Optimal randomness generation from optical Bell experiments

2015 New J. Phys. 17 022003

(http://iopscience.iop.org/1367-2630/17/2/022003)

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

OPEN ACCESS

RECEIVED
27 November 2014

REVISED
20 January 2015

ACCEPTED FOR PUBLICATION
27 January 2015

PUBLISHED
10 February 2015

FAST TRACK COMMUNICATION

Optimal randomness generation from optical Bell experiments

Alejandro Máttar[1], **Paul Skrzypczyk**[1], **Jonatan Bohr Brask**[2], **Daniel Cavalcanti**[1] **and Antonio Acín**[1,3]

[1] ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, E-08860 Castelldefels (Barcelona), Spain
[2] Department of Theoretical Physics, University of Geneva, 1211 Geneva, Switzerland
[3] ICREA-Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, E-08010 Barcelona, Spain

E-mail: alejandro.mattar@icfo.es

Keywords: quantum optics, quantum information, random number generation, device independent, Bell nonlocality

Abstract

Genuine randomness can be certified from Bell tests without any detailed assumptions on the working of the devices with which the test is implemented. An important class of experiments for implementing such tests is optical setups based on polarization measurements of entangled photons distributed from a spontaneous parametric down conversion source. Here we compute the maximal amount of randomness which can be certified in such setups under realistic conditions. We provide relevant yet unexpected numerical values for the physical parameters and achieve four times more randomness than previous methods.

1. Introduction

Quantum systems have the potential to provide a strong form of randomness which cannot be attributed to incomplete knowledge of any classical variable of the system. At the basis of such genuine randomness lies a quantitative relation between the amount by which a Bell inequality is violated [1] and the degree of predictability of the results of the test [2]. Intuitively, the violation of a Bell inequality certifies the presence of nonlocal correlations [3], and in turn, this guarantees that the outcomes of the measurements cannot be determined in advance [4, 5]. Furthermore, this genuine randomness can be certified without any detailed assumptions about the internal working of the devices used, that is, in a 'device-independent' fashion. Device independence is advantageous since it provides immunity to attacks that exploit imperfections in the physical implementation, to which device-dependent protocols are susceptible [6]. For this reason, device-independent randomness generation has recently received much attention [7–12].

An intense research effort has been devoted to the experimental realization of device-independent randomness generation. A few years ago, Pironio *et al* [2] implemented the first proof-of-principle experiment. It involved two entangled atomic ion qubits confined in two independent vacuum chambers separated by approximately 1 m. This implementation, which was based on light–matter interaction, managed to certify 42 random bits over a period of one month.

The principal challenge for a device-independent randomness generation experiment is that it must close the detection loophole [13, 14], i.e. it must provide a Bell inequality violation without post-selection on the data, since otherwise violation can be faked by classical resources [15] and no genuine randomness can be guaranteed. The detection loophole was first successfully closed on several systems relying on light–matter interaction; see for instance [16–18]. Very recently it has been closed in optical setups [19, 20], based on polarization measurements of entangled photons distributed from a spontaneous parametric down-conversion (SPDC) source. These optical implementations represent an important achievement as they enable much higher rates of genuine random bits per time unit.

Given these experimental achievements, the natural question that arises is how to generate this genuine randomness efficiently. What is the maximal amount of randomness that a given physical implementation allows for? And most importantly, how should the relevant physical parameters of the setup be tuned to provide such an optimal amount? Here we answer these questions for the case of optical implementations based on SPDC, for which a thorough physical characterization has been recently presented in [21].

We start out by constructing a general framework and methods for optimal randomness certification in Bell experiments. The idea is to keep as much information as possible by avoiding any sort of binning of outcomes, then to use the methods recently introduced in [7] to estimate randomness by constructing a device-independent guessing probability optimized over all possible Bell inequalities, and finally to optimize the latter quantity over all the tunable physical parameters of the experiment. We then narrow our focus to entirely optical polarization-based implementations (e.g. [19, 20]). We first characterize the realistic parameters of such Bell setups and then apply our methods to determine optimal amounts of global and local randomness under realistic conditions. We provide interesting bounds on the experimental parameters—some of them counter-intuitive and perhaps unexpected—and certify up to four times more randomness than what a standard analysis, based on a binning of the outcomes and on the Clauser–Horne–Shimony–Holt (CHSH) inequality [22], can achieve [2].

2. Methods

Here we describe methods that allow for optimal device-independent randomness certification. The general idea consists of three steps which are given in box 1. Since we do not make any physical characterization of the source or the devices, the results are kept general and can be applied to any bipartite Bell experiment free of the detection loophole (see [16–20]).

Box 1. General directions for optimal randomness certification.

(1) Estimate the most general behaviour $\mathbf{p}$, without any binning. (*Sections 2.1 and 2.2*)
(2) Construct $G_{\mathbf{p}}$, the device-independent guessing probability optimized over all possible Bell inequalities. (*Section 2.2*)
(3) Optimize $G_{\mathbf{p}}$ over the parameters $\mathcal{P}$ that can be adjusted in the experimental setup. (*Section 2.4 and section 3*)

2.1. Scenario

To begin, we recall the device-independent scenario [2, 7, 23]. Two parties, Alice and Bob, are located in two secure laboratories from which no unwanted classical information can leak out. At each round of the experiment, they receive a quantum state ρ_{AB} from a source S and perform on it one out of m_{A} (m_{B}) possible measurements $x = 0, 1, .., m_{\mathrm{A}} - 1$ ($y = 0, 1, \ldots, m_{\mathrm{B}} - 1$) and retrieve one out of o_{A} (o_{B}) possible outcomes $a = 0, 1, \ldots, o_{\mathrm{A}} - 1$ ($b = 0, 1, .., o_{\mathrm{B}} - 1$). We make no other assumption on ρ_{AB} other than the fact that it is a quantum state. In fact, ρ_{AB} could have any dimension, and could even be correlated with another quantum system in the possession of a malicious eavesdropper eve[4], such that $\rho_{\mathrm{AB}} = \mathrm{Tr}_{\mathrm{E}}\rho_{\mathrm{ABE}}$.

