4
Raman Noise and Random Number Generation

Both experiments presented here were performed after the author's return to Rio de Janeiro, from January 2008 until the end of the same year. They present results which aim to support quantum communications in optical fibers (as is the case of the Raman noise study), or in general as in the case of the random number generation experiment.

4.1. Simultaneous Classical and Quantum Communications in Optical Fibers

In order for QKD to move from the lab or niche commercial applications to mainstream usage, co-existence of quantum and classical channels inside the same optical fiber is of paramount importance. Severe care must be taken in the transmission of a classical signal in a fiber with a single-photon detector at the other end. The isolation of most commercial telecom components (filters, couplers, etc...) is of the order of 30 dB, which means that out of a typical signal power of 0 dBm (1mW), there are still too many photons falling on the detector. Careful filtering is therefore critical to a successful quantum and classical communication in the same optical fiber.

The work presented here came during preparation for the experiment with polarization control in the next chapter. While using the reference classical channels for polarization control, a source of noise was noticed, and originally we believed it was cross-talk from the WDM filters. We realized it was only present when the fiber was connected, therefore it must be a scattering effect of the classical channels along the fiber. We decided to investigate this effect further and the results are presented here.

Our initial attempts were to mimic the polarization control setup first used by our group in [78], in which the control channels were used in a counter-propagating manner to the quantum channel. This was reasonable enough in order to minimize cross-talk effects. As we will show, in this situation a problem which

PUC-Rio - Certificação Digital Nº 0510513/CA

can arise is Rayleigh backscattering [77] from the classical signals generated in the fiber, as well as reflection in optical connectors, from the amplified spontaneous emission (ASE) of the lasers used. Although all lasers concentrate most of their emission power on a center wavelength with narrow bandwidth, a small percentage of light is emitted as broadband noise called ASE, which can be tens of nanometers wide centered on the emission wavelength. ASE (as the center wavelength) will scatter back as it propagates along the fiber through Rayleigh backscattering, and reflect from connectors as well. The problem is that a part of the backscattered ASE will fall within the quantum channel wavelength, causing noise in the quantum transmission. Rayleigh backscattering can be removed filtering the ASE section corresponding to the quantum signal bandwidth out of the laser spectrum as shown in Fig. 25, using a fiber Bragg grating. Without the fiber connected it is possible to verify that all cross-talk and reflections from fiber optical connectors have been removed. When the fiber is added to the setup, a considerable amount of noise appears and this is what we wish to investigate.

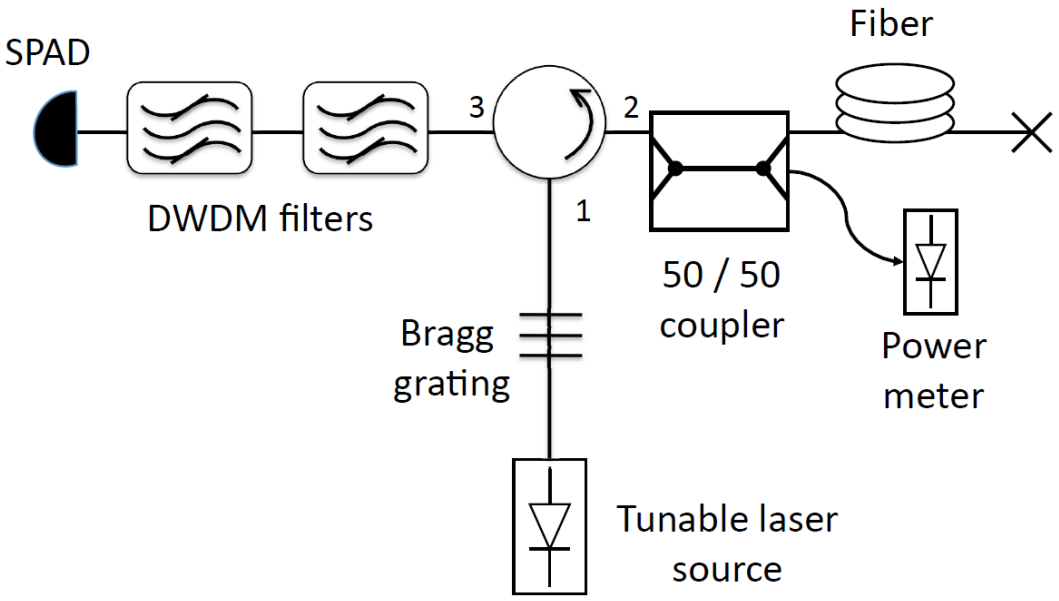


Figure 25 - Experimental setup to investigate noise generated from Raman spontaneous scattering. SPAD: Single photon avalanche detector; DWDM: Dense wavelength division multiplexer. The Bragg grating center wavelength is 1546.12 nm.

In the experiment to characterize the counter-propagating noise, we employed a tunable laser source operating in CW (continuous wave) mode to scan between 1475 and 1640 nm (a combination of two tunable lasers were used). In series with the laser is a fiber Bragg grating, designed to reflect the wavelength

PUC-Rio - Certificação Digital Nº 0510513/CA

1546.12 nm, which is the wavelength of the quantum channel in our polarization control experiment. Therefore by placing the Bragg grating, we remove the ASE component corresponding to the quantum channel from the tunable laser source spectrum. It is then connected to port 1 of an optical circulator, going to port 2 as shown in Fig. 25, is split in two by a 50/50 coupler with one output going to an optical fiber, and the other one to a power meter. The power meter is necessary to normalize the output power of the laser, since it is not constant for all wavelengths. The output of the optical fiber is an angled connector, and just before it bends of small radii were made to the fiber, to remove any reflections from the connector, which might disrupt the measurements results. Any photons returning along the fiber arrive at port 2 of the circulator and get forwarded to port 3, passing the two DWDM filters in tandem, and then arrive at the SPAD. These filters are in fact multiplexers / demultiplexers (they are passive components, so they are a multiplexer or a demultiplexer depending on which way they are connected), having one common port and 4 four wavelength ports (1545.32, 1546.12, 1546.92 and 1547.12 nm). All input light at the common port will get split in wavelength according to the other four ports, and vice-versa. In our experiment, light coming from port 3 of the circulator is connected at the common port of the first DWDM, whose 1546.12 nm port is connected to the common port of the second DWDM and finally the 1546.12 nm port is connected to the SPAD. Both DWDMs are simply used here as filters to the 1546.12 nm channel, but in the polarization control experiment the first DWDM was used to split the quantum (1546.12) and classical channels (1545.32 and 1546.92). Fig. 26 presents the measured results with 8 km of dispersion shifted (DS) and 7.5 km of standard SMF-28 optical fiber. The dark counts have been subtracted as we want to show only the effects of Raman noise.

The first thing which jumps out from this measurement is that the intensity of the counts obtained is, on average one order of magnitude higher than the dark count level of most commercial InGaAs SPADs (10^{-5} dark counts per 1 ns gate). 1 mW of input power is a typical level for many telecom systems, and sometimes it can be even more. It was verified that appropriate filtering was used by removing the fiber, and checking that we only had dark noise level for all input wavelengths. The experiment was made using two separate fiber spools, one composed of 8 km of DS fiber, and the other of 7.5 km of SMF-28. Clearly this

PUC-Rio - Certificação Digital Nº 0510513/CA

noise is being generated inside the optical fiber by the presence of a single CW classical channel. The first thought was of a non-linear effect, however as is shown in Fig. \27, the intensity of the noise varies linearly with the optical power. After considerable research and thought, we concluded that the phenomenon responsible is Raman spontaneous scattering, which is a linear effect [77]. The reason why the noise intensity is higher in the DS fiber, is due to difference in fiber core radii. Interestingly, we came to the correct conclusion that the effect we were observing is Raman spontaneous scattering independently of the works of other groups who first identified it [108, 109].

