THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 18, Special Issue 2017, pp. 389-402

TRUE RANDOM NUMBER SEQUENCES FROM GAMMA-DECAY USING FOUR EXTRACTION METHODS

Tudor PATULEANU[*], Kinga MARTON[**], Vasile CHIS[*], Sebastian TOMA[***]

[*] Babeş-Bolyai University, Faculty of Physics, Cluj-Napoca, Romania
[**] Technical University of Cluj-Napoca, Department of Computer Science, Romania
[***] "Horia Hulubei" National Institute for Physics and Nuclear Engineering, Department of Nuclear Physics, Romania
E-mail: patuleanut@gmail.com

Abstract. Randomness sources which continue to provide unpredictability even when all information about the internal states of the system is public knowledge are fundamental resources in many domains, especially cryptography. Quantum random number generators accomplish such task by displaying intrinsic randomness. However, they may be characterized by significantly reduced throughput compared to alternative sources of computational randomness. The central concept of this paper is radioactive decay as a quantum source of randomness. It is not a new concept, as other QRNGs based on alpha and beta decays have already been proposed and tested, instead it focuses on analysing the suitability of the gamma-decay phenomenon as the entropy source for generating high quality random number sequences. The employed extraction methods can easily be adapted to any event-generating stochastic phenomenon, as long as the assumptions regarding the source hold, or are taken into account by compensating for biases that might occur over time. The comparative analysis regarding the quality and throughput of the extraction methods is presented based on the results of both statistical and visual randomness assessment. Furthermore, the generated binary sequences are subject to several post-processing functions and the impact on the randomness quality is highlighted. Experimental results show a high randomness quality of the raw data altogether, the throughput being limited only by the activity of the source and the efficiency of the detector.

Key words: intrinsic randomness, quantum random number generation, radioactive gamma-decay, Poisson distribution, randomness assessment.

1. INTRODUCTION

Due to their wide-range applicability, random numbers have become an indispensable ingredient in providing many of the requirements of our modern technologized society, which heavily relies on information management. At the same time the choice of randomness sources is continually facing challenges regarding the evaluation and improvement of the randomness quality, the generation bitrate and even philosophical debates on the realism, causality and the character of nature, as thoroughly exemplified in [1, 2].

Particularly in the domain of cryptography, the need for unpredictability is critical, mainly in the generation of cryptographic keys. This necessity is fulfilled not by developing algorithms or algebraic methods, as these produce what is known as pseudo-randomness, but by observing high-entropy physical phenomena and using suitable extraction methods for producing random number sequences from the collected measurements.

Even in the case of physical observation of phenomena, the pseudorandom character does not disappear completely, as one can still predict macroscopic behaviours that are apparently random, by applying the deterministic laws of classical physics as an algorithm behind the generated dynamics. That is, classical mechanics leads only to an apparent character of randomness owned to the lack of information regarding the system involved [1, 2]. In this classical regime, the personal ignorance occurs mainly due to the high number of internal degrees of freedom, leading to *computational randomness*, where unpredictability can often be reduced by employing sufficient computing resources. And albeit

computational randomness may provide high statistical quality and throughput, there is a stringent need for *true randomness*, based on physical phenomena that can guarantee the random character even if all information regarding the entropic source is available to any unauthorized third-party in advance. Thereby, Kerckhoffs's Principle can be satisfied.

One major source of true randomness is quantum physics, and this paper focuses on the extraction and evaluation of random sequences from the quantum phenomena of radioactive γ-decay.

1.1. The suitability of Quantum Random Number Generators

Bell suggested in [4], a method for testing the completeness of quantum mechanics [5] by proposing an inequality satisfied by any scientific theory that assumes *locality* and *realism*.

The experimental results of the tests performed by Aspect [6] and of further refinements [7], which managed to close existing loopholes, showed a clear violation of Bell's inequality, which indicates that nature must be described by a theory that does not satisfy at least one or both the assumptions imposed by Bell, that is, it either allows signalling at-a-distance or hidden variables truly do not exist, the theory manifesting inherent randomness.

As quantum mechanics correctly predicts, although stochastically, the experimental results obtained by Aspect and, further, it does not allow for faster-than-light signalling [8], the second assumption involving the existence of supplementary variables that might be added in order to make the theory strictly deterministic, is violated. Other view is that realism holds, but the locality assumption is not satisfied. This signifies that hidden variables exist but are nonetheless unobservable, as their observation would allow for faster than light signalling, contradicting special relativity. In the latter case, the hidden variables have a purely computational mathematical value, so we must distinguish what we mean by "existence" as a physical quantity from the theory that can be observed experimentally, and the mathematical sense. This highlights the suitability of quantum theory for randomness extraction as it does not possess some underlying predictive mechanism we can make use of by observation.

A more productive, rather than interpretational, manner of treating a theory is to only demand to improve the theoretical predictive power. The question whether an extension of the theory can provide improved predictive power, instead of considering the removal of all indeterminism, has been raised and responded negatively by Colbeck and Renner [9].

According to our current understanding, quantum theory offers the basis for delivering theoretically-proven randomness due to its intrinsic nondeterministic character, hence, microscopic phenomena deliver genuine randomness.

1.2. Related work and personal contribution

Historically, radioactive decay-based quantum random number generators were among the first to be examined, starting around 1960s [3], and focusing mostly on the emission of electrons (β-decay) detected by Geiger-Muller counters. These types of QRNGs are still used in practice, one example being Walker's web-based random number server, HotBits [10].

Further improvements replace Geiger-Muller counters in favour of solid-state detectors such as semiconductor detectors like PIN photodiodes or high-purity germanium crystals. These are more suitable for the detection of single events, as they employ lower operating voltages, but they do require calibration before operation and possess comparative or greater dead times [3].

Alkassar, Nicolay & Rohe [11] present an electronic low-cost implementation of a physical random number generator (RNG) with PIN diode as detector, based on α-decay of ^{241}Am, isotope commonly found in household smoke detectors.

Although high quality randomness is obtained by radioactive decay of atomic nuclei, the drawbacks involve the relatively low bitrate output, the behavioural modification of the source over time, as it consumes and decomposes into daughter nuclides that can further be unstable, and also the long term damage inflicted to the detector by the incoming radiation [3]. A solution to these problems was the construction of high-speed optical QRNG's based on the two choice beam-splitting of single photons obtained from attenuated light, yielding one bit accordingly per scattered photon [12, 13, 14].

Our personal contribution, in the mentioned context, involves the study of the suitability of γ-decay phenomenon as the entropy source for generating random number sequences. The binary sequences obtained by employing four methods of extraction are subject to randomness assessment using the ENT utility program [15], a visual inspection tool [23], and the NIST Statistical Test Suite [16].

