PROCEEDINGS OF SCIENCE

Random Number Hardware Generator Using Geiger-Mode Avalanche Photo Detector

D. Beznosko[1], T. Beremkulov, A. Duspayev, A. Iakovlev
Nazarbayev University
53 Kabanbai Batyr ave, Astana, KZ
E-mail: dmitriy.beznosko@nu.edu.kz, timur.beremkulov@nu.edu.kz, alisher.duspayev@nu.edu.kz, aleksandr.iakovlev@nu.edu.kz,

The main problems with existing hardware random number generators today are either low speed and/or prohibitively high cost. The physical concept and test results of sample data of the high-speed hardware true random number generator design based Hamamatsu MPPC photo sensor are shown. Main features of this concept are the high speed of the true random numbers generation (tens of Mbt/s), miniature size and estimated lower production cost. This allows the use of such a device not only in large companies and government offices but for the end-user data cryptography, in classrooms, in scientific Monte-Carlo simulations, computer games and any other place where large number of true random numbers is required. The physics of the operations principle of using a Geiger-mode avalanche photo detector is briefly discussed and the high quality of the data collected is demonstrated.

PACS Numbers: 89.20.Bb, 89.20.Ff, 85.60.Ha

International Conference on New Photo-detectors
PhotoDet2015
6-9 July 2015
Moscow, Troitsk, Russia

[1]Speaker

1.Introduction

In everyday life, we use things that came from the High Energy Physics (HEP) research fully or in part, but found their use in other parts of life, such as microwaves, Internet protocols, data analysis tools, etc. Another whole niche that comes from HEP is related to photo detectors.

Here, the application of the photo detectors to the hardware for random number production is discussed. The use of the available in HEP for the last decade miniature photo sensors as a basis for the hardware-based random number generator (HRNG) that could be commercialized in the future. This project was carried out as part of the research program by the Nazarbayev University Cosmic Rays and Particles group.

The hardware produced random numbers are used by large companies (such as banks, communications and cell phone companies) and by many countries' government planning offices in their simulations of the economy growth and similar tasks. Due to the low speed, low availability and/or prohibitively high cost of such devices, analytical algorithms are used instead in small companies, in science or by the end-users. The downside of any analytical (e.g. software) method is that it is not truly random, thus presenting a weakness that can be exploited for malicious purpose. Therefore, a simple, robust and affordable solution is necessary.

1.1 Physics of HRNG

All HRNGs have one thing is common – there is a source of randomness in them that relies on some unpredictable process. Commonly known are thermal noise in a p-n junction, radioactive decays, atmospheric discharges, radio white noise, etc. Typically, the listed above are slow and/or give randomness of 'low quality' such as its not completely unpredictable or is easily influenced by environmental factors.

Better sources of randomness are based on optics and quantum mechanical effects. Example is the reflection from the half-transparent mirror – one could detect whether a single photon was reflected or transmitted through such as mirror, but it will be a fairly expensive device as it involves accurate single-photon counting plus a very precisely made mirror.

The physics principle used in the device described here also involves the detection of the single photons incident onto the Geiger-mode avalanche photo detector [1], which is a true random process in itself, without any additional components. These detectors have been developed recently and are available from several manufacturers, the sensors from two of these are tested in [2] and in [3], many others are available nowadays with different sensitivities, areas and price. These devices have been already used in large scale High Energy physics experiments, e.g. in the T2K-ND280 pi-zero detector [4]. The main features for these devices are high gain (~10^6), robustness, low biasing voltage, high sensitivity in optical range (~25%) and relatively low cost compared to devices that use other technologies but with similar sensitivity/gain.

2.Experimental Setup

The 400-pixel square Hamamatsu [5] MPPC photodiode [6] was used for this setup. With the biasing voltage of ~-70.10V, it provides the gain of ~7·105 and the detection efficiency of up to 30% at 500nm. A Bivar [7] SM1204PGC Light Emitting Diode (LED) in the metal box with a small opening was used as a light signal source, and a CAEN [8] DT5743 digitizer was used to record the signal. The schematic of the setup is shown in Figure 1. Since the pulse width from the MPPC is couple of tens of ns, the system can be run at several MHz (thus collecting Mbits of data per second), but we were limited by DAQ system to about 1 kHz. Note that devices with narrower pulse width exist thus higher rates could be achieved.