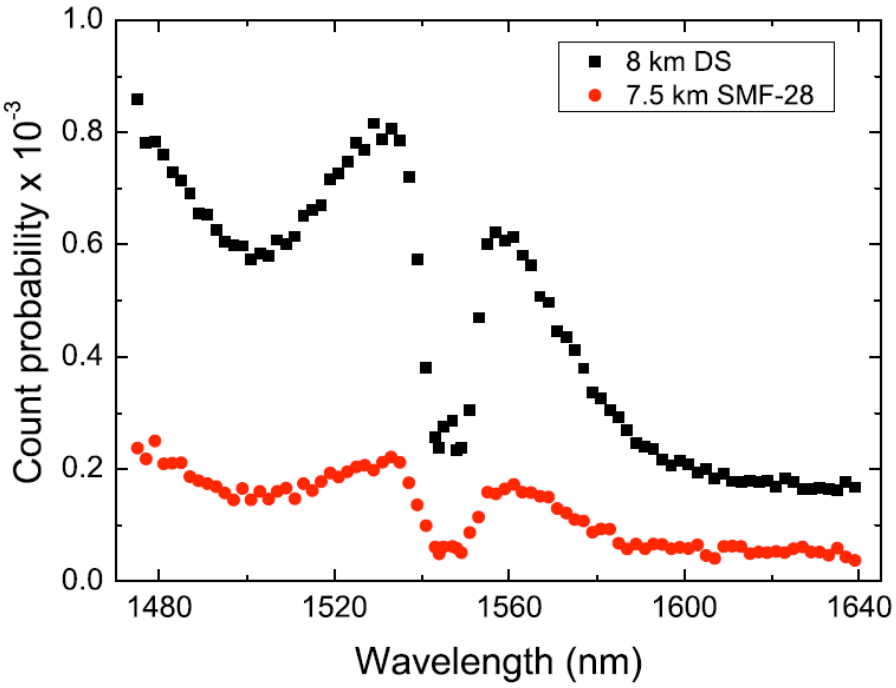


Figure 26 - Count probability per 1 ns gate for 1 mW (0 dBm) of input CW optical power as a function of the tunable laser wavelength. Dark counts have been subtracted.

Raman scattering works in the following manner: a photon while in its time-of-flight inside a fiber may be absorbed by one of the many SiO_2 atoms that make up the fiber's lattice. It is then re-emitted at a longer wavelength, and to conserve energy and momentum, a phonon is absorbed by the lattice (anti-Stokes), or is re-emitted at a shorter wavelength, and a phonon is consumed in the process (Stokes). For this reason Raman is a type of inelastic scattering, while Rayleigh's scattering is elastic since no phonons are involved and therefore no wavelength conversion takes place [77]. Another important result showing the linearity of the noise is presented in Fig. 27. As we can clearly observe the noise intensity is

PUC-Rio - Certificação Digital Nº 0510513/CA

linear through a broad variation of input power, which is among the typical levels used in telecom. The difference in intensity between different wavelengths is simply because as we observed, the Raman noise intensity is wavelength dependent.

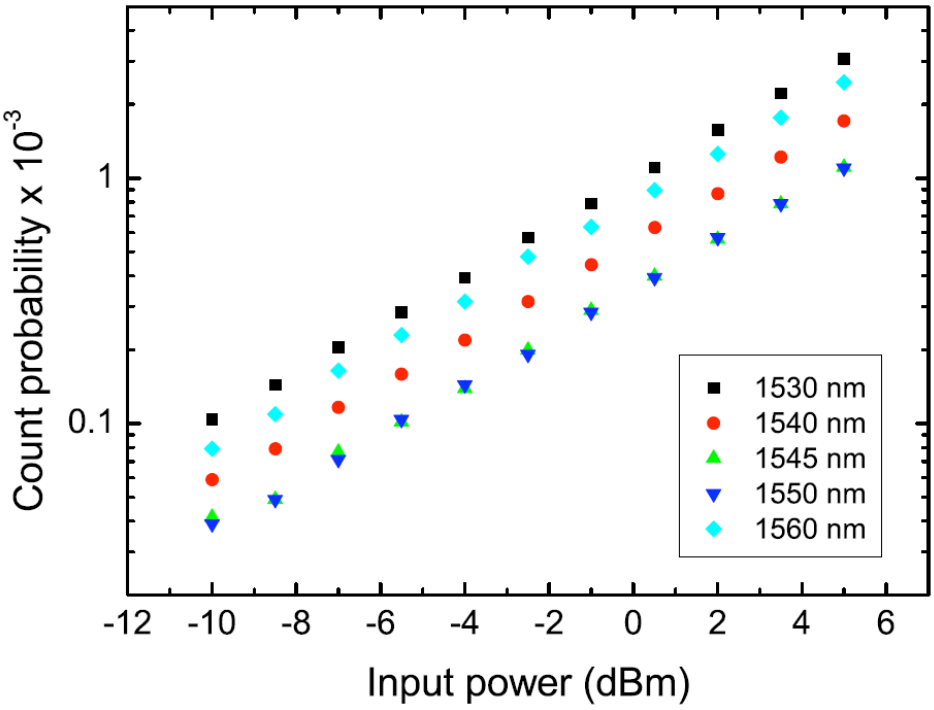


Figure 27 - Count probability per 1 ns gate for the counter propagating setup for 8 km of DS fiber as a function of input power and wavelength.

The spectra of one of the DWDMs (both of them have the same characteristics) is shown in Fig. 28. Since they are standard for the ITU-T telecom grid (0.8 nm spacing), their spectra is of the flat-top type with 0.4 nm FWHM (full-width at half maximum). The extinction ratio between adjacent channels is of the order of 40 dB. This measurement was performed using a tunable laser as the light source, and an optical spectrum analyzer connected after the DWDM. The output fiber of the laser was connected to the common port of the DWDM, and each wavelength port was connected to the spectrum analyzer in turn, yielding the four spectra shown in the figure.

So far we have seen the noise contribution in only one direction, the counter-propagating one. This is one possible configuration we may have while using a quantum channel in a fiber with live traffic being transmitted. In fact, this could have been the most sought after configuration, since it makes filtering easier due to the fact that channels are counter-propagating. Unfortunately, according to

PUC-Rio - Certificação Digital Nº 0510513/CA

the results we have shown, any quantum communication session is impossible with the noise levels presented. The first experiment presented in the previous section used a separate channel in the same fiber as the single-photons, providing trigger information for the single-photon detector. In that case there was no noise generated due to Raman scattering because the trigger signals consisted of optical pulses separated in time from the single-photons. At the end of this section comments will be made regarding to what can be done to minimize the effects of Raman spontaneous scattering induced-noise when propagating classical and quantum channels in the same fiber.