Furthermore, we provide a *comparative analysis* of the extraction methods regarding the quality of the produced randomness and the throughput of the generation method.

It is important to notice the general character of these methods that makes them suitable for application to any event generating phenomenon, as long as the assumptions regarding the source, such as that it does not degrade over time, hold, or are taken into account by compensating for biases that might occur this way.

This paper is organized as follows: **the second section** describes the counting statistics with emphasis on the stochastic process of radioactive decay and defines the extraction methods accordingly, based on the expressed probability distributions; **the third section** briefly presents the experimental setup; the **fourth section** highlights the results obtained by statistical testing the bit sequences of raw and post-processed data; finally, the **fifth section** outlines the conclusions and suggests directions for further research.

2. DESCRIPTION OF METHODS

In any event measurement of a stochastic process, we need to specify the uncertainties in the fluctuation of the measured data around the mean value. We achieve this by considering the repetition of measurements performed under the same starting conditions. The statistics of results obtained in the long term run of such repetitions allows us to assign a probability distribution for a single trial and define a random variable for that process, accordingly.

Another view for assigning a probability distribution to a random variable describing a physical process is by considering a large number of copies of the same stochastic system and performing a single measurement on all the ensemble at once and count the frequency of favorable occurrences.

In counting statistics, we are usually confronted with binary independent processes with two possible outcomes, either a success or a failure and wish to count the number of successes appearing in a given number of trials, as represented in Fig. 1. The random variable that describes this number of successes for a total number of trials follows the binomial distribution.

Fig. 1 – Binomial distribution for repeated trials (successes are represented by the blue dot).

For radioactive decay, the independence assumption holds and asserts that any particular nucleus of a species has the same probability p to decay at some specific time t, or more correctly in the next interval dt, not affected by the decay of all the other nuclei in the sample and the past history of the sample (Schweidler's assumption [17]). This decay probability can sometimes be computed by more fundamental theoretical considerations and, when not feasible, be assigned by fitting the long term statistics.

As radioactive decay is an entirely random process in the time domain, one can consider the measurement of the number of nuclei decaying over a certain time interval T when starting with N initial atoms. In this situation, one decaying nucleus in the considered time interval represents a success and the number of atoms are equivalent to the number of trials, following the second view for constructing a random variable mentioned before.

Another method used in counting statistics, preferred over the binomial distribution which is impractical to calculate with, is to consider instead the mean arrival rate of successful events over a specified time interval, where the mean is to be defined by the procedure explained above, of long run statistics of repeated trials, case referred to as a Poisson process.

Mathematically, this is expressed as a stochastic discrete process $\{N(t) \mid t \geq 0\}$, with mean rate λ, where $N(t)$ expresses the number of events that occurred up to time t, possessing the following properties:

i. *Time homogeneity*. The probability $P(k, \Delta t)$ of k arrivals is the same for all intervals of the same length Δt and is given by:

$$P(k,\Delta t) \coloneqq (P(N(t+\Delta t) - N(t) = k) = \frac{(\lambda \Delta t)^k \ * \ e^{-\lambda \Delta t}}{k!} \quad (1)$$

ii. *Independence*. The outcomes within any interval particular interval is independent of the history of arrivals outside this interval;

iii. *Small intervals probability*. The probability of more than one occurrence in a small time interval is negligible when compared to the probability of just one occurrence in the same time interval (see Fig. 2).

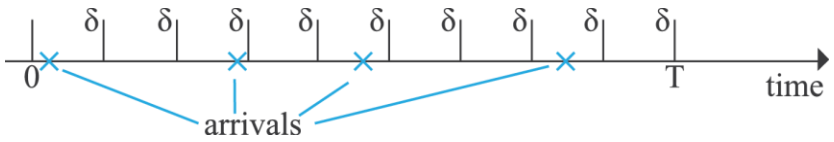


Fig. 2 – Time axis with event arrival for Poisson process.

One can derive the Poisson distribution as a limiting case of the binomial by imposing the conditions $\lambda t \ll 1, N \gg 1$ and the number of $x \ll N$ decays as specified in [17].

Counting of radioactive events can be mathematically described by both processes presented above, when making the assumption that the initial quantity of atoms changes insignificantly over the periods of time of interest. Naturally, over timescales compared to the half-life of the isotope, the activity becomes weaker as the number of atoms in the sample diminishes by exponentially decaying on average. In such a case, the Poisson mean rate incorporates the time dependence by also manifesting an exponentially decaying trend, yielding a nonhomogeneous Poisson process.

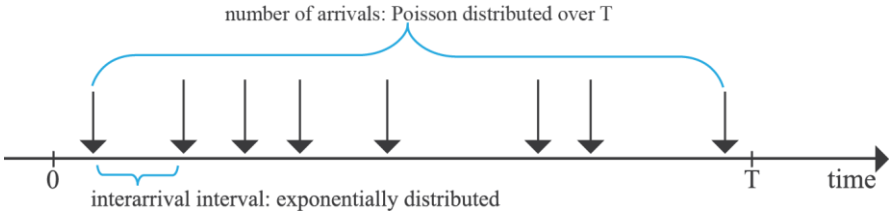


Fig. 3 – Poisson distribution and its associated inter-arrival exponential distribution.

Another important random variable associated with a homogeneous Poisson process is the inter-arrival time between events (shown in Fig. 3) and is described by the following exponential distribution:

$$P\big(next\ impulse\ occuring\ at\ time\ t\big) = \lambda \cdot e^{-\lambda t} \quad (2)$$

The following subsections describe in detail the extraction methods.

2.1. Generation by comparison of time intervals between three consecutive events (Interval comparison)

The inherent randomness in radioactive decay times implies that the inter-arrival times of consecutive decays can also be considered as a source of unpredictability. Walker, in [10], proposed the registration of the random inter-arrival times between two pairs of successive pulses. An adaptation of this method is to consider the time interval between the first two consecutive pulses, Δt_1, and compare it to the time interval between the second and third pulse, Δt_2.

The bits are extracted according to the following criterion, exemplified in Fig. 4:

- $\Delta t_1 > \Delta t_2$ yields a 1 bit;
- $\Delta t_1 < \Delta t_2$ yields a 0 bit;
- $\Delta t_1 = \Delta t_2$ neglect – this case occurs with negligible probability.

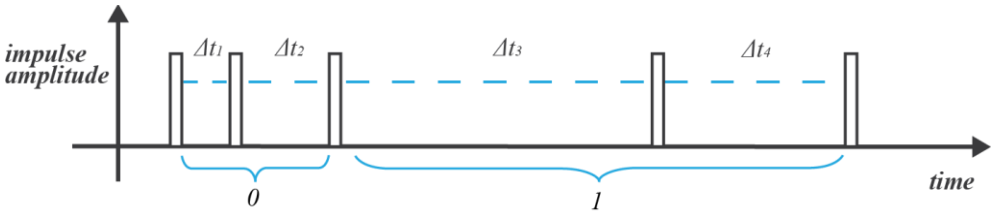


Fig. 4 – Bit extraction by comparison of successive inter-arrival times.

