Quantum Random Numbers: Certification and Generation

Alastair Avery Abbott

under the supervision of

Professor Cristian S. Calude

MSc Thesis

A thesis submitted in fulfilment of the requirements
for the degree of Master of Science
in Computer Science, The University of Auckland, 2011.

Abstract

In this thesis we study the generation of randomness from quantum mechanical sources: quantum random number generators (QRNGs). Such devices are thought to provide a better quality of randomness than is possible with classical devices, but the issue is in need of more rigorous study.

In Chapter 1 we provide the necessary background for the thesis from both computability and probability theory. We then present and extend some recent results providing a more theoretical grounding for the indeterminism in quantum mechanics. In particular, we show that sequences of quantum random bits are incomputable in the strongest sense: no bit can be provably computed in advance.

In Chapter 2 we study in detail the use of von Neumann normalisation for QRNGs producing both finite strings and infinite sequences of bits; such normalisation methods are necessary in order to counter for experimental imperfections. We show that, in the case of a constantly biased source, this normalisation gives the desired uniform distribution. The effect of this normalisation on the (in)computability of the generated sequences is studied, and it is shown that algorithmic randomness and Borel normality are preserved, but ε-randomness (and thus incomputability) is, in general, not. Experimental bounds for the extent of departure from the uniform distribution in the non-ideal case of a slowly drifting bias are derived.

In Chapter 3 we propose a new QRNG which uses entangled photon pairs and is certified to produce incomputable bits by value indefiniteness; this is the first QRNG with explicit certification by a physical principle. The effects of various experimental imperfections are discussed in detail, and care is taken to make the proposed QRNG as robust as is possible to these. A robust method based on XORing the produced bitstrings together is proposed to further improve the quality of the distribution produced by the QRNG. Various improvements and optimisations to this scheme are discussed.

Acknowledgements

I would foremost like to thank my supervisor, Cristian Calude, whose enthusiasm for this work was unfaltering. Secondly, I give particular thanks to Karl Svozil, who much of this work was conducted in cooperation with, and whose expertise and viewpoints were extremely instructive. I would like to acknowledge Marius Zimand, Sonny Datt and Tania Roblot for comments which helped improve this thesis.

I gratefully acknowledge the support of: a) the Centre for Discrete Mathematics and Theoretical Computer Science and the Department of Computer Science at the University of Auckland, for providing funding to allow me to work abroad on this research; b) the Technical University of Vienna, for their hospitality and support while visiting to work on this project; c) the University of Auckland, for the Masters scholarship which this study was conducted under.

Finally, I would like to thank my partner, Nicole, as well as my parents, Ken and Marylin, for their support throughout a busy and much-travelled year.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Randomness is a particularly elusive concept. Random numbers have been used for thousands of years, but what exactly are random numbers? Randomness is often understood through various 'symptoms', some statistical and others not. Three typical such symptoms are unpredictability, lack of bias and lack of patterns. But how can we cast this intuition more formally?

Randomness plays an essential role in probability theory, the mathematical calculus of random events. Kolmogorov axiomatic probability theory assigns probabilities to sets of outcomes and shows how to calculate with such probabilities; it assumes randomness, but does not distinguish between individually random and non-random elements. For example, under a uniform distribution of bits, the outcome of n zeros has the same probability as any other outcome of length n, namely 2^{-n}.

Algorithmic information theory (AIT) [Cha77], defines and studies individual random objects, such as finite bit-strings or infinite sequences, in order to overcome the shortcomings of the probabilistic notion of randomness. AIT shows that 'pure randomness' or 'true randomness' does not exist from a mathematical point of view. For example, there is no infinite sequence passing all tests of randomness. Furthermore, randomness cannot be mathematically proved: one can never be sure a given sequence is random, there are only forms and degrees of randomness.

Returning to our symptoms of randomness, we can relate these to formal concepts as follows:

1) *Unpredictability* is a manifestation of strong incomputability—a sequence in which it is impossible to compute the next bit is in a strong sense unpredictable;

2) *Uniform distribution* of the generated bits (not simply of individual bits, but of all n-bit strings), a manifestation of the Borel normality of a sequence;

3) *Lack of patterns*, a manifestation of algorithmic randomness (incompressibility) of a sequence.

Mathematically, it is well known that 3) $\rightarrow$ 1) and 3) $\rightarrow$ 2), but both the converse implications are false. Further, both implications 1) $\rightarrow$ 2) and 2) $\rightarrow$ 1) are false [Cal02]. As a result, it is the strong notion of algorithmic randomness defined in AIT [Cha77] which should be considered the goal for any device generating randomness.

Randomness in this strong sense is impossible to obtain through computational sources, so, given the prevailing canon that quantum mechanics is 'intrinsically random', it is not surprising that there has been a considerable degree of interest in using quantum mechanics to generate random numbers. However, this randomness is postulated rather than deduced and in need of a solid theoretical grounding.

In this thesis we tackle quantum random number generators from both an algorithmic and a probabilistic viewpoint. In order to do so it is, perhaps surprisingly, necessary to study infinite sequences (rather than finite strings) of bits produced by such a device. Firstly, we expand upon the results of Calude and Svozil [CS08] to attempt to understand the extent to which 'quantum randomness' fulfils the mathematical notion of randomness. We show via the Kochen-Specker Theorem that such sequences are strongly incomputable (and hence satisfy the symptom of unpredictability). However, it is not yet known if such sequences are algorithmically random, and further research is needed to clarify this point. Secondly, we study the probability distribution of quantum random sequences, and show that the bits are uniformly distributed (it is not known if the sequences of quantum bits are Borel normal, although uniformity guarantees this with probability one).[1] We bring these concepts together and study the effects of von Neumann normalisation procedures (which are necessary to ensure uniformity of bits) on the incomputability of quantum random bits. Finally, we propose a quantum random number generator which is explicitly certified to produce incomputable bits and in which special emphasis is put on making it robust to experimental imperfections and ensuring it produces the correct probability distribution. Many of the results of this thesis have been communicated through the papers [AC10] and [ACS10].

1.1 Preliminaries

In this section we present the main notation used throughout this thesis, as well as presenting the key background concepts from computability theory, algorithmic random-

[1] It is important to note the subtle theoretical difference between a *probability-one* event and a *provably guaranteed* event in the probability space of infinite sequences: in contrast with a provably guaranteed event whose complement is empty, the complement of a probability-one event can be not only non-empty, but even infinite [Hal74].

ness and axiomatic probability theory. We adopt the notation of [Cal02] throughout this thesis.

1.1.1 Notation

By 2^X we denote the power set of X. By $|X|$ we denote the cardinality of the set of X. Let $B = \{0, 1\}$ be the binary alphabet and denote by B^* the set of all bit-strings (λ is the empty string). If $x \in B^*$ and $i \in B$ then $|x|$ is the length of x and $\#_i(x)$ is the number of i's in x. For $x, y \in B^*$ with $|x| = |y|$ we write $d(x, y)$ for the Hamming distance between x and y, i.e the number of $1 \leq i \leq |x|$ such that $x_i \neq y_i$. We extend the exclusive-or (XOR) operation, denoted $\oplus$, from individual bits to bit-strings in the natural bit-wise fashion. That is, for two strings $x, y \in B^*$ with $|x| = |y|$, $z = x \oplus y$ has $z_i = x_i \oplus y_i$ for $1 \leq i \leq |x|$. By B^n we denote the finite set $\{x \in B^* \mid n = |x|\}$. The concatenation product of two subsets $X, Y \subseteq B^*$ is defined by $XY = \{xy \mid x \in X, y \in Y\}$. If $X = \{x\}$ then we write xY instead of $\{x\}Y$.

By B^ω we denote the set of all infinite binary sequences. For $\mathbf{x} \in B^\omega$ and natural n we denote by $\mathbf{x}(n)$ the prefix of $\mathbf{x}$ of length n. We write $w \sqsubset v$ or $w \sqsubset \mathbf{x}$ in that case that w is a prefix of the string v or the sequence $\mathbf{x}$.

1.1.2 Computability

We briefly present the required concepts from computability theory and algorithmic randomness. The reader is referred to either [Cal02] or [DH10] for a thorough review of these fields.

Definition 1. A sequence $\mathbf{x} = x_1 x_2 \cdots \in B^\omega$ is *computable* if there exists a computable function $f : \mathbb{N} \to B$ such that for all $i \in \mathbb{N}$, $f(i) = x_i$. A sequence which is not computable is called *incomputable*.

A stronger property is bi-immunity: a sequence $\mathbf{x}$ is bi-immune if only finitely many bits of $\mathbf{x}$ are computable. More formally:

Definition 2. A sequence $\mathbf{x} = x_1 x_2 \cdots \in B^\omega$ is *bi-immune* if there is no Turing machine T with infinite domain $\mathrm{dom}(T)$ such that for all $a \in \mathrm{dom}(T)$ we have $T(a) = x_a$.

A prefix-free (Turing) machine is a Turing machine whose domain is a prefix-free set of strings. The complexity of a string will be an important concept in classifying forms of randomness. Loosely speaking, the complexity of a string x with respect to a particular Turing machine T is the size of the shortest program for T which halts and outputs x.

Definition 3. The *prefix complexity* of a string σ, $H_W(\sigma)$, induced by a prefix-free machine W is $H_W(\sigma) = \min\{|p| \ : \ W(p) = \sigma\}$.

Definition 4. Fix a computable ε with $0 < \varepsilon \leq 1$. An ε-universal prefix-free machine U is a machine such that for every machine W there is a constant c (depending on U and W) such that $\varepsilon \cdot H_U(\sigma) \leq H_W(\sigma) + c$, for all $\sigma \in B^*$. If $\varepsilon = 1$ then U is simply called a *universal* prefix-free machine.

We now present the formal definition of randomness which will be crucial in this thesis. This formal definition of randomness is due to Chaitin [Cha77] and is based on the notion that a random sequence should be incompressible; it is equivalent to other definitions such as Martin-Löf's measure theoretic formulation [ML66]. The intuition behind this is that a random sequence contains no computable patterns; such patterns would allow the sequence to be compressed and would hence be symptoms of non-randomness.

Definition 5. A sequence $\mathbf{x} \in B^\omega$ is called ε-random if there exists a constant c such that $H_U(\mathbf{x}(n)) \geq \varepsilon \cdot n - c$, for all $n \geq 1$. Sequences that are 1-random are called *algorithmically random*, and when the context is clear we will simply refer to them as random.

Borel normality allows us to determine if a string or sequence is 'typical' or 'normal' in the sense that all substrings occur within it equally often. For finite strings it only makes sense to consider substrings of reasonably short length. Let $N_i^m : B^* \to \mathbb{N}$ count the number of non-overlapping occurrences of the ith (in lexicographical order) binary string of length m in a given string. We can then proceed to define Borel normality for strings [Cal94].

Definition 6. Let $x \in B^*$ be a non-empty string and $1 \leq m \leq |x|$.

1) We call x *Borel m-normal* if for every $1 \leq i \leq 2^m$ we have:

$$\left| \frac{N_i^m(x)}{\lfloor \frac{n}{m} \rfloor} - 2^{-m} \right| \leq \sqrt{\frac{\lg |x|}{|x|}}.$$

2) If for every natural $1 \leq m \leq \lg \lg |x|$, x is Borel m-normal, then we call x *Borel normal*.

Borel normality is readily generalised to infinite sequences (indeed, historically it was first formulated for infinite sequences or, equivalently, real numbers) by requiring that it holds in the limit for substrings of increasing length. Interestingly, for infinite sequences we can equally well count the number of overlapping substrings, $\mathcal{N}_i^m$, as non-overlapping substrings, N_i^m [KN74].

Definition 7. Let $\mathbf{x} \in B^\omega$ be a sequence and $m \geq 1$.

1) We call $\mathbf{x}$ *Borel m-normal* if for every $1 \leq i \leq 2^m$ we have:

$$\lim_{n\to\infty} \frac{N_i^m(\mathbf{x}(n))}{\left\lfloor \frac{n}{m} \right\rfloor} = \lim_{n\to\infty} \frac{\mathcal{N}_i^m(\mathbf{x}(n))}{n} = 2^{-m}.$$

2) If for every natural $m \geq 1$, $\mathbf{x}$ is Borel m-normal, then we call $\mathbf{x}$ *Borel normal*.

Random sequences are Borel normal, but the converse is false: there are computable (even primitive recursive) Borel normal sequences, like Champernowne's sequence. In general, ε-random sequences with $\varepsilon \in (0, 1)$ are not Borel normal [CHS10].

1.1.3 Probability spaces

A probability space is a measure space such that the measure of the whole space is equal to one [Ros06]. More precisely, we have the following definition.

Definition 8. A *(Kolmogorov) probability space* is a triple $(\Omega, \mathcal{F}, P)$ where the sample space, Ω, is a non empty set, $\mathcal{F}$ is a σ-algebra on Ω, and the countably additive function $P : \mathcal{F} \to [0, 1]$ is the probability measure satisfying $P(\Omega) = 1$ and $P(\emptyset) = 0$.

It will be important to have a quantitative measure of how 'close' two probability measures are. This is realised by the total variation distance.

Definition 9. The *total variation distance* between two probability measures P and Q over the space Ω is $\Delta(P, Q) = \max_{A\subseteq\mathcal{F}} |P(A) - Q(A)|$. We say that P and Q are ρ-close if $\Delta(P, Q) \leq \rho$.

It is well known (see, for example, [Vad11]) that the following Lemma holds.

Lemma 10. *For finite Ω we have* $\Delta(P, Q) = \frac{1}{2} \sum_{x\in\Omega} |P(\{x\}) - Q(\{x\})|$.

1.2 Random number generators

Random numbers have been used for more than 4,000 years, but never have they been in such demand as in our time. People use random numbers everywhere, in applications such as simulations [RK07], randomised algorithms [Rab76] and cryptography [MvV01]. But to have access to random numbers to use, we need a source of random numbers: a *random number generator (RNG)*.

Computers offer a fast and easy way to generate 'random numbers' produced by algorithms. However, as we have seen no computation can yield a sequence free of patterns, and, as von Neumann aptly put it, "Any one who considers arithmetical methods

of producing random digits is, of course, in a state of sin" [vN63]. Computer scientists needed a long time to realise this fact—that randomness produced by software is not random, but only *pseudorandom*. This form of randomness mimics well the human perception of randomness, but its quality is rather low because computability destroys many symptoms of randomness, e.g. unpredictability. It is not totally unreasonable to put forward that pseudorandomness rather reflects its creators' subjective 'understanding' and 'projection' of randomness.[2]

Such methods are, nonetheless, still widely used, and focus is generally put on disguising the patterns in such sequences and ensuring the bits are uniformly distributed (i.e., sampled from a uniform distribution). Most pseudorandom number generators (as such computational RNGs are referred to) make use of a short 'random' seed—whether it be from the system clock or the user's mouse movements—and use this to generate longer strings of uniformly distributed bits [Vad11]. However, such methods suffer from the same problems as any pseudorandom method does, and the issue of obtaining a random seed remains.

There is danger in over-emphasising the need for uniformity of distribution in RNGs—this is, after all, only one of many necessary symptoms of randomness—and it is possible to have simple, deterministic sequences which satisfy this requirement. For example, Champernowne's sequence, obtained by concatenating the binary representation of the integers, 011011100101110 ... is known to be Borel normal [Cha33, BC02], but clearly a device outputting consecutive bits of this sequence is not a good RNG. Further still, some simulations using pseudorandom number generators have been shown to fail as a result of the quality of pseudorandomness used [FLW92].

To counter these shortcomings it seems promising to attempt to utilise physical sources of randomness. Classically, sources such as dice throwing, apart from being slow, have been shown to have flaws too. Diaconis et al. [DHM07] showed, against all expectations, that a coin tossed with heads facing upwards will land heads approximately 51% of the time. More generally, classical sources of randomness suffer from the fact that Newtonian physics is deterministic, and hence any apparent randomness is epistemic in nature and due to our ignorance of the system or inability to accurately compute the output from our limited knowledge of the initial conditions (as is the case in chaotic systems [Hoy07]).

Quantum mechanics has appeared on the scene, and at least appears to offer an improvement over such RNGs. Consequentially, it has prompted much interest in quantum

[2]Psychologists have known for a long time that people tend to distrust streaks in a series of random bits, hence they imagine a coin flipping sequence alternates between heads and tails much too often for its own sake of 'randomness'. A simple illustration of this phenomenon, called the gambler's fallacy, is the belief that after a coin has landed on tails ten consecutive times there are more chances that the coin will land on heads at the next flip [Isa95].

random number generators—QRNGs. However, much needs to be done to understand the quality of randomness present in such devices. The field is still in its infancy and, perhaps due to the high involvement of experimental physics, has escaped thorough theoretical treatment. As a result, formally unfounded claims have re-appeared for randomness produced with physical experiments suggesting that "truly random numbers have been generated at last" [Haa10, Mer10].

1.3 Quantum randomness

1.3.1 Quantum mechanical formalism

Here we present the foundational quantum mechanical aspects required for this thesis. We make no attempt to present a more thorough overview of the basic quantum mechanical principles, and the reader is referred to [Sak94] for such an overview. Instead, we focus on the abstract Hilbert-space formalism following [Svo98].

A *Hilbert space* in quantum mechanics is a linear vector space over the complex numbers, $\mathbb{C}$, along with the scalar product $\langle\cdot|\cdot\rangle$. We denote the N-dimensional Hilbert space $\mathcal{H}_N$ and we write an element of $\mathcal{H}_N$ as the ket (vector) $|x\rangle$. The scalar product is the natural generalisation of the scalar product for real, linear vector spaces, with the specific condition that $\langle x|y\rangle = \langle y|x\rangle^*$ where '*' indicates the complex conjugate. To summarise, the important features are that for $|x\rangle, |y\rangle \in \mathcal{H}$ we have $\langle x|y\rangle = 0$ if and only if $|x\rangle$ and $|y\rangle$ are orthogonal, and $\langle x|x\rangle = \|x\|$ is the length of $|x\rangle$.

The following important relations to the physical world must be made.

A *physical state* is represented by a unit vector $|x\rangle \in \mathcal{H}$. Fixing a set of basis states $\{|x_1\rangle, \ldots, |x_N\rangle\}$, any vector $|\psi\rangle \in \mathcal{H}_N$ can be written as a *superposition* of basis states and is also a valid physical state. Specifically, we have $|\psi\rangle = \sum_{i=1}^{N} \alpha_i |b_i\rangle$ where $\alpha_i = \langle b_i|\psi\rangle$.

Definition 11. An *observable* is a measurable quantity of a state, and is represented by a hermitian operator A in the Hilbert space. That is, for $|x\rangle, |y\rangle \in \mathcal{H}_n$ we have $\langle x|A|y\rangle = \langle y|A|x\rangle^*$.

The eigenvalues of A are the states $|a_i\rangle$ which satisfy $A|a_i\rangle = a_i|a_i\rangle$.[3] Upon measurement of an observable A, a state $|\psi\rangle = \sum_i \alpha_i |a_i\rangle$ 'collapses'[4] into one of the eigenstates $|a_i\rangle$ and the result of the measurement is the corresponding eigenvalue. Eigenstates thus represent states for which the observable A has the definite value a_i. In general the state

[3] Here we assume the non-degenerate case and label eigenstates by their corresponding eigenvalues for simplicity. This can readily be extended the case that multiple eigenstates have the same eigenvalue.

[4] We make no attempt to explain the nature of this collapse. The key concept is that measurement irreversibly changes the state of the system.

$|\psi\rangle$ is not an eigenstate of A, and quantum mechanics tells us that the result of a measurement is obtained probabilistically: we measure the eigenvalue a_i with probability $|\langle a_i|\psi\rangle|^2 = |\alpha_i|^2$. This is called the *Born rule* and is due to Max Born [Bor26].

Definition 12. Two observables A and B are said to be *compatible* or *co-measurable* if they commute, i.e. $[A, B] = AB - BA = 0$.

Put differently, the values of two compatible observables can be simultaneously well defined for a state. The fact that there are observables which *cannot* be simultaneously well-defined is one of the key points in which quantum physics differs from classical physics. Incompatible observables give rise to the quantum property of *complementarity*, and are responsible for the 'wave-particle duality' of quantum mechanical objects.

1.3.2 Formalising quantum randomness

The probabilistic outcome of measurements in quantum mechanics is the primary motivation for the interest in QRNGs. It is important to realise that this is nonetheless a physical assumption and not a consequence of the foundations of quantum mechanics [Erb95]. It is easy to think that perhaps we are mistaken and that the result is not obtained probabilistically, but rather that we are merely ignorant of the true state the system is in. In other words, one can envisage some *hidden variables* which, although we are unaware of, determine the outcome of individual measurements. This indeed was the view of Einstein, Podolsky and Rosen who argued that quantum mechanics was an incomplete theory [EPR35]. However, there are several results showing that the situation is not nearly that simple, and the class of possible hidden variable theorems (as they are so called) has been significantly narrowed down. The prevalent view is now that this randomness is inherent [Zei05].

The first of these hidden variable theorems is Bell's Theorem, which shows that *local realism* is in contradiction with the statistical predictions of quantum mechanics.

Theorem 13 (Bell [Bel64])**.** *There is no hidden variable theorem which gives the same statistical predictions as quantum mechanics and satisfies the following two conditions:*

> ***Value definiteness*** *All possible observables (including non-compatible ones) simultaneously have predefined values (explicitly or indirectly via a hidden parameter);*
>
> ***Locality*** *Two space-like separated events cannot influence each other in any way.*

Note that value definiteness is sometimes referred to in the literature as counterfactual definiteness, alluding to the fact that even though non-compatible observables cannot

be simultaneously measured, there is a definite value associated with each observable which would be obtained *should the experimenter have chosen to measure it* [Cab94].

The importance of Bell's Theorem is primarily that it manifests itself in statistical inequalities—the class of which we shall call *Bell-type inequalities*—which pose a bound on the correlation between events subject to local realism, but which quantum mechanics predicts violations of. These bounds are *experimentally testable*, and many experiments have been conducted to test them [CS78, GC08]. While many such experiments have minor technical loopholes [Lar98], the results have in general been resoundingly in support of quantum mechanics and are widely seen as one of the most important results in favour of the quantum mechanical description of reality.

Bell's Theorem deals with the statistical predictions of quantum mechanics, but it turns out that no-go theorems can tell us something deeper about the outcome of *individual* quantum measurements. But first, we must introduce the concept of a measurement context.

Definition 14. A measurement *context* is a maximal set of pairwise co-measurable observables.

In a hidden variable theorem, the values associated with observables are simultaneously pre-determined via the hidden variables. For a measurement context $\mathcal{C} = \{A_1, A_2, \dots\}$, the values corresponding the observables are $v(A_1, \mathcal{C}), v(A_2, \mathcal{C}), \dots$ Note that we do not exclude the possibility that the value of an observable may depend on the context it is measured in. Indeed, Bell's theorem subtly tackles this issue by showing that a hidden variable theory must necessarily be non-local in the respect that while the value associated with an observable may be context dependent, information regarding the context measured must be able to propagate superluminally.

The hidden variable theorem which will be of most importance to us is the Kochen-Specker Theorem. This theorem deals directly with the structure of the Hilbert space used to represent quantum systems and, loosely speaking, says that in general there can be no co- or pre-existing definite values prescribable to certain sets of measurement outcomes [CS08, Svo10]. This is in many senses stronger than Bell's Theorem which assumes that such values can be prescribed, but that if this is done then the statistical predictions of quantum mechanics cannot be replicated [Hel08]. Further, no assumption of locality or separability is necessary for the Kochen-Specker Theorem to hold.

Theorem 15 (Kochen-Specker [KS67])**.** *For a quantum mechanical system represented by a Hilbert space of dimension greater than two, it is impossible for a hidden variable theory to satisfy the following two conditions:*

Value definiteness *(as stated in Theorem 13);*

> ***Non-contextuality*** *The value of an observable A, $v(A)$, is independent of the other observables measured alongside it. In other words, for two contexts $\mathcal{C}_1, \mathcal{C}_2$ with $A \in \mathcal{C}_1$ and $A \in \mathcal{C}_2$, we have $v(A) = v(A, \mathcal{C}_1) = v(A, \mathcal{C}_2)$;*

while still fulfilling the quantum mechanical prediction that if a set of pairwise compatible observables $\{A_1, A_2, \dots\}$ satisfy the relation $f(A_1, A_2, \dots) = 0$, then $f(v(A_1), v(A_2), \dots) = 0$ also.