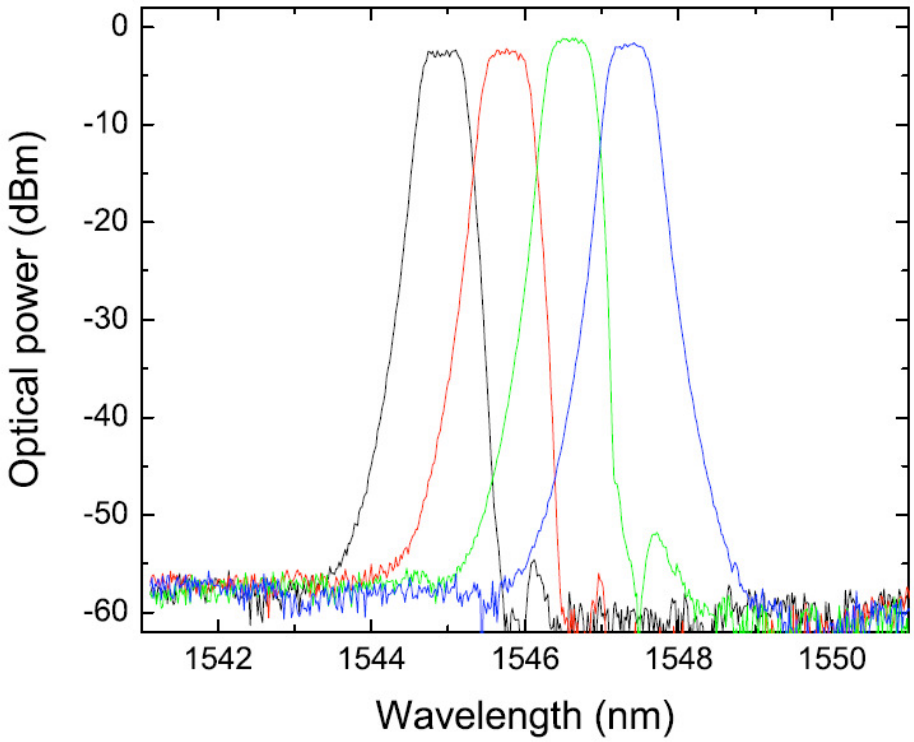


Figure 28 - Spectra of each DWDM channel measured with a tunable laser source and an optical spectrum analyzer.

What can we expect when we wish to co-propagate a classical and a quantum channel along the same fiber? This is the objective of the next measurement in respect to Raman spontaneous scattering. The experimental setup is presented in Fig. 29. We have employed the same combination of two tunable laser sources (yielding a total scan spectrum between 1475 and 1640 nm). The filtering needs to be more stringent in this configuration since we are shining a classical light source directly at the detector, which can easily saturate or even

PUC-Rio - Certificação Digital Nº 0510513/CA

destroy the SPAD. Once again, even though the center wavelength of the laser will not coincide with the quantum channel (1546.12 nm), a component of the ASE from the laser corresponding to the quantum wavelength (the power of the ASE in respect to the center peak can range between -30 to -50 dB depending on the laser quality) falls *directly* on the detector. The first step we take to protect against this is to place a fiber Bragg grating, centered at 1546.12 nm, in series with the tunable laser to remove that ASE component. The power is split in two in a fiber coupler, to monitor the tunable laser power. The signal passes through the fiber, then through a circulator connected as shown in the figure. Port number 2 of the circulator has a fiber Bragg grating connected with the end reflection of the fiber removed by using an angled connector and several small bends on the fiber just before it. This grating is centered on the quantum channel wavelength to improve the filtering provided by the two DWDMs. Finally port 3 is connected to the two DWDM filters in the same configuration as before.

PUC-Rio - Certificação Digital Nº 0510513/CA

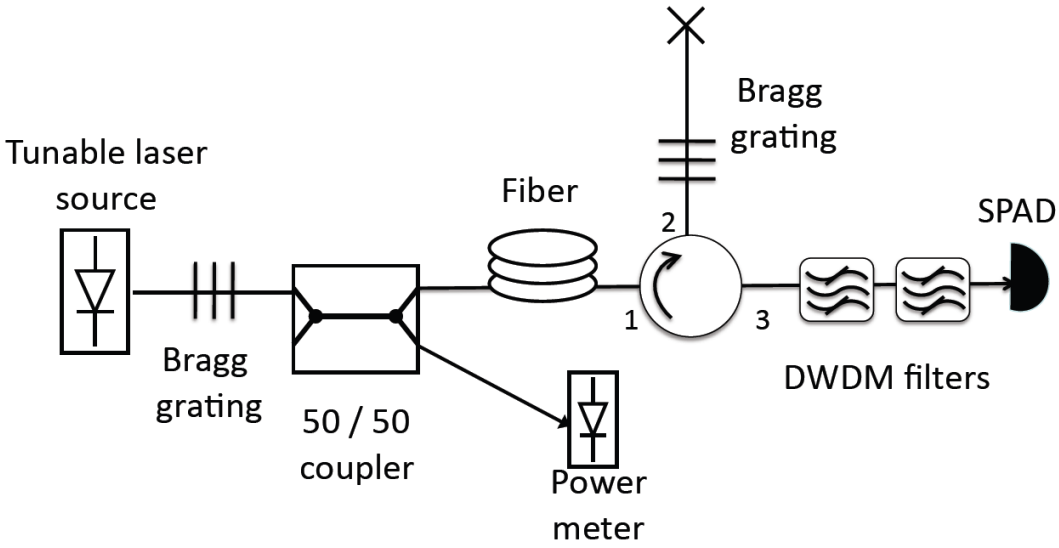


Figure 29 - Setup for characterizing co-propagating Raman noise.

Even by taking the extra filtering precautions we did not manage to remove all cross-talk noise, since we saw that without the fiber connected, the level of counts on the detector was a bit higher than the dark noise level. When we connected the fiber we could still see a noticeable difference so we nevertheless observe the effects of Raman noise. In order to correct the curves however, we did two measurements, one without the fiber, and the other with the fiber connected.

The two curves were subtracted in order to obtain only the Raman noise contribution (Fig. 30).

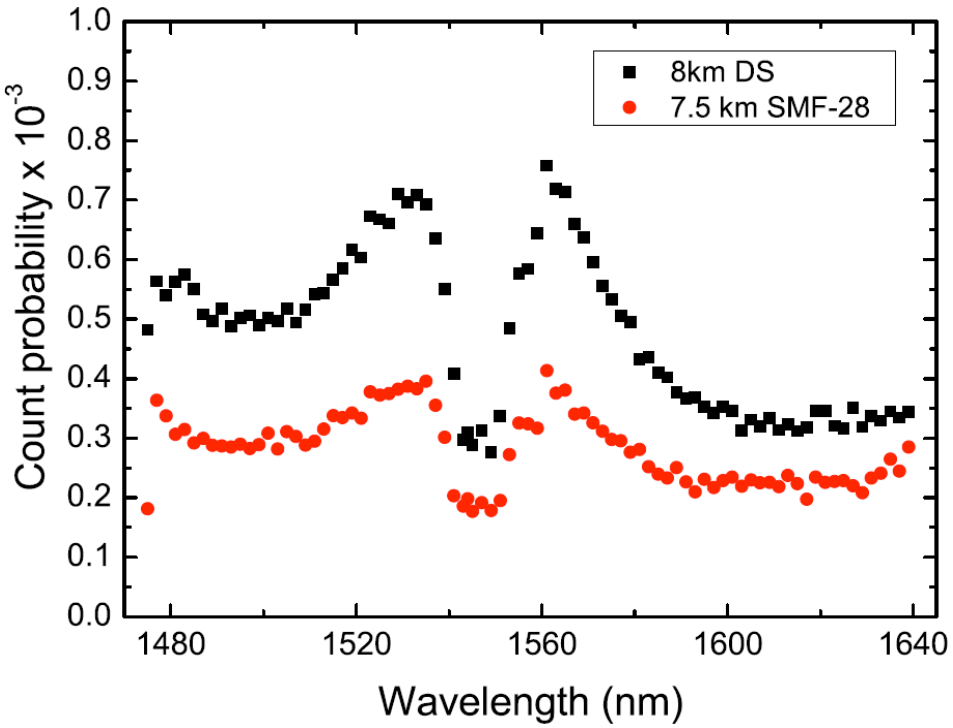


Figure 30 - Count probability per 1ns gate for 1mW (0dBm) of input CW optical power as a function of the tunable laser wavelength. Dark counts + counts from cross-talk have been subtracted.

