Ionization from Alpha Decay for Random Bit Generation

Mike Rosing, Engineering Research
Patrick Emin, Retired Prof. of Computer Science,
University of New Brunswick

Abstract

A random bit generator is described which is based on the radioactive source found in common household smoke detectors. Signals created by the ionizing alpha particles are measured and digitized using a specially designed electronic circuit. We then show that application of statistical techniques can transform the data to random numbers distributed uniformly over (0, N).

Keywords: electronic circuit, ionization, radioactive source, random number, statistical transform

Introduction

Since the late 1960's radioactive sources have been used to create random bits. Today one can get random bits from a radioactive source over the internet. Called "HotBits", this source uses the classic method described in the 1970's by Vincent [CHV1], [CHV2]. The major problem with this method is that it has a "dead time" which prevents measurements while the radiation detector recovers from a previous event. In this paper we describe a random bit generator which is also based on a radioactive source but has no dead time. Due to the vast amount of statistical methods available for testing, we can report that the generated bits are as "random" as mathematics knows how to measure.

We will first describe the radioactive source and detector, then describe the electronic circuit in enough detail that it can be easily duplicated. Following that is a discussion of radiation statistics and the method used to convert the analog signal to random bits. A description of the tests used to determine randomness is followed by a detailed analysis of the output statistics. A short discussion of ways to scale the method up to faster bit rates ends the paper.

Ionization Chamber

The radioactive source used in a smoke detector is placed in the center of a plastic block and connected to ground. Typical ion chambers are 3 to 5 cm in diameter and 2 to 3 cm tall. A metal plate with a hole in the center (a "grid" in tube parlance) is mounted in the plastic about 2 to 5 mm from the radiation source. An outer cover of steel surrounds the top and sides of the grid and source. Openings are punched into the cover to allow air to filter in.

The radioactive source used in smoke alarms is Americium 241 which has a half life of 400+ years. Americium 241 decays via a 5 MeV alpha particle. This alpha has a range of about 5 cm in air. The material holding the Americium absorbs some of the energy from decays occurring deeper within the bonder so there is a large spread of energies and ranges within the ion chamber.

The alpha particles ionize the air in the chamber creating a very weak plasma. By grounding the alpha source and charging the outer shell, positive ions drift into the source and negative ions diffuse to the shell. The floating grid sees the plasma potential. Because the plasma density is low compared to the neutral background the grid floats to a DC value half way between the source and outer cover.

For a smoke detector one only needs to see the DC level shift by a few millivolts to know that conditions have changed in the ion chamber. For use as a random bit generator we have to look a bit closer.

As an alpha particle rips through the air it strips electrons off nearby atoms with enough force to cause these electrons to ionize other atoms. This process occurs in nanosecond time scales. When the energy is expended many electrons are absorbed by the air. Some return to previously ionized atoms creating neutrals and others create negative ions and molecules in the air. This recovery time is over within microseconds.

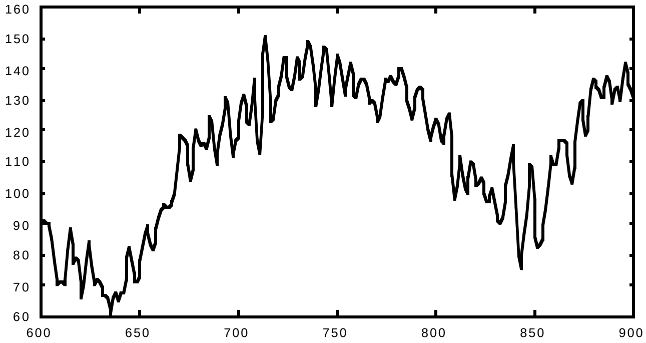
Negative charges are drawn to the outer cover and positive charges towards the radiation source. The charges diffuse at thermal sound speed due to collisions with neutrals. Because the radiation source is practically constant, the current flow in the ion chamber is constant.

Each decay event creates an electric shock wave when the free electrons interact with the background plasma. This shock is coupled to the neutral atmosphere and slowed down to a sound wave which soon hits the grid. For a basic description of waves in plasma see [FFC].

Like all natural noise sources the ion chamber exhibits an inverse amplitude to frequency relationship. There is a spike in amplitude vs. frequency proportional to the decay rate (see figure 3). The geometry of the ionization chamber determines what fraction of the source can be detected. For smoke detectors this amounts to 1/8 full solid angle ($\pi/2$ radians).

