PAPER

How random are random numbers generated using photons?

To cite this article: Aldo Solis *et al* 2015 *Phys. Scr.* **90** 074034

View the article online for updates and enhancements.

How random are random numbers generated using photons?

Aldo Solis, Alí M Angulo Martínez, Roberto Ramírez Alarcón, Hector Cruz Ramírez, Alfred B U'Ren and Jorge G Hirsch

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Apdo. Postal 70-543, México 04510 D.F.

E-mail: hirsch@nucleares.unam.mx

CrossMark

Abstract

Randomness is fundamental in quantum theory, with many philosophical and practical implications. In this paper we discuss the concept of algorithmic randomness, which provides a quantitative method to assess the Borel normality of a given sequence of numbers, a necessary condition for it to be considered random. We use Borel normality as a tool to investigate the randomness of ten sequences of bits generated from the differences between detection times of photon pairs generated by spontaneous parametric downconversion. These sequences are shown to fulfil the randomness criteria without difficulties. As deviations from Borel normality for photon-generated random number sequences have been reported in previous work, a strategy to understand these diverging findings is outlined.

Keywords: quantum randomness, single photons, Borel normality

1. Introduction

For many decades after quantum mechanics was formally established, it was relatively easy to live with the fact that it only allows the prediction of probabilities of certain results of experiments. The description and manipulation of condensed matter, molecules, atoms, atomic nuclei, and subnuclear particles was indeed very successful, and the statistical nature of the results, describing a huge number of similar processes, as in scattering and decay, fitted quite well with this probabilistic interpretation.

In the last few decades, the ability to manipulate individual quantum objects (e.g., molecules, atoms, and photons) and even to place many of them in a single quantum state, as in a Bose–Einstein condensate, has emerged. The experiments along these lines that are possible nowadays invite one to ponder the quantum mechanical description of individual systems, and to engineer them to obtain technologically useful devices.

The emerging view is, for us, both surprising and challenging. Nearly ten years ago, Anton Zeilinger wrote: 'The discovery that individual events are irreducibly random is probably one of the most significant findings of the twentieth century. ... for the individual event in quantum physics, not only do we not know the cause, there is no cause. The instant when a radioactive atom decays, or the path taken by a photon behind a half-silvered beamsplitter are objectively random' [1].

This provocative statement helps to visualize the relevance that randomness has in our description of the physical world. According to this view, reality and information are two sides of the same coin. Randomness, complementarity, and entanglement emerge from the fact that from individual measurements there is a finite amount of information available. It is postulated that an elementary system can only give a definite result in one specific measurement. Other independent measurements must then be irreducibly random [2].

The consequences of these assumptions are both philosophical and practical. Random numbers are widely employed for classical computation in science and industry. Monte Carlo and other numerical methods require the use of random numbers, which are demanded to be both efficiently generated and having proved randomness [3].

When the intrinsic quality of quantum randomness is accepted as a postulate, a world of applications emerges. Quantum key distribution is a cryptographic process whose security is guaranteed by the quantum randomness [4]. A cryptographically secure random number generator that does

not require any assumption about the internal working of the device has been proposed [5, 6]. Such a strong form of randomness generation is impossible classically and possible in quantum systems only if certified by a Bell inequality violation. 'Private randomness' is defined by the presence of correlations that cannot be reproduced with local variables. It is quantified by the violation of Bell inequalities, and is associated with the impossibility to predict a given string employing a classical computer and classical information [5–9]. Quantum contextuality has also been invoked to certify randomness in a random number generator [10].

Is the randomness of quantum phenomena a physical assumption that is testable? Many attempts have been made to answer this question. Recently a comprehensive suite of tests, developed at the National Institute of Standards and Technology (NIST) to assess the quality of computer-based random number generators [11], was employed to study the randomness of single-photon polarization measurement outcomes, using pairs of photons generated by spontaneous parametric downconversion (SPDC). No statistically significant deviations from randomness were observed [12].

Has quantum randomness been experimentally proved? The previously mentioned results suggest that a quantum-generated random sequence looks as random as a computer-generated one. But there are deep differences between these sources of random numbers. Quantum randomness can be proven incomputable; that is, it is not exactly reproducible by any algorithm, while software-generated random numbers, known as pseudo-random, can be reproduced if the computer code and the seed are known. Is it possible to distinguish between them? Calude *et al* [13] performed finite tests of randomness inspired by algorithmic information theory, analyzing algorithmic randomness, the strongest possible form of incomputability. They performed tests of randomness on pseudo-random strings (finite sequences) generated with software (Mathematica, Maple), which are cyclic (so, strongly computable), the bits of π, which are computable, but not cyclic, and strings produced by quantum measurements (with the commercial device Quantis and by the Vienna IQOQI group). They report that all tests produced evidence, with different degrees of statistical significance, of differences between quantum and nonquantum sources.