These rules can be reversed after several impulses to compensate for small systematic biases coming from the disintegration of the radioactive source that makes the second interval shorter on average by a very short time, [3]. This, however, is not motivated when the measurement interval is short in comparison with the half-life of the radioisotope.

The rate of this method is equal to half of the captured disintegration rate as two consecutive intervals between successive captured disintegrations yield a single bit.

By considering a homogeneous Poisson process, the probability of obtaining the next bit 1 can be computed if we consider for any value Δt_1 of the first interval only the values higher than Δt_1 for the second interval Δt_2. As the first interval was chosen arbitrarily and the probabilities of obtaining any pair of intervals $(\Delta t_1, \Delta t_2)$ are independent, we can integrate over Δt_1, that is:

$$\int_0^\infty \lambda \cdot e^{-\lambda t_1} \int_{t_1}^\infty \lambda \cdot e^{-\lambda t_2} dt_2 \, dt_1 = \frac{1}{2} \tag{3}$$

A similar calculation with the inversed roles of the time intervals yields the same result for obtaining the next bit 0, and as a result we expect a bit-uniform probability distribution.

2.2. Generation by binning the Poisson distribution of number of events in a fixed time interval

Silverman, in [18], proposes the counting of the number of decays occurring in each fixed time interval, Δt, and converting it into a binary value by either exploiting the fact that all the counts are necessarily integers and using the parity bit of this integers, or setting a threshold for comparison with the number of counts, providing a bit, accordingly.

Tissa & Zappa, in [19], apply Silverman's procedure of using the parity bit of the number of counts, to a single-photon avalanche photodiode in dark counting mode used as source, and further extend the procedure by truncating the last two bits of the number of counts and compute the deviations from the uniform distribution in each case.

Generalizing the arguments above, the probability distribution allows us to extract bits by dividing it into bins in order to obtain uniform probability distribution over the involved bins, as shown in Fig. 5.

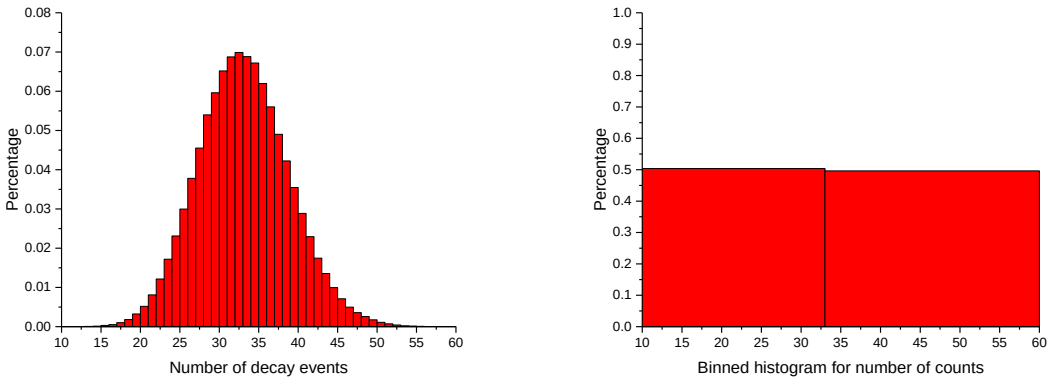


Fig. 5 – Binning procedure applied to Poissonian source. Left: Original histogram from actual experimental data. Right: Binned Histogram with negligible bin height differences.

Thus, by assigning each bin a sequence of n bits, we may deliver for each time slot Δt a sequence of $seq = (b_0 b_1 \dots b_n)$ bits, if the number of decays arriving in the considered time slot enters the bin having the mentioned sequence seq.

Nonetheless, the aim is to include as much as possible of the initial probability distribution into the involved bins in order to minimize the number of events falling outside the bins, which are discarded. This translates to the fact that n bits would be generated for each time slot Δt with a probability of the event being captured in one of the bins.

There are no restrictions imposed on the bins other than no overlap, that is they can have varying width and furthermore, the same sequence of bits can be assigned to two separated bins in order to guarantee for that sequence the same probability as for all the other possible sequences.

In order to provide a high generation rate, the number of bins are maximized over a fixed time interval Δt. Moreover, we aim at minimizing the time interval over which the binning procedure is applied. This time interval is limited by the time resolution of the detector and also by the impossibility of proper binning (with negligible probability difference from the uniform distribution over the bin sequences) the distribution due to high degree of asymmetry, Fig. 6.

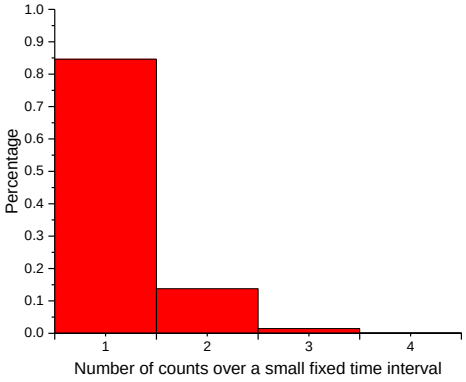


Fig. 6 – Highly asymmetric histogram impossible for binning properly (with negligible bin height difference).

The optimization process is carried out taking into consideration the specifics of the Poisson distribution:

- the symmetry of the Poisson distribution (it tends to the symmetrical normal distribution for sufficiently high enough $\lambda\Delta t$);
- the variance $\lambda\Delta t$ – a higher variance allows for more bins to be included but may require longer time intervals Δt of observation or stronger radioactive isotopes;
- the minimum bin size comes in discrete atomic units (either we have an event or don't), so the binning procedure may imply an almost uniform distribution with variation in a small interval ε, as shown in Figure 5;
- for lower values of λ, the symmetry of the distribution decreases, and that may require to split the bins corresponding to the same sequence of bits.

In practice, for any event generating source, one can perform the binning procedure either by a calibration step (observing the long-term-run statistics) or by adopting beforehand a theoretical probability distribution suitably describing the source and using it for binning. The probability distribution must be in agreement with the long-term-run statistics. This discussion does not strictly refer to Poisson distribution, because the binning procedure can be applied to any probability distribution, as further exemplified in the case of Erlang statistics.

2.3. Generation by binning the Erlang distribution of time passed until the next k^{th} event

A similar approach to the above described method is to apply the binning procedure presented above on the resulting Erlang distribution of the time random variable passed from the current event to the next k^{th} event, exemplified in Fig. 7.