Proofs of the Kochen-Specker Theorem construct sets of 'overlapping' measurement contexts, i.e. contexts which share some, but not all, observables. Non-contextuality says that such 'link' observables should have the same definite value regardless of the context they are measured within. Constructions of the contexts are such that no values can be assigned satisfying these conditions. It is worth noting that by necessity such proofs are counterfactual—they require consideration of observables which could have been measured, but instead a different, complementary observable was measured [Svo10]. Such proofs are generally informative, although rather complex: the original proof [KS67] involved 117 projection operators (observables with eigenvalues of 0 and 1 only), although this has been reduced significantly in subsequent proofs, such as Peres' which uses only 33 projection operators [Per91]. A particularly elegant and instructive proof (although it holds only in four-dimensions and higher) is that of Cabello, Estebaranz and García-Alcaine which uses an (optimal) 18 such projection operators [CEGA96].

The results we present in this thesis on the incomputability rely on the Kochen-Specker Theorem and the assumption of non-contextuality, i.e. we utilise the value indefiniteness which this implies. As such, it is important for us to justify our assumption of non-contextuality, as many of the theoretical results will not hold if we instead assumed value definiteness and accepted contextuality. As discussed in [Hel08], both value indefiniteness and non-contextuality are motivated by an 'innocuous realism', and it is difficult to construct a consistent interpretation which is value definite and contextual. It has been experimentally demonstrated in light of Bell's Theorem that the behaviour of quantum states is nonlocal in the sense that measuring an observable on a local particle can have superluminal effects on the state of a space-like separated entangled particle (i.e., a particle sharing a non-separable state with the local, measured particle) [WJS$^+$98]. However, a contextual but value definite theory would require something stronger—that choosing a different measurement to perform on the local particle would superluminally alter the hidden value corresponding to the (predetermined) observable to be measured on the distant particle. Such a result can be interpreted either as causal violation of Lorentz invariance, or as an ontological feature of the quantum world [HR83]; in either case it seems in contradiction with the realism which motivates us to attempt to hold onto our value definite view of quantum mechanics [Hel08]. As

such, we feel we are, at least, well justified in exploring the consequences that the commonly held belief of non-contextuality and value indefiniteness have on the quality of randomness in quantum mechanics [Svo10, Zei05].

Now that we have the appropriate formalism, we can examine the implications of non-contextuality and the Kochen-Specker Theorem to the results of context preparations and measurements in a QRNG. Consider a quantum state $|\psi\rangle \in \mathcal{H}_n$ and let $\mathcal{C}$ be a measurement context in $\mathcal{H}_n$. Then by the spectral decomposition theorem, $\mathcal{C}$ can be formalised by a singly maximal observable O with n eigenvalues [Svo98]. We call a measurement of the observable O on the state $|\psi\rangle$ non-trivial if all n possible outcomes have non-zero probability, i.e. for every eigenstate $|o_i\rangle$ of O, we have $\langle o_i|\psi\rangle > 0$.

Theorem 16 (Calude-Svozil [CS08], extended in [ACS])**.** *Let $\mathbf{x} = x_1x_2\cdots \in B^\omega$ be the sequence of bits obtained from the concatenation of repeated state preparations and non-trivial measurements in dimension three or greater Hilbert space by discarding all but two possible outcomes. Then, under the assumption of non-contextuality, $\mathbf{x}$ is bi-immune, and hence also incomputable.*

We say that such a sequence $\mathbf{x}$ is *certified by value indefiniteness*, indicating that it results from measurements where the Kochen-Specker Theorem applies. In fact, an even stronger result is true as a result of value indefiniteness. It is impossible to provably compute in advance what any individual bit may be. To formalise this notion more strongly, consider the following example: you have a black box outputting successive bits, and the goal is to predict the next bit. Solovay showed that for the case of Omega numbers—the halting probabilities of universal prefix-free Turing machines—there exist universal machines U such that no bit of the corresponding halting probability Ω_U can be proven to be either 0 or 1 in Zermelo-Frankael set theory with the axiom of choice (ZFC) [Sol00]. If the hypothetical black-box outputs bits of Ω_U, then no bit can be provably predicted in advance. For a sequence $\mathbf{x}$ produced by a QRNG, a similar result holds. Since, before measurement, there is no definite value associated with the observable to be measured we cannot predict what value measurement will result in. If such a value could be proven in advance, then a definite value could be associated to the observable, a contradiction.

Theorem 17. *Let $\mathbf{x} = x_1x_2\cdots \in B^\omega$ be a sequence certified by value indefiniteness as in Theorem 16. Let p_i be the statement "$x_i = 0$" and $\bar{p}_i$ be the statement "$x_i = 1$", both of which are readily formalisable in ZFC. Then if ZFC is consistent,*

$$\{i \geq 1 \mid ZFC \vdash p_i\} = \{i \geq 1 \mid ZFC \vdash \bar{p}_i\} = \emptyset.$$

Proof. Fix an $i \geq 1$, and assume that either $ZFC \vdash p_i$ or $ZFC \vdash \bar{p}_i$. Since we assume ZFC is consistent, ZFC can prove only p_i or $\bar{p}_i$, but not both. It follows that we can

assign in advance a definite value to the outcome of x_i before it is measured: 0 if $ZFC \vdash p_i$, 1 if $ZFC \vdash \bar{p}_i$. However, since the outcome of the measurement producing x_i is certified by value indefiniteness, we know no such value can exist, a contradiction. This holds for all $i \geq 1$, so the outcome of no measurement can be proven in ZFC. □

For any infinite sequence produced by the QRNG, Theorem 17 tells us that it is impossible to compute the value of *any bit* before it is measured. Even in the finite case one experiences in the laboratory this result has something to offer. We can view any finite string produced by a QRNG as the initial segment of an infinite sequence the QRNG would produce if left to run indefinitely. Hence, even in this case there is no way to compute the value of the next bit before it is measured. In light of value indefiniteness this is not unexpected, but nonetheless gives mathematical grounding to the postulated unpredictability of each individual measurement, as well as the independence of successive measurements—indeed, we can rule out any computable causal link within the system which may give rise to the measurement outcome.

1.4 Quantum random number generators

Given the implications of the Kochen-Specker Theorem and value indefiniteness, along with the shortcomings of classical RNGs, it is not surprising that many have been seduced by the prospect of ontic, rather than epistemic, randomness. The incomputable sequences which quantum mechanics seems to allow to be generated has encouraged a flurry of research into QRNGs. There are many quantum phenomena which can be used for random number generation: nuclear decay radiation sources [Sch70], the quantum mechanical noise in electronic circuits known as shot noise [STZ10] or photons travelling through a semi-transparent mirror [Svo90, SGG$^+$00, KCK09, Svo09, PAM$^+$10].

This latter method—photon based QRNGs—appears particularly promising as quantum-optical set-ups are not too difficult to implement and can be controlled with great precision. They are also, compared to the other phenomena, much simpler to describe theoretically, and as a consequence it is much easier to analyse the randomness they should generate.

However, current proposals and realisations have focused on the physical apparatus, but the quality of the randomness has largely been ignored. Analysis techniques have followed those used for pseudo-RNGs, and as a result put emphasis on the uniformity of the distribution; such techniques are designed for computable, rather than incomputable, sources of randomness. Blind faith has been put in the fact that these devices are 'truly indeterministic', but has only been justified with appeal to the inherent (postulated) indeterminism in the Born rule [SGG$^+$00]; the use of value indefiniteness has

not been used to argue for the quality of such devices. Proper analysis needs to take into account the computability aspects of quantum randomness, such as the results of Theorems 16 and 17, to ensure the QRNGs fulfil the potential which quantum mechanics offers. The most non-intuitive aspect of this is that we need to look at infinite sequences to properly discuss the incomputability and randomness of individual bits. It is this formal incomputability which corresponds to the physical notion of indeterminism in quantum mechanics—the inability *even in principle* to predict the outcome of certain quantum measurements—rather than the mathematically vacuous notion of 'true randomness'. Without such analysis and certification one cannot claim that a QRNG products better quality randomness that classical RNGs.

In this thesis we will attempt to address these issues more thoroughly. We combine both the probabilistic notions of randomness, which culminate in the uniformity of the generated distribution, with the mathematical notions of randomness associated with strong incomputability. We not only study the theoretical strength of randomness produced by QRNGs, but how this strength holds under normalisation techniques which are necessarily used to account for bias and correlations within physical devices. Further, we propose a new QRNG based on entangled photons which is explicitly certified by value indefiniteness, and continue to thoroughly analyse the device from a physical and mathematical point of view.

Chapter 2

Von Neumann Normalisation of Quantum Random Strings and Sequences

2.1 Introduction

Due to imperfections in measurement and hardware, the flow of bits generated by a QRNG will contain bias and correlation, two symptoms of non-randomness [Cal02]. Since uniformity of distribution is a necessary requirement for a good QRNG in addition to incomputability, techniques to remove bias such as von Neumann's method will need to be used for real devices. It considers pairs of bits, and takes one of three actions: a) pairs of equal bits are discarded; b) the pair 01 becomes 0; c) the pair 10 becomes 1. Contrary to wide spread claims, the technique works for some sources of bits, but not for all.

The output produced by an independent source of constantly biased bits is transformed (after reducing the number of bits produced significantly) into a flow of bits in which the frequencies of 0's and 1's are equal: 50% for each. As we shall show, a stronger property is true: the un-biasing works not only for bits but for all reasonable long bit-strings. However, if the bias is not constant the procedure does not work. Finally, we emphasise that von Neumann's procedure cannot assure 'true randomness' in its output since such a concept does not exist mathematically.

In this chapter we study the effect of von Neumann normalisation on finite strings and infinite sequences of quantum random bits. We study the effect of normalisation both on the uniformity of the distribution produced, as well as on the symptoms of randomness within the sequences. We focus on von Neumann normalisation because it

is simple, easy to implement, and (along with the more efficient iterated version due to Peres [Per92] for which the results will also apply) is widely used by current proposals for QRNGs [KCK09, MWZ+04, iQ09, SGG+00].

Up until now, QRNGs have been given largely the same mathematical treatment used for pseudo-RNGs, focusing on producing uniformly distributed bits. For real devices this primarily entails the use of randomness extractors—which von Neumann's procedure is one of—to make the source as close as possible to the uniform distribution [Vad11]. Our methods are primarily developed to address photon-based QRNGs, since these are one of the most direct and popular ways to generate quantum random numbers, but many of our mathematical results will be applicable to other QRNGs.

The main results of this chapter are the following.

In the 'ideal case', the von Neumann normalised output of an independent constantly biased QRNG is the probability space of the uniform distribution (un-biasing). This result is true for both the finite strings and infinite sequences produced by QRNGs (the QRNG runs indefinitely in the latter case).

As explained above, QRNGs do not operate in ideal conditions. We develop a model for a real-world QRNG in which the bias, rather than holding steady, drifts slowly (within some bounds). In this framework we evaluate the speed of drift required to be maintained by the source distribution to guarantee that the output distribution is as close as one wishes to the uniform distribution.

We also examine the effect von Neumann normalisation has on various properties of infinite sequences. In particular, Borel normality and (algorithmic) randomness are invariant under normalisation, but for ε-random sequences with $0 < \varepsilon < 1$, normalisation can both decrease or increase the randomness of the source.

2.2 The finite case

2.2.1 Source probability space and independence

In this section we define the QRNG source probability space and the independence property.

Consider a string of n independent bits produced by a (biased) QRNG. Let p_0, p_1 be the probability that a bit is 0 or 1, respectively, with $p_0 + p_1 = 1$, $p_0, p_1 \leq 1$.

The probability space of bit-strings produced by the QRNG is $(B^n, 2^{B^n}, P_n)$ where $P_n : 2^{B^n} \to [0, 1]$ is defined by

$$P_n(X) = \sum_{x \in X} p_0^{\#_0(x)} p_1^{\#_1(x)}, \tag{2.1}$$

for all $X \subseteq B^n$.

We can easily verify that this is indeed a probability space:

Fact 18. *The space* $(B^n, 2^{B^n}, P_n)$ *with* P_n *defined in* (2.1) *is a probability space.*

Proof. We readily verify that the Kolmogorov axioms are satisfied:

1. $P_n(\emptyset) = 0$, trivially true;
2. $P_n(B^n) = 1$, guaranteed by the Binomial Theorem;
3. For $X, Y \subseteq B^n$, $X \cap Y = \emptyset \implies P_n(X \cup Y) = P_n(X) + P_n(Y)$, trivially true. □

The space $(B^n, 2^{B^n}, P_n)$ is just the n-fold product of the single bit probability space $(B, 2^B, P_1)$. For this reason this space is often called an 'independent identically-distributed bit source' [Vad11]. The resulting space is 'independent' because each bit is independent of previous ones. But what is an 'independent probability space'?

Physically, the independence of a QRNG is usually expressed as the impossibility of extracting any information from the flow of bits $x_1, \ldots, x_{k-1}$ to improve chances of predicting the value of x_k, other than what one would have from knowing the probability space. The fact that photon-based QRNGs obey this physical independence between photons (and thus generated bits) rather well [SGG+00] is the primary motivation for our modelling of these devices. These sources (where the condition of independence still holds) are often termed 'independent-bit sources' [Vad11]. In a real device we cannot, of course, expect each bit to be identically distributed, so we study this more general case more thoroughly in Section 2.2.5.

It is important to note that independence in the mathematical sense of multiplicity of probabilities is a model intended to correspond to the physical notion of independence of outcomes [Kac59]. In order to study the theoretical behaviour of QRNGs, which are based on the *assumption of physical independence of measurements*, we must translate this appropriately into our formal model.

Formally, two events $A, B \subseteq B^n$ are independent (in a probability space) if the probability of their intersection coincides with the product of their probabilities [BLM96] (a complexity-theoretic approach was developed in [CZ10]). This motivates the definition of independence of a general source probability space given in Definition 20. But first we need the following simple property:

Fact 19. *For every bit-string* x *and non-negative integers* n, k *such that* $0 \leq k + |x| \leq n$ *we have:*

$$P_n\left(B^k x B^{n-k-|x|}\right) = p_0^{\#_0(x)} p_1^{\#_1(x)} = P_{|x|}(\{x\}). \tag{2.2}$$

Definition 20. The probability space $(B^n, 2^{B^n}, \mathrm{Prob}_n)$ is *independent* if for all $1 \leq k \leq n$ and all $x_1 \dots x_k \in B^k$ the events $x_1 x_2 \dots x_{k-1} B^{n-k+1}$ and $B^{k-1} x_k B^{n-k}$ are independent, i.e.

$$\mathrm{Prob}_n\left(x_1 x_2 \dots x_{k-1} x_k B^{n-k}\right) = \mathrm{Prob}_n\left(x_1 x_2 \dots x_{k-1} B^{n-k+1}\right) \cdot \mathrm{Prob}_n\left(B^{k-1} x_k B^{n-k}\right).$$

Fact 21. *The probability space $(B^n, 2^{B^n}, P_n)$ with P_n defined in (2.1) is independent.*

Proof. Using (2.2) we have:

$$\begin{aligned} P_n\left(x_1 x_2 \dots x_{k-1} x_k B^{n-k}\right) &= p_0^{\#_0(x_1 \dots x_k)} p_1^{\#_1(x_1 \dots x_k)} \\ &= p_0^{\#_0(x_1 \dots x_{k-1})} p_1^{\#_1(x_1 \dots x_{k-1})} p_0^{\#_0(x_k)} p_1^{x_k)} \\ &= P_n\left(x_1 x_2 \dots x_{k-1} B^{n-k+1}\right) \cdot P_n\left(B^{k-1} x_k B^{n-k}\right). \end{aligned}$$ □

As we will see later, there are other relevant independent probability spaces.

2.2.2 Von Neumann normalisation function

Here we present formally the von Neumann normalisation procedure.

We define the mapping $F : B^2 \to B \cup \{\lambda\}$ as

$$F(x_1 x_2) = \begin{cases} \lambda & \text{if } x_1 = x_2, \\ x_1 & \text{if } x_1 \neq x_2, \end{cases}$$

and $f : B \to B^2$ as

$$f(x) = x\bar{x},$$

where $\bar{x} = 1 - x$. Note that for all $x \in B$ we have $F(f(x)) = x$ and, for all $x_1, x_2 \in B$ with $x_1 \neq x_2$, $f(F(x_1 x_2)) = x_1 x_2$.

For $m \leq \lfloor n/2 \rfloor$ we define the normalisation function $VN_{n,m} : B^n \to \left(\bigcup_{k \leq m} B^k\right) \cup \{\lambda\}$ as

$$VN_{n,m}(x_1 \dots x_n) = F(x_1 x_2) F(x_3 x_4) \cdots F\left(x_{(2\lfloor \frac{m}{2} \rfloor - 1)} x_{2\lfloor \frac{m}{2} \rfloor}\right).$$

Fact 22. *For all $1 < m \leq \lfloor n/2 \rfloor$ and $y \in B^m$ there exists an $x \in B^n$ such that $y = VN_{n,m}(x)$.*

Proof. Take $x = f(y_1) f(y_2) \cdots f(y_m) 0^{n-2m}$. □

In fact we can define the 'inverse' normalisation $VN_{n,m}^{-1} : 2^{B^m} \to 2^{B^n}$ as

$$VN_{n,m}^{-1}(Y) = \left\{ u_1 f(y_1) u_2 f(y_2) \cdots u_m f(y_m) u_{m+1} v \mid y = y_1 \dots y_m \in Y, \right.$$
$$\left. u_i \in \{00, 11\}^*, v \in B \cup \{\lambda\}, |v| + 2m + \sum_{i=1}^{m+1} |u_i| = n \right\}.$$

While this isn't a 'true' inverse, for every $y \in B^m$ we have $VN_{n,n}\left(VN_{n,m}^{-1}(y)\right) = \{y\}$.

2.2.3 Target probability space and normalisation

We now construct the target probability space of the normalised bit-strings over B^m for $m \leq \lfloor n/2 \rfloor$, i.e. the probability space of the output bit-strings produced by the application of the von Neumann function on the output bit-strings generated by the QRNG.

The von Neumann normalisation function $VN_{n,m}$ transforms the source probability space $(B^m, 2^{B^m}, P_n)$ into the target probability space $(B^m, 2^{B^m}, P_{n\to m})$. The target space of normalised bit-strings of length $1 < m \leq \lfloor n/2 \rfloor$ associated with the source probability space $(B^m, 2^{B^m}, P_n)$ is the space $(B^m, 2^{B^m}, P_{n\to m})$, where $P_{n\to m} : 2^{B^m} \to [0,1]$ is defined for all $Y \subseteq B^m$ by the formula:

$$P_{n\to m}(Y) = \frac{P_n\left(VN_{n,m}^{-1}(Y)\right)}{P_n\left(VN_{n,m}^{-1}(B^m)\right)}.$$

Proposition 23. *The target space $(B^m, 2^{B^m}, P_{n\to m})$ of normalised bit-strings of length $1 < m \leq \lfloor n/2 \rfloor$ associated to the source probability space $(B^m, 2^{B^m}, P_n)$ is a probability space.*

Proof. We check the Kolmogorov axioms:

1. $P_{n\to m}(\emptyset) = 0$, trivially true;

2. $P_{n\to m}(B^m) = 1$, guaranteed by the normalisation of the probability measure;

3. For $X, Y \subseteq B^m$, $X \cap Y = \emptyset \implies P_{n\to m}(X \cup Y) = P_{n\to m}(X) + P_{n\to m}(Y)$, which is true since $VN_{n,m}^{-1}(X \cup Y) = VN_{n,m}^{-1}(X) \cup VN_{n,m}^{-1}(Y)$ and $P_n\left(VN_{n,m}^{-1}(Y) \cup VN_{n,m}^{-1}(X)\right) = P_n\left(VN_{n,m}^{-1}(Y)\right) + P_n\left(VN_{n,m}^{-1}(X)\right)$, as $VN_{n,m}^{-1}(X) \cap VN_{n,m}^{-1}(Y) = \emptyset$ because X and Y are disjoint. □

2.2.4 Normalisation of the output of a source with constant bias

We now show that von Neumann's procedure transforms the source probability space with constant bias into the probability space with the uniform distribution over B^m, i.e. the target probability space $(B^m, 2^{B^m}, P_{n\to m})$ has $P_{n\to m} = U_m$, the uniform distribution. Independence and the constant bias of P_n play a crucial role.

Theorem 24 (Von Neumann [vN63])**.** *Assume that $1 < m \leq \lfloor n/2 \rfloor$. In the target probability space $(B^m, 2^{B^m}, P_{n\to m})$ associated to the source probability space $(B^m, 2^{B^m}, P_n)$ we have $P_{n\to m}(Y) = U_m(Y) = |Y| \cdot 2^{-m}$, for every $Y \subseteq B^m$.*

Proof. Since $P_{n\to m}$ is additive it suffices to show that for any $y \in B^m$, $P_{n\to m}(\{y\}) = 2^{-m}$. Let $Z = P_n\left(VN_{n,m}^{-1}(B^m)\right)$. We have (the sums are over all $u_i \in \{00, 11\}^*$, $v \in B \cup \{\lambda\}$ such that $|v| + \sum_{i=1}^{m+1} |u_i| = n - 2m$):

$$\begin{aligned} P_{n\to m}(\{y\}) &= \frac{1}{Z} \sum_{u_i,v} p_0^{\#_0(u_1 f(y_1)\dots u_m f(y_m) u_{m+1} v)} p_1^{\#_1(u_1 f(y_1)\dots u_m f(y_m) u_{m+1} v)} \\ &= \frac{p_0^{\#_0(f(y_1)\dots f(y_m))} p_1^{\#_1(f(y_1)\dots f(y_m))}}{Z} \sum_{u_i,v} p_0^{\#_0(u_1\dots u_{m+1} v)} p_1^{\#_1(u_1\dots u_{m+1} v)} \\ &= \frac{p_0^m p_1^m}{Z} \sum_{u_i,v} p_0^{\#_0(u_1\dots u_{m+1} v)} p_1^{\#_1(u_1\dots u_{m+1} v)}, \end{aligned}$$

which is independent of y. Since $P_{n\to m}(B^m) = 1$ and for all $x_1, x_2 \in B^m$ we have $P_{n\to m}(\{x_1\}) = P_{n\to m}(\{x_2\})$ it follows that $P_{n\to m}(\{y\}) = 2^{-m} = U_m(\{y\})$. By additivity, for every $Y \subseteq 2^m$ we thus have $P_{n\to m}(Y) = U_m(Y) = |Y| \cdot 2^{-m}$. □

It is natural to check whether the independence and constant bias of the source probability space are essential for the validity of the von Neumann normalisation procedure.

Example 25. The source probability space $(B^2, 2^{B^2}, \mathrm{Prob}_2)$ where $\mathrm{Prob}_2(00) = 0$, $\mathrm{Prob}_2(01) = \mathrm{Prob}_2(10) = \mathrm{Prob}_2(11) = 1/3$ is independent and $\mathrm{Prob}_{2\to 1} = U_1$.

Example 26. The source probability space $(B^2, 2^{B^2}, \mathrm{Prob}_2)$ where $\mathrm{Prob}_2(00) = \mathrm{Prob}_2(11) = 0$, $\mathrm{Prob}_2(01) = 1/3, \mathrm{Prob}_2(10) = 2/3$ is independent but $\mathrm{Prob}_{2\to 1} \neq U_1$.

Comment. One could present the above examples in the more general framework of Theorem 24.

Theorem 27. *Let $m \geq 1$ and $n = 2m$. Consider the source probability space $(B^n, 2^{B^n}, Prob_n) = \Pi_{i=1}^m (B^2, 2^{B^2}, P_2^i)$, where $P_2^i(01) = P_2^i(10)$, for all $1 \leq i \leq m$. Then, in the target probability space $(B^m, 2^{B^m}, Prob_{n\to m})$, where $Prob_n = \Pi_{i=1}^m P_2^i$, we have $Prob_{n\to m} = U_m$.*

Proof. It is easy to check that for every $y = y_1 \dots y_m \in B^m$ we have

$$\mathrm{Prob}_{n\to m}(\{y_1 \dots y_m\}) = \prod_{i=1}^{m} \frac{P_2^i(y_i \bar{y}_i)}{\mathrm{Prob}_n(VN_{n,m}^{-1}(B^m))},$$

so $\mathrm{Prob}_{n\to m}(\{y_1 \dots y_m\})$ does not depend on y (because $P_2^i(a\bar{a}) = P_2^i(\bar{a}a)$, for every $a \in B$). Hence, $\mathrm{Prob}_{n\to m} = U_m$. □

The source probability space $(B^m, 2^{B^m}, \mathrm{Prob}_n)$ in Theorem 27 is not constantly biased and may be independent or not, but von Neumann normalisation still produces the uniform distribution under these conditions.