We clearly notice that the order of magnitude of both co and counter-propagating setups is the same, as it would be expected for Raman spontaneous scattering. We can also observe almost the same difference between using DS and SMF-28 fibers, once again owning to the difference in fiber core radii. The most important result is that since the noise contribution in both configurations is roughly the same, there is little to no benefit in using one configuration or the other in principle. Based on this conclusion, we should then prefer the counter-propagating setup, as the cross-talk filtering is easier. Unfortunately the counter-propagating setup does not support the idea of shutting down the classical channel when the single-photon is to be transmitted, eliminating all Raman noise, due to the fact that this "dark pulse" would never be in sync with the single photon (they are counter-propagating).

Our results show that the photon counts contribution from Raman for a 1 mW input CW laser power is, on average, one-two orders of magnitude higher

PUC-Rio - Certificação Digital Nº 0510513/CA

than the dark count rate of commercial InGaAs SPADs, which makes quantum communication unfeasible. There are steps we can use to minimize Raman spontaneous scattering impact while performing quantum transmission in fibers simultaneously with live commercial traffic. The first clear measure to take is to lower the input power. Since the noise contribution is linear, a 10 dB reduction in input power yields a 10 dB reduction in noise. It is possible to use -10 dBm input power depending on the transmission distance, and even so, Erbium Doped Fiber Amplifiers can be used to amplify the classical channel, after it is demultiplexed. Such an experiment was performed (in fact it was the first experiment we know of done with 0.8 nm spacing in optical fibers) [85]. The authors however do not discuss Raman noise at all, or mention any noise contributions in their system. The other alternative is to use much narrower filters, to minimize the noise (since it is broadband). Very narrow filters are possible with today's technology, such as 10 pm bandwidth [110]. Since the filters we used have a bandwidth of around 0.4 nm, we can expect a reduction by a factor of 40 in the noise. In order to use a much narrower filter the source must be equally narrow, such as an attenuated pulsed laser. On the other hand sources based on SPDC have typically many nanometers bandwidth, which would be very inefficient to use together with a very narrow filter. One final solution is to temporarily shut down the classical channel when the single-photon is transmitted. This idea can remove all Raman noise, and we successfully used it in the polarization control experiment explained in the last chapter. Unfortunately this is not a practical solution for commercial classical optical communication systems, since: a) too much adaptation of classical systems is needed to ensure synchronization is kept, no data is lost, etc\ldots; b) the lasers when turned on from zero current have a long transient time until they stabilize, making any modulation directly to zero current unfeasible at rates beyond 2.5 Gb/s [64]. Most modern fiber systems operate at 10 or 40 Gb/s.

Of course, all the above discussion was done considering one classical channel only, while many channels are routinely used in wavelength multiplexed systems in order to reuse a single optical fiber. If we only consider Raman spontaneous noise, it will grow linearly with the number of channels used. However, other effects come into play in multi-channel systems such as four-wave mixing (FWM) and cross-phase modulation (XPM). FWM is non-linear in nature, and will generate frequency combinations that depend on three frequencies

PUC-Rio - Certificação Digital Nº 0510513/CA

(e. g. $\omega_4 = \omega_1 + \omega_2 - \omega_3$), and can fall exactly in the quantum channel wavelength. The transmission of a quantum channels inside a fiber populated with several classical channels is a far from trivial feat, and further investigation is expected in this area.

4.2. Quantum Random Number Generation Protocol

One of the central requirements for security in QKD is that Alice and Bob must be able to generate their measurement basis choices in such a way that Eve with all her technological prowess cannot predict the values they choose. The only way to perform such a task is to use a random number generator. Although it sounds like a simple problem to solve, the generation of random numbers is a non-trivial matter. Let us think about this for a moment, how can we generate a random number inside a computer (the command "rand" for example, exists in many programming languages)? Once we begin to ponder it, we can see that it is impossible for a deterministic machine such as a computer, to generate a truly random value. What programming commands such as "rand" do, is to take an initial value known as a seed, perform a mathematical algorithm on it, and give the user the result as a random number [111]. In fact the only random component of these numbers is the seed. Clearly the quality of the random number generator depends upon the seed. In modern computers, a typical seed is the time of the day the rand command is run. Another seed which is sometimes used is the content of the last network packet present in the computer's Ethernet port when the rand command is typed. This last seed example is clearly more suited for the task since it is much harder to predict than the current time of the day. If we decided to use this as a random number generator for QKD, clearly we should use the last packet present in the network port as our seed choice, as the time of the day is too easy for Eve to predict. But can we be certain that Eve will not be able to guess our random numbers? The answer is no, since *in principle*, Eve could be eavesdropping in the entire network, and predict the packets being read by Alice and Bob's software-based random number generator.

The more we think about it, the more obvious it becomes that software-based random number generators are not the way to go when we want protection against an eavesdropper with unlimited access to technology. There are physical

PUC-Rio - Certificação Digital Nº 0510513/CA

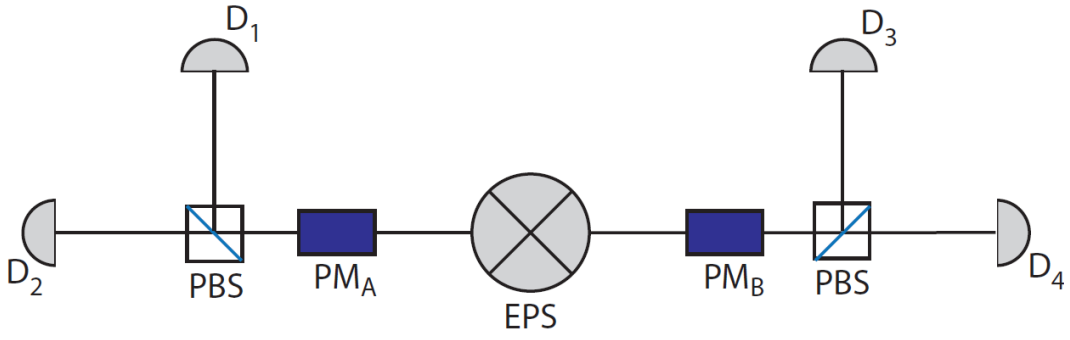
processes which can give better seeds, such as chaotic process like noise fluctuations in a resistor. This is definitely harder to predict by Eve, but once again not impossible. What can we use that Eve cannot predict? Fortunately for us there are processes in nature that are truly random, and thus unpredictable. The answer of course, is given by quantum physics. The basic foundations of quantum physics tell us that given that the wavefunction of a particle is known, we can only then speak of the particle in terms of probabilities. If we can perform a measurement on a degree of freedom of a quantum particle with maximum uncertainty, then we have ourselves a random number which cannot be predicted.

