detects non-locality for all efficiencies above the no-signalling threshold. A natural question is whether the same Bell inequality is optimal with respect to quantum correlations; i.e. $\sum_{a',b',x,y} s^{xy}_{a'b'} p^{\mathrm{Q}}_{\eta}(a'b'|xy) \nleq k$, $\forall \eta > \eta_c$.

For quantum correlations, a critical efficiency of $\eta_{\mathrm{c}} = 2/3$ is achievable via qubits using the $m_{\mathrm{A}} = m_{\mathrm{B}} = 2$ CHSH inequality [12], whilst testing ququarts with a $m_{\mathrm{A}} = m_{\mathrm{B}} = 4$ inequality allows a critical efficiency of $(\sqrt{5} - 1)/2 \approx 0.618$ [13]. The respective Bell inequalities verifying non-locality for efficiencies higher than the critical efficiency, when applied to the generalised PR box achieve the no-signalling detection threshold, η^*, for their respective scenarios. These inequalities are somewhat special in that they are 'lifted inequalities'; they are of the form

$$s_{\mathrm{F}b|xy} = s_{a_x b|xy}, s_{a\mathrm{F}|xy} = s_{ab_y|xy}, \; a_x, b_y \in \{0, 1\} \quad \forall\; x, y \tag{9}$$

i.e. facet two-outcome inequalities where F is treated identically to one of the valid outputs. In contrast, the optimal Bell inequality for $m_{\mathrm{A}} = m_{\mathrm{B}} = 3$ requires a truly new three-output inequality; something noted in [28].

In order to test whether our optimal Bell inequalities could lead to new quantum constructions, we employed the NPA hierarchy of correlations [29]. These allow one to define successively tighter outer approximations to $\mathcal{Q}$, which we label $\mathcal{Q}_1 \supset \mathcal{Q}_2 \ldots \supset \mathcal{Q}$. For a fixed η, we can then employ semidefinite programming to look for a set of correlations such that $\mathbf{p} \in \mathcal{Q}_i, \mathbf{s} \cdot \mathbf{p}_{\eta} \nleq k$, which implies $\mathbf{p}_{\eta} \notin \mathcal{L}$. It is then clear that, if for a given i, $\tilde{\eta}$ no such $\mathbf{p}$ is found, then $\{\mathbf{p} \in \mathcal{Q}, \mathbf{s} \cdot \mathbf{p}_{\tilde{\eta}} \nleq k\}$ must also be empty.

For the scenarios $m_{\mathrm{A}} = m_{\mathrm{B}} = 3$ and $m_{\mathrm{A}} = 3, m_{\mathrm{B}} = 4$, we know the quantum critical efficiency is not higher than $2/3$; since we may always embed the CHSH/qubit construction into these scenarios. Therefore, an improvement in the quantum critical efficiency would require that $\{\mathbf{p} \in \mathcal{Q}, \mathbf{s} \cdot \mathbf{p}_{2/3} \nleq k\}$ is non-empty. However, in both scenarios, choosing $\mathbf{s}$ as the optimal Bell inequality for non-locality, we find that this set is empty at level $\mathcal{Q}_2$ of the hierarchy; thus these inequalities do not help us to improve the quantum critical efficiency, η_{c}.

5. Conclusions and discussion

In this paper, we have exploited the structure of the bipartite binary-output no-signalling polytope in order to provide a lower bound on the detection loophole critical efficiency for an arbitrary number of inputs. We have done this by constructing an explicit LHV model valid for all extremal points. Numerical evidence suggests that when Alice and Bob share an equal number of inputs, this construction is optimal. An open question is whether one can find a family of Bell inequalities verifying this.

One possible extension to this work would be improve the strategy for asymmetric measurement capabilities; since we know our model does not provide a tight bound for $m_{\mathrm{A}} = 5, m_{\mathrm{B}} = 6$. A further generalisation would be to test if this approach generalises to a larger number of outputs. Unfortunately, the vertices of higher output no-signalling polytopes are not generally known, so we cannot say much about their structure. Considering the results here, one would expect the successful simulation efficiency of a construction which predicts n inputs in a k-output scenario to scale as $\frac{n}{m_\mathrm{A}}\frac{1}{k^{n-1}}$, which for $k > 2$ achieves optimal integer value only at $n = 1$. This suggests for higher output-number scenarios the construction of [25], defining equation (5), may be optimal.