Example 28. The source probability space $(B^4, 2^{B^4}, \mathrm{Prob}_4)$ as in Theorem 27 where $P_2^1(00) = P_2^1(01) = 1/3, P_2^1(10) = 1/4, P_2^1(11) = 1/12$ and $P_2^2(00) = 1/12, P_2^2(01) = 1/4, P_2^2(10) = P_2^2(11) = 1/3$ is not independent and $\mathrm{Prob}_{4\to 2} = U_2$.

The outcome of successive context preparations and measurements (which photon-based QRNGs consist of) are postulated to be independent of previous and future outcomes [Jau68]. This means there must be no causal link between one measurement and the next within the system (preparation and measurement devices included) so that the system has no memory of previous or future events. It is this physical understanding which is behind the modelling of QRNGs as independent bit sources.

The above assumption needs to be made clear as in high bit-rate experimental configurations to generate QRNs with photons, its validity may not always be clear. If the wavefunctions of successive photons 'overlap' the assumption no longer holds and (anti)bunching phenomena (the tendency for photon detections to be (anti)correlated due to the Bosonic statistics of the indistinguishable photons [HT56]) may play a role. This is an issue that needs to be more seriously considered in QRNG design and will only become more relevant as the bit-rate of QRNGs is pushed higher and higher. The physical nature of these temporal correlations (and any non-independence they may cause) will be discussed further in Chapter 3, but for now we pose the following open question.

Consider the case in which the probability of each bit depends on no more than a fixed number k of previous bits, and the difference between the conditioned and non-conditioned probability is bounded by a small constant κ.

Open Question 29. *Fix an integer $k \geq 0$ and small positive real κ. Consider the probability space $(B^n, 2^{B^n}, P_n^\dagger)$ where $P_n^\dagger$ is a modification of the probability P_n satisfying the conditions that for all $i \leq n$ and $x_i \in B$ we have $P_n(B^{i-1}x_iB^{n-i}) = P_n^\dagger(B^{i-1}x_iB^{n-i})$,*

$$\left|P_n^\dagger(B^{i-1}x_iB^{n-i}) - P_n^\dagger(B^{i-1}x_iB^{n-i} \mid B^{i-k-1}x_{i-k}\dots x_{i-1}B^{n-i-1})\right| \leq \kappa,$$

and for all $l > k$

$$\begin{aligned} &P_n^\dagger(B^{i-1}x_iB^{n-i} \mid B^{i-l-1}x_{i-l}\dots x_{i-1}B^{n-i-1}) \\ &= P_n^\dagger(B^{i-1}x_iB^{n-i} \mid B^{i-k-1}x_{i-k}\dots x_{i-1}B^{n-i-1}). \end{aligned}$$

If the output of such a source is normalised with the von Neumann procedure, how close is the resulting probability space of strings of length m to the uniform distribution (see Definition 9 for a definition of the closeness of probability spaces)?

2.2.5 Normalisation of the output of a source with non-constant bias

Now we consider the probability distribution obtained if von Neumann normalisation is applied to a string generated from an independent source with a non-constant bias—an 'independent-bit source'. We consider only a bias which varies smoothly; this excludes the effects of sudden noise which could make the bias jump significantly from one bit to the next. Such a source corresponds to a QRNG in which the bias varies slowly (drifts) from bit to bit over time, but never too far from its average point. We choose this to model photon-based QRNGs since the primary cause of variation in the bias will be of this nature. For example, the detector efficiencies may vary as a result of slow changes in temperature or power supply. While abrupt changes—which this model does not account for—are plausible, their relatively rare occurrence (in comparison with the bit generation rate in the order of MHz) will mean they have little effect on the resultant distribution.

Let $p_0, p_1 < 1$ and $p_0 + p_1 = 1$ be constant. Let $x = x_1 x_2 \dots x_n \in B^n$ be the generated string. Then define the probability of an individual bit x_i being either zero or one as

$$q_i^{x_i} = \begin{cases} p_0 - \varepsilon_i & \text{if } x_i = 0, \\ p_1 + \varepsilon_i & \text{if } x_i = 1. \end{cases} \tag{2.3}$$

The variation in the bias is bounded, so we require that for all i,

$$|\varepsilon_i| \leq \beta, \text{ with } \beta < \min(p_0, p_1).$$

Let $\gamma_i = \varepsilon_{i+1} - \varepsilon_i$. Furthermore, we assume that the 'speed' of variation be bounded, i.e. there exists a positive δ such that

$$|\gamma_i| \leq \delta, \tag{2.4}$$

for all i. Evidently we have $\delta \leq \beta$ (presumably in any real situation $\delta \ll \beta$); however, we introduce two separate constants since they correspond to two physically different (but related) concepts. Note that we will discuss in more detail the importance of these two parameters for the approximation of the uniform distribution and their relevance to calibration of the QRNG later once the analysis is completed. Indeed, the rate of change, γ_i, is more important; the need for β stems from the realisation that, even though the probabilities can fluctuate, they can only fluctuate in one direction for so long (since $q_i \in [0, 1]$), hence $|\sum_i \gamma_i| = |\varepsilon_n - \varepsilon_1| \leq 2\beta$.

For a string $y = y_1 y_k \dots y_k \in B^k$ and positive integer i we introduce, for convenience, the following notation:

$$q_i(y) = q_i^{y_1} q_{i+1}^{y_2} \cdots q_{i+k-1}^{y_k}.$$

The difference in probability between 01 and 10 depends only on γ_i, and this allows us to evaluate the effect of normalisation on such a string:

$$\begin{aligned} q_i(01) - q_i(10) &= (p_0 - \varepsilon_i)(p_1 + \varepsilon_{i+1}) - (p_1 + \varepsilon_i)(p_0 - \varepsilon_{i+1}) \\ &= (p_0 + p_1)(\varepsilon_{i+1} - \varepsilon_i) \\ &= \gamma_i. \end{aligned} \tag{2.5}$$

Let us first formally define the probability space generated by this QRNG.

Proposition 30. *The probability space of bit-strings produced by the QRNG is* $(B^n, 2^{B^n}, R_n)$ *where* $R_n : 2^{B^n} \to [0,1]$ *is defined for all* $X \subseteq B^n$ *as follows:*

$$R_n(X) = \sum_{x \in X} q_1(x). \tag{2.6}$$

Proof. We check the Kolmogorov axioms:

1. $R_n(\emptyset) = 0$, trivially true;
2. $R_n(B^n) = 1$, which holds since $q_i^0 + q_i^1 = 1$, and $R_n(B^n) = (q_1^0 + q_1^1) \cdots (q_n^0 + q_n^1)$;
3. For $X, Y \subseteq B^n$, $X \cap Y = \emptyset \implies R_n(X \cup Y) = R_n(X) + R_n(Y)$, trivially true. □

Fact 31. *For all* $i \geq 1$ *and* $x, y \in \{0,1\}^*$ *we have:* $q_i(xy) = q_i(x) q_{i+|x|}(y)$.

Fact 32. *For all* $k, n \geq 1$, $x \in \{0,1\}^*$ *with* $0 \leq k + |x| \leq n$ *we have:*

$$R_n\left(B^{n-k} x B^{n-k-|x|}\right) = q_{n-k+1}(x). \tag{2.7}$$

Proof. Using Fact 31 we get:

$$\begin{aligned} R_n\left(B^{n-k} x B^{n-k-|x|}\right) &= \sum_{y \in B^{n-k}} \sum_{z \in B^{n-k-|x|}} q_1(yxz) \\ &= \sum_{y \in B^{n-k}} \sum_{z \in B^{n-k-|x|}} q_1(y) q_{|y|+1}(x) q_{|y|+|x|+1}(z) \\ &= q_{n-k+1}(x) \sum_{y \in B^{n-k}} \sum_{z \in B^{n-k-|x|}} q_1(y) q_{|y|+|x|+1}(z) \\ &= q_{n-k+1}(x) \sum_{y \in B^{n-k}} q_1(y) \left(\sum_{z \in B^{n-k-|x|}} q_{|y|+|x|+1}(z) \right) \\ &= q_{n-k+1}(x). \end{aligned}$$

□

Fact 33. *The probability space* $(B^n, 2^{B^n}, R_n)$ *with* R_n *defined in* (2.6) *is independent.*

Proof. Using (2.7) we have:

$$\begin{aligned} R_n\left(x_1x_2\dots x_{k-1}x_kB^{n-k}\right) &= q_1(x_1x_2\dots x_{k-1}x_k) \\ &= q_1(x_1x_2\dots x_{k-1})q_k(x_k) \\ &= R_n\left(x_1x_2\dots x_{k-1}B^{n-k+1}\right)\cdot R_n\left(B^{k-1}x_kB^{n-k}\right). \end{aligned}$$ □

As with the constantly biased source, we consider the probability space $R_{n\to m}$. We first investigate the simplest case $n = 2m$. In this situation, for any $y \in B^m$ we have $VN^{-1}_{n,m}(\{y\}) = \{f(y_1)f(y_2)\cdots f(y_m)\}$ and $VN^{-1}_{n,m}(B^m) = \{f(z_1)f(z_2)\cdots f(z_m) \mid z = z_1\dots z_m \in B^m\}$.

Fact 34. *The probability space of normalised bit-strings of length $m = n/2$ is $(B^n, 2^{B^n}, R_{n\to m})$ where $R_{n\to m} : 2^{B^n} \to [0,1]$ is defined for all $Y \subseteq B^m$ as follows:*

$$R_{n\to m}(Y) = \frac{R_n(VN^{-1}_{n,m}(Y))}{R_n(VN^{-1}_{n,m}(B^m))} = \sum_{y\in Y}\prod_{i=1}^{m}\frac{q_{2i-1}(f(y_i))}{q_{2i-1}(01)+q_{2i-1}(10)}. \tag{2.8}$$

Proof. We readily verify that the Kolmogorov axioms are satisfied.

1. $R_{n\to m}(\emptyset) = 0$, trivially true;
2. $R_{n\to m}(B^m) = 1$, trivially true because of the normalisation factors for each bit;
3. For $X, Y \subseteq B^m$, $X \cap Y = \emptyset \implies R_{n\to m}(X \cup Y) = R_{n\to m}(X) + R_{n\to m}(Y)$, trivially true. □

2.2.6 Approximating of the uniform distribution

Unlike the case for a constantly biased source, we no longer have $q_i(01) = q_i(10)$; in fact by (2.5) we have $q_i(01) = q_i(10) + \gamma_i$. As a result the normalised equation is no longer the uniform distribution, but only an approximation thereof. We now explore how closely $R_{n\to m}$ approximates U_m.

The variation $\Delta(R_{n\to m}, U_m)$ depends on each γ_i and q_i (thus on p_0, p_1 and each ε_i), but we wish to calculate the worst case in terms of the bounds δ, β and p_0, p_1, i.e. using Lemma 10,

$$\max_{\gamma_i,q_i}\Delta(R_{n\to m}, U_m) = \frac{1}{2}\max_{\gamma_i,q_i}\sum_{y\in B^m}|R_{n\to m}(\{y\}) - 2^{-m}|.$$

Let us first note that we can write

$$\begin{aligned} \frac{q_{2i-1}(f(y_i))}{q_{2i-1}(01)+q_{2i-1}(10)} &= \frac{q_{2i-1}(f(y_i))}{2q_{2i-1}(f(y_i)) - (-1)^{y_i}\gamma_{2i-1}} \\ &= \frac{1}{2}\left(1 + \frac{(-1)^{y_i}\gamma_{2i-1}}{2q_{2i-1}(f(y_i)) - (-1)^{y_i}\gamma_{2i-1}}\right), \end{aligned}$$

and hence we have

$$R_{n\to m}(\{y\}) = 2^{-m}\prod_{i=1}^{m}\left(1+\frac{(-1)^{y_i}\gamma_{2i-1}}{q_{2i-1}(01)+q_{2i-1}(10)}\right).$$

We have rewritten the denominator in its original form to emphasise that only the signs $(-1)^{y_i}$ depend on y. Thus, we want to find the values of q_{2i-1} and γ_{2i-1} which maximise

$$\sum_{y\in B^m}\left|1-\prod_{i=1}^{m}\left(1+\frac{(-1)^{y_i}\gamma_{2i-1}}{q_{2i-1}(01)+q_{2i-1}(10)}\right)\right|, \tag{2.9}$$

subject to the constraints that $|\gamma_\ell| \le \delta$ and $|\varepsilon_\ell| \le \beta$ for $1 \le \ell \le n$.

Lemma 35. *The function*

$$g(c_1,\ldots,c_n) = \sum_{y\in B^n}\left|\prod_{i=1}^{n}(1+(-1)^{y_i}c_i)-1\right|$$

is strictly increasing for $0 \le c_i < 1$, $i = 1,\ldots,n$ *(note that for* $1 \le i \le n$, $g(c_1,\ldots,c_i,\ldots,c_n) = g(c_1,\ldots,-c_i,\ldots,c_n)$*).*

Proof. We take $0 \le c_i < 1$ for $1 \le i \le n$. For $y = y_1\ldots y_n \in B^n$ define $p(y,j) = \prod_{i=1,i\neq j}^{n}(1+(-1)^{y_i}c_i)$. Without loss of generality, choose a $j \le n$ and let $\varepsilon > 0$ be an (arbitrarily small) positive real with $c_j + \varepsilon \le 1$. Note that

$$g(c_1,\ldots,c_n) = \sum_{y\in B^n}\left|(1+(-1)^{y_j}c_j)p(y,j)-1\right|.$$

We partition B^n as follows:

$$\begin{aligned}
Y_1 &= \{y \mid (1-c_j-\varepsilon)p(y,j)-1 \ge 0\},\\
Y_2 &= \{y \mid (1-c_j-\varepsilon)p(y,j)-1 < 0 \text{ and } (1-c_j)p(y,j)-1 \ge 0\},\\
Y_3 &= \{y \mid (1-c_j)p(y,j)-1 < 0 \text{ and } (1+c_j)p(y,j)-1 \ge 0\},\\
Y_4 &= \{y \mid (1+c_j)p(y,j)-1 < 0 \text{ and } (1+c_j+\varepsilon)p(y,j)-1 \ge 0\},\\
Y_5 &= \{y \mid (1+c_j+\varepsilon)p(y,j)-1 < 0\}.
\end{aligned}$$

Note that for $y \in B^n$, $p(y,j) \ge 0$, and for $y_i \in Y_i$, $i = 1,\ldots,5$, we have

$$p(y_5,j) < p(y_4,j) < p(y_3,j) < p(y_2,j) < p(y_1,j),$$

and $\bigcup_{i=1}^{5} Y_i = B^n$. We have:

$$\begin{aligned}
g(c_1,\ldots,c_j+\varepsilon,\ldots,c_n) =& \sum_{i=1}^{5}\sum_{y\in Y_i} |(1+(-1)^{y_j}c_j+(-1)^{y_j}\varepsilon)p(y,j)-1| \\
=& \sum_{y\in Y_1} [(1+(-1)^{y_j}c_j)p(y,j)-1+(-1)^{y_j}\varepsilon p(y,j)] \\
&+\sum_{i=2}^{4}\sum_{y\in Y_i}(-1)^{y_j}\left[(1+(-1)^{y_j}c_j)p(y,j)-1+(-1)^{y_j}\varepsilon p(y,j)\right] \\
&+\sum_{y\in Y_5} -\left[(1+(-1)^{y_j}c_j)p(y,j)-1+(-1)^{y_j}\varepsilon p(y,j)\right] \\
=& \sum_{i=1}^{5}\sum_{y\in Y_i} |(1+(-1)^{y_j}c_j)p(y,j)-1| + 2\varepsilon\sum_{i=2}^{4}\sum_{y\in Y_i} p(y,j) \\
&-2\sum_{y\in Y_2}[(1-c_j)p(y,j)-1]+2\sum_{y\in Y_4}[(1+c_j)p(y,j)-1] \\
=& g(c_1,\ldots,c_j,\ldots,c_n) - 2\sum_{y\in Y_2}[(1-c_j-\varepsilon)p(y,j)-1] \\
&+2\sum_{y\in Y_4}[(1+c_j+\varepsilon)p(y,j)-1]+2\varepsilon\sum_{y\in Y_3} p(y,j) \\
>& g(c_1,\ldots,c_j,\ldots,c_n),
\end{aligned}$$

where the final line follows from the definition of Y_2 and Y_4. Since this holds for all $j \le n$, g is strictly increasing over $[0,1)^n$. □

Hence, in order to maximise (2.9) we need to maximise the functions

$$u_j(\varepsilon_j,\gamma_j) = \left|\frac{\gamma_j}{q_j(01)+q_j(10)}\right| \tag{2.10}$$

$$= \left|\frac{\gamma_j}{(p_0-\varepsilon_j)(p_1+\varepsilon_j+\gamma_j)+(p_1+\varepsilon_j)(p_0-\varepsilon_j-\gamma_j)}\right|, \tag{2.11}$$

for $j = 2i-1$, $1 \le i \le m$, subject to the constraints $|\gamma_j| \le \delta$, $|\varepsilon_j| \le \beta$ and $|\varepsilon_{j+1}| = |\varepsilon_j + \gamma_j| \le \beta$.

Lemma 36. *For every $j \ge 1$ we have*

$$u_j(\varepsilon_j,\gamma_j) \le \begin{cases} u_j(\beta,-\delta) = u_j(\beta-\delta,\delta) & \text{if } p_1 \ge p_0, \\ u_j(-\beta,\delta) = u_j(-\beta+\delta,-\delta) & \text{if } p_0 > p_1, \end{cases} \tag{2.12}$$

$$= \frac{\delta}{2\left[p_0p_1 - \beta(\beta-\delta) - |p_0-p_1|(\beta-\delta/2)\right]}. \tag{2.13}$$

Proof. We omit the index j as it is not needed in this context. Let

$$v(\varepsilon,\gamma) = \frac{\gamma}{(p_0-\varepsilon)(p_1+\varepsilon+\gamma)+(p_1+\varepsilon)(p_0-\varepsilon-\gamma)}.$$

Since $q(01) + q(10) > 0$, in order to maximise u we look for maxima and minima of v; clearly maxima have $\gamma > 0$ and minima have $\gamma < 0$. We use Lagrange multipliers with inequality constraints to find the critical points. We have the following six constraints: $h_1(\varepsilon, \gamma) = \varepsilon - \beta \leq 0$, $h_2(\varepsilon, \gamma) = -\varepsilon - \beta \leq 0$, $h_3(\varepsilon, \gamma) = \varepsilon + \gamma - \beta \leq 0$, $h_4(\varepsilon, \gamma) = -\varepsilon - \gamma - \beta \leq 0$, $h_5(\varepsilon, \gamma) = \gamma - \delta \leq 0$, $h_6(\varepsilon, \gamma) = -\gamma - \delta \leq 0$. We must solve the following equations:

$$\nabla_{\varepsilon,\gamma} v(\varepsilon, \gamma) + \sum_{i=1}^{6} \lambda_i \nabla_{\varepsilon,\gamma} h_i(\varepsilon, \gamma) = 0, \tag{2.14}$$

$$\lambda_i h_i(\varepsilon, \gamma) = 0 \qquad \text{for } i = 1, \ldots, 6, \tag{2.15}$$

$$h_i(\varepsilon, \gamma) \leq 0 \qquad \text{for } i = 1, \ldots, 6, \tag{2.16}$$

$$\begin{cases} \lambda_i \geq 0 & \text{for minima, } i = 1, \ldots, 6, \\ \lambda_i \leq 0 & \text{for maxima, } i = 1, \ldots, 6. \end{cases} \tag{2.17}$$

We say a constraint is inactive if $\lambda_i = 0$ and active otherwise; the condition of complimentarity (2.15) captures the notion that a critical point satisfying the constraints either occurs at $h_i(\varepsilon, \gamma) = 0$ or is also a critical point in the unconstrained problem.

Noting that $0 < p_0 - \beta \leq p_0 + \beta < 1$ and solving, we find the candidate points are:

$$(\varepsilon, \gamma) = \begin{cases} (\frac{1}{2}(p_0 - p_1) \pm \frac{\delta}{2}, \mp\delta) & \\ (\beta, -\delta), (\beta - \delta, \delta) & \text{for } p_0 - p_1 \leq 2\beta - \delta, \\ (-\beta, \delta), (-\beta + \delta, -\delta) & \text{for } p_1 - p_0 \leq 2\beta - \delta. \end{cases}$$

Note that $u(\varepsilon, \gamma) = u(\varepsilon + \gamma, -\gamma)$. Testing values shows the second case maximises $u(\varepsilon, \gamma)$ when $p_1 > p_0$ and the third cases maximises $u(\varepsilon, \gamma)$ for $p_0 > p_1$. For $p_0 = p_1$ both cases give the same value. Substituting in ε, γ and consolidating the cases we arrive at (2.13). □

Next we let

$$\alpha = \max_{\gamma_i, \varepsilon_i} u_j(\varepsilon_j, \gamma_j),$$

where $u_j(\varepsilon_j, \gamma_j)$ comes from (2.10).

Then we have

$$\begin{aligned} \max_{\gamma_i, \varepsilon_i} \Delta(R_{n\to m}, U_m) &= \frac{1}{2} \sum_{y \in B^m} \left| \prod_{i=1}^{m} \left(\frac{1}{2} + (-1)^{y_i} \frac{\alpha}{2} \right) - 2^{-m} \right| \\ &= \frac{1}{2} \sum_{k=0}^{m} \binom{m}{k} \left| \left(\frac{1}{2} + \frac{\alpha}{2} \right)^k \left(\frac{1}{2} - \frac{\alpha}{2} \right)^{m-k} - 2^{-m} \right|. \end{aligned}$$

Note that in this worst case, the normalised source acts as an independent identically-distributed source with $p_0 = 1/2 \pm \alpha/2$ and the total variation is bounded by that of

two binomial sources: one with $p_0 = 1/2$, the other with $p_0 = 1/2 \pm \alpha/2$ (the number k of successful outcomes is identified with the number of ones in y).

There are two interesting questions: a) what is the quality of the distribution produced by a QRNG, i.e. how close are $R_{n\to m}$ and U_m in terms of α? and b) given a real $\rho \in (0,1)$, how accurate does the QRNG need to be in terms of α to guarantee that $R_{n\to m}$ and U_m are ρ close?

We can take a rough approach to solve the above problems as follows. First note that

$$\begin{aligned}\Delta(R_{n\to m}, U_m) &\leq \frac{1}{2}\sum_{y\in B^m}\left|\prod_{i=1}^{m}\left(\frac{1}{2}+(-1)^{y_i}\frac{\alpha}{2}\right)-2^{-m}\right| \\ &\leq \frac{1}{2}\sum_{y\in B^m}\frac{1}{2^m}\left((1+\alpha)^m-1\right) \\ &= \frac{1}{2}\left((1+\alpha)^m-1\right).\end{aligned}$$

So given α, $R_{n\to m}$ and U_m are at most $\frac{1}{2}\left((1+\alpha)^m-1\right)$-close. Conversely, $R_{n\to m}$ and U_m are ρ close if

$$\alpha \leq (1+2\rho)^{1/m} - 1. \tag{2.18}$$

We will express further results in the latter form, focusing on question b), although both are important questions depending on the operational circumstances and results can easily be transformed from one form to the other.

So, by making α very small, $R_{n\to m}$ can be made as close as we wish to the uniform distribution. This is intuitive since $\alpha \to 0$ only as $\delta \to 0$ and we approach the constantly biased source situation.