Such a device is called a quantum random number generator (QRNG), and there have been successful devices built recently using the idea of which port a photon exits a beamsplitter [112, 113] and the time-of-arrival of a photon [114, 115]. These generators provide truly random numbers and if they are not used as a seed provider, but rather their output sequence is directly used to choose the basis (and in Alice's case to generate the actual key itself), then Eve has no way to guess the numbers. If it is used to generate seeds, and then expanded using a mathematical algorithm, the sequence generated in this way is no longer truly random and for this matter not secure [111]. Therefore to be complete foolproof against Eve, the QRNG must provide the random bits in the rate the system requires to operate, without any expansion of the random number sequence. Rates of Mbit/s is achievable with commercial QRNGs [24]. However there have been recent experiments with high-rate QKD, reaching Gbit/s data rates [54, 116], and it will likely become the standard in the near future. We present here a protocol which generates random numbers for QKD independent of the rate required, and furthermore, it only requires small modifications to the hardware, as it is based on the detectors already used in a typical QKD setup. It can fully replace QRNGs in the case of QKD with an entangled photon pair source or an HSPS, and can replace Bob's QRNG for a standard QKD setup.

We will first briefly explain how QKD works with an entangled photon pair source. Such a source produces entangled-photon pairs (we will use polarization entanglement as our example, but it works for any other degree of freedom). One photon of the pair is sent to Alice and the other to Bob as shown in Fig. 31. Let us assume the wavefunction generated by the source is

PUC-Rio - Certificação Digital Nº 0510513/CA

$|\psi\rangle = 1/\sqrt{2}\left(|H\rangle_a|V\rangle_b + |V\rangle_a|H\rangle_b\right)$ where the subscripts *a* and *b* stand for the photons going to Alice and Bob respectively. The basic idea behind this protocol relies on the fact that the photons going to Alice and Bob are entangled. It is a well known fact, that due to entanglement if Alice uses the same measurement basis as Bob (through polarization modulators PM_A and PM_B, which work as automatic wave plates) their measurement results will always be correlated, that is, for our wavefunction if Alice measures H, Bob will obtain V and vice-versa. The remarkable about this, is that *any* identical measurement basis yields perfect correlations. Ekert in 1991 realized this could be used for QKD if the measurement choices are performed at random and independently by Alice and Bob, hence each of them needs a QRNG [117]. It is possible to remove the requirement of QRNGs, by employing a passive basis choice at both Alice and Bob's stations, however four (or more) detectors are required [118]. The protocol briefly runs as follows: for each incoming photon of the pair, Alice and Bob choose a random measurement basis and record the results. After the photons are received they publicly reveal over a classical communications channel the measurement bases chosen for each detected photon. Like in BB84, they discard all values where incompatible measurement bases were used. The standard steps of BB84 now follow: Eve's presence verification, error correction and privacy amplification. This scheme has the advantage of having another security check: the violation of Bell's inequalities [1, 117]. If Eve attempts to perform an attack there will no longer be a violation of Bell's inequalities from the photon's correlations and therefore Eve is caught. In practice a modified version of Bell's inequalities called the Clauser-Horne-Shimony-Holt inequality (CHSH) [37] is used, since it allows a check of violation using directly measurable quantities. This scheme is secure even if Eve has control of the source [5].



PUC-Rio - Certificação Digital Nº 0510513/CA

Figure 31 - Scheme for QKD with an entangled-photon source based on the E91 protocol. PBS: Polarization beam splitter; PM: Polarization modulator; EPS: Entangled photon source.

Our protocol is therefore meant to be used to perform the random choices required on an Ekert based QKD protocol with active-basis choices. We will later extend it to be used with BB84 with an HSPS for both Alice and Bob, or just Bob in the case of standard BB84. A recent theoretical modification [119] shows that three bases are needed at Alice, but Bob can use two (when compared to the original case that both need three) for QKD and a full check of the CHSH inequality. It runs as follows: In our proposal both Alice and Bob have a clock generator each, working asynchronously from each other. This signal can be easily generated within modern electronics already used in a QKD setup. The main idea is to generate the random numbers based on the number of clock pulses between consecutive photon detections, that is, between consecutive generations of electrical pulses in the output of their detectors. Since the entangled source may not be trusted, the following procedure is performed: initially, both Alice and Bob block their detector inputs and wait for the first dark count, thus obtaining random and independent integers equal to the number N_{A0} (N_{B0} for Bob) of clock pulses until either detector has fired. Alice (Bob) will proceed to calculate N_{A0} mod 3 \ (N_{B0} mod 2) and choose one of the 3 (2) needed bases for the first transmitted qubit depending on the result. The detector inputs are opened, the quantum transmission starts and they count the number of pulses N_{A1} (N_{B1}) until the next detection, perform N_{A1} mod 3 (N_{B1} mod 2), choose the basis for the next qubit and so on, where a random number $N_{AT(BT)}$ will be obtained between detections *T-1* and *T*. Since the times of detection follow an exponential distribution [114,115], the generated sequence will be random. As long as we can assume Eve is not looking inside the detectors (However, Eve can still ``hear'' the detection clicks without gaining any information, in exactly the same way as in the standard BB84 protocol), she will not be able to guess the bases used, as required in all QKD protocols. The scheme is presented in Fig. 32 for the particular case of the E91 protocol with polarization coding in optical fibers. Our protocol is independent of the coding method, and can be readily applied to phase coding systems.

PUC-Rio - Certificação Digital Nº 0510513/CA

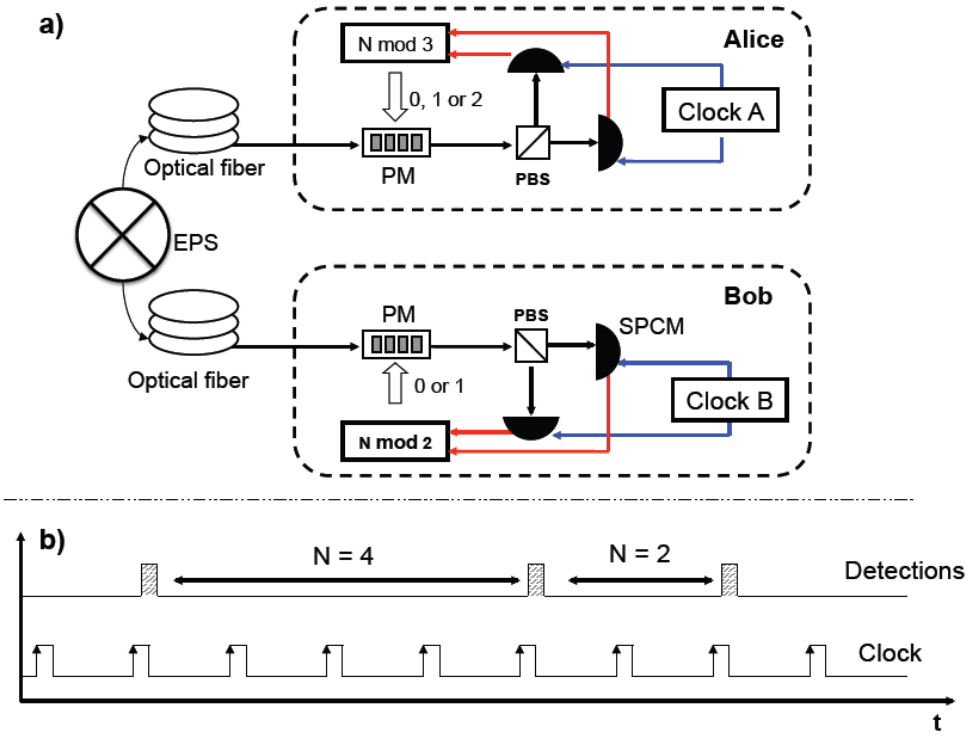


Figure 32 - Schematics of our proposal applied to the Ekert (E91) protocol. Black arrows represent optical connections, while blue and red ones depict electrical cables. EPS - Entangled photon source; PM - Polarization modulator; PBS - Polarizing beam splitter; SPCM - Single photon counting module. The master clock synchronizing Alice and Bob, as well as QKD electronics are omitted for the sake of clarity. b) Illustrative representation of the waveforms from the detection and clock pulses.