Figure 1 displays the maximal relative deviations from Borel normality reported in [13]. It shocked us to observe that the noncomputable, photon-generated random sequences depart from Borel normality far more than the pseudo random numbers obtained from Mathematica, Maple and the digits of π. In particular, the sequences produced by the Vienna group, obtained from an attenuated laser impinging on a beamsplitter, cannot be validated as random under this criterion.

Puzzled by these results, we have generated random numbers employing the detection times of photon pairs generated by SPDC and analyzed them with the tools of algorithmic randomness. We have found that, at variance with the findings of [13], they pass all the randomness tests with flying colors. These results call for a more detailed analysis, comparing different sources of single photons and different ways
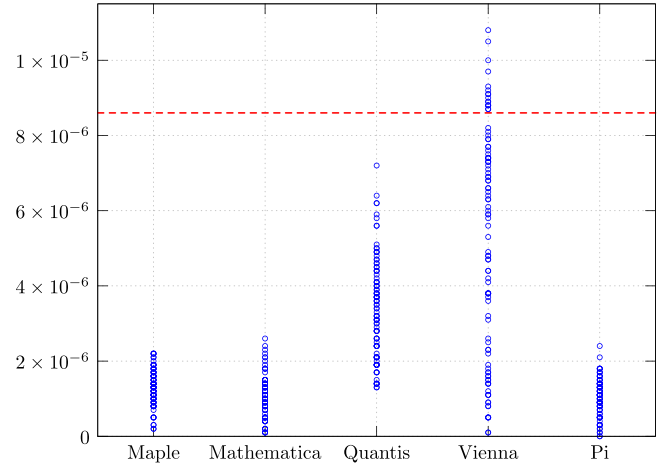

Figure 1. Maximal relative deviations from Borel normality reported in [13]. The red dashed line represents the maximum possible deviation allowed by Borel normality. The blue circles represent the (absolute value of the) deviations from $1/2^m$, for different orders m (see section 3). Some sequences produced by the Vienna group are not Borel normal.

to generate random bits from their detection times; this constitutes work in progress in our group.

In what follows we present a short review of the challenges faced in defining randomness, the algorithmic information approach, and the Borel normality test, which combined with the algorithmic complexity approach, provides a necessary but insufficient test of randomness. We describe in some detail the experimental setup, the procedure to generate the bit sequence from the photon arrival times, and the analysis. We close with some conclusions and open questions.

2. Randomness

Randomness in physics is related with two main ideas: first with the lack of information about a system; for instance, every time a coin is tossed we ignore the initial conditions of the event so that we cannot predict the result, and second, with the intrinsically unpredictable behavior of a quantum system, as in a photon impinging on a beamsplitter. In both cases we use a probabilistic approach to describe phenomena regardless of their conceptual differences, and sometimes we combine both in the density matrix, but there are situations where we want to talk about randomness itself, i.e., characterize and/or quantify randomness. We therefore need a definition of randomness and a theoretical framework that allows us to develop and use these ideas.

There have been some works trying to define randomness through different mathematical objects. For example, the concept of normal numbers due to Borel [14] formalizes the notion of a random real number. One successful attempt that has become a mathematical theory is algorithmic information theory, also known as Kolmogorov complexity. It is based on the idea of patterns in a mathematical object.

In order to provide some physical insight leading to the definition of randomness in algorithmic information theory, consider a system having a property with the two possible outcomes, 0 and 1, with equal probabilities of measuring each of them.

If we measure it repeatedly, the result may look like one of these cases:

$$0101010101010101010101010101010101 \cdots$$
$$0100011011000010100111001011110111 \cdots$$
$$0100010110100101111010101001001011 \cdots$$

These sets are noticeably different. The first one does not look random at all, because it has a strong pattern: after a 0 the next result will be a 1. We can sum up this set of data with the recipe 'repeat "01" indefinitely'. The second case looks random but actually has a pattern. It was generated by writing 0 and 1, then the combinations 00 01 10 11, then the combinations with three symbols 000 001 … etc, and is known as the Champernowne's constant. It serves as a cautionary example of the difficulty in identifying a random sequence, which can look random despite being defined by a very simple pattern. How can we determine if the third one has a pattern?

It is possible to formalize the intuitive relationship between randomness and the lack of patterns in a sequence in a definition like this:

'A sequence without patterns is random.'

Although this looks like a very simple definition, we need to define what a pattern is. In the previous examples we described a pattern in such a way that anyone can reproduce the sequence just by following the steps, so everyone could use this description, in principle, in the same way that a computer executes a program.

Suppose that we have a program that generates a given sequence. Does that mean that the sequence has a pattern? As a useful illustration, the program

PRINT '0100010110100101111010101001001011'

evidently can reproduce the third sequence, and in general the same technique can be used for every sequence simply by placing the actual sequence as part of the code. However, this implies that the program is longer than the sequence itself and it does not make much sense considering it as a pattern, so we define a pattern as a program whose output is a given sequence in such a way that this program is shorter than the sequence.

The previous considerations lead us to a simple definition:

There are no patterns in a random sequence, i.e., every program that outputs a random sequence is longer than the sequence itself.