For the low cost solution, it is possible to realize the whole setup on a single controller chip with LED and MPPC connected to it directly, thus making the possible cost of such a device much lower than comparable ones on the market. The clear separation of signal detection from no-signal for MPPC (and similar devices) allows for a digital readout of the signal, thus eliminating the most costly component of the setup – ADC. With a simple comparator, the controller can read digital output from the MPPC directly.

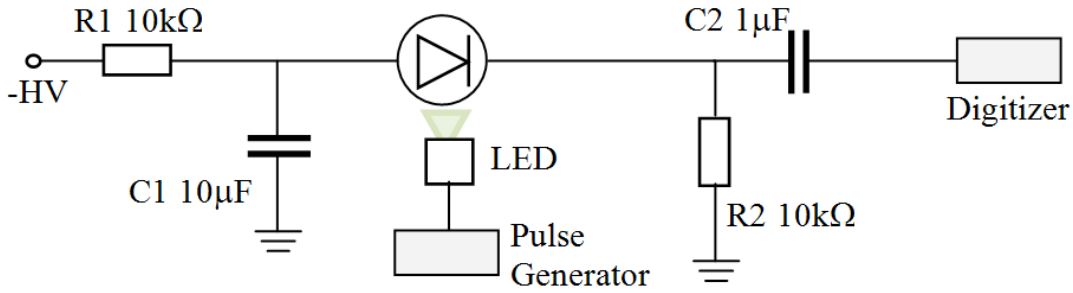

Figure 1: Randorm Generator Test Setup Schematics.

3.Experimental Procedure

The experimental basis of the idea uses the fact that the MPPC photodiode has a clear photoelectron (PE) separation both in amplitude and in charge. The light from the LED via a small opening in the metal housing is incident onto the photodiode producing the signal output shown in Figure 2 for both the amplitude (left) and area (right). The rightmost peak is the pedestal (that is, zero photons detector), the next peak to the left is the 1st PE, next is the 2nd, etc. since the signal amplitudes are negative.

Since a simple comparator will work with the amplitudes rather than areas, only amplitude will be used further to emulate it. Note that the signal powering the LED is also the trigger generator for the data collection in order to reduce the probability of the external noise to enter the signal. The detection gate width used was 100ns and was centered on the MPPC average response position. The area is the algebraic sum of all amplitudes within the gate.

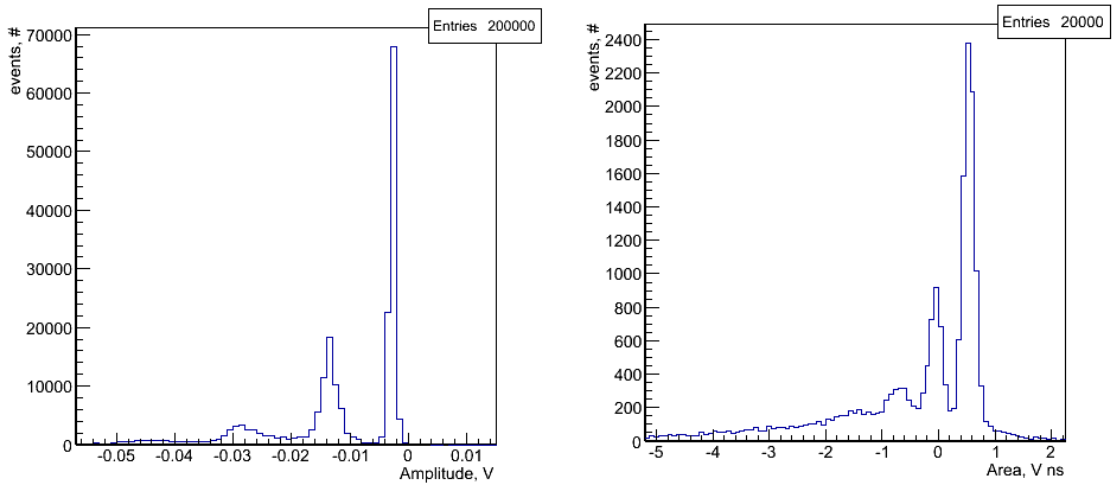

Figure 2: MPPC PE signal separation in amplitude (left) and charge (right).

As the area of the pedestal in Figure 2 (left) has approximately 50% of the events in it, the threshold was chosen to be at -0.005V. No extra care was taken for the pedestal to be 50% exactly or the area to be stable long-term due to the analysis method used. Any after-pulse effects of the MPPC pixel did not affect the result as only the detection fact that matters.