There are, unfortunately, some issues with this bound. First, as $m \to \infty$ the bound on the variation becomes infinite too. This is unreasonable as by definition we should have $\Delta(R_{n\to m}, U_m) \leq 1$. It only makes sense to talk about $\rho \leq 1$, although in any useful situation we will require ρ to be small (close to 0) so it is only of real importance that the bound is good in this situation. However, (2.18) requires α to be significantly smaller than we really require for the two probabilities to be ρ close. Even for small ρ the bound is no-way near tight enough (see Figure 2.2). Further, it would be instructive to examine more correctly the behaviour for large m and investigate fully the nature of the relationship between α, m and ρ.

To rectify this and find a more reasonable bound, we carry out a finer analysis making use of the previous observation that this is the same problem as finding the variation between two binomial distributions. Let us denote a binomial probability distribution function for n trials and probability of success p as $S_{n,p} : \{0,\ldots,n\} \to [0,1]$ where for

each $A \subseteq \{0, \dots, n\}$,

$$S_{n,p}(A) = \sum_{k \in A} \binom{n}{k} p^k (1-p)^{n-k}.$$

For $0 \leq p, p' \leq 1$, we then have

$$\Delta(S_{n,p}, S_{n,p'}) = \frac{1}{2} \sum_{k=0}^{n} \binom{n}{k} \left| p^k (1-p)^{n-k} - (p')^k (1-p')^{n-k} \right|,$$

and

$$\max_{\gamma_i, \varepsilon_i} \Delta(R_{n \to m}, U_m) = \Delta(S_{m,1/2(1\pm\alpha)}, S_{m,1/2}).$$

Fact 37. *For $0 \leq p, p' \leq 1$ we have $\Delta(S_{n,p}, S_{n,p'}) = \Delta(S_{n,1-p}, S_{n,1-p'})$.*

The total variation between two binomial distributions can be given in terms of regularised incomplete beta functions [AJ06].

Definition 38. The *incomplete beta function* is defined as

$$\mathrm{B}_\ell(a, b) = \int_0^\ell u^{a-1} (1-u)^{b-1} du.$$

For $\ell = 1$ we write $\mathrm{B}_1(a, b) = \mathrm{B}(a, b)$ for the *complete* beta function, or just *beta function*. The *regularised incomplete beta function* is defined as

$$I_\ell(a, b) = \frac{\mathrm{B}_\ell(a, b)}{\mathrm{B}(a, b)}.$$

Theorem 39. *Let $0 \leq p \leq 1$, $q = 1 - p$ and $0 \leq x \leq q$. The total variation between two binomial distributions with probability of success p and $p + x$ is*

$$\begin{aligned} \Delta(S_{n,p}, S_{n,p+x}) &= n \int_p^{p+x} S_{n-1,u}(\ell - 1) du \\ &= n \binom{n-1}{\ell-1} \int_p^{p+x} u^{\ell-1} (1-u)^{n-\ell} du \\ &= I_{p+x}(\ell, n - \ell + 1) - I_p(\ell, n - \ell + 1), \end{aligned}$$

where

$$\lceil np \rceil \leq \ell := \ell(n, p, x) = \left\lceil \frac{-n \log(1 - x/q)}{\log(1 + x/p) - \log(1 - x/q)} \right\rceil \leq \lceil n(p + x) \rceil.$$

Proof. The first line is from Adell and Jodrá [AJ06]. The rest follows from the well known properties of the beta functions: $\mathrm{B}_\ell(a, b) = \mathrm{B}_\ell(b, a)$ and

$$\binom{n}{k} = \frac{1}{(n+1)\mathrm{B}(n-k+1, k+1)}.$$

□

Theorem 40. *The total variation is bounded by*

$$\begin{aligned}\Delta(R_{n\to m}, U_m) &\leq \Delta(S_{m,1/2}, S_{m,1/2(1+\alpha)}) \\ &= I_{1/2(1+\alpha)}(\ell, m-\ell+1) - I_{1/2}(\ell, m-\ell+1), \\ &= F(m-\ell; m, 1/2-\alpha/2) - F(m-\ell; m, 1/2)\end{aligned}$$

where

$$\lceil m/2 \rceil \leq \ell = \ell(m, 1/2, \alpha/2) = \left\lceil \frac{-m\log(1-\alpha)}{\log(1+\alpha) - \log(1-\alpha)} \right\rceil \leq \lceil m(1+\alpha)/2 \rceil ,$$

and

$$F(k; n, p) = \sum_{x=0}^{k} S_{n,p}(x)$$

is the cumulative distribution function for the binomial distribution.

Proof. This follows directly from Theorem 39 and Fact 37. The last line follows from well known properties of the binomial distribution. □

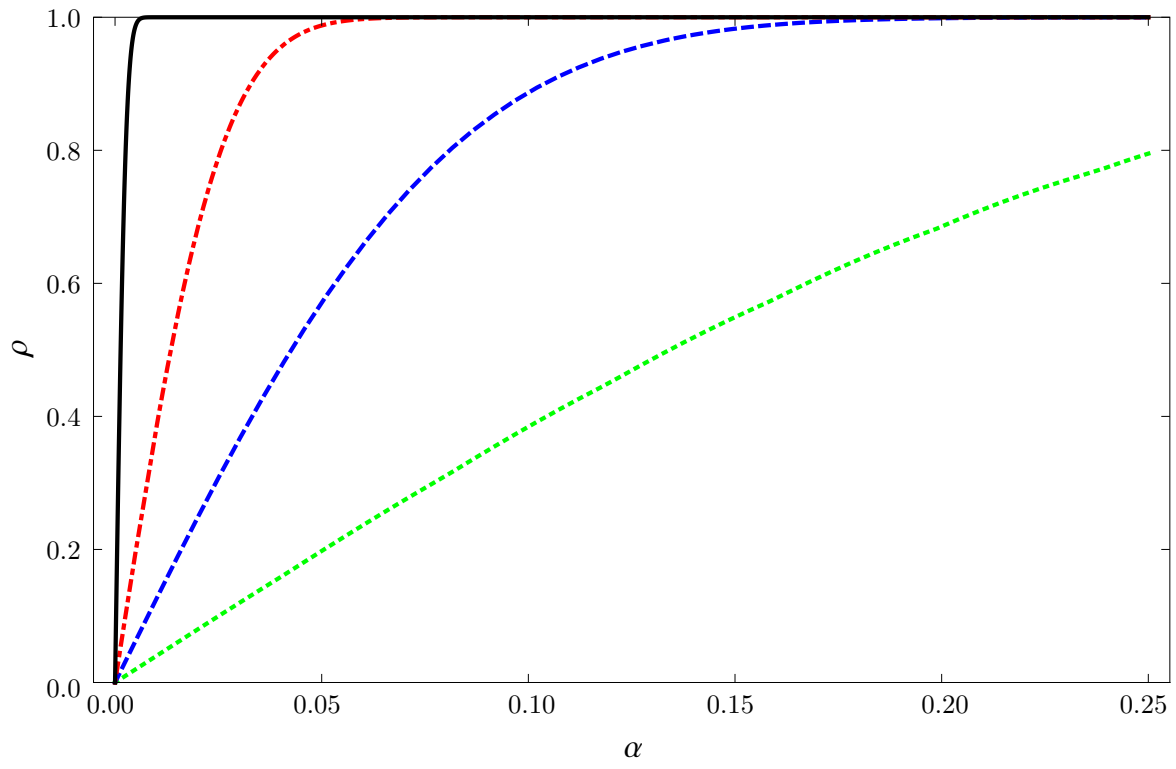

Figure 2.1: (Colour online) Plot of ρ against α using the bound in Theorem 40 for four values of m: 100 (green, dotted), 1,000 (blue, dashed), 10,000 (red, dot-dashed) and 1,000,000 (black, solid).

This bound is exact (under the extrema given by Lemma 36), and we easily verify that $\Delta(R_{n\to m}, U_m) \leq 1$ since $I_p(a, b) \leq 1$ for all a, b and $p \leq 1$, and for $p' \geq p$ we have

$I_{p'}(a,b) \geq I_p(a,b)$ (with equality only for $p = p'$). Unfortunately this bound on the variation has no simple closed form, so we can not easily relate α, m and ρ like we did in (2.18). The shape and nature of this relationship can be seen for various values of m in Figure 2.1. In practice, with m fixed and given ρ it is easy to compute (via numerical methods) α such that $\Delta(R_{n\to m}, U_m) \leq \rho$. For relatively small ρ however, we can find a simple and fairly good bound which is easy to work with for rough approximations.

Theorem 41. *Assume that $m = n/2$. Consider the probability spaces $(B^m, 2^{B^m}, R_{n\to m})$ and $(B^m, 2^{B^m}, U_m)$. For every real ρ such that $0 \leq \rho < 1$, if*

$$\alpha \leq \rho\sqrt{\frac{2\pi(1-\frac{2}{m})}{m+1}}, \tag{2.19}$$

then $\Delta(R_{n\to m}, U_m) \leq \rho$.

Proof. We will take a first order (linear) approximation of $\Delta(S_{m,1/2}, S_{m,1/2(1+\alpha)})$ around $\alpha = 0$. From Theorem 39 and the Fundamental Theorem of Calculus we have

$$\Phi(\alpha) := \frac{d}{d\alpha}\Delta(S_{m,1/2}, S_{m,1/2(1+\alpha)}) = m\binom{m-1}{\ell-1}2^{-m}(1+\alpha)^{\ell-1}(1-\alpha)^{m-\ell}.$$

Since $\ell \geq \lceil m/2 \rceil$ we have

$$\Phi(\alpha) \leq \Phi(0),$$

so our first order upper bound is given by

$$\Delta(S_{m,1/2}, S_{m,1/2(1+\alpha)}) \leq \alpha\Phi(0) = \alpha m\binom{m-1}{\ell-1}2^{-m}.$$

Since the central binomial coefficient (i.e. $\binom{n}{\lfloor n/2 \rfloor}$) is the largest, for $k \leq m-1$ we have

$$\binom{m-1}{k} \leq \binom{m-1}{\lfloor \frac{m-1}{2} \rfloor} = \binom{m-1}{\lceil \frac{m}{2} \rceil - 1},$$

which can easily be shown by taking the two cases of m odd and m even. Since $\ell \geq \lceil m/2 \rceil$ we have that

$$\Phi(0) \leq 2^{-m}m\binom{m-1}{\lceil \frac{m}{2} \rceil - 1} = 2^{-m}m\frac{\lceil \frac{m}{2} \rceil}{m}\binom{m}{\lceil \frac{m}{2} \rceil} = 2^{-m}\lceil m/2 \rceil \binom{m}{\lceil \frac{m}{2} \rceil}.$$

Using the bounds given in [Sta01, Corollary 2.3], and writing $m = a\lceil m/2 \rceil$ where $a \leq 2$, we have

$$\begin{aligned}
\binom{a\left\lceil \frac{m}{2} \right\rceil}{\left\lceil \frac{m}{2} \right\rceil} &< \frac{1}{\sqrt{2\pi \left\lceil \frac{m}{2} \right\rceil}} \frac{a^{m+\frac{1}{2}}}{(a-1)^{(a-1)\left\lceil \frac{m}{2} \right\rceil + \frac{1}{2}}} \\
&= \frac{1}{\sqrt{2\pi \left\lceil \frac{m}{2} \right\rceil}} \frac{m^{m+\frac{1}{2}}}{\left\lfloor \frac{m}{2} \right\rfloor^{\left\lfloor \frac{m}{2} \right\rfloor + \frac{1}{2}} \left\lceil \frac{m}{2} \right\rceil^{\left\lceil \frac{m}{2} \right\rceil}} \\
&\leq \frac{1}{\sqrt{2\pi \left\lceil \frac{m}{2} \right\rceil}} \frac{m^{m+\frac{1}{2}}}{\left((\frac{m}{2}+\frac{1}{2})(\frac{m}{2}-\frac{1}{2})\right)^{\left\lfloor \frac{m}{2} \right\rfloor} (\frac{m}{2}-\frac{1}{2})^{\frac{1}{2}} (\frac{m}{2})} \\
&\leq \frac{1}{\sqrt{2\pi \left\lceil \frac{m}{2} \right\rceil}} \frac{2^{m+\frac{1}{2}}}{\left(1-\frac{1}{m^2}\right)^{\left\lfloor \frac{m}{2} \right\rfloor} (1-\frac{1}{m})^{\frac{1}{2}}} \\
&\leq \frac{1}{\sqrt{\pi \left\lceil \frac{m}{2} \right\rceil}} \frac{2^{m}}{\left(1-\frac{1}{2m}\right) (1-\frac{1}{m})^{\frac{1}{2}}} \\
&\leq \frac{2^m}{\sqrt{\pi \left\lceil \frac{m}{2} \right\rceil (1-\frac{2}{m})}}.
\end{aligned}$$

Hence, we have

$$\Phi(0) \leq \sqrt{\frac{\left\lceil \frac{m}{2} \right\rceil}{\pi(1-\frac{2}{m})}} \leq \sqrt{\frac{m+1}{2\pi(1-\frac{2}{m})}}. \qquad \square$$

This bound is much much better than the bound given in (2.18), and for small α is extremely good. It has the desired properties that as $\alpha \to 0$, the bound on the variation tends to 0 also. Obviously this bound is not less than one for all α, but for small ρ the bound is very good, as can be seen in Figure 2.2.

Another interesting question refers to the possibility of manipulating the parameter α for fine calibration of the QRNG. For $R_{n\to m}$ to become closer to U_m we need to make α smaller, but this can be done by adjusting both δ and β. As previously discussed, both are reasonable physical parameters, and which one is the most suitable (or easiest) to decrease experimentally will to a large extent depend on the QRNG set-up itself. However, adjusting δ has a larger effect on α than adjusting β does, and $R_{n\to m}$ will only approach U_m arbitrarily close as $\delta \to 0$, as even with $\beta = \delta$ (recall $\delta \leq \beta$) we do not have $\alpha = 0$ unless $\delta = 0$.

These results can be extended to all $m \leq n/2$, although the analysis is rather elaborated. The key difference is that in the definition of $R_{n\to m}$ in (2.8) the set $VN_{n,m}^{-1}(Y)$ no longer has the same size as Y, so an additional summation is needed in the right hand side of (2.8). However, the total variation will still be maximised under the same conditions as in Lemmata 35 and 36, and the same relation as in Theorem 40 holds.

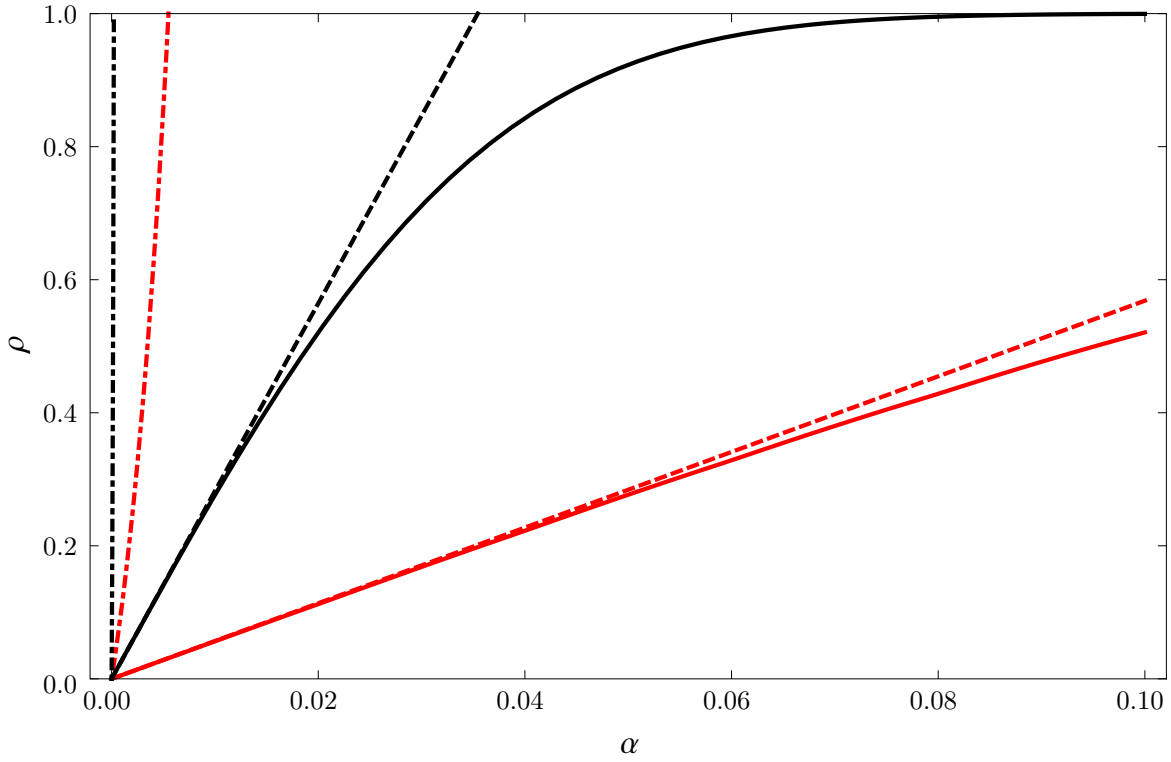

Figure 2.2: Plot of upper bounds on the variation between $R_{n\to m}$ and U_m for $m = 200$ (red curves) and $m = 5000$ (black curves). The three bounds shown are the exact bound $\Delta(S_{m,1/2}, S_{m,1/2(1+\alpha)})$ (solid curves), the improved bound (2.19) (dashed curves) and the naïve bound (2.18) (dot-dashed curves).

It is worth noting that the conditions which maximised the variation in (2.12) correspond to every ε_i being the same up to a small variation δ. Physically this would indicate that p_0, p_1 have been incorrectly stated, but that the device is actually rather accurate except for a small drift in probabilities of no more than δ. Since the parameters ε_i are supposed to physically account for the amount the probability is allowed to drift, which will normally be much more than the drift between individual bits (the γ_i), if the device is calibrated so that p_0 and p_1 are centred so that the ε_i are distributed around them, then the variation will not be nearly as bad as in this worst case. However, the bound on the variation remains valid as it is not necessarily meaningful (or useful) to look into the physical situation under which the worst case bound is achieved.

We briefly wish to point out that other methods for dealing with independent-bit sources have been proposed. For example, grouping bits into blocks of size ℓ and taking the parity of these bits for the 'normalised' bit produces a string of length n/ℓ [Vad11]. With this method each bit becomes unbiased exponentially fast in ℓ. However, the bound in Theorem 41 is asymptotically tighter than the corresponding bound that can be obtained by the parity method if the block size ℓ is fixed; if ℓ scales polynomially with

n then this method produces a better bound, but at a substantial cost to the number of bits produced [Vad11, Proposition 6.5]. The reason the von Neumann normalisation outperforms the parity method is due to the fact that the bias is required to vary slowly.

2.3 The infinite case

The extension of the above results to infinite sequences of bits produced by QRNGs is fairly straightforward, but forces us to address a few unexpected problems. It is an important step in determining the effect of von Neumann normalisation on the incomputability of sequences of quantum random bits. First, we must extend the definition of the normalisation function $VN_{n,m}$ to sequences. We define $VN : B^\omega \to B^\omega \cup B^*$ as

$$VN(\mathbf{x} = x_1 \dots x_n \dots) = F(x_1x_2)F(x_3x_4)\cdots F(x_{2\lfloor \frac{n}{2} \rfloor - 1}x_{2\lfloor \frac{n}{2} \rfloor})\cdots .$$

For convenience we also define $VN_n : B^\omega \to \left(\bigcup_{k\leq n} B^k\right) \cup \{\lambda\}$ as

$$VN_n(\mathbf{x}) = F(x_1x_2)F(x_3x_4)\cdots F(x_{2\lfloor \frac{n}{2} \rfloor - 1}x_{2\lfloor \frac{n}{2} \rfloor}) = VN_{n,n}(x_1 \dots x_n).$$

Secondly, we introduce the probability space of infinite sequences as in [Cal02]. Let $A_Q = \{a_1, \dots, a_Q\}$, $Q \geq 2$ be an alphabet with Q elements. We let $\mathcal{P} = \{xA_Q^\omega \mid x \in A_Q^*\} \cup \{\emptyset\}$ and $\mathcal{C}$ be the class of all finite mutually disjoint unions of sets in $\mathcal{P}$; the class $\mathcal{P}$ can be readily shown to generate a σ-algebra $\mathcal{M}$. Using Theorem 1.7 from [Cal02], the probabilities on $\mathcal{M}$ are characterised by the functions $h : A_Q^* \to [0, 1]$ satisfying:

1. $h(\lambda) = 1$,
2. $h(x) = h(x_{a_1}) + \cdots + h(x_{a_Q})$, for all $x \in A_Q^*$.

If $Q = 2$ so that $A_2 = B$, and for $x \in B^n$ we take $h(x) = P_n(\{x\})$ with P_n as defined in Fact 18, then the above conditions are satisfied. This induces our probability measure μ_P on $\mathcal{M}$, which satisfies $\mu_P(XB^\omega) = P_n(X)$ for $X \subseteq B^n$. Hence the suitable extension of the finite case probability space to infinite generated sequences is the space $(B^\omega, \mathcal{M}, \mu_P)$. In the special case when $p_0 = p_1$ we get the Lebesgue probability $\mu_{P_L}(XB^\omega) = \sum_{x\in X} 2^{-|x|}$.

In general, if $Q \geq 2$, $p_i \geq 0$ for $i = 1, \dots, Q$ are reals in [0,1] such that $\sum_{i=1}^{Q} p_i = 1$, we can take $h_Q(x) = p_1^{\#_{a_1}(x)} \dots p_Q^{\#_{a_Q}(x)}$ to obtain the probability space $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$ in which $\mu_{P_Q}(xA_Q^\omega) = h_Q(x)$, for all $x \in A_Q^*$.

The first result notes that there exist sequences $\mathbf{x} \in B^\omega$ such that $VN(\mathbf{x}) \in B^*$. In fact every string can be produced via von Neumann normalisation from a suitable sequence.

Theorem 42. *For every string $y \in B^*$ there exists an uncountable set $R \subset B^\omega$ of μ_P measure zero such that for all $\mathbf{x} \in R$, $VN(\mathbf{x}) = y$.*

Proof. Let $y = y_1 \dots y_n \in B^*$ and $D = \{00, 11\}$, the two-bit blocks which are deleted by von Neumann normalisation and $y' = f(y_1) \dots f(y_n)$. Then every sequence $\mathbf{x} \in y'D^\omega$ satisfies $VN(\mathbf{x}) = VN_{2n}(\mathbf{x})VN(x_{2n+1}x_{2n+2} \dots) = y$ since $VN_{2n}(\mathbf{x}) = VN_{2n,2n}(y') = y$ and for all $\mathbf{z} \in D^\omega$ we have $VN(\mathbf{z}) = \lambda$. Obviously, the set $R = y'D^\omega$ is uncountable and has μ_P measure zero as the set of Borel normal sequences has measure one [Cal02]. □

Corollary 43. *The set $Q = \{\mathbf{x} \in B^\omega \mid VN(x) \in B^*\}$ has μ_P measure zero.*

Proof. We simply note that the union of countably many measure zero sets also has measure zero. □

It is interesting to note that the 'collapse' in the generated sequence produced by von Neumann normalisation in Theorem 42 is not due to computability properties of the sequence. In particular, there are random sequences that collapse to any string, so to strings which are not Borel normal.

In the following we need a measure-theoretic characterisation of random sequences, so we present a few facts from constructive topology and probability.

Consider the compact topological space (A_Q^ω, τ) in which the basic open sets are the sets wA_Q^ω, with $w \in A_Q^*$. Accordingly, an open set $G \subset A_Q^\omega$ is of the form $G = VA_Q^\omega$, where $V \subset A_Q^*$.