This definition isn't fully formalized. The notion of a program needs to be written in terms of universal Turing machines, and we need to specify the symbols allowed in the code. A full treatment of the concept of algorithmic randomness is beyond the scope of this informal exposition. The reader interested in a complete description is invited to read [15] and [16].

Unfortunately this definition holds some surprises for us. Suppose that we wish to determine whether or not a sequence is random; this is indeed a very natural question to ask. In principle, we would need to run all programs with a shorter length than the length of the sequence. If any of these programs were to give us the desired sequence as the output, an underlying pattern would then have been determined to exist, and we would conclude that the sequence is not random. But we don't know in advance if a given program will halt or not. There is a close relationship between algorithmic randomness, the halting problem, and Gödel's incompleteness [17].

The halting problem [18] in a computer can be summarized as follows: suppose that we have an algorithm designed to determine whether a given program will halt; then we could build a new algorithm that halts if the program does not halt and does not halt if the program does. What will happen if we feed this new program with its own code? It will halt if and only if it doesn't halt, which is an evident paradox. Therefore there is no algorithm capable of deciding whether a sequence is random or not.

3. Borel normality

We may be disappointed because algorithmic randomness is difficult to apply in a real case, but there are other ways to approach this definition of randomness.

In the previous sequences the 'probability' of getting 0 or 1 is equal; this is the case for the sequence 010101⋯. What about the probability of getting 01, following a subdivision of the string into symbols composed of two digits? In this case, this probability is one and the probabilities of getting 00,10 and 11 are in all three cases zero. In contrast, we expect a random sequence to lead to equal probabilities for all these cases:

$$P(0) = P(1) = \frac{1}{2}$$
$$P(00) = P(01) = P(10) = P(11) = \frac{1}{4}$$
$$\vdots$$

We can generalize this restriction on the probabilities as:

$$P(\langle m \text{ bits sequence}\rangle) = \frac{1}{2^m}. \tag{1}$$

A set of numbers satisfying this property is called the Borel normal. It is closely related with the normality concept in real numbers developed by Borel [14]. Naturally, there must be a restriction on m because in a string of 4 symbols we cannot try sequences longer than 4 symbols.

We can analyze finite sequences that may not fulfil exactly the conditions noted earlier but are close to doing so,

for example:

$$P(00) = 0.251$$

This can be expressed mathematically as [15]

$$\left| P(00) - \frac{1}{2^2} \right| < \epsilon$$

where ϵ is a 'small' number.

Although there is no algorithm that can determine the randomness of sequences, it is possible to relate algorithmic randomness with the parameters m and ϵ. In [19] it is shown that almost all algorithmic random strings are Borel normal, satisfying

$$\left| P(\langle m \text{bits sequence}\rangle) - \frac{1}{2^m} \right| < \sqrt{\frac{\log n}{n}} \tag{2}$$

where n is the length of the complete sequence and

$$m \leqslant \log \log n. \tag{3}$$

We will refer to this condition as Borel normality. This is not a sufficient condition for randomness but allows us to discard some sequences as clearly not random. Its advantage is that checking this condition is an algorithmic procedure that can be applied to any sequence.

In [13] this condition is applied to binary sequences obtained from a quantum experiment, based on an attenuated laser beam impinging on a beamsplitter. The authors found that these sequences 'were outside the expected range for $m = 3$ and $m = 4$'. In this work we analyze a related experimental setup, where we employ the differences in arrival times of SPDC photon pairs as a means to generate random bits.

4. Experimental setup

Our experimental work is based on the process SPDC in which a laser pump beam illuminates a crystal with a $\chi^{(2)}$ nonlinearity, leading to the annihilation of pump photons and the emission of photon pairs, typically referred to as signal and idler [20]. In the case of a continuous-wave pump at frequency ω_p, signal and idler photons are spectrally anti-correlated so that if a signal photon is detected with frequency ω, the conjugate idler must have a frequency $\omega_p - \omega$. Like-wise, in the idealized situation of a plane-wave pump (which we may approximate through a Gaussian beam with a large beam radius at the beamwaist), photon pairs are anti-correlated in a transverse wavevector, i.e., if a signal photon is detected with transverse wavevector $\mathbf{k}^\perp$, the conjugate idler photon must have transverse wavevector $-\mathbf{k}^\perp$.

The quantum state of the emitted photon pairs can be written as $|\Psi\rangle = |\text{vac}\rangle + \eta|\Psi_2\rangle$ in terms of the vacuum $|\text{vac}\rangle$, the two-photon component $|\Psi_2\rangle$, and of a constant η related to the conversion efficiency. Under the assumptions a continuous-wave, plane-wave pump $|\Psi_2\rangle$ may be expressed as
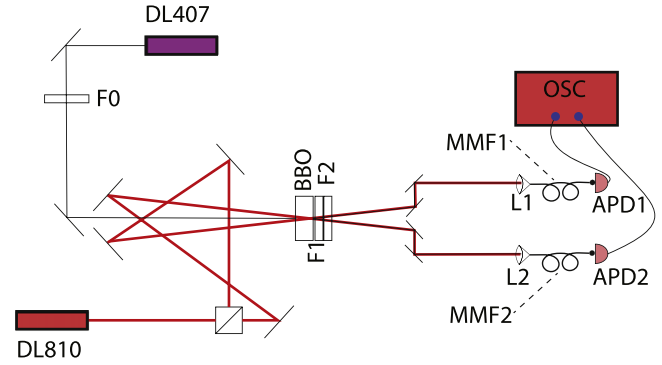