We can also extend this idea to the BB84 protocol if Alice has a HSPS, since she will have one detector yielding a trigger signal to pinpoint that an idler photon was created which is sent to Bob. Once again they also each have a local clock generator which they use as a timing reference. Alice performs her basis choices calculating N_{AT} mod 4 and converting the 4-valued number into two random bits. Bob performs the same procedure as before, he begins to count the N_{B0} number of clocks before any photons are transmitted until he detects a dark count (once again he blocks the detector input), and calculates N_{B0} mod 2 to determine the first basis to be used. He then proceeds calculating N_{BT} mod 2 for each received photon. Finally if the standard BB84 scheme is to be used with an attenuated pulsed laser source, then only Bob can use our protocol to generate random numbers. Since Alice does not detect the photon she does not have available to her a quantum event to be used as a random generator. In principle she can use a quantum non-demolition measurement (QND) and verify if a photon

PUC-Rio - Certificação Digital Nº 0510513/CA

is emitted by her laser, and with a local clock generator use this event to create a random number. The technology to perform such a measurement is not yet practical, therefore we shall only leave this case here as a possibility.

Let us now discuss what are Eve's options in cracking such a scheme. Her objective is to try to predict which bases Alice and Bob will use in the E91 protocol, or just Bob in BB84. Short of intercepting the photons (in which case her presence is revealed anyway through non-violation of the CHSH inequality), the simpler attack she can attempt is to perform a QND predicting the time instants a photon is passing through the fiber links to Alice and Bob. Assuming the QND is successful, Eve can try to predict the basis choices Alice and Bob will choose. Several factors come into play now. First of all detectors are not perfect, and if we assume real detectors are used (quantum efficiencies less than unity) then detections are randomized. This may not be enough to counter Eve, and it is not future-proof since detectors with extremely high efficiencies may appear in the future. What Alice and Bob can do is to use a local clock with a much higher resolution than the gate width of the detector (2.5 ns is a typical value for InGaAs SPADs). This way they add an extra degree of randomness to the system by detecting where in the gate pulse the photon arrives. If the coherence length of the photons is longer than the gate width (tricky in SPDC based sources), then Alice and Bob can be sure the time of arrival of the photon inside the gate window is truly random. In addition to this, the photons need to be single-mode inside the detection window, such that they are spatially indistinguishable. This is the only way to be immune against Eve. If this is not possible the best that can be done is to use a local clock, not only with a high-resolution, but also with jitter (in practice all clock signals exhibit some degree of jitter), to further randomize results.

If we now consider the poissonian probability of existing n photons inside a detection window with μ photons per pulse $p(n) = e^{-\mu}\mu^n / n!$, the probability of detecting a photon on the N_{th} gate window can be written as:

$$P(N) = (1-\mu\eta)^{N-1} \cdot \mu\eta \tag{3.7}$$

PUC-Rio - Certificação Digital Nº 0510513/CA

where η is the detection efficiency. From the poissonian distribution we can rewrite the probability of detecting a photon as:

$$P = 1 - e^{-\mu\eta} \tag{3.8}$$

Equation (3.8) can be interpreted simply as having a perfect non-resolving photon number detector with an external loss given by the detection efficiency. We can now use Eq. (3.8) to write the probability of detection as:

$$P(N) = \left(e^{-\mu\eta}\right)^{N-1} \cdot \left(1 - e^{-\mu\eta}\right) \tag{3.9}$$

Finally from Eq. (3.9) we write the probabilities of detecting photons in odd and even detection windows, summing over *N* for both cases:

$$P_{EVEN} = \frac{1}{1 + e^{\mu\eta}} \tag{3.10}$$

$$P_{ODD} = \frac{1}{1 + e^{-\mu\eta}} \tag{3.11}$$

From these two equations we note that the probability of odd and even detection events are different. This means the sequence generated from *N* mod 2 will be unbalanced in terms of zeros and ones. The above formulation was done using the gate window itself as the time unit. Intuitively it is easy to see that the bias will diminish if a higher resolution clock is used, vanishing completely at the limit of infinite high resolution, which reinforces the idea of using such a clock as explained before. Post-processing classical procedures can be used to balance the sequence [113], however the price we have to pay for this is a shortened sequence.

We can also extend the previous discussion to the case with light sources with thermal statistics. In this case the probability to find *N* photons per detection window with μ photons per window on average is [16]:

$$p(n)\frac{\mu^n}{(\mu+1)^{n+1}} \tag{3.12}$$

PUC-Rio - Certificação Digital Nº 0510513/CA

Using the same line of reasoning as before, we can arrive at the following probabilities for odd and even events:

$$P_{EVEN} = \frac{1}{\mu\eta + 2} \tag{3.13}$$

$$P_{ODD} = \frac{\mu\eta + 1}{\mu\eta + 2} \tag{3.14}$$

The probabilities for odd and even events once again differ, but the unbalance in the sequence can be made to vanish if a higher-resolution clock is used as explained before. Two proof-of-principle experiments were performed to demonstrate the idea, the first one uses a simple setup of a fiber pigtailed CW semiconductor laser (λ = 1549.32 nm) in series with a calibrated attenuator, which is then connected to an InGaAs SPAD, to which we triggered at a constant frequency using an external pulse generator. The results of this experiment have been presented in [120]. The output of the APD is connected to a fast A/D card (analog to digital) plugged into a computer. The trigger output from the pulse generator is connected to a second input of the same A/D card, and the data acquisition software we wrote counts the number of trigger pulses between each consecutive detection. Therefore we assemble the statistics of arrival times of single photons.

We employed a trigger frequency of 100 kHz, with an average photon number per detection window of 0.1, a typical value for QKD. The gate width used was 20 ns, which is quite wide compared to many QKD experiments. The reason for this is due to a timing problem in our APD, which reduces the quantum efficiency for narrow gate widths. Therefore, 20 ns was used in order to increase the quantum efficiency. Dark counts were also increased (10^{-4} per 1ns gate) but it is not as major issue as this is a proof-of-principle experiment. We obtained 500 x 10^3 counts and plotted the histogram of the times of arrival (Fig. 33). As expected the probability distribution of the arrival times is exponential since the probability to obtain *n* photons in a gate window follow a poissonian law [112].