From now on we assume that the reals $p_i, 1 \leq i \leq Q$ which define the probability μ_{P_Q} are all computable. A constructively open set $G \subset A_Q^\omega$ is an open set $G = VA_Q^\omega$ for which $V \subset A_Q^*$ is computably enumerable (c.e.). A constructive sequence of constructively open sets, c.s.c.o. sets for short, is a sequence $(G_m)_{m\geq 1}$ of constructively open sets $G_m = V_m A_Q^\omega$ such that there exists a c.e. set $X \subset A_Q^* \times \mathbf{N}$ with $V_m = \{x \in A_Q^* \mid (x, m) \in X\}$, for all natural $m \geq 1$. A constructively null set $S \subset A_Q^\omega$ is a set for which there exists a c.s.c.o. sets $(G_m)_{m\geq 1}$ with $S \subset \bigcap_{m\geq 1} G_m$, $\mu_{P_Q}(G_m) \leq 2^{-m}$. A sequence $\mathbf{x} \in A_Q^\omega$ is random in the probability space $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$ if $\mathbf{x}$ is not contained in any constructively null set in $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$. For the case of the Lebesgue probability μ_{P_L} the measure-theoretic characterisation of random sequences holds true: $\mathbf{x}$ is random if and only if $\mathbf{x}$ is not contained in any constructively null set of $(A_Q^\omega, \mathcal{M}, \mu_{P_L})$ [ML66, Cal02].

We continue with another instance in which von Neumann normalisation decreases randomness.

Proposition 44. *There exist (continuously many) infinite 1/2-random sequences $\mathbf{x} \in B^\omega$ such that $VN(\mathbf{x}) = 000 \dots 00 \dots$.*

Proof. Consider a random sequence $\mathbf{x} = x_1x_2\ldots x_n\ldots$ and construct the sequence $\mathbf{x}' = 0x_10x_2\ldots 0x_n\ldots$. Clearly, $\mathbf{x}'$ is 1/2-random, but $VN(\mathbf{x}') = 000\ldots00\ldots$ because there exist infinitely many 1's in $\mathbf{x}$. □

We follow this with instances for which the converse is true: von Neumann normalisation conservers or increases randomness.

Proposition 45. *There exist (continuously many) infinite 1/2-random sequences $\mathbf{x} \in B^\omega$ such that $VN(\mathbf{x})$ is random.*

Proof. Consider a random sequence $\mathbf{x} = x_1x_2\ldots x_n\ldots$ and construct the sequence $\mathbf{x}' = x_1\bar{x}_1x_2\bar{x}_2\ldots x_n\bar{x}_n\ldots$. Clearly, $\mathbf{x}'$ is 1/2-random and $VN(\mathbf{x}') = \mathbf{x}$. □

Comment. Both Proposition 44 and 45 are true for the more general case of ε-random sequences, where $0 < \varepsilon < 1$ is computable.

Theorem 46. *Let $\mathbf{x} \in B^\omega$ be Borel normal in $(B^\omega, \mathcal{M}, \mu_{P_L})$. Then $VN(\mathbf{x})$ is also Borel normal in $(B^\omega, \mathcal{M}, \mu_{P_L})$.*

Proof. Note that $VN(\mathbf{x}) \in B^\omega$ because $\mathbf{x}$ contains infinitely many occurrences of 01 on even/odd positions. Let $D = \{00, 11\}$, $\mathbf{x}^*(n) = VN_{n,n}(\mathbf{x}(n))$, $n' = |\mathbf{x}^*(n)|$. We have

$$\lim_{n'\to\infty} \frac{N_i^m(\mathbf{x}^*(n))}{n'} = \lim_{n'\to\infty} \left(\frac{n}{n'}\right)\left(\frac{N_i^m(\mathbf{x}^*(n))}{n}\right),$$

but as $n \to \infty$, $n' \to \infty$. We thus have

$$\begin{aligned}\lim_{n'\to\infty} \frac{n'}{n} &= \lim_{n'\to\infty} \frac{N_0^1(\mathbf{x}^*(n)) + N_1^1(\mathbf{x}^*(n))}{n} \\ &= \lim_{n\to\infty} \frac{\mathcal{N}_{01}^2(\mathbf{x}(n)) + \mathcal{N}_{10}^2(\mathbf{x}(n))}{\lfloor n/2 \rfloor} \\ &= 2^{-1}\end{aligned}$$

by the normality of $\mathbf{x}$. The number of occurrences of each $i = i_1\ldots i_m \in B^m$ in $\mathbf{x}^*(n)$ is the number of occurrences of $i' = f(i_1)y_1f(i_2)\ldots y_{m-1}f(i_m)$ in $\mathbf{x}(n)$, summed over all

$y_1, \ldots, y_{m-1} \in D^*$. Viewing i' as a string over B^2 we have:

$$\begin{aligned}
\lim_{n' \to \infty} \frac{N_i^m(\mathbf{x}^*(n))}{n} &= \lim_{n \to \infty} \frac{\sum_{y_1,\ldots,y_{m-1}} N_{i'}^{|i'|}(\mathbf{x}(n))}{n} \\
&= \sum_{y_1 \in D^*} \sum_{y_2 \in D^*} \cdots \sum_{y_{m-1} \in D^*} 2^{-2|i'|} \\
&= \sum_{|y_1|=0}^{\infty} 2^{|y_1|} \sum_{|y_2|=0}^{\infty} 2^{|y_2|} \cdots \sum_{|y_{m-1}|=0}^{\infty} 2^{|y_{m-1}|} 2^{-2|i'|} \\
&= 2^{-2m} \sum_{|y_1|=0}^{\infty} 2^{-|y_1|} \sum_{|y_2|=0}^{\infty} 2^{-|y_2|} \cdots \sum_{|y_{m-1}|=0}^{\infty} 2^{-|y_{m-1}|} \\
&= 2^{-2m} 2^{m-1} \\
&= 2^{-(m+1)}.
\end{aligned}$$

Hence, both limits exist and we have

$$\begin{aligned}
\lim_{n' \to \infty} \frac{N_i^m(\mathbf{x}^*(n))}{n'} &= \lim_{n' \to \infty} \left(\frac{n}{n'}\right) \left(\frac{N_i^m(\mathbf{x}^*(n))}{n}\right) \\
&= \frac{\lim_{n' \to \infty} \frac{N_i^m(\mathbf{x}^*(n))}{n}}{\lim_{n' \to \infty} \frac{n'}{n}} \\
&= \frac{2^{-(m+1)}}{2^{-1}} \\
&= 2^{-m}.
\end{aligned}$$

Since this holds for all m, i we have that $VN(\mathbf{x})$ is Borel normal. □

Let $A_Q = \{a_1, \ldots, a_Q\}$, $Q \geq 3$. Let $\sum_{i=1}^{Q} p_i = 1$ where $p_i \geq 0$ for $i = 1, \ldots, Q$ and $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$ be the probability space defined by the probabilities p_i. Let $A_{Q-1} = \{a_1, \ldots, a_{Q-1}\}$ and $(A_{Q-1}^\omega, \mathcal{M}, \mu_{P_{Q-1}^T})$ be the probability space defined by the probabilities

$$p_i^T = p_i \left(1 + \frac{p_Q}{\sum_{j=1}^{Q-1} p_j}\right) = \frac{p_i}{1 - p_Q},$$

with $1 \leq i \leq Q-1$. Let $T : A_Q^* \to A_{Q-1}^*$ be the monoid morphism defined by $T(a_i) = a_i$ for $1 \leq i \leq Q-1$, $T(a_Q) = \lambda$; $T(x) = T(x_1)T(x_2)\cdots T(x_n)$ for $x \in A_Q^n$. As T is prefix-increasing we naturally extend T to sequences to obtain the function $T : A_Q^\omega \to A_{Q-1}^\omega$ given by $T(\mathbf{x}) = \lim_{n \to \infty} T(\mathbf{x}(n))$ for $\mathbf{x} \in A_Q^\omega$.

Lemma 47. *The transformation T is $(\mu_{P_Q}, \mu_{P_{Q-1}^T})$-preserving, i.e. for all $w \in A_{Q-1}^*$ we have $\mu_{P_Q}\left(T^{-1}(wA_{Q-1}^\omega)\right) = \mu_{P_{Q-1}^T}\left(wA_{Q-1}^\omega\right)$.*

Proof. Take $w = w_1 \dots w_m \in A_{Q-1}^m$. We have:

$$\begin{aligned}
\mu_{P_Q}\left(T^{-1}(wA_{Q-1}^\omega)\right) &= \mu_{P_Q}\left(\{\mathbf{x} \in A_Q^\omega \mid w \sqsubset T(\mathbf{x})\}\right) \\
&= \mu_{P_Q}\left\{a_Q^{i_1} w_1 a_Q^{i_2} w_2 \dots a_Q^{i_m} w_m \mathbf{z} \mid \mathbf{z} \in A_Q^\omega\right\} \\
&= \sum_{i_1,\dots,i_m=0}^{\infty} h_Q\left(a_Q^{i_1} w_1 a_Q^{i_2} w_2 \dots a_Q^{i_m} w_m\right) \\
&= \sum_{i_1,\dots,i_m=0}^{\infty} h_{Q-1}(w) \cdot p_Q^{i_1+\dots+i_m} \\
&= h_{Q-1}(w) \cdot \frac{1}{1-p_Q} \\
&= h_{Q-1}^T(w) \\
&= \mu_{P_{Q-1}^T}\left(wA_{Q-1}^\omega\right).
\end{aligned}$$

□

Proposition 48. *If $\mathbf{x} \in A_Q^\omega$ is random in $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$ and T is the transformation defined in Lemma 47, then $T(\mathbf{x})$ is random in $(A_{Q-1}^\omega, \mathcal{M}, \mu_{P_{Q-1}^T})$.*

Proof. We generalise a result in [CHJW01] stating that, for the Lebesgue probability, measure-preserving transformations preserve randomness. Assume that $\mathbf{x}$ is random in $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$ but $T(\mathbf{x})$ is not random in in $(A_{Q-1}^\omega, \mathcal{M}, \mu_{P_{Q-1}^T})$, i.e. there is a constructive null set $R = (G_m)_{m\geq 1}$ containing $T(\mathbf{x})$. Assume that $G_m = X_m A_{Q-1}^\omega$, where $X_m \subset A_{Q-1}^\omega$ is c.e. and has the measure $\mu_{P_{Q-1}^T}(X_m A_{Q-1}^\omega)$ smaller than 2^{-m}. Define $S_m = T^{-1}(X_m A_{Q-1}^\omega) \subset A_Q^\omega$ and note that S_m is open because it is equal to $\bigcup_{w\in X_m} V_w A_Q^\omega$ with $V_w = \{v \in A_Q^\omega \mid w \sqsubset T(v)\}$ and, using Lemma 47, has the measure smaller than 2^{-m}:

$$\begin{aligned}
\mu_{P_Q}(S_m) &= \mu_{P_Q}\left(\bigcup_{w\in X_m} V_w A_Q^\omega\right) \\
&\leq \sum_{w\in X_m} \mu_{P_Q}\left(V_w A_Q^\omega\right) \\
&= \sum_{w\in X_m} \mu_{P_Q}\left(T^{-1}\left(wA_{Q-1}^\omega\right)\right) \\
&= \mu_{P_{Q-1}^T}\left(X_m A_{Q-1}^\omega\right) \\
&\leq 2^{-m}.
\end{aligned}$$

We have proved that $\mathbf{x}$ is not random in $(A_Q^\omega, \mathcal{M}, \mu_{P_Q})$, a contradiction. □

Let us define $VN^{-1} : 2^{B^*} \to 2^{B^*}$ for $x = x_1 \dots x_m \in B^m$ as

$$\begin{aligned}
VN^{-1}(x) = \{y \mid\ & y = u_1 f(x_1) u_2 \dots u_m f(x_m) u_{m+1} v \text{ and} \\
& u_i \in \{00, 11\}^* \text{ for } 1 \leq i \leq m, v \in B \cup \{\lambda\}\} \\
= \bigcup_{n=0}^{\infty} & VN_{n+2m,m}^{-1}(x),
\end{aligned}$$

and for $X \subseteq B^*$ as

$$VN^{-1}(X) = \bigcup_{x \in X} VN^{-1}(x).$$

For all $x \in B^*$ and $\mathbf{y} \in VN^{-1}(x)B^\omega$ we then have $x \sqsubset VN(\mathbf{y})$.

For the cases that $VN(\mathbf{x}) \in B^\omega$, the probability space $(B^\omega, \mathcal{M}, \mu_{P_{VN}})$ induced by von Neumann normalisation is endowed with the measure $\mu_{P_{VN}}$. The measure $\mu_{P_{VN}}$ is defined on the sets xB^ω with $x \in B^*$ by

$$\mu_{P_{VN}}(xB^\omega) = \frac{\mu_P(VN^{-1}(x)B^\omega)}{\mu_P(VN^{-1}(B^{|x|})B^\omega)}.$$

By noting that $VN^{-1}(B^{|x|}) \subset VN^{-1}(B^*)$ it is clear to see that $\mu_{P_{VN}}$ satisfies the Kolmogorov axioms for a probability measure. While the set $VN^{-1}(B^{|x|})$ contains sequences for which normalisation produces a finite string, from Corollary 43 we know that the set of such sequences have measure zero, so the definition of $\mu_{P_{VN}}$ is a good model of the target probability space.

Theorem 49. *Let $\mathbf{x} \in B^\omega$ be random in $(B^\omega, \mathcal{M}, \mu_P)$. Then $VN(\mathbf{x}) \in B^\omega$ is also random in $(B^\omega, \mathcal{M}, \mu_{P_{VN}})$.*

Proof. We write the random sequence $\mathbf{x}$ as $\mathbf{x} = x_1x_2\ldots x_n\ldots = (x_1x_2)\ldots(x_{2n-1}x_{2n})\cdots \in \{00, 01, 10, 11\}^\omega$. Renaming $a = 00, A = 01, B = 10, b = 11$ and consistently deleting first all occurrences of a we get a random sequence $\mathbf{x}_{A,B,b}$ on the alphabet $\{A, B, b\}$, then deleting all occurrences of b we get a random sequence $\mathbf{x}_{A,B}$ on the alphabet $\{A, B\}$. The result follows from the fact that $VN(\mathbf{x}) = \mathbf{x}_{0,1}$ and Proposition 48 stating that $\mathbf{x}_{A,B}$ is random. □

Corollary 50. *If $\mathbf{x} \in B^\omega$ is random in $(B^\omega, \mathcal{M}, \mu_P)$ then $VN(\mathbf{x})$ is Borel normal in $(B^\omega, \mathcal{M}, \mu_{P_{VN}})$.*

Proof. From Theorem 49 it follows that $VN(\mathbf{x})$ is Borel normal provided $\mathbf{x}$ is random [Cal02]. □

Theorem 51. *The probability space $(B^\omega, \mathcal{M}, \mu_{P_{VN}})$ induced by von Neumann normalisation is the uniform distribution $(B^\omega, \mathcal{M}, \mu_{P_L})$, where μ_{P_L} is the Lebesgue measure.*

Proof. By Lemma 47, von Neumann normalisation is measure preserving, so for $x \in B^*$ we have

$$\begin{aligned}\mu_{P_{VN}}(xB^\omega) &= \mu_P(VN^{-1}(x)B^\omega) \\ &= p_0^{|x|}p_1^{|x|}\sum_{d_i \in D^*} p_0^{\#_0(d_1\ldots d_{|x|})}p_1^{\#_1(d_1\ldots d_{|x|})}.\end{aligned}$$

The key point, as in the finite case, is that this only depends on $|x|$ not x itself. By using the fact that for any n, $\sum_{x\in B^n} \mu_{P_{VN}}(xB^\omega) = 1$, we have

$$\mu_{P_{VN}}(xB^\omega) = 2^{-|x|}$$

for all $x \in B^*$, and hence $\mu_{P_{VN}} = \mu_{P_L}$, the Lebesgue measure. □

This can easily be extended from the case when $VN(\mathbf{x})$ is infinite, to the case in which it is finite. To do so, note that if $\mathbf{y} \in B^\omega$ and $VN(\mathbf{x}) = y \in B^n$, then the probability space induced by von Neumann normalisation is $(B^n, 2^{B^n}, P_n^*)$. We then have

$$P_n^*(x) = \frac{\mu_P(VN^{-1}(x)D^\omega)}{\mu_P(VN^{-1}(B^n)D^\omega)},$$

and since the denominator is constant for all $x \in B^n$, we can proceed as for above, and $P_n^* = U_n$ as desired.

Theorem 52. *The set* $\{\mathbf{x} \in B^\omega \mid VN(\mathbf{x}) \in B^* \text{ or } VN(\mathbf{x}) \in B^\omega \text{ is computable}\}$ *has measure zero with respect to the probability space* $(B^\omega, \mathcal{M}, \mu_P)$.

Proof. By Theorem 49 we deduce that

$$\{\mathbf{x} \in B^\omega \mid VN(\mathbf{x}) \in B^\omega \text{ is computable}\} \subset \{\mathbf{x} \in B^\omega \mid \mathbf{x} \text{ is not random in } (B^\omega, \mathcal{M}, \mu_P)\},$$

which has measure zero [ML66]. To complete the proof, note that we know from Corollary 43 that the set $\{\mathbf{x} \in B^\omega \mid VN(\mathbf{x}) \in B^*\}$ also has measure zero. □

Both unpredictability and uniformity of distribution are independent symptoms of randomness, and it is important that any method to remove bias and ensure uniformity does not decrease unpredictability. Von Neumann's method preserves randomness and Borel normality, but fails to preserve incomputability in general. It is not known whether sequences of bits from a QRNG are random or even Borel normal with respect to the space $(B^\omega, \mathcal{M}, \mu_P)$, so it follows that such sequences may well loose unpredictability when normalised since only strong incomputability is known to be guaranteed. However, the incomputability produced by the QRNG is stronger than that present in the cases we have found for which incomputability is not preserved, and it remains an *open question* to determine if strong incomputability is preserved.

Open Question 53. *Given a sequence* $\mathbf{x} \in B^\omega$ *produced by a QRNG certified by value indefiniteness, is* $VN(\mathbf{x})$ *strongly incomputable in the sense of Theorem 17 in the same way the* $\mathbf{x}$ *is?*

Even if strong incomputability isn't preserved, it may still be the case that the unpredictability can be guaranteed to be preserved, but further theoretical characterisation of such sequences is needed to examine such issues. Fortunately, any 'damage' that may be caused by normalisation is limited in measure: it holds only with probability zero.

2.4 Role of probability spaces for QRNGs

The treatment of QRNGs as entirely probabilistic devices is grounded purely on the probabilistic treatment of measurement in quantum mechanics which originated with Born's decision to "give up determinism in the world of atoms" [Bor26], a viewpoint which has become a core part of our understanding of quantum mechanics. This is formalised by the Born rule, but the probabilistic nature of *individual* measurement is nonetheless postulated and tells us nothing about *how* the probability arises. Along with the assumption of independence this allows us to predict the probability of *successive* events, as we have done.

No-go theorems such as the Kochen-Specker Theorem tell us something stronger: that the unpredictability in non-trivial quantum systems is not merely due to ignorance of the system being measured. Indeed, since there are in general no definite values associated with the measured observable it is surprising there is an outcome at all [Svo04]. While this does not answer the question as to where the unpredictability arises from, it does tell us something stronger than the Born Rule does. This result is used in Theorem 16 to show that it is *impossible* for a QRNG to output a computable sequence. The set of computable numbers has measure zero with respect to the probability space of the QRNG, but the impossibility of producing such a sequence is much stronger than, although not in contradiction with, the probabilistic results.

In the finite case every string is, of course, obtainable, and we would expect the distribution to be that predicted by the probability space derived from the Born Rule. However, as was discussed in Section 1.3.2, Theorem 16 still shows that the QRNG provides unpredictable bits in a very real sense in this case.

The results we have presented in this chapter, however, describe thoroughly the distribution of strings/sequences produced by QRNGs. With the distributions known we can create more intelligent tests of the quality of output of a QRNG [CDDS10]. Current statistical tests for analysing RNGs are designed with pseudo-RNGs in mind and are not necessarily the best way to test the quality of QRNGs. The effects of normalisation on strings generated by QRNGs can help us design QRNGs which are more robust to experimental imperfection and exhibit the desired behaviour. It will further aid in developing new normalisation techniques designed to produce the expected (ideal) theoretical distribution even in the absence of experimental imperfections.

2.5 Summary

The analysis developed in this chapter involves the probability spaces of the source and output of a QRNG and the effect von Neumann normalisation has on these spaces.

In the 'ideal case', the von Neumann normalised output of an independent constantly biased QRNG is the probability space of the uniform distribution (un-biasing). This result is true for both for finite strings and for the infinite sequences produced by QRNGs (the QRNG runs indefinitely in the second case).

For a real-world QRNG in which the bias, rather than holding steady, drifts slowly, we evaluated the speed of drift required to be maintained by the source distribution to guarantee that the output distribution is arbitrarily close to the uniform distribution. It is an *open question* to study the more realistic case when, instead of the bits being independent, the probability for each bit depends on a finite number of preceding bits (for example, because of the high bit-rate of the experiment).

We have also examined the effect von Neumann normalisation has on various properties of infinite sequences. In particular, Borel normality and (algorithmic) randomness are invariant under normalisation, but for ε-random sequences with $0 < \varepsilon < 1$, normalisation can both decrease or increase the randomness of the source. It is an *open question* whether von Neumann normalisation preserves randomness and Borel normality for finite strings.

Finally, we reiterate that a successful application of von Neumann normalisation—in fact, any un-biasing transformation—does exactly what it promises, *un-biasing*, one (among infinitely many) symptoms of randomness; it will not produce 'true' randomness.

Chapter 3

Quantum Random Number Generator Design

3.1 Quantum random number generators

In this chapter we propose a QRNG which uses an entangled photon pair in a Bell singlet state, and is certified explicitly by value indefiniteness. While 'true randomness' is a mathematical impossibility, the certification by value indefiniteness ensures the quantum random bits are incomputable in the strongest sense. This is the first QRNG setup in which a physical principle (in this case, value indefiniteness) guarantees that no single quantum bit produced can be classically computed, the mathematical form of bitwise physical unpredictability.

The effects of various experimental imperfections are discussed in detail, particularly those related to detector efficiencies, context misalignment and temporal correlations between bits. The analysis is, to a large extent, relevant for the construction of any QRNG based on beam-splitters. By measuring the two entangled photons in maximally misaligned contexts and utilising the fact that two bit-strings rather than one are obtained, more efficient and robust unbiasing techniques can be applied. A robust and efficient procedure based on XORing the bit-strings together, essentially using one as a one-time-pad for the other, is proposed to extract random bits in the presence of experimental imperfections, as well as a more efficient modification of the von Neumann procedure for the same task.

3.1.1 Existing QRNGs

Quantum random number generators based on beam splitters [Svo90, ROT94] have been realised by the Zeilinger group in Innsbruck and Vienna [JAW+00] and applied for the sake of violation of Bell's inequality under strict Einstein locality conditions [WJS+98].

The Gisin group in Geneva [SGG+00], and in particular its spin-off *id Quantique*, produces and markets a commercial device called *Quantis* [iQ09]. In order to eliminate bias, the device employs Peres' more efficient iterated version of von Neumann normalisation [Per92].

A group in Shanghai and Beijing [WLL06] has utilised a Fresnel multiple prism as a polarising beam splitter. As a crude normalisation technique, previously generated experimental sequences have been used as one time pad to 'encrypt' random sequences.

QRNGs based on entangled photon pairs have been realised by a second Chinese group in Beijing and Ji'nan [HQSMD+04], who used spontaneous parametric down-conversion to produce entangled pairs of photons. One of the photons has been used as trigger, mostly to allow a faster data production rate by eliminating double counts. Again, von Neumann normalisation has been applied in an attempt to eliminate bias.

A group from the Hewlett-Packard Laboratories in Palo Alto and Bristol [FSS+07] has used entangled photon pairs in the Bell basis state $|H_1V_2\rangle + |V_1H_2\rangle$ (note that this is not a singlet state and attains this form only for one polarisation direction; in all the other directions the state contains also V_1V_2 as well as H_1H_2 contributions), where the states H_1, V_1 and H_2, V_2 are horizontal and vertical polarisation states for the two photons. In analogy to von Neumann normalisation, the coincidence events H_1V_2 and V_1H_2 have been mapped into 0 and 1, respectively. Thereby, as the authors have argued, the 2-qubit space of the photon pair is effectively restricted to a two-dimensional Hilbert subspace described by an effective-qubit state.

A more recent rendition of a QRNG [PAM+10], although not based on photons and beam splitters, utilises Bell-type setups 'secured by' Bell-type inequality violations in the spirit of quantum cryptographic protocols [Eke91, BPP00]. This provides some indirect 'statistical verification' of value indefiniteness (again under the assumption of non-contextuality), but falls short of providing certification of strong incomputability via value indefiniteness [CS08, Svo09].