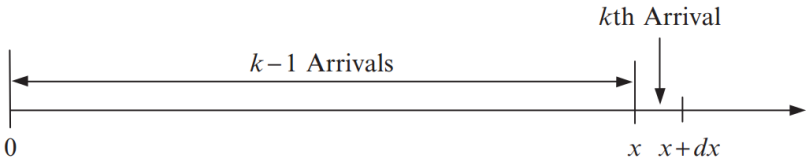


Fig. 7 – Inter-arrival times between successive k[th] events, from [20].

Considering constant mean rate, the k-Erlang distribution has the probability distribution function:

$$P\left(next\ k^{th}\ decay\ occuring\ at\ time\ t\right) = \frac{\lambda^k \cdot t^{k-1} \cdot e^{-\lambda t}}{(k-1)!} \tag{4}$$

An example of binning procedure applied to the 3-Erlang distribution is presented in Fig. 8.

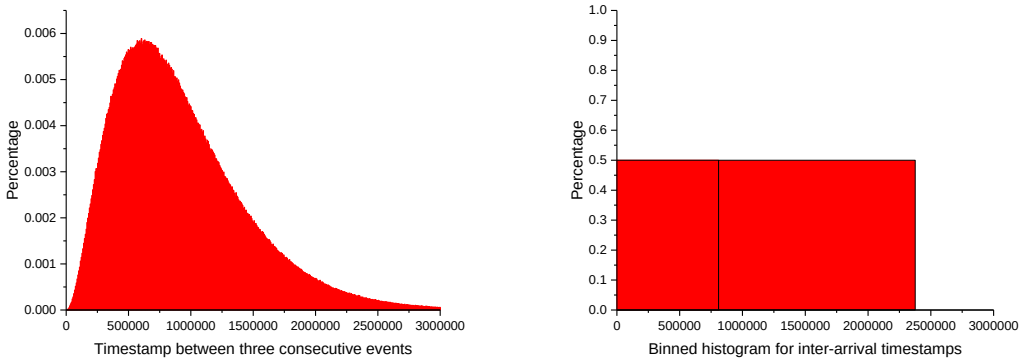


Fig. 8 – Binning procedure applied to 3-Erlang distribution of time between three consecutive events. Left: Original histogram from actual experimental data. Right: Binned histogram with negligible bin height differences.

This method provides a varying bitrate, generating each sequence of bits according to the random time that has passed since the k[th] previous decay.

Similar to the Poisson distribution binning method, the events having inter-arrival times falling outside the existing bins are discarded (events having timestamp above 2300000 in Fig. 8), and thus bit generation depends on the probability that the inter-arrival times fall inside any of the predefined bins.

2.4. Least significant bits extraction from timestamp intervals between consecutive events

This method, proposed by the authors following the line of logic of Tissa & Zappa in [19], involves the extraction of a number of least significant bits from each timestamp corresponding to a decay event, as exemplified in Fig. 6.

We will later refer to this method as n-LSB when extracting a number of n bits from each timestamp.

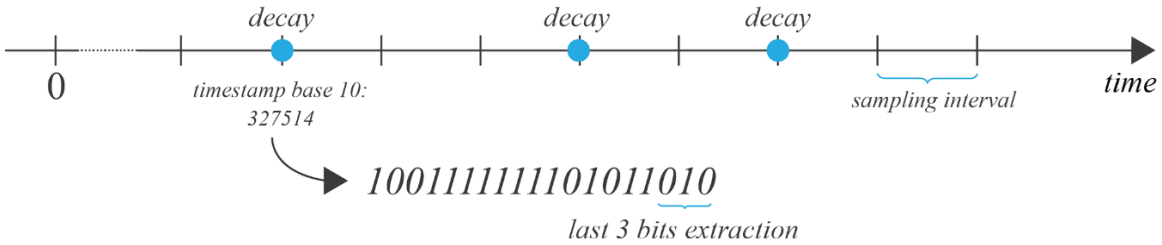


Fig. 9 – Least significant bits extraction exemplified.

The timestamp of each event can be measured in the following three ways:

- from the start of the measurements (n-LSB/start),
- from the previous event (n-LSB/prev),

- from any k^{th} previous event (n-LSB/prev-k).

The first case has the disadvantage that the most significant bits change less than the least significant bits. In consequence, as more significant bits are truncated, although this provides a higher bitrate, there is an increase in the degree of cross-bit correlation. The second case, for a homogeneous Poissonian source, would have a bias for shorter intervals due to the fact that the inter-arrival times follow an exponential distribution, but a tradeoff between the rate of change and bias for short intervals may be found in the third case, where the inter-arrival times follow an Erlang distribution.

In practice, the choice of the number of bits selected is carried out as a compromise between the quality of randomness and the required generation rate. The selection of only the parity bit gives a rate as high as the rate of decay of the material, which may suffice for some purposes.

Generally speaking, this extraction method provides a throughput equal to the rate of decay multiplied by the number of least significant bits selected from each timestamp corresponding to a single decay.

3. EXPERIMENTAL RESULTS

The experimental setup comprises a Canberra [21] HPGe clover γ-ray detector and a ^{152}Eu standard calibration source. The measurements, collected over a period of 43847.375s (approximately 12.17 hours), were carried out at the "Horia Hulubei" National Institute for R&D in Physics and Nuclear Engineering, Bucharest.

The data acquisition was performed with a TNT2-D NIM-based card [22] with a 100 MSPS sampling rate and 14-bit resolution, recording for each γ-decay event the timestamp and channel, which in the calibration step was converted to energy. The calibration procedure for determining the energy as a function of the channel number was carried out with the TUC software [22], by assuming a first-order polynomial dependence obtained by identifying the energy peaks and performing least-square fitting.

In the case of all extraction methods, the bitrate depends on the activity of the isotope (amount and decay rate) and the detector performance (efficiency, dead time and resolution when peak selection is used).

The dimensions of the raw data files (without post-processing) for each extraction method described above, are specified in Table 1 (the files can be accessed on http://bit.ly/true-random-sequences).

Table 1

The dimensions of the raw files obtained by applying the extraction methods to the measurement data

Method			**Raw File Size (kB)**	
			121 keV line	**Entire energy spectrum**
HotBits adaptation			873	8.656
Poisson Binning (2 bins for 0.01s distribution)			515	-
LSB Method	**Start**	**1bit**	1.746	17.321
		2bits	3.492	34.641
		4bits	6.983	69.282
		8bits	13.965	138.564
	Prev	**1bit**	1.746	17.321
		2bits	3.492	34.641
		4bits	6.983	69.282
		8bits	13.965	138.564
1-Erlang Binning (2 bins)			1.746	-
3-Erlang Binning (2 bins)			574	-

In order to comparatively assess the quality of randomness generated by these four extraction methods we employed both statistical testing – using the ENT utility program [15] and the NIST STS [16], as well as visual analysis using the software tool FileSeer+, described in [23].