Moreover, Alice and Bob do not trust the devices they use to measure ρ_{AB}. These devices can be thought of as measurements characterized by positive operator-valued measures (POVMs) with elements $\{M_{a|x}\}$ and $\{M_{b|y}\}$ acting on ρ_{AB}. Their probabilistic behaviour is given by Born's rule

$$p(ab\,|xy) = \mathrm{Tr}\left[\rho_{\mathrm{AB}} M_{a|x} \otimes M_{b|y}\right]. \tag{1}$$

There are a total of $m_{\mathrm{A}} m_{\mathrm{B}} o_{\mathrm{A}} o_{\mathrm{B}}$ such probabilities, which can be seen as the components of a vector $\mathbf{p} = \{p(ab|xy)\} \in \mathbb{R}^{m_{\mathrm{A}} m_{\mathrm{B}} o_{\mathrm{A}} o_{\mathrm{B}}}$. We call $\mathbf{p}$ the *behaviour* associated with the *quantum realization* Q defined by the state ρ_{AB} and the measurements with elements $\{M_{a|x}\}$ and $\{M_{b|y}\}$. In all what follows, we consider that the behaviour $\mathbf{p}$ that Alice and Bob observe is a perfect estimate, in the sense that it is assumed to be derived from an asymptotic regime of infinitely many copies.

2.2. Bounding the device-independent guessing probability

The optimal amount of randomness that Alice and Bob can certify from an observed quantum behaviour $\mathbf{p}$ is measured here by the min-entropy of the *device-independent guessing probability* $G_{\mathbf{p}}$ [7], i.e. $h = -\log_2(G_{\mathbf{p}})$. Considering that for some round of the experiment Alice and Bob have chosen and performed some measurements $x = x^\star$ and $y = y^\star$ on ρ_{AB}, it can be shown that $G_{\mathbf{p}}(x^\star, y^\star)$, the average probability that Eve correctly guesses the output of Alice and Bob boxes using an optimal strategy, is the solution to the following conic linear program [7, 8]:

[4] We consider that Eve is limited by the laws of quantum mechanics. We also assume that the behaviour of the boxes is independent and identically distributed from one round to another, though, interestingly, the bound (3) has been proved secure under less demanding assumptions (see [24]).

$$\begin{cases} G_{\mathbf{p}}\left(x^{\star}, y^{\star}\right) = \max\limits_{\{\mathbf{p}^e\}} \quad \sum\limits_e p^e\left(e \middle| x^{\star}, y^{\star}\right), \\ \text{s.t.} \sum\limits_e \mathbf{p}^e = \mathbf{p} \text{ and } \mathbf{p}^e \in \widetilde{Q}, \ \forall \ e = 0,..,\left(o_{\mathrm{A}} - 1\right)\left(o_{\mathrm{B}} - 1\right). \end{cases} \tag{2}$$

The intuition behind this program is that whenever Eve obtains the output $e = (a^{\star}, b^{\star})$ she then guesses that Alice's (Bob's) outcome was a^{*} (b^{*}). Our motivation for defining $G_{\mathbf{p}}$ according to (2) comes from the fact that this expression corresponds to the relevant figure of merit of security proofs in several device-independent cryptographic scenarios [24, 25]; in particular, this bound holds for the case where the memory of Eve is limited in time, the so called bounded quantum storage model [24]. Thus, it follows that for our analysis any strategy z of Eve can be seen as a POVM measurement with $o_{\mathrm{A}}\, o_{\mathrm{B}}$ elements $\{M_{e|z}\}$ that she applies on her reduced state $\rho_{\mathrm{E}} = \mathrm{Tr}_{\mathrm{E}}\rho_{\mathrm{ABE}}$ [7].

Each $\mathbf{p}^e$ is an un-normalized behaviour 'prepared' for Alice and Bob and conditioned on the outcome e of the measurement with POVM elements $\{M_{e|z}\}$ performed by Eve. Hence, the probability that $\mathbf{p}^e$ is prepared is the probability that Eve obtains the corresponding outcome e, i.e. $p(e|z) = \mathrm{Tr}\,[\rho_{\mathrm{E}} M_{e|z}]$. To be precise, $\mathbf{p}^e = \{p^e(a, b|x, y)\} \in \mathbb{R}^{m_{\mathrm{A}} m_{\mathrm{B}} o_{\mathrm{A}} o_{\mathrm{B}}}$, and $\widetilde{Q}$ is the set of all such un-normalized quantum behaviours. The first constraint in the program translates the fact that the behaviours $\mathbf{p}^e$ should on average reproduce Alice and Bob's observed behaviour $\mathbf{p}$. The second constraint demands that every behaviour should be quantum[5]. The program maximizes the success of Eve's strategy over all possible $\{\mathbf{p}^e\,|e = 0, \ldots, (o_{\mathrm{A}} - 1)(o_{\mathrm{B}} - 1)\}$ decompositions.