With regard to value indefiniteness, the difference between Bell-type inequalities versus Kochen-Specker-type value indefiniteness is this: in the Bell-type case, the breach of value indefiniteness needs not happen at every single particle, whereas in the Kochen-Specker-type case this must happen *for every particle* [Svo10]. Pointedly stated, the Bell-type violation is a statistical property and cannot guarantee any property of *every individual* context preparation and measurement. Hence, because a Bell-type inequality

violation does not guarantee that every bit is certified by value indefiniteness, one could potentially produce sequences containing infinite computable subsequences 'protected' by Bell-type inequality violations. Further, given the ability to simulate maximal violation of Bell-type inequalities with Mealy automata [Cab10], such 'certification' fails to truly certify the incomputability of the QRNG; we can have computable sequences which maximally violate Bell-type inequalities. Indeed, given that such criticisms seem also to hold for the statistical verification of value indefiniteness [PBD+00, HLZ+03, Cab08], it seems unlikely that statistical tests of the measurement outcomes alone can fully certify such a QRNG.

3.1.2 Shortcomings of current QRNGs

It is clear that any QRNG claiming to produce a better quality of randomness than existing RNGs has to produce at least an infinite incomputable sequence of outputs, preferably a strongly incomputable one. Do the current proposals of QRNGs generate 'in principle' strongly incomputable sequences of quantum random bits? To answer this question one has to check whether the QRNG is 'protected' by value indefiniteness; in most cases the answer is either negative or cannot be verified because of lack of information about the mechanism of the QRNG.

In [CDDS10], tests based on algorithmic information theory were used to analyse and compare quantum and non-quantum bit-strings. Ten strings of length 2^{32} bits each from two quantum sources (the commercial *Quantis* device [iQ09] and the Vienna Institute for Quantum Optics and Quantum Information group) and three classical sources (Mathematica, Maple and the binary expansion of π) were analysed. No distribution was assumed for any of the sources, yet a test based on Borel normality was able to distinguish between the quantum and non-quantum sources of 'random' numbers. It is known that all algorithmically random strings are Borel normal [Cal02], although the converse is not true. Indeed, the tests found the quantum sources to be less normal than the pseudorandom ones. Is this a property of quantum randomness, or evidence of flaws in the tested QRNGs?

In Chapter 2, the probability distribution for an ideal QRNG was discussed and found to be the uniform distribution. Using the same strings as in [CDDS10] we conducted tests to see if the strings indeed appear to be sampled from the uniform distribution; this is in contrast to the tests in [CDDS10] which assumed no prior distribution. For bit-strings of length n sampled from the uniform distribution, the probability distribution function for the number of occurrences of substrings of length k can be seen to follow a multinomial distribution: we view each bit-string as being over the alphabet B^k with length $\lfloor n/k \rfloor$, and view this as $\lfloor n/k \rfloor$ tosses of a fair 2^k sided dice. Borel normality

QRNG	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
Maple	0.79	0.15	0.83	0.47	0.97
Mathematica	0.18	0.38	0.35	0.45	0.99
π	0.38	0.27	0.05	0.62	0.21
Quantis	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$
Vienna	0.12	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$	$\mathbf{< 10^{-10}}$

Table 3.1: p-values for the χ^2 test that the bit-string is sampled from the uniform distribution. Bold values indicate statistically significant evidence (at a $p = 0.01$ significance level) that the strings are not sampled from the uniform distribution.

tells us we should test this for $k \leq \lg \lg n$, i.e. $k \leq 5$. For the strings used in [CDDS10] we have $n = 10 \cdot 2^{32}$, and we can approximate the multinomial distribution extremely well by a multivariate normal distribution, and use a χ^2-test to test the null-hypothesis that the dice is fair, i.e. that each substring of length k has equal probability. Results of the analyses are presented in Table 3.1.

We find that for all the non-quantum sources (Mathematica, Maple, π) we have $p \geq 0.05$, so we have no evidence against the hypothesis that each substring is equally probable at the 5% significance level. Certainly these sources are not being 'sampled' in the same sense as the quantum strings, but we note that nonetheless any pseudo-random number generator should satisfy the requirement of uniformity of distribution—it is a necessary symptom of randomness. Indeed, for longer bit-strings ($k = 5$ for Maple, Mathematica), we have $p > .95$, indicating that these sources are potentially *too normal*, trying too hard to be 'random'. We also note that for π, any evidence against it being from the uniform distribution would be evidence against normality and would run against our knowledge of π, even though it is not proven to be Borel normal (such proofs are notoriously hard in general).

For the quantum sources we find, except for one notable exception, with extremely significant results ($p < 10^{-17}$) that the strings are not sampled from the uniform distribution. The exception to these findings are the Vienna bits which, when viewed at the single-bit level, appear unbiased. It appears that the good performance at the 1-bit level has been achieved (perhaps through experimental feedback control) at the sacrifice of the performance at the $k \geq 2$ level, a property much harder to control without post-processing. The *Quantis* QRNG uses Peres' normalisation in an attempt to unbias the output; in light of the findings in Chapter 2, the fact that this is not completely successful indicates either a significant variation in bias over time, or non-independence of successive bits.

These results highlight the need to pay extra attention in the design process to

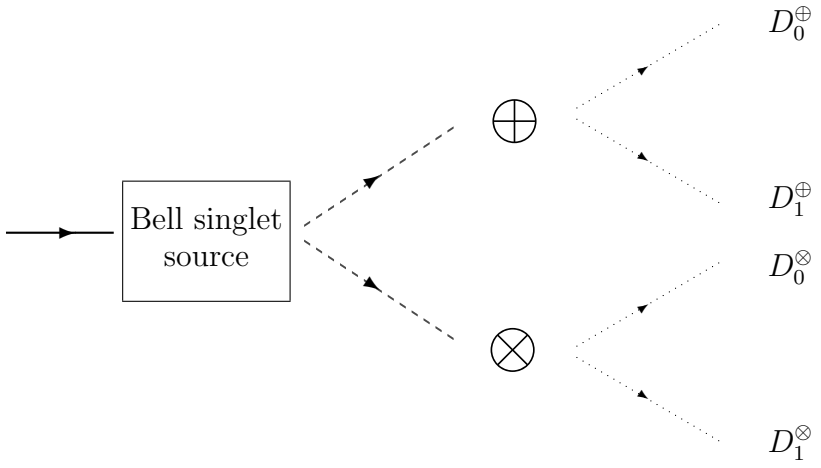

Figure 3.1: Scheme of a quantum random number generator [Svo09].

the distribution produced by a QRNG. Normalisation techniques are an effective way to remove bias, but to have the desired effect assumptions about independence and constancy of bias must be satisfied. While experiments will never realise the ideal QRNG, one needs to be aware of how much affect experimental imperfections have. Any credible QRNG should take these issues into account, as well as the need of explicit certification of randomness by some physical law, e.g. value indefiniteness.

3.2 The scheme under ideal conditions

In what follows, a proposal for a QRNG depicted in Figure 3.1, previously put forward in [Svo09], will be discussed in detail. It utilises the singlet state of two two-state particles (e.g., photons of linear polarisation) proportional to $|H_1V_2\rangle - |V_1H_2\rangle$, which is form invariant in all measurement directions.

A single photon light source (presumably an LED) is attenuated so more than one photons are rarely in the beam path at the same time. These photons impinge on a source of singlet states of photons (presumably by spontaneous parametric down-conversion in a nonlinear medium). The two resulting entangled photons are then analysed with respect to their linear polarisation state at some directions which are $\pi/4$ radians 'apart', symbolised by '$\oplus$' and '$\otimes$', respectively.

Due to the required four-dimensional Hilbert space, this QRNG is 'protected' by Bell- as well as Kochen-Specker-type value indefiniteness.[1] The protocol utilises all three principal types of quantum indeterminism: the postulated indeterminacy of individual outcomes via the Born rule, quantum complementarity (due to the use of conjugate

[1]Note that this is not the case for current QRNGs based on beam-splitters, which mostly operate in a Hilbert space of dimension two.

variables), and value indefiniteness. This, essentially, is the same experimental configuration as the one used for a measurement of the correlation function at the angle of $\pi/4$ radians (45°). Whereas the correlation function averages over 'a large number' of single contributions, a random sequence can be obtained by concatenating these single pairs of outcomes via addition modulo 2.

Formally, suppose that for the ith experimental run, the two outcomes are $O_i^{\oplus} \in B$ corresponding to $D_0^{\oplus}$ or $D_1^{\oplus}$, and $O_i^{\otimes} \in B$ corresponding to $D_0^{\otimes}$ or $D_1^{\otimes}$. These two outcomes $O_i^{\oplus}$ and $O_i^{\otimes}$, which themselves form two sequences of random bits, are subsequently combined by the XOR operation, which amounts to their parity, or to the addition modulo 2. (In what follows, depending on the formal context, XOR refers to either a binary function of two binary observables, or to the bit-wise logical operation; see Section 1.1.1.) Stated differently, one outcome is used as a *one time pad* to 'encrypt' the other outcome, and vice versa. As a result, one obtains a sequence $x = x_1 x_2 \dots x_n$ with

$$x_i = O_i^{\oplus} + O_i^{\otimes} \bmod 2. \tag{3.1}$$

For the XOR'd sequence to still be certifiably incomputable (via value indefiniteness), one must prove this certification is preserved under XORing—indeed strong incomputability itself is *not* necessarily preserved. By necessity any QRNG certified by value indefiniteness must operate non-trivially in a Hilbert space of dimension $n \geq 3$. To transform the n-ary (incomputable) sequence into a binary one, a function $f : \{0, 1, \dots, n-1\} \to \{0, 1, \lambda\}$ must be used; to claim certification, the strong incomputability of the bits must still be guaranteed after the application of f. This is a fundamental issue which has to be checked for existing QRNGs such as that in [PAM+10]; without it one cannot claim to produce truly indeterministic bits. In general incomputability itself is not preserved by f; however by consideration of the value indefiniteness of the source the certification can be seen to hold under XOR as well as when discarding bits [ACS].

3.3 'Random' errors and systematic errors

In what follows we shall discuss possible 'random' (no pun) or systematic errors in experimental realisations of this QRNG (many of these errors may appear in other types of photon-based QRNGs). Our aim is to draw attention to the specific nature of such errors and how they affect the resulting bit-strings. A good QRNG must, in addition to the necessary certification (e.g. by value indefiniteness), take into account the nature of these errors and be carefully designed (along with any subsequent post-processing) so that the resultant distribution of bit-strings the QRNG samples from is

as close as possible to the expected uniform distribution. Both the uniformity of the source and incomputability are 'independent symptoms' of randomness, and care must be taken to obtain both properties.

3.3.1 Double counting

One conceivable problem is that the detectors analysing the different polarisation directions do not respond to photons of the same pair, but to two photons belonging to different pairs. This seems to be no drawback for the application of the XOR operation since (at least in the absence of temporal correlations between bits) the postulates of quantum mechanics state that the individual outcomes occur independently and indeterministically (the last property is mathematically modelled by strong incomputability). If, however, events are not independent then more care is needed. However, correlation between events is an undesirable property in itself, and as long as care is made, it is unlikely to be made worse by double counting.

3.3.2 Non-singlet states

The state produced by the spontaneous parametric down-conversion may not be exactly a singlet. This may give rise to a systematic bias of the combined light source-analyser setup in a very similar way as for beam splitters.

3.3.3 Non-alignment of polarisation measurement angles

No experimental realisation will attain a 'perfect anti-alignment' of the polarisation analysers at angles $\pi/4$ radians apart. Only in this ideal case are the bases conjugate and the correlation function will be exactly zero. Indeed, 'tuning' the angle to obtain equibalanced sequences of zeroes and ones may be a method to properly anti-align the polarisers. However, one has to keep in mind that any such 'tampering' with the raw sequence of data to achieve Borel normality (e.g. by readjustments of the experimental setup) may introduce unwanted (temporal) correlations or other bias [CDDS10].

Incidentally, the angle $\pi/4$ is one of the three points at angles 0, $\pi/4$ and $\pi/2$ in the interval $[0, \pi/2]$ in which the classical and quantum correlation functions coincide. For all other angles, there is a higher ratio of different or identical pairs than could be expected classically. Thus, ideally, the QRNG could be said to operate in the 'quasi classical' regime, albeit fully certified by quantum value indefiniteness.

Quantitatively, the expectation function of the sum of the two outcomes modulus 2 can be defined by averaging over the sum modulo 2 of the outcomes $O_i^0, O_i^\theta \in B$ at angle

θ 'apart' in the ith experiment, over a 'large number' of experiments; i.e.,

$$E_{\oplus}(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left(O_i^0 \oplus O_i^\theta\right).$$

This is related to the standard correlation function,

$$C(\theta) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} O_i^0 \cdot O_i^\theta$$

by

$$E_{\oplus}(\theta) = \frac{|C(\theta) - 1|}{2},$$

where

$$O_i^0 \cdot O_i^\theta = \begin{cases} 1 & \text{if } O_i^0 = O_i^\theta, \\ -1 & \text{if } O_i^0 \neq O_i^\theta. \end{cases}$$

A detailed calculation yields the classical linear expectation function $E_{\oplus}^{\mathrm{cl}}(\theta) = 1 - 2\theta/\pi$, and the quantum expectation function $E_{\oplus}(\theta) = (1/2)(1 + \cos 2\theta)$.

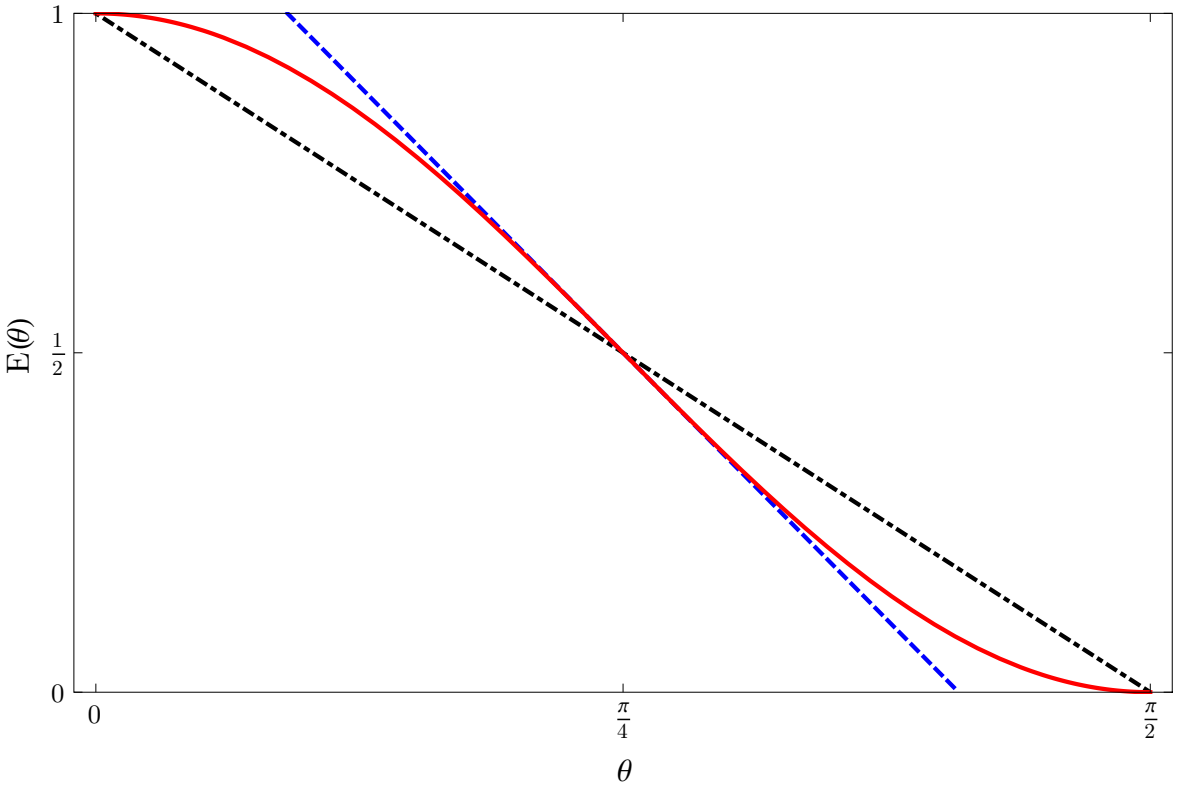

Figure 3.2: (Colour online) The classical (black, dot-dashed) and quantum (red, solid) expectation functions and the linear quantum approximation (blue, dashed) around $\pi/4$.

Thus, for angles 'far apart' from $\pi/4$, the XOR operation actually *deteriorates* the two random signals taken from the two analysers *separately.* The deterioration is even *greater quantum mechanically than classically,* as the entangled particles are more correlated and thus 'less independent'. Potentially, this could be utilised to ensure a $\pi/4$ mismatch more accurately than possible through classical means. This will be discussed in Section 3.4 below.

In order to avoid this negative feature while generating bits, instead of XORing outcomes of *identical* partner pairs, one could XOR time-shifted outcomes; e.g., instead of the expression in (3.1) one may consider

$$x_i = O_i^0 + O_{i+j}^\theta \bmod 2, \text{ with } j > 0. \tag{3.2}$$

One should make j large enough so that, taking into account double counting, there is no chance of accidentally causing two offset but correlated outcomes to be XOR'd together. Theoretical analysis of the effects of experimental imperfections and the XOR operation are discussed later in the chapter, and XORing shifted pairs is an efficient and effective procedure for reducing such errors.

3.3.4 Different detector efficiencies

Differences in detector efficiencies result in a bias of the sequence. This complicating effect is separate from non-perfect misalignment of polarisation context. Suppose that the probabilities of detection are denoted by p_{H_1}, p_{H_2}, p_{V_1}, p_{V_2}. Since $p_{H_1} + p_{V_1} = p_{H_2} + p_{V_2} = 1$, the probability to find pairs adding up to 0 and 1 modulo 2 are $p_{H_1}p_{H_2} + p_{V_1}p_{V_2} = 1-(p_{H_1}+p_{H_2})+2p_{H_1}p_{H_2}$ and $p_{H_1}p_{V_2}+p_{V_1}p_{H_2} = p_{H_1}+p_{H_2}-2p_{H_1}p_{H_2}$, respectively (adding up to 1). If both $p_{H_1} \neq p_{V_1}$ and $p_{H_2} \neq p_{V_2}$ then the resulting XOR'd sequence is biased. The two obtained sequences could be unbiased before or after XORing by the von Neumann normalisation, although any temporal correlations would violate the condition of independence required by this method. One should keep in mind, however, that the von Neumann normalisation procedure necessarily discards many bits (although methods such as Peres' [Per92] are more efficient). The efficiency can be increased by utilising both strings more carefully, and such a method is discussed in Section 3.5.4.

3.3.5 Unstable detector bias

Von Neumann type normalisation procedures will only remove bias due to detector efficiencies if the bias remains constant over time. As we saw in Chapter 2, if the bias drifts over time due to instability in the detectors then the resulting normalised sequence will not be unbiased but instead will simply be less biased. It is difficult to

overcome this, as experimental instability is inevitable. However, bounds on the bias of the normalised sequence, as in Theorem 40, can be used to determine the length for which the source samples 'closely enough' from the uniform distribution.

If the bias varies independently between detectors, the XORing process should serve to reduce the impact of varying detector efficiencies and applying von Neumann normalisation to the XOR'd bit-string is advantageous compared working with a single bit-string from a source of varying bias.

3.3.6 Temporal correlations, photon clustering and 'bunching'

Due to the Hanbury-Brown-Twiss effect [HT56], the photons may be temporally correlated and thus arrive clustered or 'bunched'. Temporal correlations appear also in double-slit type experiments in the time domain [LSW$^+$05], in which the role of the slits is played by windows in time of attosecond duration. This can, to an extent, be avoided by ensuring successive photons are sufficiently separated, although this poses a limit on the bit-rate of such a device. However, since the case where two or more singlet pairs are in the beam path at once is potentially of sufficient importance, this effect needs further careful consideration.

Another conceivable source of temporal correlations is due to the detector dead-time, T_d, during which the detector is inactive after measurement [SGG$^+$00]. If we measure $O_i^{\oplus} = 0$, the detector $D_0^{\oplus}$ corresponding to 0 is unable to detect another photon for a small amount of time, significantly increasing the chance of detecting a photon at the other detector during this time, obtaining a 1. This leads to higher than expected chances of 01 and 10 being measured. This is problematic as such a correlation will not be removed by XORing, even with an offset of j. However, this can be avoided by discarding any measurements within time T_d from the previous measurement.

In view of conceivable temporal correlations, it would be interesting to test the quality of the random signal as j is varied in (3.2). As previously mentioned, any temporal correlations will violate the condition of independence needed for von Neumann normalisation making it difficult to remove any bias in the output; if the dependence can be bounded then unbiasing techniques such as that proposed by Blum [Blu86] could be used instead of von Neumann's procedure. It seems desirable and simpler to avoid temporal correlations with carefully designed experimental methodology as opposed to post-processing where possible.

3.3.7 Fair sampling

As in most optical tests of Bell's inequalities, the inefficiency of photon detection requires us to make the *fair sampling assumption* [GM87, Lar98, Pea70, BJSR10]: the loss is independent of the measurement settings, so the ensemble of detected systems provides a fair statistical sample of the total ensemble. In other words, we must exclude the possibility of a 'demon' in the measuring device conspiring against us in choosing which bits to reject.

The strength of the proposed QRNG relies crucially on value indefiniteness, so without this fair sampling assumption we would forfeit the assurance of bitwise incomputability of the generated sequence. As an example, let us consider the extreme case that the detection efficiency is less that 50%; our supposed demon could reject all bits detected as 0 and be within the bounds given by this efficiency, while the produced sequence would be computable. In the more general case, for any efficiency $\rho < 1$ the demon could reject bits to ensure every $(1/(1-\rho))$'th bit is a zero. This would introduce an infinite computable subsequence, a property violating the strong incomputability of the output bit-string produced by our QRNG, and still be consistent with the detection efficiency.

Note that this condition is stronger than the fair sampling assumption required in tests for violation of Bell-type inequalities because, without this assumption, *any* inefficiency can lead to a loss of randomness.

3.4 Better-than-classical operationalisation of spatial orthogonality

As has already been pointed out, for no temporal offset and in the regime of relative spatial angles around $\pi/4$—i.e., at almost half orthogonal measurement directions—the classical linear expectation function $E^{\mathrm{cl}}_{\oplus}(\theta) = 1 - 2\theta/\pi$, for $0 < \theta < \pi/4$ is strictly *smaller*, and for $\pi/4 < \theta < \pi/2$ is strictly *greater* than the quantum expectation function $E_{\oplus}(\theta) = (1/2)(1 + \cos 2\theta)$. This can be demonstrated by rewriting $\theta = \pi/4 \pm \Delta\theta$, and by considering a Taylor series expansion around $\pi/4$ for small $\Delta\theta \ll 1$, which yields $E_{\oplus}(\pi/4 \pm \Delta\theta) \approx (1/2) \mp \Delta\theta$, whereas $E^{\mathrm{cl}}_{\oplus}(\pi/4 \pm \Delta\theta) = (1/2) \mp (2/\pi)\Delta\theta$ (see Figure 3.2).

Phenomenologically this indicates less-than-classical numbers of equal pairs of outcomes '0–0' as well as '1–1', and more-than-classical non-equal pairs of outcomes '0–1' as well as '1–0', respectively, for the quantum case in the region $0 < \theta < \pi/4$; as well as the reverse behaviour in the region $\pi/4 < \theta < \pi/2$. This in turn results in 'less zeroes' and 'more ones' of the resulting sequence obtained by XORing the pairs of outcomes

in the region $0 < \theta < \pi/4$, as well as in 'more zeroes' and 'less ones' in the region $\pi/4 < \theta < \pi/2$ as compared to classical non-entangled systems [Per78]. Hence, with increasing aberration from misalignment $\Delta\theta$ the quantum device 'drifts off' into biasedness of the output 'faster' than any classical device. As a result, Borel normality is expected to be broken more strongly and quickly quantum mechanically than classically.

This effect could in principle be used to operationalise spatial orthogonality through the fine-tuning of angular directions yielding Borel normality. In the resulting protocols, quantum mechanics outperforms any classical scheme due to the differences in the correlation functions.