PUC-Rio - Certificação Digital Nº 0510513/CA

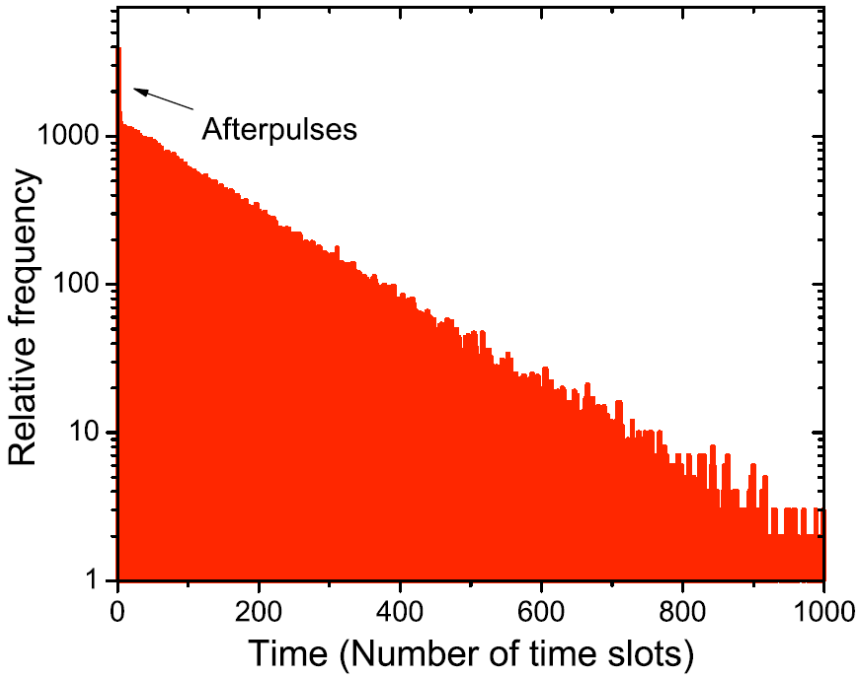


Figure 33 - Measured histogram of the number of time slots between two consecutive successful detections for μ = 0.1.

PUC-Rio - Certificação Digital Nº 0510513/CA

Observing the figure we notice two things. The first is that the reason for the unbalance comes from the exponential shape of the distribution itself, since it is always more probable for an odd event to occur than for an even one, and the slope of the distribution depends on the average photon number. The other is the clear point out of the curve in the first bin (the bins have been graphically enlarged to make the first bin stand out). Through simulations and other experimental runs, we have deduced the first bin is a result of afterpulsing. There is a current study going on to characterize afterpulses in SPADs using this technique [121].

The sequence which generated the histogram is used to create our random sequence through the $N \bmod 2$ operation, giving us a sequence of zeros and ones. The first question that comes to mind is, how random is our sequence. The first test we attempt is to calculate and plot the normalized auto-correlation function of the sequence. One of the requirements for randomness is that the sequence does not present patterns, and the auto-correlation function is a good test against repetitions (Fig. 34). As we can observe from the measurements, the function only displays a single central peak, with the rest being uncorrelated, a good sign that the sequence has no patterns. This is not a full guarantee that our

sequence is random, but it already points in the right direction. The sequence balance ratio is of 0.974 between zeros and ones.

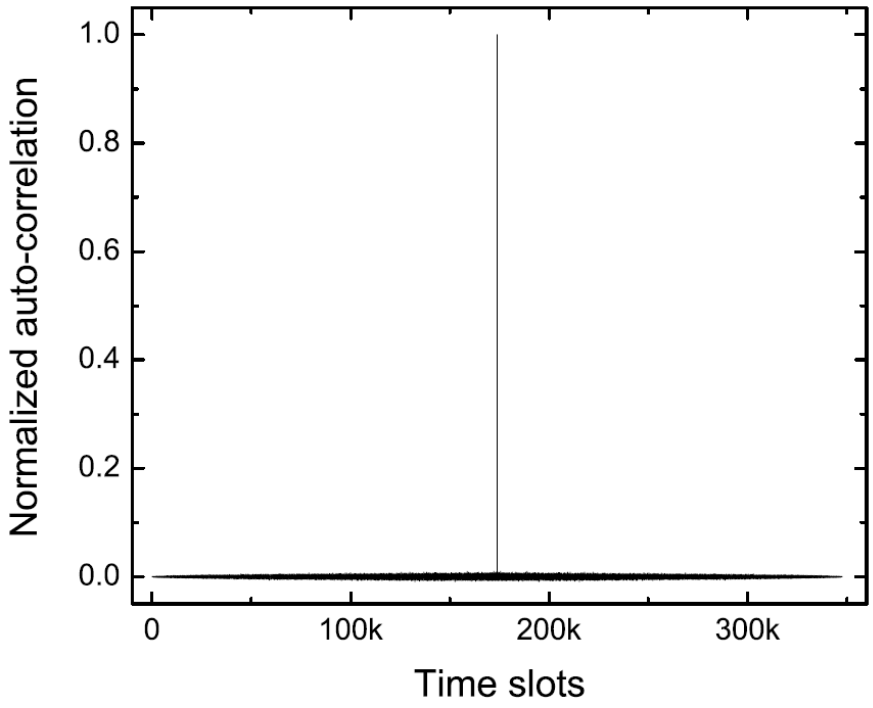


Figure 34 - Normalized auto-correlation for the random sequence generated from the distribution from Fig. 33.

There is a standard test suite to verify the randomness of a number developed by NIST and freely available on the Internet [122]. It is composed of 13 tests to be applied to a sequence, and each test gives a result called a p-value. As long as the p-value is larger than the confidence value α (for cryptographic applications, $\alpha = 0.01$), then the sequence is random with a very high probability [122]. The test expects a balanced random sequence, and if that is not the case, many of the tests fail. We generated two sequences of 1 million bits each (the number of bits required for a single run of the test) in order to perform the test twice. We increased the average photon number to $\mu = 0.4$ in order to speed up the measurement, otherwise the time taken for the data acquisition would be too long. The reason why time seems to be of an issue here is due to inefficiency in our data acquisition setup. The sequence balance ratio dropped to 0.944 due to the photon number increase, and as such, it does not pass in the NIST test. We balanced it using the simple procedure of XORing our sequence with a 0101… sequence, which is equivalently to flipping the bit assignments at each detection (even stops being zero and becomes one). The sequence was successfully balanced having

PUC-Rio - Certificação Digital Nº 0510513/CA

now a ratio of 0.9996 which is good enough for the NIST test. The results are presented in the table below:

NIST test	P-value 1	P-value 2
Frequency	0.496504	0.972877
Block-frequency	0.186886	0.618073
Cumulative-sums forward	0.712049	0.496736
Cumulative-sums reverse	0.300269	0.524704
Runs	0.060592	0.948969
Longest runs	0.130525	0.425877
Rank	0.365876	0.981847
DFFT	0.600927	0.129989
Universal	0.792907	0.087627
Apen	0.110230	0.406134
Serial 1	0.447454	0.566796
Serial 2	0.867164	0.399266
Linear complexity	0.956506	0.424975

These results clearly show our sequence is random according to the NIST test demonstrating the feasibility of our scheme. As a second proof-of principle experiment we used a SPDC process to create the photons used for the random number generation. This experiment represents what would be used in an Ekert-type QKD protocol. The results of this experiment have been published with a few improvements in [123]. We used a 20 mm long Periodically-Poled Lithium Niobate (PPLN) crystal pumped by a 532 nm CW Nd:YAG laser. The crystal is identical to the ones used in the two experiments of Chapter III, except that it is shorter. Therefore it provides type - I quasi-phase matching for $532 \rightarrow 809 + 1555\,\text{nm}$ when the crystal is heated at approximately 90ºC, this temperature is slightly different than the one used for the longer crystal in the previous experiments. The experimental setup is presented in Fig. 35. The pump light passes through an optical attenuator, and then goes through a half-wave plate to adjust the pump's polarization before the crystal. It is then focused on the 20 mm long crystal using an achromatic doublet lens L with a 100 mm long focal length. The beam is focused in the middle of the crystal, and the output beam is

PUC-Rio - Certificação Digital Nº 0510513/CA

collimated by an identical lens before the prism P, used to spatially split the generated beams and the pump. A bulk filter (RG 715) is placed before the fiber coupler (FC) to remove any residual pump light from the detector. An aspheric lens (f = 11 mm) is used to focus the signal beam on a standard 780 nm single mode fiber (SMF), mounted on a multi-axis translation stage. The fiber is connected to a Si Avalanche Photo Detector (Id Quantique ID100-MMF50). The output from the detector is finally connected to an A/D card (20 MSamples/s), which is attached to a personal computer to process the data. This time, we used the internal clock of the A/D card as our timing clock. The optical attenuator was adjusted so that a 150 kHz detection rate was obtained with optimized coupling. The measured input optical pump power on the crystal for that rate is 9.8 mW. The experiment we perform here is a perfect representation of Alice in the case where she uses a heralded single-photon source. It also represents the basic building block of the E91 protocol, with the source located somewhere between Alice and Bob. Our setup can therefore be easily upgraded into a single-crystal entangled photon pair source [84,124].