The program presented in (2) is in general intractable due to the lack of a precise characterization of $\widetilde{Q}$, but semi-definite programming (SDP) relaxations similar to the ones presented in [26] can be used tu put bounds on $G_{\mathbf{p}}$. One then defines a convergent hierarchy of convex sets having a precise characterization and being such that $\widetilde{Q}_1 \supseteq \widetilde{Q}_2 \supseteq \ldots \supseteq \widetilde{Q}$ [7, 26]. This hierarchy approximates the quantum set $\widetilde{Q}$ from the outside, and thus one can relax the difficulty of the problem (to the order k) by replacing $\widetilde{Q}$ in (2) by $\widetilde{Q}_k$. The solution $G_{\mathbf{p}}^k$ of the kth SDP program sets an upper bound on the guessing probability $G_{\mathbf{p}}$, which in turn sets a lower bound $h^k = -\log_2(G_{\mathbf{p}}^k)$ on the number h of *global* random bits that are certified from $\mathbf{p}$ and from the measurements $(x^{\star}, y^{\star})$.

It is worth mentioning that the methods presented so far can be adapted straightforwardly for *local* randomness evaluation. In this case, the situation is considered from Alice's perspective, for example, and a program equivalent to (2) is derived to obtain the local guessing probability $G_{\mathbf{p}}(x^{\star})$. Computationally speaking, local randomness is appealing as the number of POVM elements of Eve's strategies gets reduced from $o_{\mathrm{A}}\, o_{\mathrm{B}}$ to o_{A}.

The bound presented in (2) is obtained for a fixed pair of settings. Actually, these settings are pre-established before the Bell test is implemented and can even be considered as public knowledge. In [8], a technique to average $G_{\mathbf{p}}(x^{\star}, y^{\star})$ over all possible settings was introduced and shown to be advantageous (with respect to the fixed settings technique). However, to our knowledge, all existing device-independent cryptographic protocols are based on the fixed settings technique. Furthermore, averaging over all the outcomes implies that the eavesdropper can only share classical correlations with Alice and Bob; instead, here we assess a stronger scenario in which Eve's system can even be entangled with the users. For this reason we fix the settings that generate randomness, although, in the future perspective, it would be very relevant to look for a bound on the device-independent guessing probability $G_{\mathbf{p}}$ which is independent of x^{*} and y^{*}.

To conclude this section notice that the optimal Bell inequality which yields $G_{\mathbf{p}}^k$ can be accessed from the dual formulation of (2). The advantage with respect to previous methods (which assess the problem via a fixed Bell inequality, e.g. [2]) has been found to be significant in both [7] and [8–10].

2.3. Keeping as much data as possible

In section 2.2 we discussed how to quantify the maximal amount of randomness available for Alice and Bob from an observed behaviour $\mathbf{p}$. Still, there are several degrees of freedom in $\mathbf{p}$ that can be further optimized to provide even more randomness. More precisely, tailoring these degrees of freedom always leads to different behaviours, which in turn yields different—and hopefully higher—amounts of randomness. We can distinguish two types of such degrees of freedom; those that require adjustments in the experimental setup (e.g. increasing the efficiency of the detectors), and those which do not. Here we will deal with the latter, and leave the former for section 2.4.

In particular, the numbers of outcomes o_{A} and o_{B} can be adjusted without much experimental effort. All Bell experiments so far, which have managed to close the detection loophole, have relied violation of the CHSH inequality [22] (or similar ones [27]). This assumes the local observation of two outcomes per party. However,

[5] A behaviour $\mathbf{p}$ is said to be quantum whenever there exists a realization Q (i.e. a quantum state + measurements) which reproduces $\mathbf{p}$ through Born's rule (1).

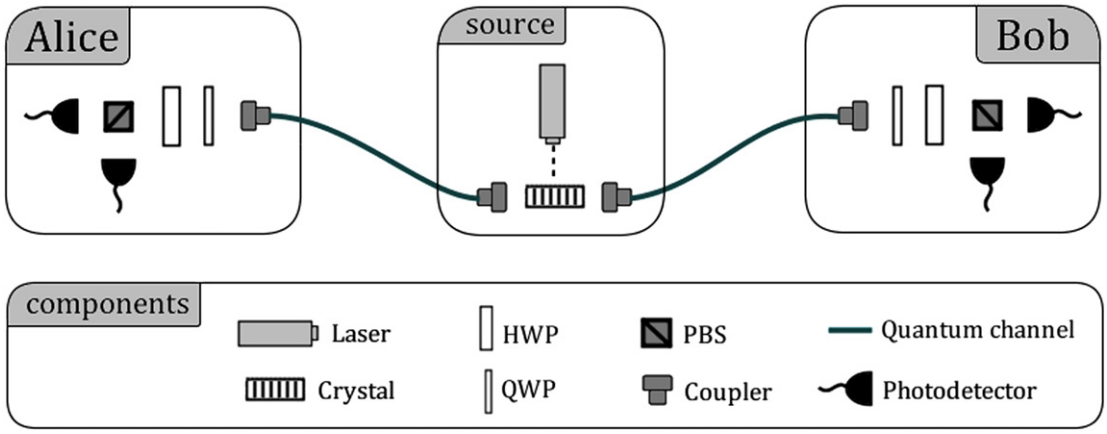

Figure 1. Experimental setup for optical Bell experiments based on SPDC.

in addition to the two good outcomes, loss and imperfections lead to events where no detector clicks, resulting in a third outcome per party; this means that a local binning process was applied in all these experiments to reduce the size of the original behaviour to two outcomes.

It is intuitive to expect that more randomness can be certified when binning strategies are avoided; any binning strategy represents a loss of potentially useful information. Still, it could be the case that the amount of certifiable randomness would not get diminished for some particular binning. Our results in section 4 show that this is not the case in general. In fact, in appendix A we explicitly show how any binning strategy applied to CHSH correlations with inefficient detectors will systematically decrease the amount of certifiable randomness. Hence, to certify optimal amounts of randomness, Alice and Bob must ensure that the number of outcomes o_{A} and o_{B} is kept as high as possible.