3.5 Theoretical analysis on generated bit-strings

Here we analyse the output distribution of the proposed QRNG and the ability to extract uniformly distributed bits from the two generated bit-strings in the presence of experimental imperfections.

3.5.1 Probability space construction

With reference to Figure 3.1 for the setup, we write the generated Bell singlet state with respect the top ('$\oplus$') measurement context (this is arbitrary as the singlet is form invariant in all measurement directions) as $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. The lower ('$\otimes$') polariser is at an angle of θ to the top one. After beam splitters we have the state

$$\frac{1}{\sqrt{2}}\left[\cos\theta(|00\rangle - |11\rangle) - \sin\theta(|01\rangle + |10\rangle)\right],$$

so we measure the same outcome in both contexts with probability $\cos^2\theta$ and different outcomes with probability $\sin^2\theta$.

More formally, the QRNG generates two strings simultaneously, so the probability space contains pairs of strings of length n. Let $e_x^{\oplus}, e_y^{\otimes}$ for $x, y = 0, 1$ be the detector efficiencies of the $D_x^{\oplus}$ and $D_y^{\otimes}$ detectors respectively. For perfect detectors, i.e $e_x^{\oplus} = e_y^{\otimes}$, we would expect a pair of bits (a, b) to be measured with probability $2^{-1}(\sin^2\theta)^{a\oplus b}(\cos^2\theta)^{1-a\oplus b}$; non-perfect detectors alter this probability depending on the values of a, b.

Fact 54. *For any $k > 0$, $0 \leq \theta < 2\pi$ and $x \in B^k$ we have*

$$\sum_{y\in B^k} (\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{k-d(x,y)} = 1.$$

Proof. We have

$$\begin{aligned}\sum_{y\in B^k}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{k-d(x,y)} &= \sum_{i=0}^{k}\binom{k}{i}(\sin^2\theta)^i(\cos^2\theta)^{k-i}\\ &= (\cos^2\theta+\sin^2\theta)^k\\ &= 1.\end{aligned}$$

□

Proposition 55. *The probability space of bit-strings produced by the QRNG is* $(B^n\times B^n, 2^{B^n\times B^n}, P_{n^2})$*, where the probability* $P_{n^2}: 2^{B^n\times B^n}\to[0,1]$ *is defined for all* $X\subseteq B^n\times B^n$ *as follows:*

$$P_{n^2}(X)=\frac{1}{Z_n}\sum_{(x,y)\in X}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}(e_0^{\oplus})^{\#_0(x)}(e_1^{\oplus})^{\#_1(x)}(e_0^{\otimes})^{\#_0(y)}(e_1^{\otimes})^{\#_1(y)},$$

where the term

$$\begin{aligned}Z_n &= \sum_{(x,y)\in B^n\times B^n}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}(e_0^{\oplus})^{\#_0(x)}(e_1^{\oplus})^{\#_1(x)}(e_0^{\otimes})^{\#_0(y)}(e_1^{\otimes})^{\#_1(y)}\\ &= \left[(\sin^2\theta(e_0^{\oplus}e_1^{\otimes}+e_1^{\oplus}e_0^{\otimes})+\cos^2\theta(e_0^{\oplus}e_0^{\otimes}+e_1^{\oplus}e_1^{\otimes})\right]^n\end{aligned}$$

ensures normalisation.

Proof. We can check easily that this satisfies the Kolmogorov axioms.

1. $P_{n^2}(\emptyset)=0$, trivially true;
2. $P_{n^2}(B^n\times B^n)=1$ by by definition of Z_n;
3. For $X,Y\subseteq B^n\times B^n$, $X\cap Y=\emptyset \implies P_{n^2}(X\cup Y)=P_{n^2}(X)+P_{n^2}(Y)$. □

Note that for equal detector efficiencies we have

$$Z_n=(e^{\oplus})^n(e^{\otimes})^n\sum_{(x,y)\in B^n\times B^n}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}=2^n(e^{\oplus})^n(e^{\otimes})^n,$$

and hence the probability has the simplified form

$$P_{n^2}(X)=\sum_{(x,y)\in X}2^{-n}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}.$$

Given that the proposed QRNG produces two (potentially correlated) strings, it is worth considering the distribution of each string taken separately. Given the rotational invariance of the singlet state this should be uniformly distributed. However, because the detector efficiencies may vary in each detector, this is not, in general, the case.

Proposition 56. *For every bit-string $x \in B^n$ we have*

$$P_{n^2}(\{x\} \times B^n) = \frac{1}{Z_n}\left(e_0^{\oplus}(e_1^{\otimes}\sin^2\theta + e_0^{\otimes}\cos^2\theta)\right)^{\#_0(x)}\left(e_1^{\oplus}(e_0^{\otimes}\sin^2\theta + e_1^{\otimes}\cos^2\theta)\right)^{\#_1(x)}.$$

Proof. We have

$$\begin{aligned}
P_{n^2}(\{x\} \times B^n) &= \frac{1}{Z_n}\sum_{y\in B^n}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}(e_0^{\oplus})^{\#_0(x)}(e_1^{\oplus})^{\#_1(x)}(e_0^{\otimes})^{\#_0(y)}(e_1^{\otimes})^{\#_1(y)} \\
&= \frac{(e_0^{\oplus})^{\#_0(x)}(e_1^{\oplus})^{\#_1(x)}}{Z_n}\sum_{y\in B^n}(\sin^2\theta)^{d(x,y)}(\cos^2\theta)^{n-d(x,y)}(e_0^{\otimes})^{\#_0(y)}(e_1^{\otimes})^{\#_1(y)} \\
&= \frac{1}{Z_n}\left(e_0^{\oplus}(e_1^{\otimes}\sin^2\theta + e_0^{\otimes}\cos^2\theta)\right)^{\#_0(x)}\left(e_1^{\oplus}(e_0^{\otimes}\sin^2\theta + e_1^{\otimes}\cos^2\theta)\right)^{\#_1(x)}. \quad \square
\end{aligned}$$

We see that each bit-string taken separately appears to come from a constantly biased source where the probabilities that a bit is 0 or 1, p_0, p_1, are given by the formulae

$$p_0 = e_0^{\oplus}(e_1^{\otimes}\sin^2\theta + e_0^{\otimes}\cos^2\theta)/Z_1,\ p_1 = e_1^{\oplus}(e_0^{\otimes}\sin^2\theta + e_1^{\otimes}\cos^2\theta)/Z_1.$$

This can alternatively be viewed as the distribution obtained if we were to discard one bit-string after measurement. Note that if either $e_0^{\otimes} = e_1^{\otimes}$ or we have perfect misalignment (i.e. $\theta = \pi/4$) then the probabilities have the simpler formulae:

$$p_x = e_x^{\oplus}/(e_0^{\oplus} + e_1^{\oplus}),\ x \in B.$$

In this case, if we further have that $e_0^{\oplus} = e_1^{\oplus}$, we obtain the uniform distribution by discarding one string after measurement.

The analogous result for the symmetrical case $P_{n^2}(B^n \times \{y\})$ also holds.

3.5.2 Independence of the QRNG probability space

If we were to discard one bit-string it is clear the other bit-string is generated independently in a statistical sense since the probability distribution source producing it is constantly biased and independent. However, we would like to extend our notion of independence defined in Definition 20 to this 2-bit-string probability space.

Definition 57. We say the probability space $(B^n \times B^n, 2^{B^n \times B^n}, R_{n^2})$ is *independent* if for all $1 \le k \le n$ and $x_1, \ldots, x_k,\ y_1, \ldots, y_k \in B$ we have

$$\begin{aligned}
R_{n^2}(x_1 \ldots x_k B^{n-k} \times y_1 \ldots y_k B^{n-k}) = {} & R_{n^2}(x_1 \ldots x_{k-1} B^{n-k+1} \times y_1 \ldots y_{k-1} B^{n-k+1}) \\
& \times R_{n^2}(B^{k-1} x_k B^{n-k} \times B^{k-1} y_k B^{n-k}).
\end{aligned}$$

Lemma 58. *For all $x, y \in B^{|x|}$ and $0 \le k + |x| \le n$ we have*

$$P_{n^2}(B^{n-k} x B^{n-k-|x|} \times B^{n-k} y B^{n-k-|x|}) = P_{|x|^2}((x, y)).$$

Proof. Indeed, using the additivity of the Hamming distance and the $\#_i$ functions, e.g. $d(x_1 \dots x_k, y_1 \dots y_k) = d(x_1 \dots x_{k-1}, y_1 \dots y_{k-1}) + d(x_k, y_k)$, we have:

$$\begin{aligned} P_{n^2}(B^{n-k}xB^{n-k-|x|} \times B^{n-k}yB^{n-k-|x|}) = & \sum_{a_1,a_2 \in B^{n-k}} \sum_{b_1,b_2 \in B^{n-k-|x|}} P_{n^2}\left((a_1xb_1, a_2yb_2)\right) \\ =& P_{|x|^2}((x,y)) \\ & \times \sum_{a_1,a_2 \in B^{n-k}} \sum_{b_1,b_2 \in B^{n-k-|x|}} P_{(n-|x|)^2}\left((a_1b_1, a_2b_2)\right) \\ =& P_{|x|^2}((x,y)) P_{(n-|x|)^2}(B^{n-|x|} \times B^{n-|x|}) \\ =& P_{|x|^2}((x,y)). \end{aligned}$$ □

In light of Lemma 58 the following Fact is evident.

Fact 59. *As a direct consequence we deduce that the probability space P_{n^2} defined in Proposition 55 is independent.*

Proof. From Lemma 58 and additivity of the Hamming distance and $\#_i$ functions, we have

$$\begin{aligned} P_{n^2}(x_1 \dots x_k B^{n-k} \times y_1 \dots y_k B^{n-k}) =& P_{k^2}((x_1 \dots x_k, y_1 \dots y_k)) \\ =& P_{(k-1)^2}((x_1 \dots x_{k-1}, y_1 \dots y_{k-1})) \times P_{1^2}((x_1, y_1)) \\ =& P_{n^2}(x_1 \dots x_{k-1} B^{n-k+1} \times y_1 \dots y_{k-1} B^{n-k+1}) \\ & \times P_{n^2}(B^{k-1}x_k B^{n-k} \times B^{k-1}y_k B^{n-k}). \end{aligned}$$ □

3.5.3 XOR application

We now consider the situation where the two output bit-strings x and y are XOR'd against each other (effectively using one as a one-time pad for the other) to produce a single bit-string, and we investigate the distribution of the resulting bit-string. Rather than only considering the effect of XORing paired (and potentially correlated) bits, we also consider XORing outcomes shifted by $j > 0$ bits as described in Section 3.3.3.

For $j \geq 0$ and $x, y \in B^{n+j}$ define the offset-XOR function $X_j : B^{n+j} \times B^{n+j} \to B^n$ as $X_j(x,y) = z$ where $z_i = x_i \oplus y_{i+j}$ for $i = 1, \dots, n$. For $z \in B^n$ the set of pairs (x,y) which produce z when XOR'd with offset j is

$$A_j(z) = \{(x,y) \mid x, y \in B^{n+j}, X_j(x,y) = z\} = \{(ua, b(u \oplus z) \mid u \in B^n, a, b \in B^j\}.$$

Proposition 60. *The probability space of the output produced by the QRNG is $(B^n, 2^{B^n}, Q_{n,j})$, where $Q_{n,j} : 2^{B^n} \to [0,1]$ is defined for all $X \subseteq B^n$ as:*

$$Q_{n,j}(X) = \sum_{z \in X} P_{(n+j)^2}(A_j(z)). \tag{3.3}$$

Proof. We check that the Kolmogorov axioms are satisfied. Note that $|A_j(z)| = 2^{n+2j}$.

1. $Q_{n,j}(\emptyset) = 0$, trivially true;

2. Since all $A_j(z)$ are disjoint we have $|\bigcup_z A_j(z)| = 2^n 2^{n+2j} = (2^{n+j})^2$, so $\bigcup_z A_j(z) = B^{n+j} \times B^{n+j}$. Hence,

$$\begin{aligned} Q_{n,j}(B^n) &= \sum_{z \in B^n} P_{(n+j)^2}(A_j(z)) = P_{(n+j)^2}\left(\bigcup_z A_j(z)\right) \\ &= P_{(n+j)^2}\left(B^{n+j} \times B^{n+j}\right) = 1; \end{aligned}$$

3. For disjoint $X, Y \subseteq B^n$ we have $Q_{n,j}(X \cup Y) = Q_{n,j}(X) + Q_{n,j}(Y)$. □

We now explore the form of the XOR'd distribution $Q_{n,j}$ for $j = 0$ and $j > 0$. Let $z \in B^n$ and $j \geq 0$. By $z[m,k]$ we denote the substring $z_m \dots z_k, 1 \leq m \leq k \leq n$. We have

$$\begin{aligned} Q_{n,j}(z) =& P_{(n+j)^2}(A_j(z))) \\ &= \sum_{a,b \in B^j} \sum_{u \in B^n} P_{(n+j)^2}((ua, b(u \oplus z)) \\ &= \sum_{u \in B^n} P_{(n-j)^2}\left((u[j+1,n], (u \oplus z)[1,n-j])\right) \\ &\quad \cdot \sum_{a \in B^j} P_{j^2}\left((a, (u \oplus z)[n-j+1,n])\right) \sum_{b \in B^j} P_{j^2}\left((u[1,j], b)\right). \end{aligned} \tag{3.4}$$

Theorem 61. *For $j = 0$ the XOR'd distribution acts as a constantly biased source with*

$$Q_{n,0}(z) = \frac{1}{Z_n}\left(\sin^2\theta(e_0^{\oplus}e_1^{\otimes} + e_1^{\oplus}e_0^{\otimes})\right)^{\#_1(z)}\left(\cos^2\theta(e_0^{\oplus}e_0^{\otimes} + e_1^{\oplus}e_1^{\otimes})\right)^{\#_0(z)}.$$

Proof. We note that for $j = 0$ we have $d(u, u \oplus z) = \#_1(z)$, and thus

$$\begin{aligned} Q_{n,0}(z) &= \sum_{u \in B^n} P_{n^2}\left((u, (u \oplus z))\right) \\ &= \frac{1}{Z_n}(\sin^2\theta)^{\#_1(z)}(\cos^2\theta)^{\#_0(z)} \sum_{u \in B^n} (e_0^{\oplus})^{\#_0(u)}(e_1^{\oplus})^{\#_1(u)}(e_0^{\otimes})^{\#_0(u \oplus z)}(e_1^{\otimes})^{\#_1(u \oplus z)} \\ &= \frac{1}{Z_n}\left(\sin^2\theta(e_0^{\oplus}e_1^{\otimes} + e_1^{\oplus}e_0^{\otimes})\right)^{\#_1(z)}\left(\cos^2\theta(e_0^{\oplus}e_0^{\otimes} + e_1^{\oplus}e_1^{\otimes})\right)^{\#_0(z)}. \end{aligned}$$ □

We recognise this as a constantly biased source where

$$p_0 = \cos^2\theta(e_0^{\oplus}e_0^{\otimes} + e_1^{\oplus}e_1^{\otimes})/Z_1,\ p_1 = \sin^2\theta(e_0^{\oplus}e_1^{\otimes} + e_1^{\oplus}e_0^{\otimes})/Z_1.$$

It is interesting to compare the form of $Q_{n,0}$ to the distribution of the constantly biased source in Proposition 56 obtained by discarding one output string—the former is more

sensitive to misalignment, the latter to differences in detection efficiencies. In the case of perfect/equal detector efficiencies (but non-perfect misalignment), discarding one string produces uniformly distributed bit-strings, whereas XORing does not.

We now look at the case where $j > 0$.

Theorem 62. *For $j > 0$, the distribution $Q_{n,j}$ takes the following form.*

1. *In the case of $\theta = \pi/4$ we have $Q_{n,j} = Q_{n,0}$,*
2. *In the case of equal detector efficiencies, i.e $e_0^{\oplus} = e_1^{\oplus}$ and $e_0^{\otimes} = e_1^{\otimes}$, we have $Q_{n,j} = U_n$.*

Proof. We prove each situation separately.

1. Working from (3.4) and Proposition 55, we see that (note that $\sin^2(\pi/4) = \cos^2(\pi/4) = 1/2$)

$$\sum_{a\in B} P_{1^2}\left((a, u_n \oplus z_n)\right) \sum_{b\in B} P_{1^2}\left((u_1, b)\right) = \frac{2^{-2}}{Z_2} e^{\otimes}_{u_n\oplus z_n} e^{\oplus}_{u_1} (e_0^{\oplus} + e_1^{\oplus})(e_0^{\otimes} + e_1^{\otimes})$$
$$= e^{\otimes}_{u_n\oplus z_n} e^{\oplus}_{u_1} \cdot \frac{2^{-1}}{Z_1}.$$

   It is then easy to see that

$$\begin{aligned} Q_{n,1}(z) &= \frac{2^{-(n-1)}}{Z_{n-1}} \left(e_0^{\oplus} e_1^{\otimes} + e_1^{\oplus} e_0^{\otimes}\right)^{\#_1(z)} \left(e_0^{\oplus} e_0^{\otimes} + e_1^{\oplus} e_1^{\otimes}\right)^{\#_0(z)} \cdot \frac{2^{-1}}{Z_1} \\ &= \frac{2^{-n}}{Z_n} \left(e_0^{\oplus} e_1^{\otimes} + e_1^{\oplus} e_0^{\otimes}\right)^{\#_1(z)} \left(e_0^{\oplus} e_0^{\otimes} + e_1^{\oplus} e_1^{\otimes}\right)^{\#_0(z)} \\ &= Q_{n,0}(z). \end{aligned}$$

2. Note that we have $Z_{n+j} = 2^{n+j}$ for this case. For any $z \in B^n$ we have

$$\begin{aligned} Q_{n,j}(z) =& 2^{-n-j} \sum_{u_n\in B} \cdots \sum_{u_{n-j}\in B} (\sin^2\theta)^{u_n\oplus z_{n-j}\oplus u_{n-j}} (\cos^2\theta)^{1-u_n\oplus z_{n-j}\oplus u_{n-j}} \cdots \\ & \times \sum_{u_1\in B} (\sin^2\theta)^{u_j+1\oplus z_1\oplus u_1} (\cos^2\theta)^{1-u_j+1\oplus z_1\oplus u_1} \\ =& 2^{-n-j} \sum_{u_n\in B} \cdots \sum_{u_{n-j}\in B} (\sin^2\theta + \cos^2\theta) \cdot \sum_{u_1\in B} (\sin^2\theta + \cos^2\theta) \\ =& 2^{-n-j} \sum_{u_{n-j+1}\ldots u_n \in B^j} 1 \\ =& 2^{-n-j} 2^j \\ =& 2^{-n} \\ =& U_n(z). \end{aligned}$$

□

x	bin(174)	bin(487)	bin(973)
$Q_{10,0}(x)$	5.90×10^{-4}	9.70×10^{-4}	1.64×10^{-4}
$Q_{10,1}(x)$	9.75×10^{-4}	9.71×10^{-4}	9.71×10^{-4}
$Q_{10,2}(x)$	9.78×10^{-4}	9.70×10^{-4}	9.70×10^{-4}
$U_{10}(x)$	9.77×10^{-4}	9.77×10^{-4}	9.77×10^{-4}

Table 3.2: Empirical evidence for the quality of XORing with $j > 0$ compared to $j = 0$ and configuration settings of $\theta = \pi/5$, $e_0^{\oplus} = 0.30$, $e_1^{\oplus} = 0.33$, $e_0^{\otimes} = 0.29$, $e_1^{\otimes} = 0.30$—this is probably much worse (further from the ideal case) than one would expect in an experimental setup. The (small) value of $n = 10$ has been used as, unfortunately, the distribution is very costly to calculate numerically. Here bin(m) denotes the (10-bit zero-extended) binary representation of m. For example, bin(1) = 0000000001, bin(2) = 0000000010, etc.

$\Delta(Q_{10,0}, U_{10})$	0.770271
$\Delta(Q_{10,1}, U_{10})$	0.00441399
$\Delta(Q_{10,1}, U_{10})$	0.00440061

Table 3.3: The variation from the uniform distribution of the distributions $Q_{10,j}$, using the same parameters as Table 3.2.

Of course, the third and most general case of non-equal detector efficiencies and non-ideal context misalignment. In this situation, the distribution is no longer independent, although in general is much closer to the uniform distribution than the $j = 0$ case. It is indeed this 'closeness'—formalised as the total variation distance—which is the important quantity. However, since $Q_{n,j}$ for $j > 0$ is not independent, von Neumann normalisation cannot be applied to guarantee the uniform distribution; indeed the dependence is not even bounded to a fixed number of preceding bits. As a result, there is unfortunately no simple form for $Q_{n,j}$ in this general case.

3.5.4 Criticisms and alternative operationalisations

This given, one may ask why not simply discard one string to give the distribution in Proposition 56 and apply von Neumann normalisation to obtain uniformly distributed bit-strings? There are two primary answers to this question.

Firstly, as discussed previously, the effect of drift in bias and temporal correlations will ensure this method will not produce the uniform distribution anyway. Indeed, the distribution $Q_{n,j}$ for $j > 0$ should be more robust to those effects ($Q_{n,j}$ is, for example, less sensitive to detector bias than the distribution in Proposition 56). It is extremely plausible that $Q_{n,j}$ gives as good results as discarding one string in practice; it is indeed

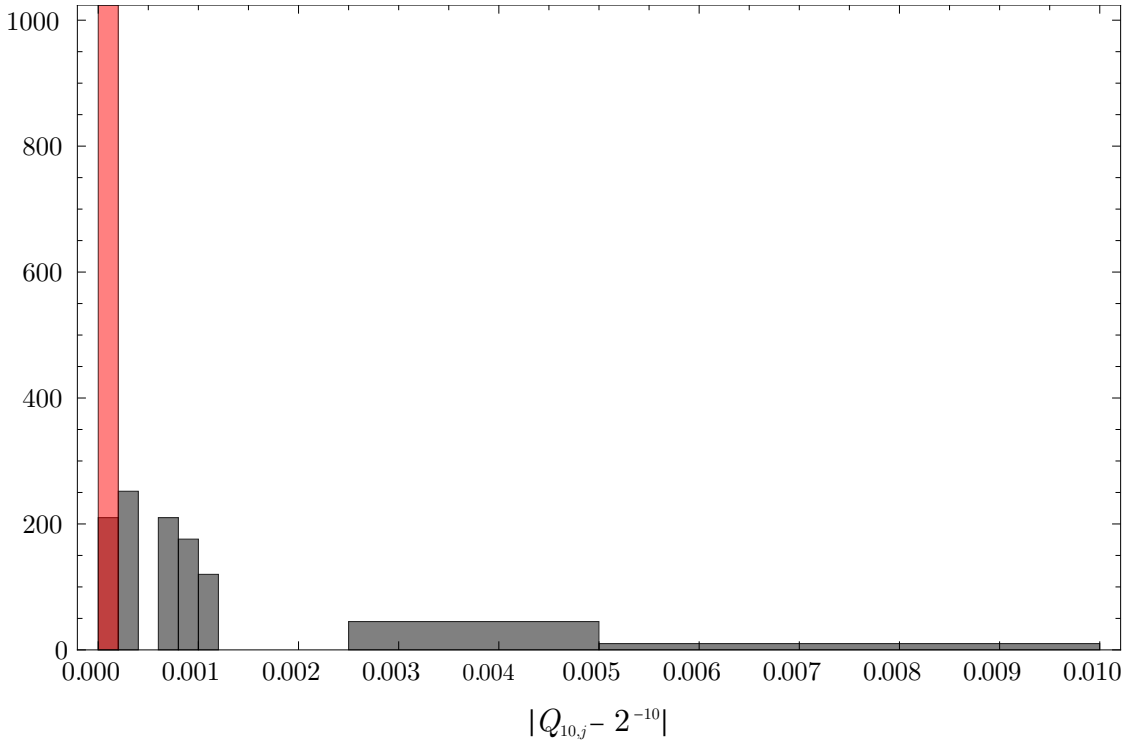

Figure 3.3: (Colour Online) A histogram of $|Q_{10,j} - 2^{-10}|$ for $j = 0$ (black) and $j = 1$ (red) showing the marked advantage of XORing with $j = 1$. The $j = 1$ case is, however, not quite the uniform distribution as we can see from Figure 3.4.

very close to the uniform distribution as can be seen from Table 3.3 and Figures 3.3 and 3.4. To compare properly the distributions, the following *open question* must be answered:

Open Question 63. *What is the bound ρ (depending on $e_x^{\oplus}, e_y^{\otimes}$ and θ) such that $\Delta(U_n, Q_{n,j}) \leq \rho$? How does it compare to that given Theorem 41 for normalisation of a source with varying bias?*

Further, $Q_{n,j}$ produces bit-strings of length n, whereas applying von Neumann to a single string produces a string with expected length at most $n/4$ bits. This is a significant increase in efficiency, making the shifted XORing process extremely appealing for a high bit-rate, un-normalised QRNG. Even the $j = 0$ case with von Neumann applied after XORing would often be preferable to discarding one string, since it is less sensitive to detector efficiency (the hardware limit) and more sensitive to misalignment (which is controlled by the experimenter).