Since a 1 microcurie source (standard amount listed on all smoke alarms) generates 3.7 x 10^4 alphas per second we expect to see 4.6 x 10^3 events per second. Some of these events will be from low energy alphas. The distribution of Americium within the source material determines how many radioactive decays have enough energy to ionize the entire length of the ion chamber.

Upon applying the discrete Fourier transform to a number of samples it was found that the actual observable mean rate of decay was 2.685 x 10^3 per second. This is approximately 58% of the theoretical expectation and is reasonable under the circumstances. An example of the signal is presented in Figure 1.



Horizontal axis:
offset in file
Vertical axis:
8 bit unsigned value

Figure 1: Amplified and digitized ion chamber signal (count = 257)

Circuit Description

The electronic circuit built to extract random bits from the radioactive source is constructed in several sections. The sections are isolated by distance or aluminum shielding. There are two sections of power supply and two main sections consisting of analog and digital components.

The schematic is shown on the next page. The first power supply section is a standard unshielded commercial switching DC supply. To help eliminate AC line coupling this supply was located about 1m from a set of voltage regulators used in the second stage.

In addition to ±9 volt regulation for the amplifiers, an independent 9 volt supply is used for the ionization chamber and its buffer amplifier. This eliminates feedback from the amplifier stages which have an overall gain of about 1 million.

The second stage supply is mounted in a small cast aluminum box. The box is mounted to a large aluminum sheet metal box. A hole drilled through both boxes delivers reasonably noise free power to the analog stage.

The analog section was constructed on a “breadboard”. Surface mount op-amps were soldered onto “surf boards” and plugged into the breadboard. The +9 and ground lines were attached to the two bus lines on the breadboard. In addition, the ground line was also attached to the large aluminum shield.

The ionization chamber is drawn as a gas filled tube in the schematic. The floating grid is buffered with a single supply instru-