2.4. Taking experimental parameters into account

The observed quantum behaviour $\mathbf{p}$ possesses physical degrees of freedom that can be adjusted in the experimental setup to produce higher amounts of randomness. The solution of (2) can be minimized over all the possible realistic values that such parameters (which we label $\mathcal{P}$) can take. In this way, the optimal amount of randomness that can be certified to the order k is the solution of:

$$\begin{cases} G^k\left(x^\star, y^\star\right) = \min\limits_{\mathcal{P}} G^k_{\mathbf{p}}\left(x^\star, y^\star\right), \\ \text{s.t. } G^k_{\mathbf{p}}\left(x^\star, y^\star\right) \text{ solves the } k\text{th SDP of (2).} \end{cases} \tag{3}$$

In particular, notice that this program optimizes $G^k_{\mathbf{p}}(x^\star, y^\star)$ over the number of measurements m_{A} and m_{B}, which are implicit quantities in $\mathcal{P}$ (see also section 4.1).

3. Realistic optical implementations

The methods presented above are general and can be adjusted to any bipartite Bell experiment. We focus and describe in the following the architecture of optical implementations based on polarization measurements of entangled photons distributed from an SPDC source (see figure 1), which was thoroughly analysed in [21]. The source is characterized by three adjustable quantities: two squeezing parameters g_1 and g_2 and a total number of modes N onto which the photons may be distributed. Each mode locally splits into two orthogonal polarizations. In terms of bosonic creation operators, the un-normalized state produced by S is given by [21]:

$$\prod_{k=1}^{N} \exp\left[\tanh(g_1)\, a_k^\dagger b_{k\perp}^\dagger - \tanh(g_2)\, a_{k\perp}^\dagger b_k^\dagger\right]|0\rangle, \tag{4}$$

where $|0\rangle$ is the vacuum state associated to the $4N$ bosonic operators $a_1^\dagger, \ldots, a_{N\perp}^\dagger, b_1^\dagger, \ldots, b_{N\perp}^\dagger$, and the a-modes (b-modes) are distributed to Alice (Bob).

All the different types of losses including detectors inefficiencies are modelled, without loss of generality, by two beam-splitters (not shown in figure 1) placed at any point between the users and the source. The transmittance η of these beam-splitters is the overall detection efficiency of the experiment.

The measurements are performed with polarizing beam-splitters (PBS) and half-wave plates (HWP) and quarter-wave plates (QWP) which allow splitting the orthogonal modes along arbitrary directions [19–21]. Each measurement u is fully characterized by two angles $(\theta_u,\ \phi_u)$ defining a projection in the Bloch sphere. Each of the parties holds two detectors, which do not resolve photon number. Hence, for each detector only the outcomes '0 = no click' and '1 = click' can be distinguished, and the maximal number of local outcomes (without binning) is $o_{\mathrm{A}} = o_{\mathrm{B}} = 4$.

4. Results

In this section we apply the methods presented in section 2 to the optical setup described in section 3.

4.1. Constructing $\mathcal{P}$, p and G

Considering that Alice and Bob respectively perform m_{A} and m_{B} measurements, the experiment is characterized by $4 + 2(m_{\mathrm{A}} + m_{\mathrm{B}})$ physical parameters, which are: $N, g_1, g_2, \eta, \theta_1^{\mathrm{A}}, \phi_1^{\mathrm{A}}, \ldots, \theta_{m_{\mathrm{B}}}^{\mathrm{B}}$ and $\phi_{m_{\mathrm{B}}}^{\mathrm{B}}$. All of these parameters are adjustable within some range of realistic values, except η which, as discussed above, represents the main restriction for an optical implementation. Hence, the adjustable parameters read:

$$\mathcal{P} = \left(N,\ g_1,\ g_2,\ \theta_1^{\mathrm{A}},\ \phi_1^{\mathrm{A}},\ \ldots,\ \theta_{n_{\mathrm{B}}}^{\mathrm{B}},\ \phi_{n_{\mathrm{B}}}^{\mathrm{B}}\right). \tag{5}$$

The analytic expression of $\mathbf{p}$ as a function of $\mathcal{P}$ and η is at first only computed for the first measurements of Alice and Bob, $(\theta_1^{\mathrm{A}},\ \phi_1^{\mathrm{A}})$ and $(\theta_1^{\mathrm{B}},\ \phi_1^{\mathrm{B}})$. In this case $\mathcal{P}$ consists of seven parameters, i.e. $\mathcal{P} = (N,\ g_1,\ g_2,\ \theta_1^{\mathrm{A}},\ \phi_1^{\mathrm{A}},\ \theta_1^{\mathrm{B}},\ \phi_1^{\mathrm{B}})$. Since the number of outcomes are kept as high as possible $(o_{\mathrm{A}} = o_{\mathrm{B}} = 4)$, this expression is obtained by solving a linear system of $4 \times 4 = 16$ equations; 15 of these equations correspond to the 'no-click' probabilities of all the detectors, which can be found in the supplementary material of [21]. The remaining equation is a normalization condition.

Next, this expression (obtained only for the first measurements) is generalized for arbitrary $(m_{\mathrm{A}},\ m_{\mathrm{B}})$. One only needs to concatenate all the individual behaviours:

$$\mathbf{p} = \left\{\mathbf{p}\left(N,\ g_1,\ g_2,\ \eta,\ \theta_i^{\mathrm{A}},\ \phi_i^{\mathrm{A}},\ \theta_j^{\mathrm{B}},\ \phi_j^{\mathrm{B}}\right)\ \middle|\ 1 \leqslant i \leqslant m_{\mathrm{A}} \quad \text{and } 1 \leqslant j \leqslant m_{\mathrm{B}}\right\}. \tag{6}$$

In particular, all the individual behaviours have the same analytical structure as the behaviour obtained for the first measurements, and hence one only needs to substitute $\theta_1^{\mathrm{A}} \leftarrow \theta_i^{\mathrm{A}}, \theta_1^{\mathrm{B}} \leftarrow \theta_j^{\mathrm{B}}, \phi_1^{\mathrm{A}} \leftarrow \phi_i^{\mathrm{A}}$ and $\phi_1^{\mathrm{B}} \leftarrow \phi_j^{\mathrm{B}}$ for each i and j in (6). This yields the desired $m_{\mathrm{A}}\, m_{\mathrm{B}}\, o_{\mathrm{A}}\, o_{\mathrm{B}}$-sized quantum behaviour (see section 2.1).