Acknowledgments

This work was supported, in part, by the Quantum Valley Lower Saxony (QVLS), the DFG through SFB 1227 (DQ-mat), the RTG 1991, and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy-EXC-2123 Quantum Frontiers-390837967. We would like to thank Tobias J Osborne, Reinhard F Werner and Le Phuc Thinh for useful discussions. The publication of this article was funded by the Open Access Fund of the Leibniz Universität Hannover.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Local weight linear program

In order to calculate the linear weight of an arbitrary distribution $\mathbf{q}$, we solve the following problem:

$$\text{Maximise}\sum_i \alpha_i, \text{subject to}: \sum_i \alpha_i \mathbf{q}_i^{\mathcal{L}} \leqslant \mathbf{q}, \quad \alpha_i \geqslant 0.$$

where $\mathbf{q}_i^{\mathcal{L}}$ are the extremal points of the polytope $\mathcal{L}$. By rearranging the inequality, we see that the leftover distribution $\mathbf{q}' := \mathbf{q} - \sum_i \alpha \mathbf{q}_i^{\mathcal{L}}$ has all positive entries, and satisfies the no-signalling constraints since so too do $\mathbf{q}, \mathbf{q}_i^{\mathcal{L}}$. Therefore it is a valid (sub-normalised) distribution. This linear program therefore looks to optimise the total weight of the local extremal points over all decompositions of $\mathbf{q}$.

It is also worth mentioning that every linear program has a dual with the same optimal value [30]. The dual of the above function gives us a vector b such that:

$$\mathbf{b}^\mathrm{T}\mathbf{q} = \sum_i \alpha_i, \;\; \mathbf{b}^\mathrm{T}\mathbf{q}_i^{\mathcal{L}} \geqslant 1 \quad \forall\, i$$

we see immediately that if $\sum_i \alpha_i < 1$, this gives us a Bell inequality violated by $\mathbf{q}$.

Appendix B. Bell inequalities which verify the threshold

In this supplemental file, the optimal Bell inequalities are provided to achieve the detection loophole threshold for the generalised PR box. They are presented in matrix format:

$$S = \left(\begin{array}{ccc|c|ccc}
s^{0,0}_{0,0} & \cdots & s^{0,0}_{0,n_\mathrm{B}} & \cdots\;\cdots\;\cdots\;\cdots & s^{0,m_\mathrm{B}-1}_{0,0} & \cdots & s^{0,m_\mathrm{B}-1}_{0,n_\mathrm{B}} \\
\vdots & \ddots & \vdots & \cdots\;\cdots\;\cdots\;\cdots & \vdots & \ddots & \vdots \\
s^{0,0}_{n_\mathrm{A},0} & \cdots & s^{0,0}_{n_\mathrm{A},n_\mathrm{B}} & \cdots\;\cdots\;\cdots\;\cdots & s^{0,m_\mathrm{B}-1}_{n_\mathrm{A},0} & \cdots & s^{0,m_\mathrm{B}-1}_{n_\mathrm{A},n_\mathrm{B}} \\
\hline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
\hline
s^{m_\mathrm{A}-1,0}_{0,0} & \cdots & s^{m_\mathrm{A}-1,0}_{0,n_\mathrm{B}} & \cdots\;\cdots\;\cdots\;\cdots & s^{m_\mathrm{A}-1,m_\mathrm{B}-1}_{0,0} & \cdots & s^{m_\mathrm{A}-1,m_\mathrm{B}-1}_{0,n_\mathrm{B}} \\
\vdots & \ddots & \vdots & \cdots\;\cdots\;\cdots\;\cdots & \vdots & \ddots & \vdots \\
s^{m_\mathrm{A}-1,0}_{n_\mathrm{A},0} & \cdots & s^{m_\mathrm{A}-1,0}_{n_\mathrm{A},n_\mathrm{B}} & \cdots\;\cdots\;\cdots\;\cdots & s^{m_\mathrm{A}-1,m_\mathrm{B}-1}_{n_\mathrm{A},0} & \cdots & s^{m_\mathrm{A}-1,m_\mathrm{B}-1}_{n_\mathrm{A},n_\mathrm{B}}
\end{array}\right), \qquad (10)$$

where the solid lines delineate different inputs. All presented inequalities have local bound $\geqslant$1. Note that there are $n_\mathrm{A} + 1$ ($n_\mathrm{B} + 1$) outputs to account for the additional output F.