Figure 2. Experimental setup used for obtaining random sequences from a SPDC photon-pair source

[21]

$$
\begin{aligned}
|\Psi_2\rangle = \int d\omega \int d\mathbf{k}^\perp F\left(\omega, \mathbf{k}^\perp\right) |\omega, \mathbf{k}^\perp\rangle_s \\
\times |\omega_p - \omega, -\mathbf{k}^\perp\rangle_i,
\end{aligned} \tag{4}
$$

written in terms of a joint amplitude function $F(\omega, \mathbf{k}^\perp)$, and where $|\omega, \mathbf{k}^\perp\rangle_\mu$ represents a single-photon Fock state with frequency ω and transverse wavevector $\mathbf{k}^\perp$ for mode μ, with $\mu = s, i$ for the signal (s) and idler (i). In writing the two-photon state, we have assumed that the parametric down-conversion process is in the spontaneous regime, so that the appearance of multiple-pair events can be neglected. This assumption is valid if the parametric gain is sufficiently low; experimentally, we restrict the pump power so that the process remains spontaneous. In all likelihood, a similar experiment and analysis carried out in the high-gain, stimulated regime would yield different results from those presented in this paper.

Note that the state in equation (4) is entangled since it cannot be factored into a direct product of separate states $|S\rangle$ (for the signal photon) and $|I\rangle$ (for the idler photon) as $|\Psi\rangle = |S\rangle|I\rangle$. While many experimental works have focused on the presence of quantum entanglement in photon pairs, in this paper we exploit another important aspect of SPDC photon pairs: they are emitted at random times.

Our experimental setup is shown in figure 2. We have used as a pump a beam from a diode laser (DL407) centered at 407 nm with ~ 60 mW power, and as a nonlinear medium a β barium borate (BBO) crystal of 1 mm length. A Schott BG-39 coloured glass filter (F0) is used to remove non ultraviolet photons from the pump beam. The BBO crystal, which is negative uniaxial, was cut so that the angle subtended by the optic axis with respect to pump beam axis is $\theta_{pm} = 29.2°$ which yields phasematching for the generation of frequency-degenerate, noncollinear photon pairs. Signal and idler photons are emitted on diametrically opposed portions of an emission cone centred on the pump beam axis, in our case, with a 3.6° half-opening angle. Pump photons are suppressed by transmitting the signal and idler modes through a long-pass filter, which transmits wavelengths $\lambda > 488$ nm (F1),

followed by a bandpass filter, centred at 800 nm with a 40 nm bandwidth (F2).

Note that in order to set up and correctly align the fiber collection modes, the signal and idler paths are initially simulated using a separate diode laser centred at 810 nm (DL810). The beam from this laser is split into two branches, and each one is reflected with a mirror so as to meet on the crystal's centre plane in such a way that the paths of these two branches emerge from the crystal in the directions expected for the emitted signal and idler photons. Alignment of the collection lenses and fibers is significantly easier with these classical beams than with SPDC light.

Collection of the signal and idler photons can be carried out on any two diametrically opposed locations on the emission ring. Each of the signal and idler collection modes is defined by an aspheric lens with $f = 8$ mm focal length (L1 and L2), which focuses incoming light into the core of a multimode fiber (MMF) with a 50 μm diameter (MMF1 and MMF2). The plane defined by the two collection fibers is chosen for convenience to be parallel to the optical table.

Each of the two photon-collection fibers leads to a silicon-based avalanche photodiode (APD1 and APD2), which emits an electronic pulse for each detection event, discriminated on its rising edge, resulting in a 4-ns-long standard nuclear instrumentation module (NIM) pulse. Because we are interested in the time series that results from the detection times, we connect the two detector outputs to a 2.5 GHz digital oscilloscope (OSC). We program the oscilloscope to subdivide a detection span of 0.512 s into 256×10^6 time bins, so that each bin has a duration of 2 ns. The voltage from the APD output is recorded at each time bin for each of the signal and idler channels, thus obtaining two separate time series composed of voltage values. These times series are post-processed so that those bins with a voltage V that satisfies $|V| > V_{th}$ with a threshold value of $V_{th} = 450$ mV are assigned a value of 1, while those bins for which $|V| < V_{th}$ are assigned a value of 0. This leads to two separate (for the signal and idler) strings s_n and i_n of 256×10^6 digits, each with value 0 or 1. Likewise, for each pair of signal and idler time series, we generate a third time series defined as $c_n = s_n \times i_n$, which corresponds to those bins for which there are coincident detection events in the two channels. We have observed by averaging over several hundred experimental runs that the average number of single-channel detection events during the detection span of 0.512 s is around 8.5×10^5, while the average number of coincidence detection events is around 8×10^4.