Both the light yield of the LED and the detection efficiency of the MPPC are weekly dependent upon the temperature and photodiode bias [9]. However, both are stable on a small time scale of several tens of seconds, which is the timescale of the data collection runs (clarified in next chapter); thus, we assume the stability of the conditions. The biasing voltage was within ±0.002V and changes of temperature were not detectable by external thermometer during the test setup operations.

4.Theoretical Considerations

The amount of photons that falls onto the photodiode follows the Poisson distribution. The light detection efficiency of the photodiode is constant (changing very weakly with applied bias value) and is on the order of ~30%. Thus, the photodiode sometimes detects no photons at all, sometimes one or more. This is evident from Figure 2, where the areas (right) and amplitudes (left) of each peak follow the distribution. However, if we set a threshold (at ~ -0.005V) that separates the pedestal from the signal, the resultant is the integral probability of signal being detected (or not).

There is also an issue of dark noise and afterpulse of the detector. In this case, the detection window was about 100ns so most afterpulses would fall within it and not cause a false detection in the next window. The dark noise has the random nature itself and is rare (meaning it's occurrence within the gate) and does not affect the overall data quality. This also answers the possible question as to why use the light pulses and not the dark noise itself. With the pulse schema, the trigger is external and constant, it sets the rate and is independent from the environment. The dark noise would be correlated with temperature and would be much more vulnerable to outside noise as some type of self-triggering scheme may need to be implemented.

As the data is acquired, the values below the threshold will be converted into the bit with the value of 1 since we are walking about negative amplitude pulses, and above the threshold to a bit with a value of 0. The resultant output file is the binary data written as a sequence of 1s and 0s without the separation into bytes or words, e. g. a stream of bits.

Final data, while being random, does not exactly contain even amounts of ones and zeros. The reason for that is the extreme difficulty to constrain all physical conditions of the system to keep the generator producing the exact 50-50 output all the time. However, there is a simple method to even out the data without loosing the randomness of it called the AMLS [10], it is a randomness extraction algorithm with the code implementation of it taken from [11]. The code output is the bits sequence with the same amount of 0s and 1s, but it requires an initial source of randomness to function. Thus this is the solution for the local stability concept: collect data for a few seconds (or even less then one second, depending on the collection rate), then run the AMLS code on the sample, collect again. This way, data samples are not correlated to each other and to any external slow changes in environment.

5.Testing of Generated Random Numbers

The testing of random numbers seems to be a large topic nowadays with many tools in existence to do that. The tests and their explanations are available in [12], including the 'birthday' test. Here, we include the most illustrative ones for completeness.

Typically, the results are dependent on the data sample size and give excellent results only on infinite datasets. Here, the sample of 10 Mbyte from the HRNG above is used. Overall, even these two simple tests indicate the quality of the obtained data. The tests were conducted using the ENT software suit [13] for random number testing. First, the short test description is given followed by the result.

First test is entropy. There are 8 bits to a byte; each can be 0 or 1. For a complete random sample, this value should be as close to 8 as possible. Below is the result for the current sample:
-Entropy = 7. 999888 bits per byte. The value of 8 is possible only for infinite data set.

Compression methods (such as zip, rar, jpg, mp3 and others) work by looking up the patterns in the data and replacing them with smaller codes. If there are no patterns, the compression will yield no reduction in the file size. For the sample above:
-Optimum compression would reduce the size of this file by 0.0000 percent.

Chi^2 distribution shows the deviation of the datapoints from the fit or a model. An important part here is the spread of it and the fact that it is not being exceeded over the nominal value for the given data. The program low limit is <0.01 even for the theoretically perfect dataset.
-Chi^2 distribution for current sample is 911.93, and randomly would exceed this value less than 0.01 percent of the times.

An arithmetic mean of the 8bit integers, starting from 0, should be 127.5 for the infinitely large dataset.

-Arithmetic mean value of data bytes of the current sample is 127. 4651 that for the amount of the data available is a very good value.

Yet another method to check randomness is to try and use the sample for the Monte Carlo (MC) simulation (e.g. using the random 'throws') for a simple known result, e.g. trying to calculate the value of Pi.
-For the current random number sample, MC value for Pi is 3.141567813 (error <0.001 percent). Note that this test is also very sensitive both to the size of the sample and to the quality of it. Any non-randomness quickly causes the MC result to diverge away from the expected value towards 4, and here we see no such divergence.