The quality analysis was performed on all the energy spectrum and also, in order to neglect disturbing external influences, on one isolated mode of decay, specifically the 121.8 keV line (28.7% intensity [24]) by

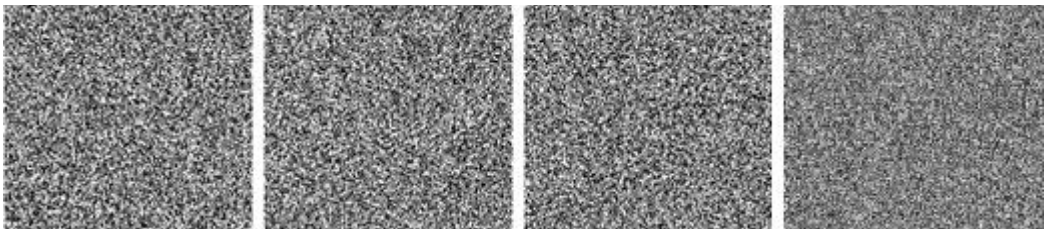
selecting the decays with energies in the 121-123 keV window. In order not to discard measurements, there can also be considered several lines at once, as the sum of Poisson processes corresponding to each individual decay mode is still Poisson.

The binning for the Poisson histogram over 0.01s was carried out, in order to simulate a calibration procedure, with measurements taken over 60 minutes, resulting in two split bins covering 99.9% of the distribution, providing a bit rate of approximately 100 bits/second.

3.1. Visual randomness assessment

In visual randomness testing the sequence of random numbers is graphically represented as an image, and the human perceptual system is employed to extract statistical properties of the set of pixels which form the image.

The grayscale image representation of the generated sequences, depicted in Fig. 10, does not reveal any patterns or bias towards certain shades of grey, suggesting a uniform distribution of values. The uniformity is also confirmed, in Fig. 11, by the graphical representation of the histogram of one bit and two bit values appearing in the file generated with the 1-LSB/start method.



a. b. c. d.

Fig. 10 – Greyscale images representation for extraction methods: adapted HotBits (a), Poisson Binning (b), LSB (c), 1-Erlang (d).

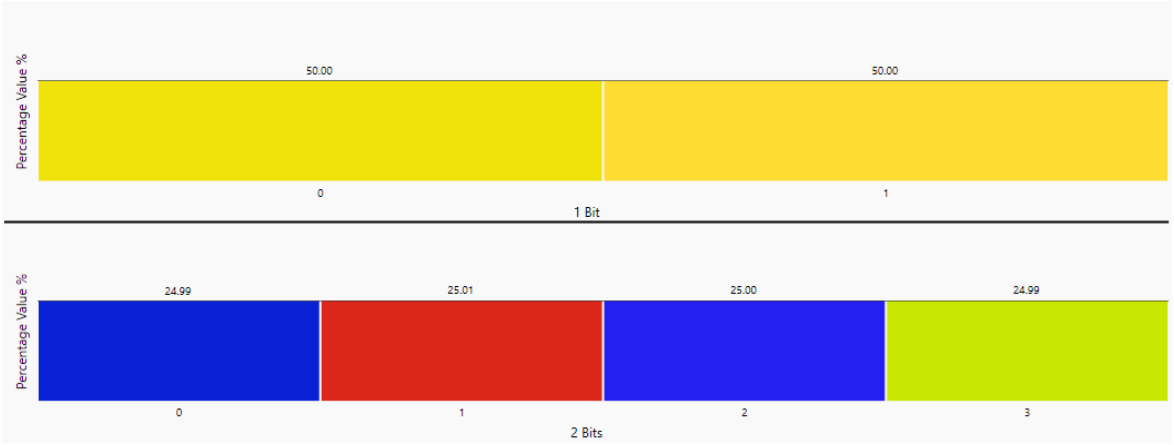


Fig. 11 – Frequency representation for 1-LSB/start method for the 121keV line.

3.2. Simple testing with ENT

The ENT Utility [15] is applied on the raw data for each extraction method, and the results, even in the absence of any post-processing, yield high quality randomness, expressed in the tests as:

- maximum entropy per bit, for all methods;
- the chi square value is randomly exceeded less than 80% and more than 20%, thus satisfying Knuth criterion [25], in all the above cases except 8-LSB/prev (where a decrease in quality is expected) and 1-Erlang method;
- the arithmetic mean is close to 0.5;
- the Monte Carlo value for Pi is within 0.2% error;
- the serial correlation coefficient is in close to 0.

According to the results presented in Table 2, the ENT Utility qualifies the majority of raw sequences as having a high degree of randomness.

Table 2

The results of ENT Utility applied on the raw data for 121 keV line

Method			ENT Tests				
			Entropy (bits/bit)	**Chi-square Value exceeded randomly less than**	**Arithmetic Mean**	**Monte Carlo value for pi**	**Serial Correlation Coefficient**
Adapted HotBits			1.000000	1.44 23.02% times	0.4998	3.145722841 (0.13% error)	-0.000114
Poisson Binning (2 bins for 0.01s distribution)			1.000000	1.53 21.58% times	0.4997	3.141439206 (0.00% error)	-0.000116
LSB Method	**Start**	**1bit**	1.000000	0.40 52.81% times	0.5001	3.138314049 (0.10% error)	0.000025
		2bits	1.000000	1.49 22.29% times	0.5001	3.144048780 (0.08% error)	0.000187
		4bits	1.000000	0.97 32.39% times	0.5001	3.141893763 (0.01% error)	-0.000106
		8bits	1.000000	1.42 23.37% times	0.5001	3.138841184 (0.09% error)	0.000038
	Prev	**1bit**	1.000000	0.10 75.14% times	0.5000	3.135883777 (0.18% error)	0.000323
		2bits	1.000000	0.45 50.30% times	0.4999	3.143055190 (0.05% error)	0.000306
		4bits	1.000000	1.63 20.10% times	0.4999	3.142541611 (0.03% error)	0.000107
		8bits	1.000000	3.54 5.99% times	0.4999	3.140536329 (0.03% error)	0.000200
1-Erlang Binning (2 bins)			1.000000	0.00 99.75% times	0.5000	3.139771808 (0.06% error)	0.000028

For a more rigorous verification, we further assess the quality with the NIST Statistical Test Suite.

3.3. NIST Statistical Test Suite

The NIST STS [16] comprises 15 statistical tests, and computes in total 188 p-values for each tested sequence, as several tests are repeatedly applied using different parameter configurations. Therefore, the following results show the percentage of passed tests (the fraction of p-values within the significance level) considering all tests in the suite in default configuration (with the recommended statistical parameters).