5. Obtaining bits from data

Obtaining random bits from data has a considerable complexity. The binary strings in the previous section have a bias because most of the time bins will have value of '0' corresponding to no detected photon. In other words, there are many more zeroes than ones.
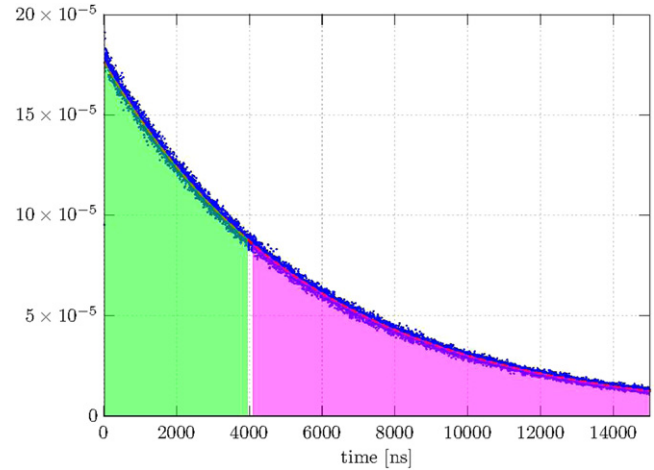

Figure 3. Distribution of time differences between subsequent detection events. Each event in the green zone yields a 0 value, while those in the purple zone yield a 1 value.

In order to obtain a sequence that looks more random, we first compute the time difference between subsequent detection events. This is equivalent to counting the number of zeroes between ones, to which we can clearly assign an integer number. These time intervals are described by an exponential distribution that can be understood in mathematical terms like this: let λ be the probability per unit time of observing a detection at any instant of time; then the probability of not observing a detection in a short time interval δt is $1 - \lambda \delta t$; if we consider a finite time interval t and divide it into n intervals with duration $\frac{t}{n}$ where n is large, the probability of not having a detection in t is equal to the product of the probabilities for each short time interval, then:

$$\left(1 - \frac{\lambda t}{n}\right)\left(1 - \frac{\lambda t}{n}\right) \cdots = \left(1 - \frac{\lambda t}{n}\right)^n.$$

Note that the right side of the previous expression becomes $e^{-\lambda t}$ as n tends to infinity. It is relevant to remark that the result relies on the absence of correlations between detections at different times.

Figure 3 shows an example of an experimentally measured time interval distribution for our SPDC photon pairs detected in coincidence, along with a fit to an exponential distribution. Note that our APDs have a dead time of around $T_d = 20$ ns, so that following a detection event in any of the two channels, the detector is unable to register further events during a time interval of T_d duration. Of course, this will impact the time interval distribution for short times. We have based our analysis later on a truncated time interval distribution, so as to exclude the previous features, which appear at short times smaller than $t_0 = 2 T_d$. As the time intervals follow an exponential distribution, removing the short time intervals is equivalent to redefining them as $t \rightarrow t - t_0$. Note that while the interval distribution reaches times greater than 15 μs, with a mean time of around 4 μs, the omission of the first 40 ns of this temporal range is expected to have only a minor effect on our procedure for obtaining random sequences.

Table 1. Standard deviation of $P(i)$ values from $1/2^m$ (see equation 6), for different values of m, computed for four different sequences.

	Standard deviation σ_m			
m	Sequ. 1	Sequ. 2	Sequ. 3	Sequ. 4
1	0.000980	0.000405	0.000542	0.000760
2	0.001314	0.000750	0.000494	0.000564
3	0.000633	0.000312	0.000535	0.000492
4	0.000619	0.000421	0.000460	0.000451

Thus, using our knowledge about the distribution, random bits can be obtained by dividing the possible time values into two bins: those lower than x and those greater than x, where x is the mean time such that

$$\int_0^x \lambda\, e^{-\lambda t} dt = \int_x^\infty \lambda\, e^{-\lambda t} dt = \frac{1}{2}.$$

Each individual time difference t obtained from the experiment allows us to generate a random bit 0 if the $t < x$ (green zone in figure 3) and 1 if $t > x$ (purple zone). The fraction of 0's and 1's in each string departs from the exact $1/2$ value due to the fluctuations around the average exponential distribution. These deviations from $1/2$ are quantified by the standard deviation shown in the $m = 1$ row of table 1. Using this method, we generated 10 sequences of 10^6 bits using the string of coincident detection events c_n obtained from the quantum source described in the experimental setup. Note that this method could be improved upon by dividing the possible time values into more bins, i.e., 4, 8, 16, so we can obtain more bits per detection; see, for example, [22–24].

6. Results

We have experimentally generated ten sequences, which have an equal count of 0's and 1's with a maximum discrepancy of 0.1% after a careful adjustment of the mean difference time x. From condition (3) we have

$$m < log\ log\ 10^6 = 4.3. \tag{5}$$

Therefore, the maximum possible value of m in our case is 4.