Secondly, if one insists on a perfect theoretical distribution in the presence of non-ideal misalignment and unequal detector efficiencies, or perhaps the $Q_{n,j}$ distribution is not sufficient for particular requirements, then one can still operationalise both strings

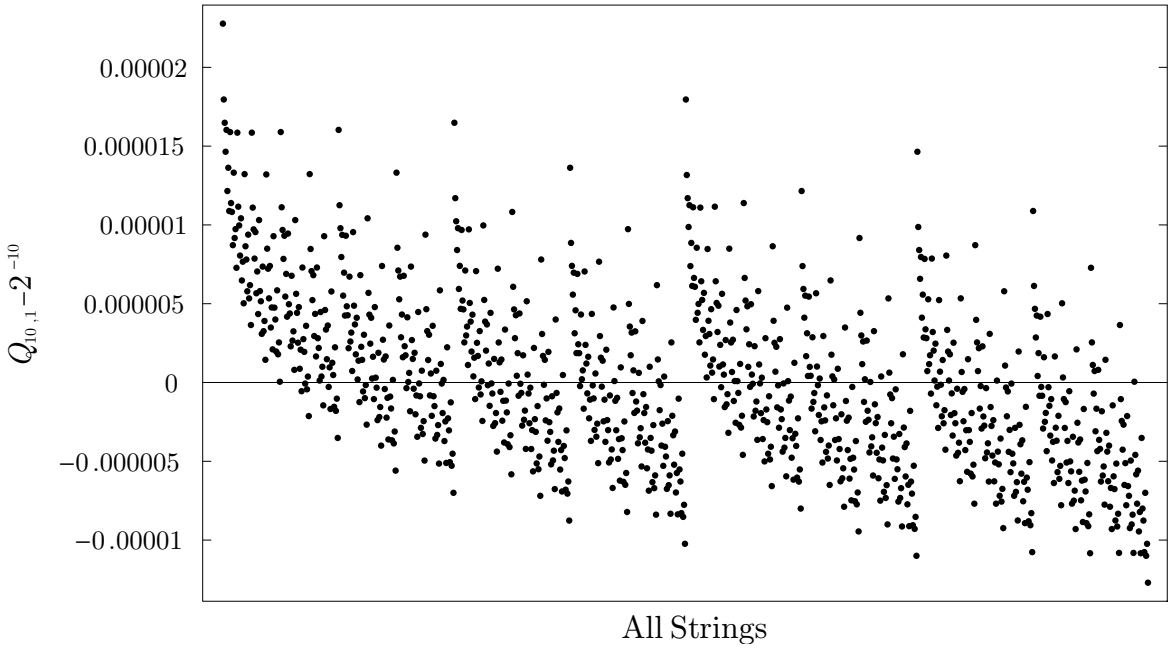

Figure 3.4: A plot of $Q_{10,1} - 2^{-10}$ for each of the 2^{10} strings of length 10 showing that the probabilities are very close (but not equal) to what would be expected from the uniform distribution. The same experimental configuration as in Table 3.3 has been used.

to improve the efficiency of the QRNG over discarding a single string by a simple modification of von Neumann's procedure. To do so, note that the pair of pair (a_1a_2, b_1b_2) has the same probability as the pair (a_2a_1, b_2b_1). By mapping those with $a_1b_1 < a_2b_2$ (lexicographically) to 0, those with $a_1b_1 > a_2b_2$ to 1, and discarding those with $a_1b_1 = a_2b_2$, one will obtain the uniform distribution as for von Neumann's procedure. The key advantage is that this will obtain strings of expected length up to $3n/8$, while maintaining the desired property of sampling from the uniform distribution.

The problem of determining how best to obtain the maximum amount of information from the QRNG is largely a problem of randomness extractors [Gab11], and is a trade off between the number of uniformly distributed bits obtained and the processing cost—a suitable extractor needs to operate in real-time for most purposes. As we have seen, the fact that two (potentially correlated) bit-strings are obtained allows more efficient operation than a QRNG using single-photons. We have shown how the proposed QRNG can be operationalised in more than one way: either by using shifted XORing of bits to sample from a distribution which is close to (equal to in the ideal limit) the uniform distribution and efficient and robust to various errors, or by utilising both produced bit-strings to allow a more efficient normalisation procedure giving (in absence of the aforementioned temporal effects) the uniform distribution. Many more operationalisations are undoubtedly possible.

3.6 Summary

In this chapter we have proposed a QRNG which, by utilising an entangled photon singlet-state in four-dimensional Hilbert space, is certified by value indefiniteness which implies strong incomputability, the mathematical property corresponding to physical indeterminism. While this is an ingredient of fundamental importance in any reasonable QRNG, we have recognised that experimental imperfections will always prevent the QRNG from producing exactly the theoretical uniform probability distribution, another essential symptom of randomness (independent of incomputability). The form and effects of these conceivable experimental errors have been discussed, and care has been taken to make the proposed QRNG robust to these effects.

Since this QRNG produces two bit-strings, we have proposed XORing the bit-strings produced, using one as a one-time pad for the other, to obtain better protection against experimental imperfections. We further propose XORing bits from temporally offset measurements in order to mitigate the effects of temporal correlations between adjacent bits. We leave it as an *open question* to improve upon the time-shifted XOR method and find a technique to extract bits which are provably uniformly distributed and is more efficient than the improved von Neumann method we also discussed.

Analyses of sequences generated by the proposed QRNG should be conducted, utilising the knowledge of the expected uniform distribution, as in Section 3.1.2 and [CDDS10]. In particular, the quality of both the individual strings produced should be compared with that of the XOR'd sequence, both with and without von Neumann normalisation applied, as well as the sequence produced by our improved von Neumann method.

Further, in view of conceivable temporal correlations between bits, the quality of the random bits should be tested as j is varied in (3.3). Since this has little effect on the bias of the resulting string (and normalisation can subsequently remove this), it would allow investigation of the effect and significance of these conceivable temporal correlations.

Chapter 4

Conclusions

In this thesis we have provided a solid theoretical framework for the purposes of analysing the quality of QRNGs. The potential offered by QRNGs is great, but current work has focused on details of implementation while taking the quality of randomness provided by such devices for granted. We have summarised existing results on the quality of quantum randomness from the perspective of algorithmic information theory, and strengthened some of the results in the process. In particular, we show that no bit of a sequence produced by a QRNG is provably computable in ZFC. This gives, for the first time, a proper mathematical understanding to the notion of 'irreducible randomness' which is often taken for granted in quantum experiments.

In Chapter 2 we studied in detail the effect of von Neumann normalisation on both finite strings and infinite sequences of quantum random bits. Application of such normalisation techniques is unavoidable in experimental situations where bias is present, so it is important to understand the effect of this on the strength of randomness. We showed that while Borel normality and algorithmic randomness are preserved by normalisation, ε-randomness for $0 < \varepsilon < 1$ (and thus incomputability) is in general not preserved. We also analysed the effect of normalisation on a source in which the bias varies slowly, and derived bounds for how close the resulting distribution is to the uniform distribution.

In Chapter 3 we proposed a new QRNG which is explicitly certified by value indefiniteness, and thus provides the strong incomputability we have shown to be possible with QRNGs. Special emphasis was put on making the design robust to experimental imperfections in order to make it a good practical device. We discussed in detail the effect of various such imperfections on our QRNG, but the discussion is largely applicable to pre-existing QRNGs based on photons as well. We analysed the distribution produced by our proposed QRNG, and examined techniques to make use of the

experimental setup in order to make the resulting distribution closer to the uniform distribution than would otherwise be the case.

4.1 Future work

From a theoretical point of view, what we have presented is merely the tip of the iceberg, the setting of a framework. While we have shown that a sequence of quantum random bits is strongly incomputable, it is not known whether or not it is algorithmically random, or even Borel normal. Further work is needed to classify the level of randomness which can be certified by value indefiniteness. This is important since the question of what level of randomness we can generate with QRNGs is not only interesting from a theoretical and philosophical point of view, but from a practical point of view too. A QRNG producing incomputable bits is already breaking the Turing barrier, but it is natural to want to know what level of randomness can be produced by quantum mechanics. Furthermore, since incomputability is not, in general, preserved by normalisation, it needs to be determined if the sequences produced by a QRNG maintain their certification under normalisation. Without such a proof, it cannot be rigorously claimed that a QRNG using such techniques is as strong as one which does not.

Since the proof of the incomputability of quantum randomness relies on the physical assumption of value indefiniteness, reasonable as it may be, it is worth asking what is the case in the alternative scenario. The Kochen-Specker Theorem assures us we either have value indefiniteness *or* contextuality, so it is an open avenue of research to explore the strength of quantum randomness under the assumption of contextuality. Does all randomness disappear and leave us with computable sequences, or can we still guarantee incomputability? If the latter case were true, this would add significant weight to the claim of incomputability in nature.

To complement the theoretical work which is necessary, more experimental testing is needed. While one cannot have finite tests for randomness, it is important to test the physical assumptions of independence between successive bits. Experimental imperfections will always assure we never fully obtain uniformly distributed bits, so experimental work is required to test techniques which aim to get us as close as possible to this goal.

All in all, quantum random number generators offer huge potential as the next generation of random number generators. We are already at the stage where we can physically implement such devices, we are just in need of the theoretical foundations to allow us to make good the promise of the technology.

Bibliography

[AC10] A. A. Abbott and C. S. Calude. Von Neumann normalisation of a quantum random number generator. *CDMTCS Research Report*, 392, 2010. Submitted for publication to *Theoretical Computer Science C*, January 2011. [2]

[ACS] A. A. Abbott, C. S. Calude, and K. Svozil. Unpublished work on the incomputability of quantum randomness, in preparation. [11, 48]

[ACS10] A. A. Abbott, C. S. Calude, and K. Svozil. A quantum random number generator certified by value indefiniteness. *CDMTCS Research Report*, 396, 2010. [2]

[AJ06] J. Adell and P. Jodrá. Exact Kolmogorov and total variation distances between some familiar discrete distributions. *Journal of Inequalities and Applications*, (64307), 2006. [29]

[BC02] D. H. Bailey and R. E. Crandall. Random generators and normal numbers. *Experimental Mathematics*, 11(4):527–546, 2002. [6]

[Bel64] J. S. Bell. On the Eistein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964. [8]

[BJSR10] D. W. Berry, H. Jeong, M. Stobińska, and T. C. Ralph. Fair-sampling assumption is not necessary for testing local realism. *Physical Review A*, 81(1):012109, Jan 2010. [53]

[BLM96] P. Busch, P. J. Lahti, and P. Mittelstaedt. *The Quantum Theory of Measurement*. Lecture Notes in Physics: New Series m, Monographs; 2. Springer-Verlag, Berlin, Heidelberg, 2nd edition, 1996. [17]

[Blu86] M. Blum. Independent unbiased coin flips from a correlated biased source—a finite state Markov chain. *Combinatorica*, 6(2):97–108, 1986. [52]

[Bor26] M. Born. Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*, 38:803–837, 1926. English translation by J. A. Wheeler and W. H. Zurek, in *Quantum Theory and Measurement*, Chapter I.2. Princeton University Press, 1983. [8, 41]

[BPP00] H. Bechmann-Pasquinucci and A. Peres. Quantum cryptography with 3-state systems. *Physical Review Letters*, 85(15):3313–3316, 2000. [44]

[Cab94] A. Cabello. A simple proof of the Kochen-Specker Theorem. *European Journal of Physics*, 15(179–183), 1994. [9]

[Cab08] A. Cabello. Experimentally testable state-independent quantum contextuality. *Physical Review Letters*, 101(21):210401, 2008. [45]

[Cab10] A. Cabello. Memory cost of simulating quantum mechanics. In Hélia Guerra, editor, *Physics and Computation 2010. 3rd International Workshop. Luxor/Aswan, Egypt, August 30-September 6, Pre-proceedings*, pages 119–125, 2010. [45]

[Cal94] C. S. Calude. Borel normality and algorithmic randomness. In G. Rozenberg and A. Salomaa, editors, *Developments in Language Theory*, pages 113–129. World Scientific, Singapore, 1994. [4]

[Cal02] C. S. Calude. *Information and Randomness: An Algorithmic Perspective.* Springer-Verlag, Berlin, 2nd edition, 2002. [2, 3, 15, 34, 35, 39, 45]

[CDDS10] C. S. Calude, M. J. Dinneen, M. Dumitrescu, and K. Svozil. Experimental evidence of quantum randomness incomputability. *Physical Review A*, 82(022102), 2010. [41, 45, 46, 49, 63]

[CEGA96] A. Cabello, J. M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker Theorem: A proof with 18 vectors. *Physics Letters A*, 212:183–187, 1996. [10]

[Cha33] D. G. Champernowne. The construction of decimals normal in the scale of ten. *Journal of the London Mathematical Society*, 8:254–260, 1933. [6]

[Cha77] G. J. Chaitin. Algorithmic information theory. *IBM Journal of Research and Development*, 21:350–359, 496, 1977. Reprinted in G. J. Chaitin. *Information, Randomness and Incompleteness.* World Scientific, Singapore, 2nd edition, 1990. [1, 2, 4]

[CHJW01] C. S. Calude, P. Hertling, H. Jürgensen, and K. Weihrauch. Randomness on full shift spaces. *Chaos, Solutions & Fractals*, 12(3):491–503, 2001. [38]

[CHS10] C. S. Calude, N. J. Hay, and F. Stephan. Representation of left-computable ε-random reals. *Journal of Computer and System Sciences*, 2010. [5]

[CS78] J. F. Clauser and A. Shimony. Bell's Theorem: experimental tests and implications. *Reports on Progress in Physics*, 41:1881–1926, 1978. [9]

[CS08] C. S. Calude and K. Svozil. Quantum randomness and value indefiniteness. *Advanced Science Letters*, 1(165–168), 2008. [2, 9, 11, 44]

[CZ10] C. S. Calude and M. Zimand. Algorithmically independent sequences. *Information and Computation*, 208:292–308, 2010. [17]

[DH10] R. Downey and D. Hirschfeldt. *Algorithmic Randomness and Complexity.* Theory and Applications of Computability. Springer, 2010. [3]

[DHM07] P. Diaconis, S. Holmes, and R. Montgomery. Dynamical bias in the coin toss. *SIAM Review*, 49(2):211–235, 2007. [6]

[Eke91] A. K. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67:661–663, Aug 1991. [44]

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777, 1935. [8]

[Erb95] T. Erber. Testing the randomness of quantum mechanics: Nature's ultimate cryptogram? *Annals of the New York Academy of Sciences*, 755:748–756, 1995. [8]

[FLW92] A. M. Ferrenberg, D. P. Landau, and Y. J. Wong. Monte Carlo simulations: Hidden errors from "good" random number generators. *Physical Review Letters*, 69(23):3382–3384, 1992. [6]

[FSS+07] M. Fiorentino, C. Santori, S. M. Spillane, R. G. Beausoleil, and W. J. Munro. Secure self-calibrating quantum random-bit generator. *Physical Review A*, 75(3):032334, 2007. [44]

[Gab11] A. Gabizon. *Deterministic Extraction from Weak Random Sources*. Springer-Verlag Berlin Heidelberg, Berlin Heidelberg, 2011. [62]

[GC08] J. C. Garrison and R. Y. Chiao. *Quantum Optics*. Oxford University Press, Oxford, 2008. [9]

[GM87] A. Garg and D. N. Mermin. Detector inefficiencies in the Einstein-Podolsky-Rosen experiment. *Phys. Rev. D*, 35(12):3831–3835, 1987. [53]

[Haa10] M. Haahr. True random number generator. http://www.random.org, 2010. [7]

[Hal74] P. R. Halmos. *Measure Theory*. Springer-Verlag, New York, 1974. [2]

[Hel08] C. Held. The Kochen-Specker Theorem. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Winter 2008 edition, 2008. `http://plato.stanford.edu/archives/win2008/entries/kochen-specker/`. [9, 10]

[HLZ+03] Y.-F. Huang, C.-F. Li, Y.-S. Zhang, J.-W. Pan, and G.-C. Guo. Experimental test of the Kochen-Specker theorem with single photons. *Physical Review Letters*, 90(25):250401, Jun 2003. [45]

[Hoy07] M. Hoyrup. Dynamica systems: stability and simulability. *Mathematical Structures in Comptuer Science*, 17:247–259, 2007. [6]

[HQSMD+04] M. Hai-Qiang, W. Su-Mei, Z. Da, C. Jun-Tao, J. Ling-Ling, H. Yan-Xue, and W. Ling-An. A random number generator based on quantum entangled photon pairs. *Chinese Physics Letters*, 21(10):1961–1964, 2004. [44]

[HR83] P. Heywood and M. L. G. Redhead. Nonlocality and the Kochen-Specker paradox. *Foundations of Physics*, 13(5):481–499, 1983. [10]

[HT56] R. Hanbury Brown and R. Q. Twiss. A test of a new type of stellar interferometer on Sirius. *Nature*, 178:1046–1048, 1956. [21, 52]

[iQ09] id Quantique. Quantis — quantum random number generators. `http://idquantique.com/products/quantis.htm`, 12/08/2009. [16, 44, 45]

[Isa95] R. Isaac. *The Pleasures of Probability*. Springer-Verlag, New York, 1995. [6]

[Jau68] J. M. Jauch. *Foundations of Quantum Mechanics*. Addison-Wesley, Reading, MA, 1968. [21]

[JAW+00] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger. A fast and compact quantum random number generator. *Review of Scientific Instruments*, 71:1675–1680, 2000. [44]

[Kac59] M. Kac. *Statistical Independence in Probability, Analysis and Number Theory*. The Carus Mathematical Monographs. The Mathematical Association of America, 1959. [17]

[KCK09] O. Kwon, Y. Cho, and Y. Kim. Quantum random number generator using photon-number path entanglement. *Applied Optics*, 48(9):1774–1778, 2009. [12, 16]

[KN74] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. John Wiley & Sons, New York, 1974. [4]

[KS67] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17:59–87, 1967. Reprinted in E. Specker. *Selecta*. Brikhäuser Verlag, Basel, 1990. [9, 10]

[Lar98] J.-Å. Larsson. Bell's inequality and detector inefficiency. *Physical Review A*, 57(5):3304–3308, May 1998. [9, 53]

[LSW+05] F. Lindner, M. G. Schätzel, H. Walther, A. Baltuška, E. Goulielmakis, F. Krausz, D. B. Milošević, D. Bauer, W. Becker, and G. G. Paulus. Attosecond double-slit experiment. *Physical Review Letters*, 95(4):040401, Jul 2005. [52]

[Mer10] Z. Merali. A truth test for randomness. *Nature News*, April 2010. [7]

[ML66] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9(6):602–619, 1966. [4, 35, 40]

[MvV01] A. J. Menezes, P. C. van Oorschot, and S. A. Vantsone. *Handbook of Applied Cryptography*. CRC Press, 5th edition, 2001. [5]

[MWZ+04] H. Ma, S. Wang, D. Zhang, J. Change, L. Ji, Y. Hou, and L. Wu. A random-number generator based on quantum entangled photon pairs. *Chinese Physics Letters*, 21(19):1961–1964, 2004. [16]

[PAM+10] S. Pironia, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmchenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by Bell's Theorem. *Nature*, 464(09008), 2010. [12, 44, 48]

[PBD+00] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger. Experimental test of quantum nonlocality in three-photon Greenberger-Horne-Zeilinger entanglement. *Nature*, 403:515–519, 2000. [45]

[Pea70] P. M. Pearle. Hidden-variable example based upon data rejection. *Phys. Rev. D*, 2(8):1418–1425, Oct 1970. [53]

[Per78] A. Peres. Unperformed experiments have no results. *American Journal of Physics*, 46:745–747, 1978. [54]

[Per91] A. Peres. Two simple proofs of the Kochen-Specker Theorem. *Journal of Physics A: Mathematical and General*, 24:L175–L178, 1991. [10]

[Per92] Y. Peres. Iterating von Neumann's procedure for extracting random bits. *The Annals of Statistics*, 20(1):590–597, 1992. [16, 44, 51]

[Rab76] M. O. Rabin. Probabilistic algorithms. In J. F. Traub, editor, *Algorithms and Complexity, New Directions and Recent Results*, pages 21–39. Academic Press, New York, 1976. [5]

[RK07] R. Y. Rubinstein and D. P. Krose. *Simulation and the Monte Carlo Method*. John Wiley & Sons, New York, 2nd edition, 2007. [5]

[Ros06] J. S. Rosenthal. *A First Look at Rigorous Probability Theory*. World Scientific Publishing Co. Pte. Ltd., second edition, 2006. [5]

[ROT94] J. G. Rarity, M. P. C. Owens, and P. R. Tapster. Quantum random-number generation and key sharing. *Journal of Modern Optics*, 41:2435–2444, 1994. [44]

[Sak94] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, revised edition, 1994. [7]

[Sch70] H. Schmidt. Quantum-mechanical random-number generator. *Journal of Applied Physics*, 41(2):462–468, 1970. [12]

[SGG+00] A. Stefanov, N. Gisin, O. Guinnard, L. Guinnard, and H. Zbinden. Optical quantum random number generator. *Journal of Modern Optics*, 47(4):595–598, 2000. [12, 16, 17, 44, 52]

[Sol00] R. M. Solovay. A version of Ω for with ZFC can not predict a single bit. In C. S. Calude and G. Păun, editors, *Finite Versus Infinite. Contributions to an Eternal Dilemma*, pages 323–334. Springer-Verlag, London, 2000. [11]

[Sta01] P. Stanică. Good lower and upper bounds on binomial coefficients. *Journal of Inequalities in Pure and Applied Mathematics*, 2(3):30, 2001. [32]

[STZ10] Y. Shen, L. Tian, and H. Zou. Practical quantum random nubmer generator based on measuring the shot noise of vacuum states. *Physical Review A*, 81(063814), 2010. [12]

[Svo90] K. Svozil. The quantum coin toss — testing microphysical undecidability. *Physics Letters A*, 143(9):433–437, 1990. [12, 44]

[Svo98] K. Svozil. *Quantum Logic*. Springer-Verlag, Singapore Pte. Ltd., 1998. [7, 11]

[Svo04] K. Svozil. Quantum information via state partitions and the context translation principle. *Journal of Modern Optics*, 51:811–819, 2004. [41]

[Svo09] K. Svozil. Three criteria for quantum random-number generators based on beam splitters. *Physical Review A*, 79(5):054306, 2009. [12, 44, 47]

[Svo10] K. Svozil. Quantum value indefiniteness. *Natural Computing*, in press, 2010. [9, 10, 11, 44]

[Vad11] S. Vadhan. *Pseudorandomness*. Foundations and Trends in Theoretical Computer Science. now publishers, 2011. [5, 6, 16, 17, 33, 34]

[vN63] J. von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards Applied Math Series*, **12** (1951), 36–38. In A. H. Traub, editor, *John von Neumann, Collected Works*, pages 768–770. MacMillan, New York, 1963. [6, 19]

[WJS+98] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger. Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters*, 81:5039–5043, 1998. [10, 44]

[WLL06] P. X. Wang, G. L. Long, and Y. S. Li. Scheme for a quantum random number generator. *Journal of Applied Physics*, 100(5):056107, 2006. [44]

[Zei05] A. Zeilinger. The message of the quantum. *Nature*, 438:743, 2005. [8, 11]