B.1. Optimal inequality for two-inputs

As mentioned in the main body of the paper, this inequality is a 'lifting' of the CHSH inequality. For all measurements failure to output is treated identically to 0. Since other liftings of the same CHSH inequality achieve the optimal value, we can see generally there is not a single unique inequality that witnesses the threshold.

$$\left(\begin{array}{ccc|ccc}
0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 \\
\hline
0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1
\end{array}\right) \tag{11}$$

B.2. Optimal inequality for three-inputs

Unlike the previous case; this inequality is a 'true' three-input, three-output inequality; it cannot be created from lifting a previous, lower dimensional inequality. What is interesting to note is that, for the first two inputs for each party, failure is again treated identically to 0—it is only the final input which treats failure differently.

$$\left(\begin{array}{ccc|ccc|ccc}
0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 \\
\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\
0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 \\
\hline
0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 \\
\frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\
0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 \\
\hline
0 & \frac{2}{3} & 0 & 0 & \frac{2}{3} & 0 & \frac{2}{3} & 0 & \frac{2}{3} \\
\frac{2}{3} & 0 & \frac{2}{3} & 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{2}{3} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{2}{3} & \frac{2}{3}
\end{array}\right) \tag{12}$$

B.3. Optimal inequality for four-inputs

This inequality is also a lifting of a four-input, two-output inequality; however in this instance the choice of treating failure as 0 or 1 depends on the input.

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\
\frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} \\
0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\
\hline
0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\
0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & \frac{1}{4} & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
\frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\
\frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\
\hline
0 & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} & 0 \\
\frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} \\
0 & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} & 0
\end{array}\right) \tag{13}$$

B.4. Optimal inequality for an asymmetric case: Alice 3 inputs, Bob 4 inputs

For this asymmetric case we can again provide a Bell inequality which achieves the optimal threshold for the generalised PR Box. Like the previous cases, the inequality we provide here is a *facet* inequality; that is a maximally dimensional face of the local polytope. This is the first inequality provided where the failure outcome is treated differently from the valid outcomes for *all* input choices; we leave open the question whether this is necessary, or an artefact of this particular inequality.

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
0 & \frac{1}{3} & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\
\frac{5}{12} & 0 & \frac{1}{4} & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{3} & 0 & \frac{1}{12} & \frac{1}{2} & 0 & \frac{1}{4} \\
\frac{1}{2} & \frac{5}{12} & \frac{5}{12} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{12} & 0 & 0 \\
\hline
0 & \frac{1}{3} & \frac{1}{12} & 0 & 0 & 0 & 0 & \frac{7}{12} & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{6} \\
\frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{12} \\
0 & \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{4} & \frac{1}{6} & 0 & \frac{1}{4} & 0 \\
\hline
0 & \frac{1}{6} & \frac{1}{12} & 0 & 0 & 0 & \frac{3}{4} & 0 & \frac{1}{6} & \frac{5}{12} & 0 & \frac{5}{12} \\
\frac{1}{3} & 0 & \frac{1}{6} & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{6} \\
0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} & \frac{1}{4} & 0 & 0 & 0
\end{array}\right) \tag{14}$$

B.5. Optimal inequality for five-inputs

The previous inequalities provided were all calculated using exact arithmetic. Unfortunately this takes much longer than floating point methods, particularly as the dimension increases. Therefore, we are only able to provide a Bell inequality here which is accurate up to 6 s.f. and moreover, not a facet inequality. However, it still verifies the detection loophole threshold, and is included for completeness.