For a given m, there are 2^m different strings of length m forming the set S_m. Thus, $S_2 = \{00, 01, 10, 11\}$, $S_3 = \{000, 001, 010, 100, 011, 101, 110, 111\}$, and so on. We evaluate the probability of occurrence of a given string $i \in S_m$ as $P(i) = N(i)/N_m$, where $N(i)$ is the number of times that the string i is present in a sequence subdivided into segments of m symbols, and $N_m = Int[n/m]$ is the total number of strings of length m in the sequence of length n. For instance, in the sequence 01010101 we find that $N(01) = 4$ and $N(00) = N(10) = N(11) = 0$. Note that we define the strings i sequentially in such a way that no two strings overlap; for the specific example noted earlier, this means that the string '10' does not appear.
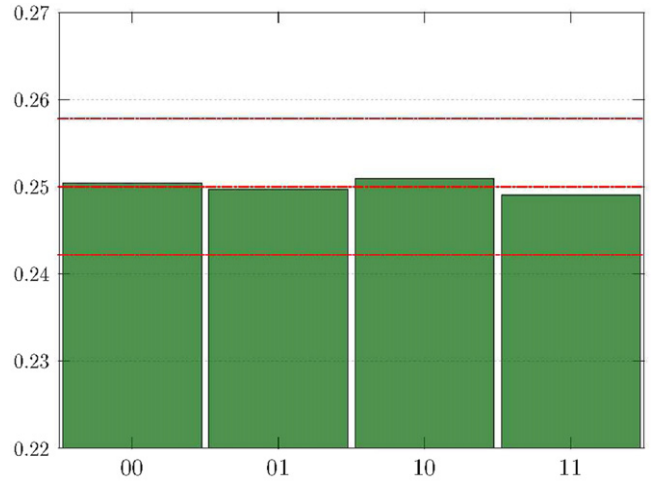

Figure 4. Borel normality analysis for $m = 2$ using sequence 2. The top and bottom red lines are the maximum deviations allowed by Borel normality.
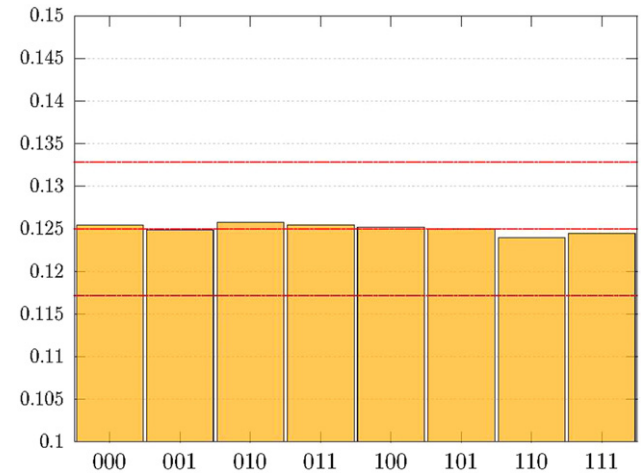

Figure 5. Borel normality analysis for $m = 3$ using sequence 2. The top and bottom red lines are the maximum deviations allowed by Borel normality.
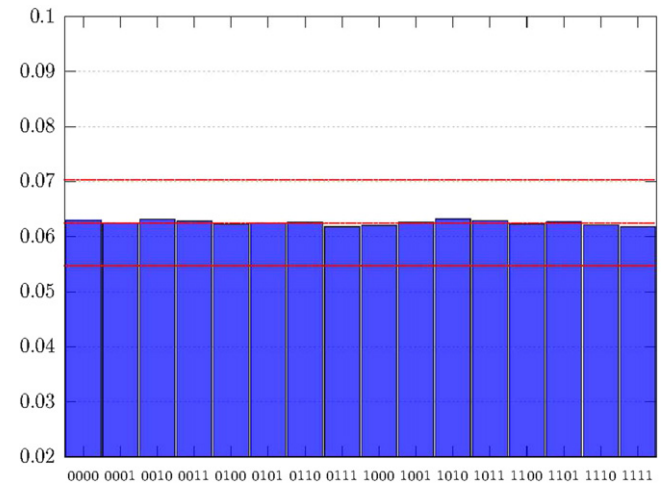

Figure 6. Borel normality analysis for $m = 4$ using sequence 2. The top and bottom red lines are the maximum deviations allowed by Borel normality.
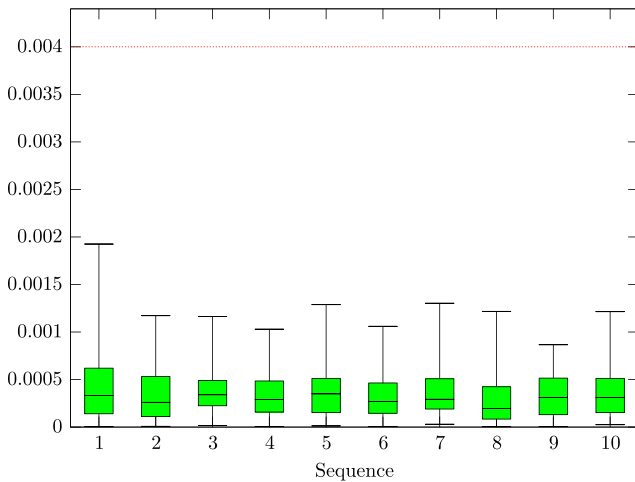
Figure 7. Deviations from the mean value for each sequence. The red line represents the maximum deviation allowed by Borel normality.