In order to benchmark the randomness quality of the extraction methods we comparatively present the NIST STS results obtained on the raw data together with results on filtered data, after applying several post-processing functions.

The employed post-processing functions are as follows: error correcting code proposed by Lacharme (ECC), Fibonacci Linear Feedback Shift Register (Fib-LFSR), the hash functions SHA-1, SHA-256 and SHA-256 with compression (SHA-256Compress), the Von Neumann filter (VN), and the XOR function on consecutive bytes (XOR) and on consecutive 4 bits (XOR-nibbles).

The statistical results, computed as a percentage indicating the pass rate, are provided in the following. Fig. 12 and Fig. 13 present the comparative evaluation of the n-LSB method - using 1, 2, 4 and 8 bits from the least significant part of each timestamp, considering only the 121 keV line.

NIST STS indicates that the randomness quality of raw data is very good, and it very slightly decreases as more bits are extracted from each timestamp.

The graphical representations highlight that in this configuration, filtering the raw data will not necessarily bring better statistical results. On the other hand, when considering the whole spectrum, the quality of randomness in the raw data is considerably decreased, as shown in Fig. 14 and Fig 15, hence post-processing becomes necessary for obtaining a better quality, even with the cost of reducing the throughput of the generation method.

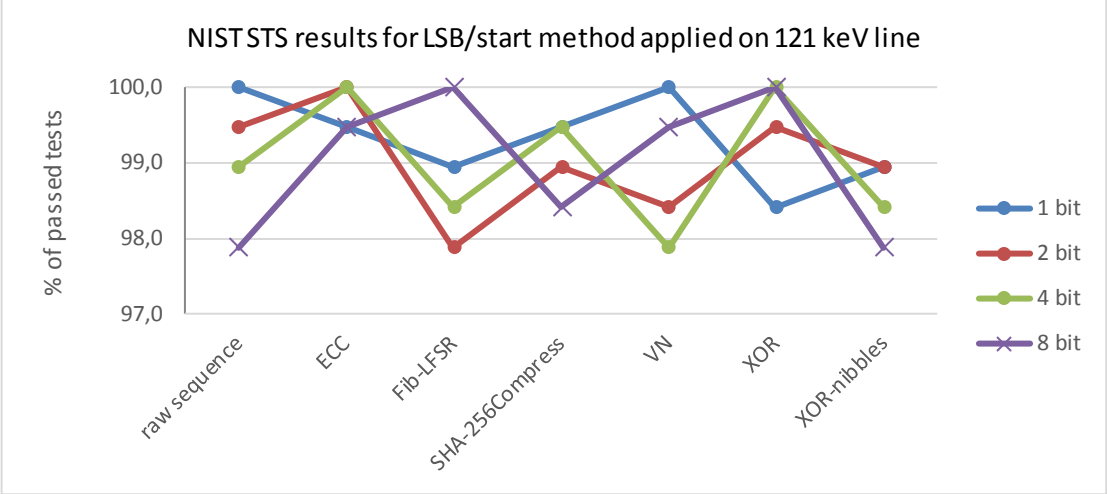


Fig. 12 – NIST STS results for LSB/start method applied on 121 keV line.

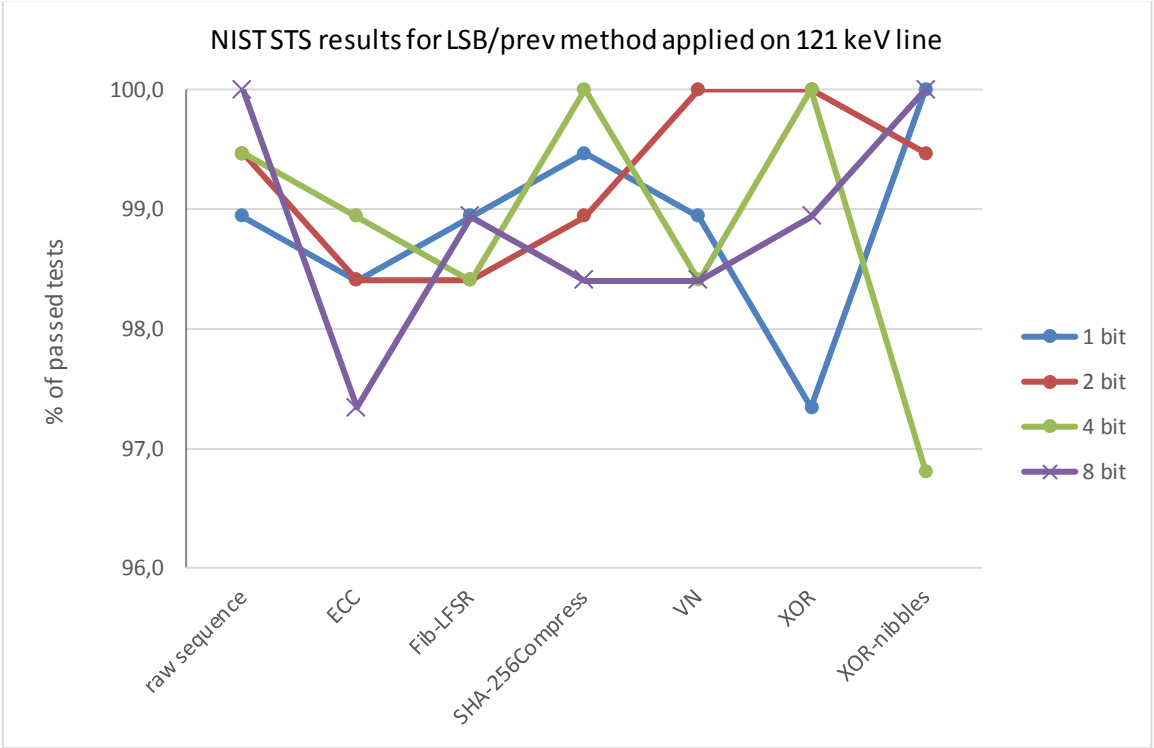


Fig. 13 – NIST STS results for LSB/prev method applied on 121 keV line.

These results highlight both the importance of selecting suitable energy lines corresponding to particles which are known to be provided by the Poissonian entropy source, and moreover the various possible

impacts filtering functions can have on the statistical quality of the produces sequences, including negative impact, such as shown by XOR filtering in Fig. 14, due to a high sensitivity to certain correlations between the most significant part of consecutive 8-LSB values.