Finally, it is necessary to set realistic limits on $\mathcal{P}$; otherwise, the minimization in (3) is unbounded. We let $1 \leqslant N \leqslant 100$, $-1/2 \leqslant g_1,\ g_2 \leqslant 1/2$ (corresponding to about 4.3 dB of squeezing) and we let all the measurement angles vary in a 2π-length interval.

4.2. Optimal randomness for $m_{\mathrm{A}} = m_{\mathrm{B}} = 2$

Optimal randomness is retrieved from (3) upon optimization of all adjustable parameters, which include the number of measurements in the experiment. Optimizing G^k over m_{A} and m_{B} is of particular relevance for the setup that we consider as distinct rotation directions of the incoming modes can be achieved by adjusting the HWP and QWP, i.e. without the need of further experimental resources. Still, to illustrate the performance of our methods we consider here the simplest case $m_{\mathrm{A}} = m_{\mathrm{B}} = 2$.

We find[6] that whenever the parties are restricted to $o_{\mathrm{bin}} = 2$ outcomes, more global randomness is certified when no specific Bell inequality is considered. This was to be expected following section 2.2 and the line of research of [7–9] (see dashed and dotted curves in figure 2). However, we improve considerably this expected result by suppressing the binning of the outcomes and letting $o = 4$, as we explained in section 2.3 (solid curve in figure 2). For $\eta = 1$ our methods certify 0.74 bits of global randomness per source use, four times more than the 0.19 bits that are certified from the CHSH inequality (we provide the Bell inequality that certifies this improvement in appendix B). The numerical values of the optimal parameters $\mathcal{P}$ are given in figure 3 for several values of η. Intuitively, the ratio $t = \tanh(g_1)/\tanh(g_2)$ quantifies the degree of entanglement of the source, as (4) shows. For $\eta = 1$ optimal randomness is obtained from a 'maximally entangled' state, i.e. $t = 100\%$, but as η decreases t also decreases. This was to be expected for the lower values of η, where nonlocality can only be certified with non-maximally entangled states [27]. Interestingly, for $\eta \approx 1$ the optimal measurements are not similar to the ones that intuitively maximize the violation of the CHSH inequality on two maximally entangled qubits (e.g. they are not mutually unbiased); see appendix B for the exact expressions. That is, the optimal measurements for optimal randomness certification are not the same as those maximizing the CHSH violation.

[6] All our results were obtained at the order $k = 1 + \mathrm{AB}$. This corresponds to an intermediate stage $\widetilde{Q}_1 \supseteq \widetilde{Q}_{1+\mathrm{AB}} \supseteq \widetilde{Q}_2$; see [26] for details.

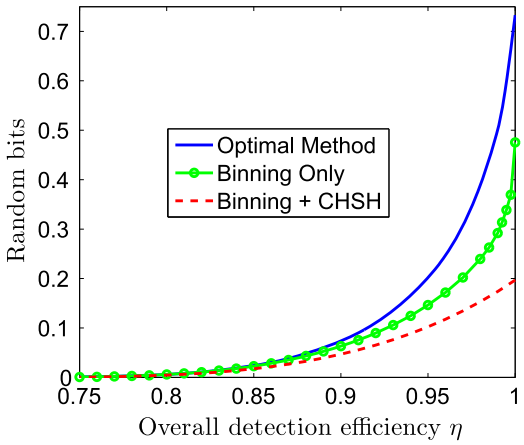

Figure 2. Global randomness for the case $m_A = m_B = 2$. For the three curves, the parameters $\mathcal{P}$ are optimized at each point, as explained in section 2.4. The solid curves are the min-entropy of the solution of program (3) for $o_A = o_B = 4$ (optimal) and for $o_A = o_B = 2$ (binning). The dashed and dotted curves were obtained following the binning strategy presented in [21].

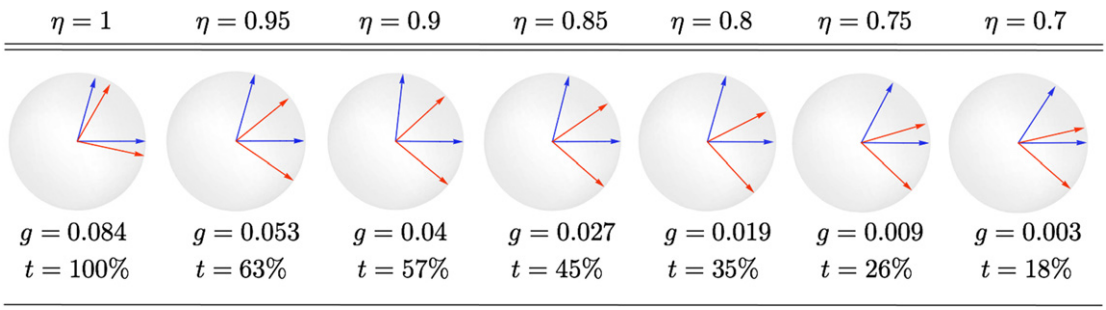

Figure 3. Optimal parameters $\mathcal{P}$ for different values of η. t is the ratio between $\tanh(g_1)$ and $\tanh(g_2)$, while $g = \max(g_1, g_2)$. N always reaches 100. Blue (Red): optimal measurements for Alice (Bob) in the Bloch sphere representation. All these quantities were obtained after solving program (3).