Table 2. NIST test suite results for our photon-generated random sequences.

Test	P-value	Pass
Frequency	0.191687	100/100
Block Frequency	0.021999	100/100
Cumulative Sums	0.171867	100/100
Cumulative Sums	0.319084	100/100
Runs	0.383827	98/100
Longest Run	0.224821	100/100
Rank	0.019188	99/100
FFT	0.867692	100/100
Nonoverlapping Template	0.507021	99/100
Overlapping Template	0.304126	99/100
Approximate Entropy	0.319084	99/100

In figures 4, 5 and 6 we present the results of the analysis for one sequence. Each histogram shows $P(i)$. It is represented in a different range because the mean value is 0.25 in 4, 0.125 in 5, and 0.0625 in 6. The top and bottom dashed red lines represent the maximum and minimum possible values allowed by equation (2), respectively, and the central red line corresponds to the mean value expected. It can be seen that all cases easily satisfy the Borel normality condition

$$\left| P(i) - \frac{1}{2^m} \right| < \sqrt{\frac{log\ 10^6}{10^6}} = 0.00441.$$

Because we are mainly interested in the deviations from the mean value for every string, it makes sense to look at the standard deviation (σ) of the probabilities. The standard deviation is defined as

$$\sigma_m^2 = \frac{1}{2^m} \sum_{i \in S_m} \left(P(i) - \frac{1}{2^m} \right)^2 \qquad (6)$$

It follows from equation (2) that for the sequence to be considered Borel normal, all σ_m values must satisfy the condition

$$\sigma_m < \frac{log\ n}{n} = 0.00441. \qquad (7)$$

In table 1 we display the values of the standard deviations of four sequences. All of them fulfill the previous condition for $m = 1, 2, 3, 4$, being about an order of magnitude smaller than the limit imposed by algorithmic randomness on Borel normality.

In figure 7 we present for each of the ten sequences their deviations from the expected mean value. The box-and-whisker plots, inspired by the ones used in [13], display in short horizontal lines, from bottom to top, the minimum value, first quantile, median, third quantile, and maximum value of the difference $|P(i) - \frac{1}{2^m}|$, including the results for $m = 2, 3,$ and 4. The whisker-and-box plot allows us to see the large difference between the maximum deviations from Borel normality and the value imposed by condition (2). The

maximum deviations from the mean value reach up to 34% of the limit set by condition (2) for $m = 1$, 47% for $m = 2$, 40% for $m = 3$, and 27% for $m = 4$.

Returning to the motivation for the present investigation, we were puzzled by the results reported in [13], where the sequences of pseudo-random numbers passed without difficulty the Borel normality test, while the sequences of random numbers built employing the photon detections produced by the Vienna group failed, having maximum deviations from the mean value reach up to 27% of the limit set by condition (2) for $m = 1$, 127% for $m = 2$, 103% for $m = 3$, and 105% for $m = 4$.

We have based our analysis on the coincident-event sequences c_n. Note that since SPDC photons are born in pairs, ideally the single-channel sequences s_n and i_n would be identical to c_n. Realistic experiments are affected by optical losses and by spurious detection events from noise sources and from dark counts. We did in fact perform the Borel normality analysis on the single-channel sequences, with the result that sequences s_n and i_n do satisfy the bounds for Borel normality imposed by equation (2), albeit with larger deviations from the expected fractions (81% for $m = 2$, 54% for $m = 3$, and 37% for $m = 4$) as compared to the sequence c_n.

In [13] the random bits were obtained with a different experimental setup. The signals of the Vienna group were generated with photons from a weak blue light-emitting diode (LED) light source, which impinged on a non polarising beamsplitter with two output ports associated with the values '0' and '1', respectively. There was no pre- or post-processing of the raw data stream; however, the output was constantly monitored. The signals of the QUANTIS device are produced in a similar way, but due to hardware imbalances, which are difficult to overcome at this level, QUANTIS processes these raw data by unbiasing the sequence by a von Neumann-type normalization. The sequences employed had $2^{32} \approx 4 \times 10^9$ bits.

Finally, we have used the NIST statistical test suite [11] to check the (intuitive or statistical) randomness of our generated bits. Even though these tests are not directly related to algorithmic randomness, it is worthwhile comparing these

results with those of the Borel normality test. So as to have at least 100 sequences on which to run the NIST tests, we divided each string into ten sub strings of equal length, ending up with 100 strings of 10^5 bits each. Each test returns a P-value that must be greater than 0.01 (in our case) to pass the test; this value was calculated using the 100 sequences of random bits, and at least 97 of the sequences need to pass the test individually. As expected, our sequences pass each of the tests in the suite; the results are presented in table 2.