$$\begin{pmatrix}
0 & 144.708 & 0.047\,6284 & 0 & 147.002 & 0.056\,0145 & 0 & 147.002 & 0.056\,0145 & 0 \\
147.002 & 0.056\,0145 & 0 & 148.982 & 0.050\,4226 & 144.708 & 0 & 0.047\,6284 & 147.002 & 0 \\
0.056\,0145 & 147.002 & 0 & 0.056\,0145 & 147.002 & 0 & 0.056\,0145 & 148.982 & 0 & 0.050\,4226 \\
0.047\,6289 & 0.047\,6289 & 0.100\,928 & 0.049\,4062 & 0.049\,4062 & 0.032\,6552 & 0.049\,4062 & 0.049\,4062 & 0.032\,6552 & 0.049\,4062 \\
0.049\,4062 & 0.032\,6552 & 0.071\,2469 & 0.071\,2471 & 0.000\,105\,974 & 0 & 147.002 & 0.049\,4056 & 147.860 & 0 \\
0.058\,2717 & 0 & 149.205 & 0.051\,0420 & 147.860 & 0 & 0.058\,2717 & 0 & 145.578 & 0.057\,3425 \\
147.002 & 0 & 0.049\,4056 & 0 & 147.860 & 0.058\,2717 & 149.205 & 0 & 0.051\,0421 & 0 \\
147.860 & 0.058\,2717 & 145.578 & 0 & 0.057\,3426 & 0.056\,0145 & 0.056\,0145 & 0.032\,6547 & 0.058\,2717 & 0.058\,2717 \\
0.033\,6107 & 0.051\,0419 & 0.051\,0421 & 0.022\,1489 & 0.058\,2717 & 0.058\,2717 & 0.033\,6107 & 0.057\,8340 & 0.057\,8340 & 0.070\,8747 \\
0 & 147.002 & 0.049\,4056 & 0 & 149.205 & 0.051\,0420 & 147.860 & 0 & 0.058\,2717 & 147.860 \\
0 & 0.058\,2717 & 0 & 145.578 & 0.057\,3425 & 147.002 & 0 & 0.049\,4056 & 149.205 & 0 \\
0.051\,0421 & 0 & 147.860 & 0.058\,2717 & 0 & 147.860 & 0.058\,2717 & 145.578 & 0 & 0.057\,3426 \\
0.056\,0145 & 0.056\,0146 & 0.032\,6547 & 0.051\,0419 & 0.051\,0421 & 0.022\,1489 & 0.058\,2717 & 0.058\,2717 & 0.033\,6107 & 0.058\,2717 \\
0.058\,2717 & 0.033\,6107 & 0.057\,8340 & 0.057\,8340 & 0.070\,8747 & 0 & 147.002 & 0.049\,4056 & 147.860 & 0 \\
0.058\,2717 & 147.860 & 0 & 0.058\,2717 & 0 & 149.205 & 0.051\,0420 & 0 & 145.578 & 0.057\,3425 \\
147.002 & 0 & 0.049\,4056 & 0 & 147.860 & 0.058\,2717 & 0 & 147.860 & 0.058\,2717 & 149.205 \\
0 & 0.051\,0421 & 145.578 & 0 & 0.057\,3426 & 0.056\,0145 & 0.056\,0146 & 0.032\,6547 & 0.058\,2717 & 0.058\,2717 \\
0.033\,6107 & 0.058\,2717 & 0.058\,2717 & 0.033\,6107 & 0.051\,0419 & 0.051\,0421 & 0.022\,1489 & 0.057\,8340 & 0.057\,8340 & 0.070\,8747 \\
0 & 148.982 & 0.071\,2492 & 0 & 145.578 & 0.057\,8340 & 0 & 145.578 & 0.057\,8340 & 0 \\
145.578 & 0.057\,8340 & 149.070 & 0 & 0.066\,1534 & 148.982 & 0 & 0.071\,2493 & 145.578 & 0 \\
0.057\,8341 & 145.578 & 0 & 0.057\,8341 & 145.578 & 0 & 0.057\,8341 & 0 & 149.070 & 0.066\,1534 \\
0.050\,4204 & 0.050\,4204 & 0.000\,1060\,05 & 0.057\,3429 & 0.057\,3430 & 0.070\,8752 & 0.057\,3429 & 0.057\,3430 & 0.070\,8752 & 0.057\,3429
\end{pmatrix} \tag{15}$$