The number of modes attains the maximal value that we allow ($N = 100$) whenever η is greater than $2\sqrt{2} - 2$. For η smaller than this value, the single mode case $N = 1$ is sufficient to obtain maximal randomness; this fact was noticed in [21] for the maximization of the CHSH inequality violation. Finally, we have found that the improvement obtained when increasing the number of modes beyond ≈25 is very small.

4.3. Optimal randomness with more than two measurements

Our next goal is to see whether deploying more measurements yields an improvement in the number of random bits. In the previous subsection we considered the case $m_A = m_B = 2$; however, by adjusting the HWP and QWP located in front of their PBS, Alice and Bob can measure their incoming subsystem along any arbitrary polarization direction of the Bloch sphere. These adjustments can thus be obtained with relatively low experimental cost, the main drawback being a non-negligible increase in the amount of statistical data (the size of the observed behaviour $\mathbf{p}$ increases with $m_A m_B$).

Our results in table 1 show that more measurements certify more randomness, even in scenarios for which a binning strategy had to be considered and $\mathcal{P}$ could not be fully optimized due to computational limitations. The time required to solve (3) becomes large as the number of measurements increases, since the total number of SDP variables describing the behaviours $\mathbf{p}^e$ in (2) increases as $(m_A m_B)^2$. The increase is less dramatic when *local* randomness is certified e.g. from Alice's perspective, as there are only o_A (instead of $o_A o_B$) SDP matrices in (2) for each choice of $\mathcal{P}$.

In particular, with four measurements per party we certify 0.557 local random bits. This is three times more than the amount that is certified from the CHSH inequality (≈0.17 bits) under the same considerations.

4.4. Experiments with only one detector per side

The setup depicted in figure 1 has been hitherto central in our analysis as it captures the general architecture for Bell experiments with entangled photons. Unfortunately, state-of-the-art superconducting detectors, i.e. those which achieve detection efficiencies above 70% and thus enable a true Bell violation without post-selection, represent an extremely high experimental cost nowadays.

Table 1. Local randomness certified for different scenarios for $\eta = 1$. The scenario specifies the couple (m_A, m_B). The * symbol is used when full optimization was not possible, and instead: (i) the optimization was only carried over the number of modes, with $g_1 = g_2 = 0.1$; (ii) the measurements were inspired from the chained inequality [28] and (iii) we considered 3 outcomes per party by locally binning the 'no click–no click' and the 'click–click' outcomes.

SCENARIO	(2, 2)	(3, 2)	(3, 3)	(4, 3)	(4, 4)	(5, 5)
Total SDP variables	1348	3340	8392	15 748	29 620	$\sim 10^5$
Local random bits	0.454	0.459	0.519*	0.523*	0.557*	N/A

This situation can be alleviated (the cost can be reduced by half) by realizing that a Bell test can still be carried on with the use of only one detector on each arm of the experiment [19, 20]. Given the techniques that we have shown so far, it is interesting to see how the optimal amount of randomness is affected. For a fixed overall detection efficiency η, how does the optimal amount of randomness that can be certified in an experiment with only one detector compare to the optimal amount of randomness that can be certified with two detectors?

The statistics of an experiment with only one detector are straightforwardly obtained from the statistics of an experiment with two detectors (those which we presented in 4.1). As discussed in 3 the possible local outcomes of an experiment with two detectors are 00, 01, 10 and 11 where the first (second) number labels the outcome of the first (second) detector'0 = no click' and '1 = click'. Then, applying the local binning $\mathcal{B}_{1\text{Det}} = \{00 \to 0', 01 \to 0', 10 \to 1', 11 \to 1'\}$ on Alice and Bob's sides yields the statistics of the experiment without the second detector.

We observe that for $\eta \lesssim 0.8$ no disadvantage occurs if the second detector is removed: the optimal amount of local and global randomness than can be certified in both cases is $\sim 6 \times 10^{-4}$ bits. On the other hand, as η becomes close to 1 removing a detector negatively affects the optimal amount of randomness: for $\eta = 1$ the optimal amount of local (global) random bits certified with two detectors is ≈ 0.45 (≈ 0.73) bits, while with only one detector the optimal amount is ≈ 0.31 (≈ 0.34) bits.

5. Discussion

Summarizing, in the present article we have explicitly shown the benefits of optimizing randomness in a Bell experiment over all possible inequalities, and the negative consequences that occur when information is lost through a binning of the resulting outcomes. We carefully analysed and characterized optical setups based on SPDC and certified up to four times more randomness when all of the physical parameters were optimized.

To put it in a nutshell, here are the important facts to be aware of in order to retrieve optimal amounts of randomness from an optical Bell implementation based on SPDC *(and their experimental cost)*:

(1) Keep the whole statistics and avoid binning the outcomes. *(No cost.)*

(2) Use as many polarization measurements as possible. *(Small cost.)*

(3) Use many modes to distribute the entangled photons. *(High cost in principle, but keep in mind that more than* ≈ 25 *modes will provide little improvement.)*

(4) For $\eta \approx 1$, the optimal measurements for randomness extraction are not the ones that maximize the violation of the CHSH inequality. *(No cost.)*

(5) For $\eta \lesssim 0.8$ it is enough to use a single mode to distribute entanglement and use a single detector per side. *(No cost.)*

We hope that this work will be useful for the future development of Bell-type randomness generation experiments.

Note added: while finishing this work, we became aware of another Bell-type randomness generation setup based on SPDC, which, interestingly, considers other type of measurements [29]. For this setup a CHSH optimization—as in [21]—was derived; it could be interesting to apply our methods to derive the optimal amount of randomness that this recent experimental setup based on SPDC allows for.