PUC-Rio - Certificação Digital Nº 0510513/CA

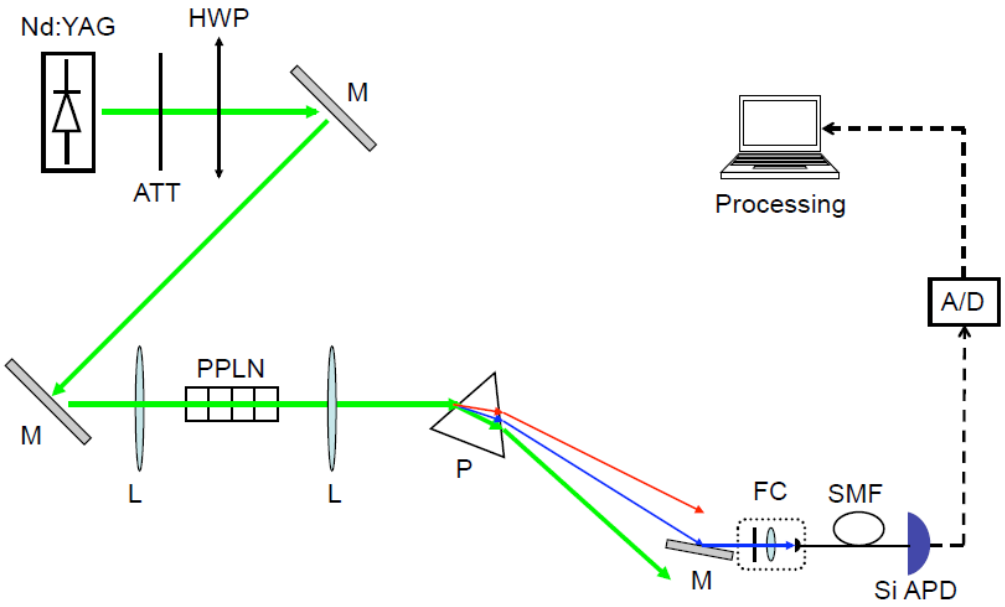


Figure 35 - Experimental setup: ATT: Optical attenuator; HWP: Half-wave plate; M: Mirror; L: Lens; PPLN: Periodically-poled lithium niobate; P: Prism; FC: Fiber coupler, here consisting of a multi-axis translation stage (not shown here), RG 715 high-pass filter, 11 mm focal length aspheric lens and fiber holder; SMF: 780 nm single mode optical fiber; APD: Avalanche photon detector; A/D: Analog to digital converter. The green, red

and blue arrows represent the pump, idler and signal beams respectively. The dashed lines represent electrical cables.

The next step is the NIST set of randomness tests. We generated a sequence of 20 million bits at 150 kHz count rate, using the previously described *N* mod 2 procedure. Since the clock resolution was limited by the A/D card sampling rate (20 MHz) it was not enough to entirely remove the bias, as discussed before. We applied the same balancing procedure with the XOR operation and increased the balance from 0.9968 to 0.9995, which was enough to pass the tests. The results for the 13 tests are presented in Fig. 36, which indicates that our sequence is random with a very good degree of confidence.

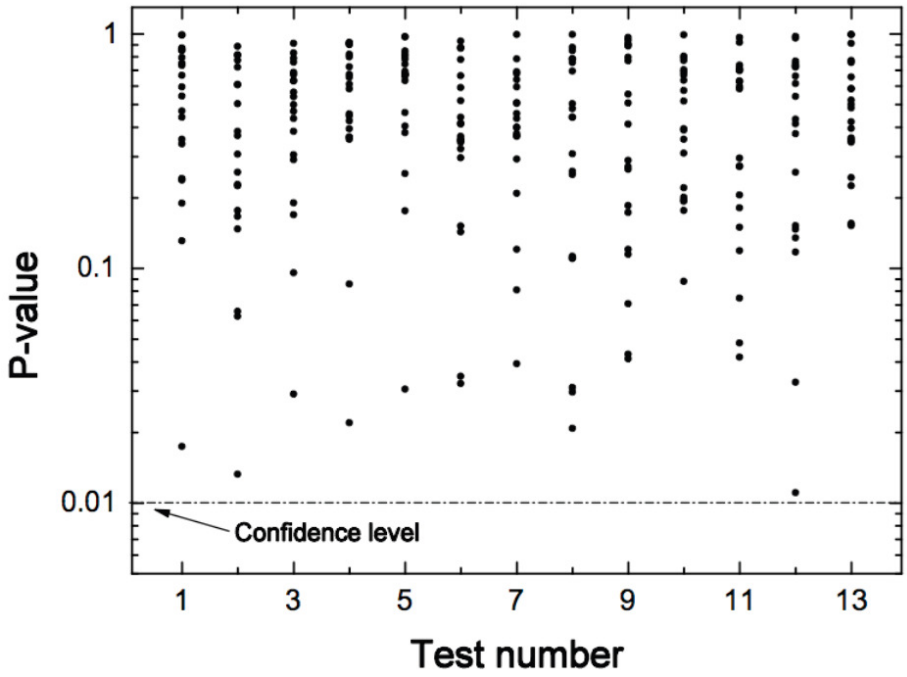


Figure 36 - P-values plotted for the NIST test suite individual tests for the 20 million bit generated sequence after bias removal. Each dot represents a run of 1 million bits for a particular test. The results are all above the confidence value for cryptography applications. The tests are: 1 - Frequency; 2 - Block frequency; 3 - Cumulative-sums forward; 4 - Cumulative-sums reverse; 5 - Runs; 6 - Longest runs; 7 - Rank; 8 - DFFT; 9 - Universal; 10 - Approximate entropy; 11 - Serial 1; 12 - Serial 2; 13 - Linear complexity.

A scheme was presented to perform true random basis choices for the E91 protocol which is based on the hardware which is already present in any QKD system. It may also be extended to be used by Bob in the BB84 protocol, or even Alice, if she uses a heralded single photon source. Our proposal has the advantage

PUC-Rio - Certificação Digital Nº 0510513/CA

of being readily scalable ``on-the-fly'' with the transmission rate without any active changes from the user, as long as high-resolution local clocks are used. Therefore, it could be used in future entangled based QKD networks [125] or any quantum cryptography system which employs a variable key rate. The protocol can be implemented with simple modifications and it replaces true RNGs for the active basis choices, decreasing the building cost of a practical QKD setup. We have shown that the generated sequence is indeed random like we expected, and supports our idea of random number generation in QKD systems. If Alice and Bob employ asynchronous clock generators with a timing resolution higher than the detection jitter of the single-photon counting modules, and the light source coherence time is longer than the detection window, then Eve cannot gain any information on the basis selection.

PUC-Rio - Certificação Digital Nº 0510513/CA