ORCID iDs

Thomas Cope https://orcid.org/0000-0003-4571-2872

References

[1] Pearle P M 1970 Hidden-variable example based upon data rejection *Phys. Rev.* D **2** 1418–25

[2] Clauser J F and Horne M A 1974 Experimental consequences of objective local theories *Phys. Rev.* D **10** 526–35

[3] Fine A 1982 Some local models for correlation experiments *Synthese* **50** 279

[4] Mayers D and Yao A 1998 Quantum cryptography with imperfect apparatus *Proc. 39th Annual Symp. on Foundations of Computer Science (FOCS-98)* (Los Alamitos, CAU) (IEEE Computer Society) pp 503–9

[5] Barrett J, Hardy L and Kent A 2005 No signalling and quantum key distribution *Phys. Rev. Lett.* **95** 010503

[6] Acín A, Gisin N and Masanes L 2006 From Bell's theorem to secure quantum key distribution *Phys. Rev. Lett.* **97** 120405

[7] Vazirani U and Vidick T 2014 Fully device-independent quantum key distribution *Phys. Rev. Lett.* **113** 140501

[8] Colbeck R 2007 Quantum and relativistic protocols for secure multi-party computation *PhD Thesis* University of Cambridge

[9] Pironio S *et al* 2010 Random numbers certified by Bell's theorem *Nature* **464** 1021–4

[10] Colbeck R and Kent A 2011 Private randomness expansion with untrusted devices *J. Phys. A: Math. Theor.* **44** 095305

[11] Miller C A and Shi Y 2014 Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices *STOC '14 Proc. 46th Annual ACM Symp. on Theory of Computing* (New York: ACM) pp 417–26

[12] Eberhard P H 1993 Background level and counter efficiencies required for a loophole-free Einstein–Podolsky–Rosen experiment *Phys. Rev.* A **47** R747–50

[13] Vértesi T, Pironio S and Brunner N 2010 Closing the detection loophole in Bell experiments using qudits *Phys. Rev. Lett.* **104** 060401

[14] Bell J S 1964 On the Einstein Podolsky Rosen paradox *Phys. Phys. Fiz.* **1** 195–200

[15] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **23** 880–4

[16] Froissart M 1981 Constructive generalization of Bell's inequalities *Il Nuovo Cimento* B **64** 241–51

[17] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 Bell nonlocality *Rev. Mod. Phys.* **86** 419

[18] Aspect A, Grangier P and Roger G 1981 Experimental tests of realistic local theories via Bell's theorem *Phys. Rev. Lett.* **47** 460–3

[19] Tittel W, Brendel J, Gisin B, Herzog T, Zbinden H and Gisin N 1998 Experimental demonstration of quantum correlations over more than 10 km *Phys. Rev.* A **57** 3229–32

[20] Giustina M *et al* 2015 Significant-loophole-free test of Bell's theorem with entangled photons *Phys. Rev. Lett.* **115** 250401

[21] Hensen B *et al* 2015 Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres *Nature* **526** 682–6

[22] Shalm L K *et al* 2015 Strong loophole-free test of local realism *Phys. Rev. Lett.* **115** 250402

[23] Popescu S and Rohrlich D 1994 Quantum nonlocality as an axiom *Found. Phys.* **24** 379–85

[24] Tsirelson B 1993 Some results and problems on quantum Bell-type inequalities *Hadronic J. Suppl.* **8** 329

[25] Massar S and Pironio S 2003 Violation of local realism vs detection efficiency *Phys. Rev.* A **68** 062109

[26] Cope T and Colbeck R 2019 Bell inequalities from no-Signalling distributions *Phys. Rev.* A **100** 022114

[27] Jones N and Masanes L 2005 Interconversion of nonlocal correlations *Phys. Rev.* A **72** 052312

[28] Wilms J, Disser Y, Alber G and Percival I 2008 Local realism, detection efficiencies, and probability polytopes *Phys. Rev.* A **78** 032116

[29] Navascués M, Pironio S and Acín A 2008 A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations *New J. Phys.* **10** 073013

[30] Boyd S and Vandenberghe L 2004 *Convex Optimization* (Cambridge: Cambridge University Press)