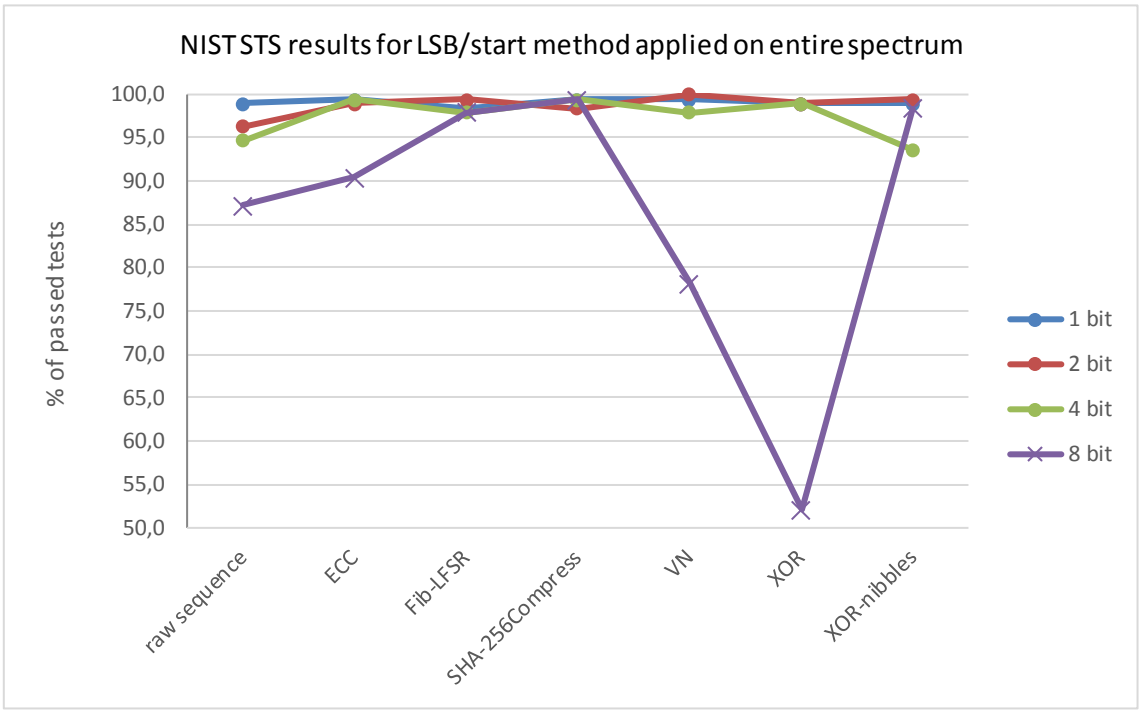


Fig. 14 – NIST STS results for LSB/start method applied on entire energy spectrum.

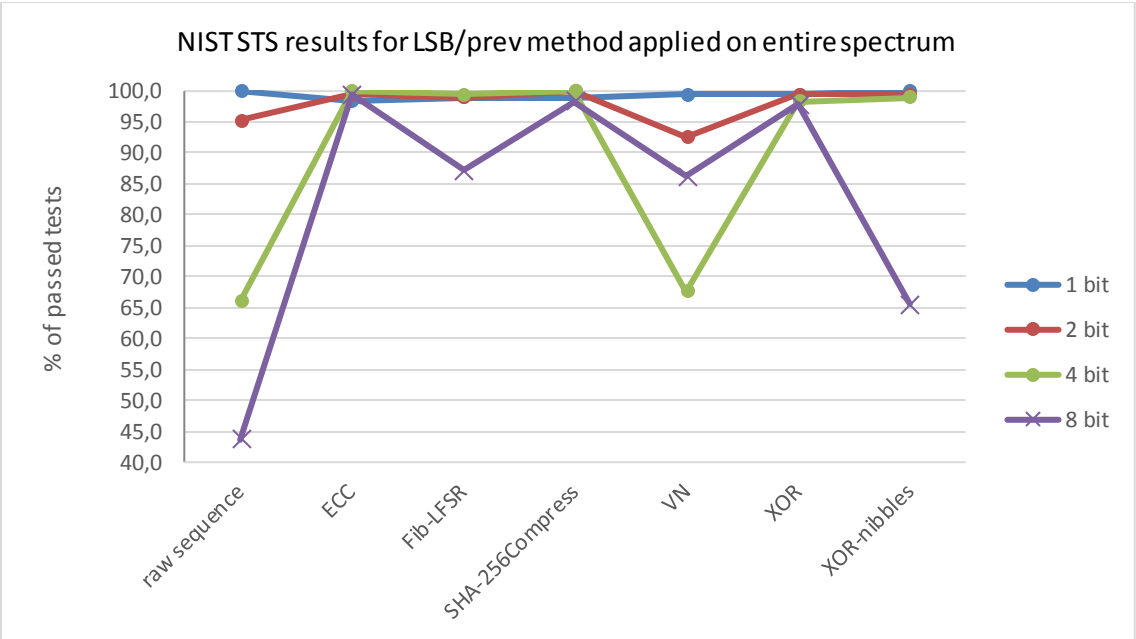


Fig. 15 – NIST STS results for LSB/prev method applied on entire energy spectrum.

Table 3 provides a more complete set of NIST STS results for the raw and post-processed sequences for each generation method.

Table 3

NIST STS results on raw and post-processed sequences

		Percentage of NIST STS passed tests						
Methods Applied		**Adapted HotBits**	**Poisson Binning**	**LSB/start Method**				**3-Erlang Method**
Selection		*all spectrum*	*121 keV line*	*all spectrum*	*all spectrum*	*all spectrum*	*all spectrum*	*121 keV Line*
Specifications		-	*2 bins over 0.01 s*	*1 bit*	*2 bits*	*4 bits*	*8 bits*	*2 bins*
Post-processing Algorithms	**Raw sequence**	99.4680	98.9361	98.9361	96.2765	94.6808	87.2340	98.9361
	ECC	100.0000	100.0000	98.9361	96.2765	94.6808	87.2340	100.0000
	Fib-LFSR	99.4680	98.4042	99.4680	98.9361	99.4680	90.4255	98.4042
	SHA	99.4680	98.4042	98.4042	99.4680	97.8723	97.8723	98.4042
	SHA256	98.9361	98.9361	99.4680	100.0000	99.4680	99.4680	98.9361
	SHA-256Compress	98.4042	98.9361	100.0000	98.4042	95.7446	100.0000	98.9361
	VN	98.9361	96.8085	100.0000	98.9361	100.0000	98.4042	96.8085
	XOR	99.4680	98.9361	99.4680	100.0000	97.8723	78.1914	98.9361
	XOR-nibbles	98.4042	98.4042	98.9361	98.9361	98.9361	52.1276	98.4042

4. CONCLUSION

This paper applies four extraction methods of random numbers on the phenomenon of radioactive gamma-decay whose description based on quantum mechanics renders it intrinsically random.