One can also try to find the correlations between the numbers, but again, 0 will be achieved only for an infinite perfect sample of random numbers.
-For the current sample, serial correlation coefficient is 0.00082.

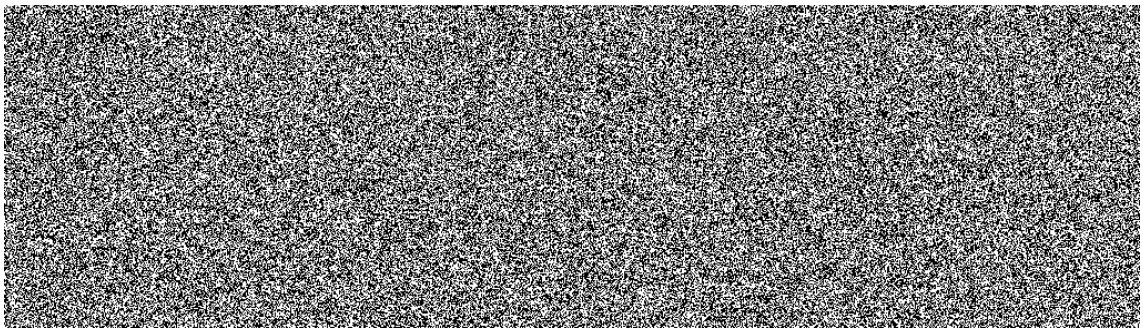

Figure 3: Visual data representation

Figure 3 is the visual presentation of data, where 8bit numbers are formed from the data sample, and these numbers are the brightness of each pixel. This is a good demonstrative test as the eye is very good in catching any patterns and non-randomness in the visual information.

6.Conclusion

The results from the conducted tests show that the data collected using the presented design of the hardware random number generator is of high quality. Thus, the simple generator setup using the MPPC photosensor and the LED in conjuncture with the AMLS un-biasing algorithm presents a possibility for the future use in this capacity (patent pending). The future plans include the building of the miniature functioning prototype of this device based on the small micro-controller with USB connectivity.

References

[1] D. Beznosko, G. Blazey, D. Chakraborty, A. Dyshkant, K. Francis, D. Kubik et al., *Investigation of a Solid State Photodetector. NIM A* **545** (727)

[2] D. Beznosko, G. Blazey, A. Dyshkant, V. Rykalin, V. Zutshi, *Effects of the Strong Magnetic Field on LED, Extruded Scintillator and MRS Photodiode. NIM A **553*** (438)

[3] Beznosko, D. *Novel Multi-pixel Silicon Photon Detectors and Applications in T2K*. physics.ins-det/0910.4429.

[4] 4. S. Assylbekov et al., *The T2K ND280 Off-Axis Pi-Zero Detector*. *NIMA* **686** (48) [physics.ins-det/1111.5030v1]

[5] Hamamatsu Corporation. 360 Foothill Road, PO Box 6910, Bridgewater, NJ 08807-0919, USA; 314-5,Shimokanzo, Toyooka-village, Iwatagun,Shizuoka-ken, 438-0193 Japan.

[6] K.Abe et al., *The T2K Experiment*. *Nucl. Instrum. Meth. A* **659** (106) [physics.ins-det/1106.1238].

[7] Bivar Inc. 4 Thomas, Irvine, CA 92618, USA.

[8] CAEN S.p.A. Via della Vetraia, 11, 55049 Viareggio Lucca, Italy. http://caen.it.

[9] D. Beznosko et al., *Investigation of a Solid State Photodetector, NIM A* **545** (727)

[10] Peres, Yuval. *Iterating von Neumann's Procedure for Extracting Random Bits. The Annals of Statistics*, 1992, (590).

[11] 11. Crowley, Paul. Random data unbiasing. http://www.ciphergoth.org/software/unbiasing/.

[12] D. Beznosko, T. Beremkulov, A. Duspayev, A. Iakovlev, A. Tailakov, M. Yessenov, *A Physical Principle for Fast and Miniature Random Number Hardware Generator Using MPPC Photo Detector. Journal of Advances in Physics [Online]* **7.3** (1968) [physics.ins-det/1501.05521]

[13] Walker, John. A Pseudorandom Number Sequence Test Program. http://www.fourmilab.ch/random/