7. Conclusions

Random number generators are usually assessed using a battery of tests (such as the one from NIST [11]), which are quite practical but not based on a formal definition of randomness. The Borel normality test employed here is based on algorithmic information theory, which provides a mathematical framework that can formalize randomness.

In [13] it was reported that quantum random number sequences built from photon detection events failed in some cases to pass the Borel normality test, while the pseudo-random sequences generated with computer codes had no problems fulfilling the Borel normality requirements demanded by algorithmic randomness.

In this contribution we report an analysis of Borel normality of sequences of random numbers generated from the time intervals between successive detection events in a photon-pair source based on SPDC. They pass comfortably the Borel normality test.

Before definite conclusions can be extracted from this comparison, we plan to carry out further experiments with longer photon-derived random sequences. We also plan to carry out experiments with an attenuated laser instead of SPDC light, and with a beamsplitter introduced both in one arm of an SPDC source and on the path of an attenuated laser. We hope that these steps will help to clarify if, and under what circumstances, it is experimentally posible to distinguish random number sequences created using quantum sources from computer-generated pseudo-random numbers.

Acknowledgments

References

[1] Zeilinger A 2005 The message of the quantum *Nature* **438** 743
[2] Zeilinger A 1999 A foundational principle for quantum mechanics *Found. Phys.* **29** 631–43
[3] Katzgraber H G 2010 Random numbers in scientific computing: An introduction arXiv: 1005.4117
[4] Nielsen M A and Chuang I L 2002 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[5] Scarani V 2010 Guaranteed randomness *Nature* **464** 988–9
[6] Pironio S *et al* 2010 Random numbers certified by Bell's theorem *Nature* **464** 1021–4
[7] Christensen B G *et al* 2013 Detection-loophole-free test of quantum nonlocality, and applications *Phys. Rev. Lett.* **111** 130406
[8] Bancal J-D, Sheridan L and Scarani V 2014 More randomness from the same data *New J. Phys.* **16** 033011
[9] Nieto-Silleras O, Pironio S and Silman J 2014 Using complete measurement statistics for optimal device-independent randomness evaluation *New J. Phys.* **16** 013035
[10] Um M *et al* 2013 Experimental certification of random numbers via quantum contextuality *Sci. Rep.* **3** 1627
[11] Bassham L E III 2010 *Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications* NIST Special Publication) 800–22 Revision 1a (Gaithersburg, MD: National Institute of Standards and Technology)
[12] Branning D and Bermudez M 2010 Testing quantum randomness in single-photon polarization measurements with the NIST test suite *J. Opt. Soc. Am.* B **27** 1594–602
[13] Calude C S, Dinneen M J, Dumitrescu M and Svozil K 2010 Experimental evidence of quantum randomness incomputability *Phys. Rev.* A **82** 022102
[14] Émile Borel M 1909 Les probabilités dénombrables et leurs applications arithmétiques *Rendiconti del Circolo Matematico di Palermo* **27** 247–71
[15] Calude C 2002 *Information and Randomness: An Algorithmic Perspective* 2nd edn (New York: Springer)
[16] Li M and Vitnyi P M 2008 *An Introduction to Kolmogorov Complexity and Its Applications* 3rd edn (New York: Springer)
[17] Calude C 2005 Algorithmic randomness, quantum physics, and incompleteness *Machines, Computations and Universality* (Lecture Notes in Computer Science) vol 3354 (Berlin: Springer) pp 1–17
[18] Turing A 1937 On computable numbers, with an application to the entscheidungsproblem *Proceedings of the London Mathematical Society* **42** 230–65
[19] Calude C 1994 Borel normality and algorithmic randomness *Developments in Language Theory* ed G Rozenberg and A Salomaa (Singapore: World Scientific) pp 113–29
[20] Burnham D C and Weinberg D L 1970 Observation of simultaneity in parametric production of optical photon pairs *Phys. Rev. Lett.* **25** 84
[21] Vicent L E, U'Ren A B, Rangarajan R, Osorio C I, Torres J P, Zhang L and Walmsley I A 2010 Design of bright, fiber-coupled and fully factorable photon pair sources *New J. Phys.* **12** 093027
[22] Li S, Wang L, Wu L-A, Ma H-Q and Zhai G-J 2013 True random number generator based on discretized encoding of the time interval between photons *J. Opt. Soc. Am.* A **30** 124–7
[23] Wayne M A, Jeffrey E R, Akselrod G M and Kwiat P G 2009 Photon arrival time quantum random number generation *J. Mod. Opt.* **56** 516–22
[24] Nie Y-Q, Zhang H-F, Zhang Z, Wang J, Ma X, Zhang J and Pan J-W 2014 Practical and fast quantum random number generation based on photon arrival time relative to external reference *Appl. Phys. Lett.* **104** 051110