A comparative analysis on the bit sequences obtained from the methods involved is performed, based on the relative rate of the generated bits and also on quality testing with ENT, a visual inspection and the NIST Statistical Test Suite.

In the case of probability distribution binning method, an appropriate binning procedure given either by a calibration step or by theoretical considerations is required, while for the other methods no external intervention or imposed conditions are necessary.

Due to the fact that even the raw measurement datasets extracted from the physical phenomenon yield high quality randomness by passing the visual assessment, the ENT tests and a high percentage of the NIST tests, the post-processing algorithms may not even be required, as in some cases it is actually seen that the post-processing decreases the quality. The aim, in general, is to provide raw randomness, that is, to apply only the extraction method to the physical phenomenon, avoiding other further interventions or corrections to increase the quality, which would introduce unwanted computational time expended in this process (decreasing the rate), and most importantly, introduce pseudo-randomness in the context.

The criterion of statistical tests that check for specific patterns (like the uniformity of the distributed outputs, serial correlation, etc.) is insufficient alone to characterize randomness, as there still may be patterns hidden in the generated sequence which may be revealed later by further tests [26], as would be the case, for instance, of a pseudorandom generator which tends to pass successfully all known tests but in the long run statistics, an appropriate test would reveal an algorithm. This would apparently mean that an infinite number of statistical tests is needed to attest the truly random character [27]. However, it is only necessary to ensure the independence (of the next bit) of the information available in advance. The absence of this predictability of behavior by means of prior information that the system might be correlated to [26] can be certified by physical principles, which is the case of the output of a QRNG, considered in the current article.

Taking into consideration the fact that the statistical tests employed in the current paper display a high quality of randomness (a necessary but not sufficient condition), together with the inherent randomness of quantum mechanics, that ensures the absence of correlation to information available in advance, we conclude that the γ-decay phenomenon offers the possibility of high quality raw randomness extraction with a bitrate depending on the activity of the radioisotope, the detector efficiency and the extraction method involved.

A further research direction of interest would be a low-cost, less-maintenance implementation of gamma QRNGs, developed around scintillator detectors and high-speed memory and digitizers, with focus on temporal resolution.

ACKNOWLEDGEMENTS

The first author highly acknowledges financial support from the Babeş-Bolyai University through a Performance Scholarship during the academic year 2016-2017, contract number 37406/23.11.2016.

This work was made possible with the support of "Horia Hulubei" National Institute for R&D in Physics and Nuclear Engineering, Magurele, where the measurements were carried out.

REFERENCES

1. A. ACIN, *Randomness in Quantum Physics*, Euresis Journal, **b**, pp. 35-37, 2014.
2. J. KOFLER, A. ZEILINGER, *Quantum Information and Randomness*, European Review, **18**, pp. 469-471, 2010.
3. M. HERRERO-COLLANTES, J. C. GARCIA-ESCARTIN, *Quantum Random Number Generators*, Rev. Mod. Phys. **89**, 2017.
4. J. BELL, *On the Einstein Podolsky Rosen Paradox*, Physics, **1**, 195-200, 1964.
5. A. EINSTEIN, B. PODOLSKY, N. ROSEN, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*, Physical Review, pp. 777-780, 1935.
6. A. ASPECT, P. GRANGIER, G. ROGER, *Experimental Tests of Realistic Local Theories by Bell's Theorem*, Phy. Rev. Lett., **47**, 1981.
7. B. HENSEN. ET AL., *Loophole-free Bell Inequality Violation using Electron Spins Separated by 1.3 Kilometres*, Nature, **526**, pp. 682–686, 2015.
8. S. BARNETT, *Quantum Information*, Oxford Master Series, pp. 118, 2009.
9. R. COLBECK, R. RENNER, *No Extension of Quantum Theory Can Have Improved Predictive Power*, Nature Communications **2**, 2011.
10. J. WALKER, *HotBits: Genuine random numbers, generated by radioactive decay*, http://www.fourmilab.ch/hotbits/, 1996.
11. A. ALKASSAR, T. NICOLAY, M. ROHE, *Obtaining True-Random Binary Numbers from a Weak Radioactive Source*, Lecture Notes in Computer Science, **3481**, 2005.
12. A. STEFANOV, N. GISIN, O. GUINNARD, L. GUINNARD, H. ZBINDEN, *Optical Quantum Random Number Generator*, J. Modern Optics, **47**, 595-598, 2000.
13. M. FURST ET AL., *High speed optical quantum random number generation*, Opt. Express, **18**, 13029-13037, 2010.
14. D. BEZNOSKO, T. BEREMKULOV, A. DUSPAYEV, A. IAKOVLEV, *Random Number Hardware Generator Using Geiger-Mode Avalanche Photo Detector*, Proceedings of Science, 2015.
15. J. WALKER, *A Pseudorandom Number Sequence Test Program*, http://www.fourmilab.ch/random/, 2008.
16. L. E. BASSHAM ET AL., *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, Special Publication (NIST SP), 2010.
17. G. FRIEDLANDER, J. W. KENNEDY, E. S. MACIAS, J. M. MILLER, *Nuclear and Radiochemistry*, Wiley, 1981.
18. M. SILVERMAN, *A Universe of Atoms, An Atom in the Universe*, Springer, 2002, pp. 305-307.
19. S. TISA, F. ZAPPA, *One-chip Quantum Random Number Generator*, Proc. SPIE, Quantum Communications Realized II, **7236**, 2009.
20. O. IBE, *Fundamentals of Applied Probability and Random Processes*, Academic Press, 2014, pp. 128.
21. G. DUCHENE ET. AL., T*he Clover: A new generation of composite Ge detectors, Nuclear Instruments and Methods*, in Physics Research A, **432**, pp. 90-110, 1998.
22. M. RICHER, C. BONNIN, C. SANTOS, *TNT2 Digital Pulse Processor Functionalities & TUC control software*, http://www.iphc.cnrs.fr/IMG/pdf/TNT_Functionalities.pdf, 2013.
23. K. MARTON, D. PATRASCU, A. SUCIU, *Perceptual Evaluation of Random Number Sequences using FileSeer+*, Studia Universitatis Babes-Bolyai–Series Informatica, **LX**, *1*, pp. 98-110, 2015.
24. T. JOHNSON, B. BIRKY, *Health Physics and Radiological Health*, Lippincott Williams and Wilkins, 2011, pp. 1139.
25. D. KNUTH, *The Art of Computer Programming*, **2**, Seminumerical Algorithms, Addison-Wesley, 1969, pp. 35-40.
26. D. FRAUCHIGER, R. RENNER, M. TROYER, *True randomness from realistic quantum devices*, arXiv, 2013.
27. M. STIPCEVIC, *Quantum number generators and their use in cryptography*, arXiv, 2011.