Acknowledgments

We thank M Hoban and S Pironio for interesting discussions and for the proof presented in appendix A. We also thank N Sangouard for sharing with us the exact expressions of the no-click probabilities discussed in section 4.1. The SDP calculations were performed using the code QMBOUND written by J D Bancal. This work

was supported by by the EU projects QITBOX and SIQS and the John Templeton Foundation. JB was supported by the Swiss National Science Foundation (QSIT director's reserve) and SEFRI (COST action MP1006), DC by the Beatriu de Pinós fellowship (BP-DGR 2013), AM by the Mexican CONACYT graduate fellowship porgram, and PS by the Marie Curie COFUND action through the ICFOnest program.

Appendix A. CHSH correlations with inefficient detection

Here we show that it is always advantageous to keep the 'no-click' outcome in a CHSH test with inefficient detection. Assume that at each round of the experiment Alice and Bob receive a perfect singlet, i.e. a maximally entangled state of two qubits, on which with equal probability Alice measures σ_z and σ_x, while Bob measures $(\sigma_z + \sigma_x)/2$ and $(\sigma_z - \sigma_x)/2$. If the measurement processes have non-unit η efficiency, the possible outcomes that the users observe are 0, 1 and 2 (here the outcome two labels the no-click outcome). Under the assumption that losses occur independently, the observed quantum behaviour can be written as

$$\mathbf{p}_\eta = \left(\begin{array}{ccc|c} \eta^2\, c & \eta^2 s & \dfrac{\eta(1-\eta)}{2} & \vdots \\ \eta^2 s & \eta^2\, c & \dfrac{\eta(1-\eta)}{2} & \vdots \\ \dfrac{\eta(1-\eta)}{2} & \dfrac{\eta(1-\eta)}{2} & (1-\eta)^2 & \\ \hline \multicolumn{3}{c|}{\cdots} & \ddots \end{array}\right) \tag{A.1}$$

with $c,\ s = (2 \pm \sqrt{2})/8$. In this expression each of the 4 blocks describes the joint probability $P(a,\ b|x,\ y)$ for a choice of measurements of Alice and Bob. The first block corresponds to $(x = 0,\ y = 0)$ and so on. Blocks 2 and 3 are equal to block 1, while a swap between c and s transforms block 1 into block 4. For each choice of x, any *physical* binning is a deterministic map from the outcomes $\{(a = 0,\ a = 1,\ a = 2)\}$ into the binned outcomes $(a' = 0,\ a' = 1)$, and the same applies to each choice of y with b. Up to local relabelings, there are only three relevant binning strategies (three ways to bin a local trit to a bit) which are, with a slight abuse of notation, $\mathcal{B} = \{0 \to 0',\ 1 \to 1',\ 2 \to 0'\}, \mathcal{B}' = \{0 \to 0',\ 1 \to 1',\ 2 \to 1'\}$ and $\mathcal{B}'' = \{0 \to 0',\ 1 \to 0',\ 2 \to 1'\}$. However, $\mathcal{B}''$ is not relevant as it erases all non-local data. Hence we are left with two local binning strategies which in turn generate four possible quantum behaviours for Alice and Bob:

$$\mathbf{p}_{BB} = \left(\begin{array}{cc|c} \eta^2 c + 1 - \eta & \eta^2 s + \dfrac{\eta(1-\eta)}{2} & \vdots \\ \eta^2 s + \dfrac{\eta(1-\eta)}{2} & \eta^2\, c & \vdots \\ \hline \multicolumn{2}{c|}{\cdots} & \ddots \end{array}\right);$$

$$\mathbf{p}_{B'B'} = \left(\begin{array}{cc|c} \eta^2\, c & \eta^2 s + \dfrac{\eta(1-\eta)}{2} & \vdots \\ \eta^2 s + \dfrac{\eta(1-\eta)}{2} & \eta^2 c + 1 - \eta & \vdots \\ \hline \multicolumn{2}{c|}{\cdots} & \ddots \end{array}\right) \tag{A.2}$$

and

$$\mathbf{p}_{BB'} = \left(\begin{array}{cc|c} \eta^2 c + \dfrac{\eta(1-\eta)}{2} & \eta^2 s + 1 - \eta & \\ \eta^2 s & \eta^2 c + \dfrac{\eta(1-\eta)}{2} & \vdots \\ \hline \multicolumn{2}{c|}{\cdots} & \ddots \end{array}\right);$$

$$\mathbf{p}_{B'B} = \left(\begin{array}{cc|c} \eta^2 c + \dfrac{\eta(1-\eta)}{2} & \eta^2 s & \\ \eta^2 s + 1 - \eta & \eta^2 c + \dfrac{\eta(1-\eta)}{2} & \vdots \\ \hline \multicolumn{2}{c|}{\cdots} & \ddots \end{array}\right). \tag{A.3}$$

Notice from (A.2) that whenever Alice and Bob apply the same binning strategy the two resulting probability distributions have the same values up to a permutation of the elements. The same occurs in (A.3) whenever they apply a different binning. It is therefore sufficient to evaluate the optimal randomness available from $\mathbf{p}_{BB}$ and from $\mathbf{p}_{BB'}$, for example. In figure A1 we plot the percentage by which the guessing probability for these quantum

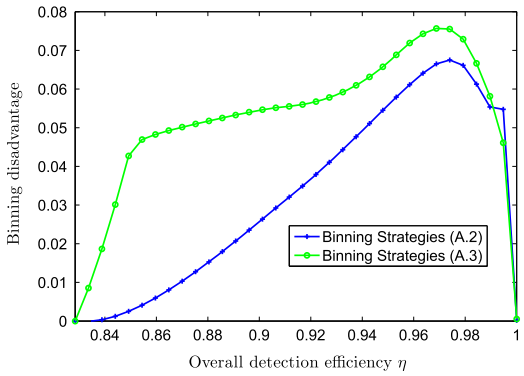

Figure A1. The binning disadvantage is the difference in the number of bits that are certified from either $\mathbf{p}_{BB}$ or $\mathbf{p}_{BB'}$ with respect to $\mathbf{p}_{\eta}$.

behaviours is increased with respect to the guessing probability obtained from $\mathbf{p}_{\eta}$. We find that for any $2\sqrt{2} - 2 < \eta < 1$ it is always advantageous to keep the no-click outcome.

Appendix B. Bell inequality and relevant parameters expressions

As explained in the main text, the dual formulation of (2) yields the expression of the Bell inequality that certifies the optimal amount randomness [7]. It is therefore possible to retrieve the Bell inequality associated to the optimal parameters. One first solves the program (3) for η fixed; this yields some optimal parameters $\mathcal{P} = \mathcal{P}^{\star}$. Then one comes back to solve the dual program of (2) using as input $\mathbf{p}(\mathcal{P}^{\star})$. In the Collins–Gisin parametrization, the 7×7 Bell inequality which certifies 0.74 bits of global randomness (see section 4.2) is:

$$I_{\eta=1}^{1+\mathrm{AB}} = \left(\begin{array}{c|ccc|ccc} 1 & 8.02 & 8.18 & 8.18 & 8.11 & 12.38 & 12.37 \\ \hline 8.02 & -8.07 & 8.13 & 8.13 & -8.11 & 7.11 & 7.11 \\ 8.18 & 8.13 & -2.80 & 6.68 & 7.53 & 19.63 & -20.54 \\ 8.18 & 8.13 & 6.68 & -2.80 & 7.53 & -20.54 & 19.64 \\ \hline 8.11 & -8.11 & 7.53 & 7.53 & -7.98 & 7.77 & 7.77 \\ 12.37 & 7.11 & 19.64 & -20.54 & 7.77 & 3.92 & -6.71 \\ 12.37 & 7.11 & -20.54 & 19.64 & 7.77 & -6.71 & 3.92 \end{array}\right), \tag{B.1}$$

and the optimal parameters which enable this realization are (cf (5)):

$$\mathcal{P} = (100,\ 0.084,\ 0.084,\ 2.088,\ 1.116,\ 1.473,\ 1.117,\ 1.36,\ 1.117,\ 1.976,\ 1.116). \tag{B.2}$$

References

[1] Bell J 1964 *Physics* **1** 195–200
[2] Pironio S *et al* 2010 *Nature* **464** 1021–4
[3] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 *Rev. Mod. Phys.* **86** 419–78
[4] Ekert A K 1991 *Phys. Rev. Lett.* **67** 661–3
[5] Colbeck R 2007 *PhD Thesis* University of Cambridge
[6] Gerhardt I, Liu Q, Lamas-Linares A, Skaar J, Kurtsiefer C and Makarov V 2011 *Nat. Commun.* **2** 349
[7] Nieto-Silleras O, Pironio S and Silman J 2014 *New J. Phys.* **16** 013035
[8] Bancal J D, Sheridan L and Scarani V 2014 *New J. Phys.* **16** 033011
[9] Bancal J D and Scarani V 2014 *9th Conf. on the Theory of Quantum Computation, Communication and Cryptography* vol 27 p 1
[10] Law Y Z, Le Phuc T, Bancal J D and Scarani V 2014 *J. Phys. A: Math. Theor.* **47** 424028
[11] Dhara C, de la Torre G and Acin A 2014 *Phys. Rev. Lett.* **112** 100402
[12] De la Torre G d, Hoban M J, Dhara C, Prettico G and Acin A 2014 (arXiv: 1403.3357)
[13] Pearle P M 1970 *Phys. Rev.* D **2** 1418–25
[14] Santos E 1992 *Phys. Rev.* A **46** 3646–56
[15] Gerhardt I, Liu Q, Lamas-Linares A, Skaar J, Scarani V, Makarov V and Kurtsiefer C 2011 *Phys. Rev. Lett.* **107** 170404
[16] Rowe M A, Kielpinski D, Meyer V, Itano W M, Monroe C and Wineland D J 2001 *Nature* **409** 791–4
[17] Ansmann M *et al* 2009 *Nature* **461** 504–6
[18] Hofmann J, Krug M, Ortegel N, Gérard L, Weber M, Rosenfeld W and Weinfurter H 2012 *Science* **337** 72–75
[19] Christensen B G *et al* 2013 *Phys. Rev. Lett.* **111** 130406
[20] Giustina M *et al* 2013 *Nature* **497** 227–30
[21] Vivoli V C, Sekatski P, Bancal J D, Lim C, Christensen B, Thew A M R, Zbinden H, Gisin N and Sangouard N 2015 *Phys. Rev.* A **91** 012107

[22] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880–4
[23] Acin A, Brunner N, Gisin N, Massar S, Pironio S and Scarani V 2007 *Phys. Rev. Lett.* **98** 230501
[24] Pironio S, Masanes L, Leverrier A and Acin A 2013 *Phys. Rev.* X **3** 031007
[25] Masanes L, Pironio S and Acin A 2011 *Nat. Commun.* **2** 238
[26] Navascués M, Pironio S and Acin A 2007 *Phys. Rev. Lett.* **98** 010401
[27] Eberhard P H 1993 *Phys. Rev.* A **47** 747–50
[28] Braunstein S L and Caves C M 1990 *Ann. Phys.* **202** 22
[29] Vivoli V C, Sekatski P, Bancal J D, Lim C, Christensen B, Thew A M R, Zbinden H, Gisin N and Sangouard N 2014 (arXiv: 1409.8051)