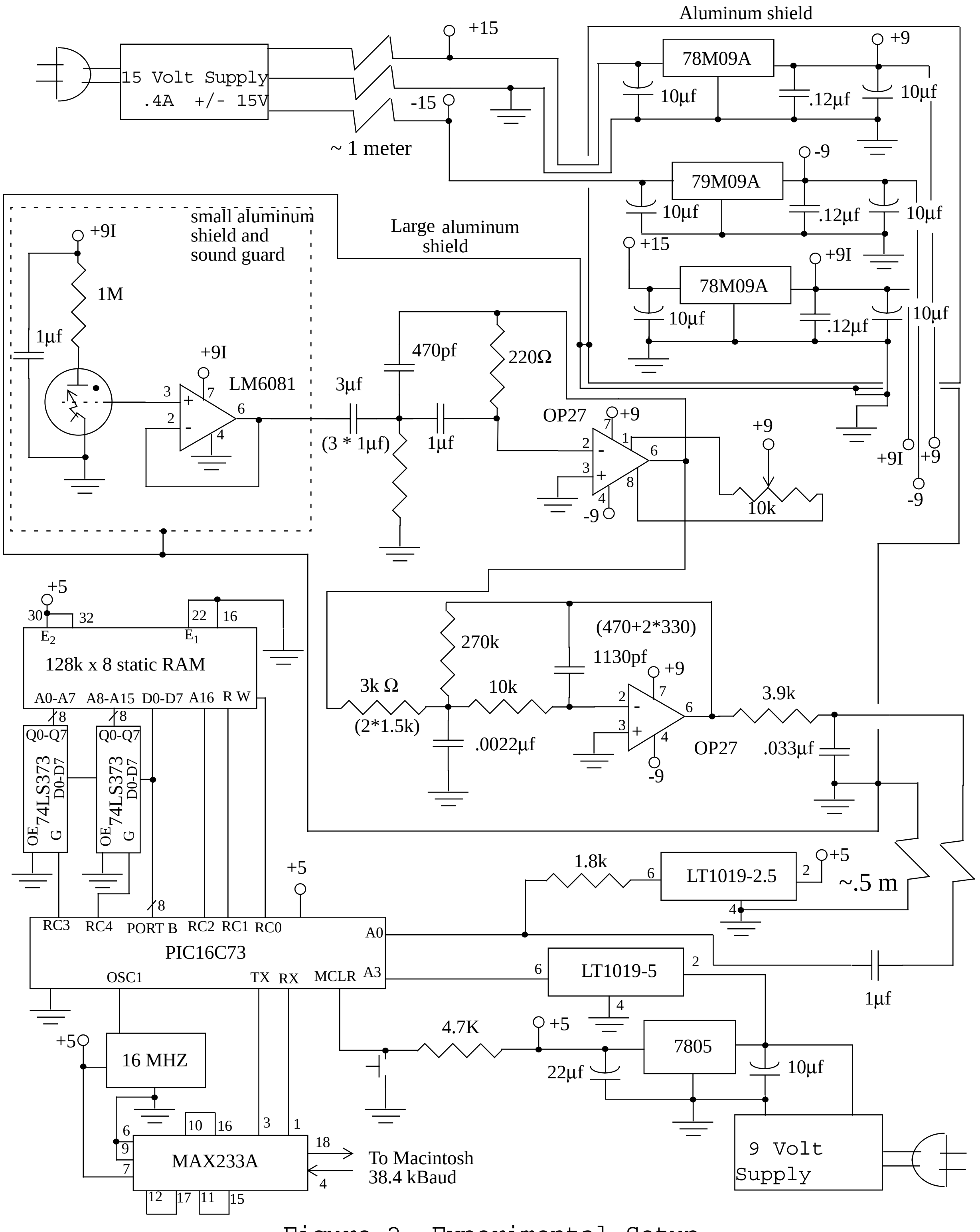


Figure 2: Experimental Setup
Experimental Setup

mentation op-amp. The op-amp is mounted within the shield of the ionization chamber. The chamber is insulated with a layer of electrical tape and that in turn is wrapped in several layers of aluminum foil. The power supply, ground and signal lines are all twisted together and brought to the amplifier breadboard.

In addition to electrical shielding the ionization chamber was surrounded by dense packing foam to help remove external sound signals. The only external signal visible at the amplifier output was from pounding the table which holds the equipment. The ionization chamber is a poor microphone but some external sounds may have been coupled to the signal. Whether or not this is significant requires further testing.

The analog amplifier section consists of a high pass first stage and a low pass second stage with an overall gain of 5.7×10^5. The filters are second order overdamped. The op-amps were chosen for low noise and small output offset voltage. While the offset adjust was included on the first stage amplifier it did not seem to be necessary.

The first stage parameters include a gain of 6400, corner frequency of 1.5 kHz and Q of 2.5×10^{-4}. The second stage has a gain of 90, corner frequency of 1.9 kHz and a Q of .06. A final passive stage follows with a single order low pass filter with corner frequency of 1.2 kHz. Frequencies above 10 kHz have amplitudes 40 dB below the peak at 1.5 kHz and are undetectable by the 8 bit digitizer.

The output signal ranges between ±4.6 Volts but this amplitude is mostly below 100 Hz. The input capacitor to the digital stage acts as a high pass filter and reduces this to ±2 Volts, well within the digitizer’s range.

The amplified signal is passed thorough a small hole in the side of the aluminum shield and wrapped around a ground wire connecting the shield with the digitizer section ground. The digital section is located about 50 cm from the amplifier section. It is mounted on a 10 cm x 12 cm perfboard, all connections are wirewrap.

The analog input to the microcontroller has a 5 Volt band gap voltage reference for a maximum and 2.5 volt band gap voltage reference as a DC offset for the input signal. With the +9I supply removed from the ion chamber and all other supplies to the amplifier and digitizer turned on, the output from the A/D converter in the microcontroller is split between codes 0x7F and 0x80 about 75/25 per cent.

The static RAM chip is used to store sampled data for analysis. This section is not used when generating random bits. The data rate between the digitizer and host computer is only 3.8 kilobytes per second and the sample rate ranged from 15 kHz to 25 kHz.

Generating Uniform Random Data

Ideally, we seek to generate samples which resemble white noise [GMJ,SMK]. However it is clear from the signal shown in Figure 2 that the ionization events and the shape of each corresponding bump are similar. The frequency spectrum is clearly not flat (see figure 3 below) therefore some massaging of the data is required to realize a uniform random bitstream.

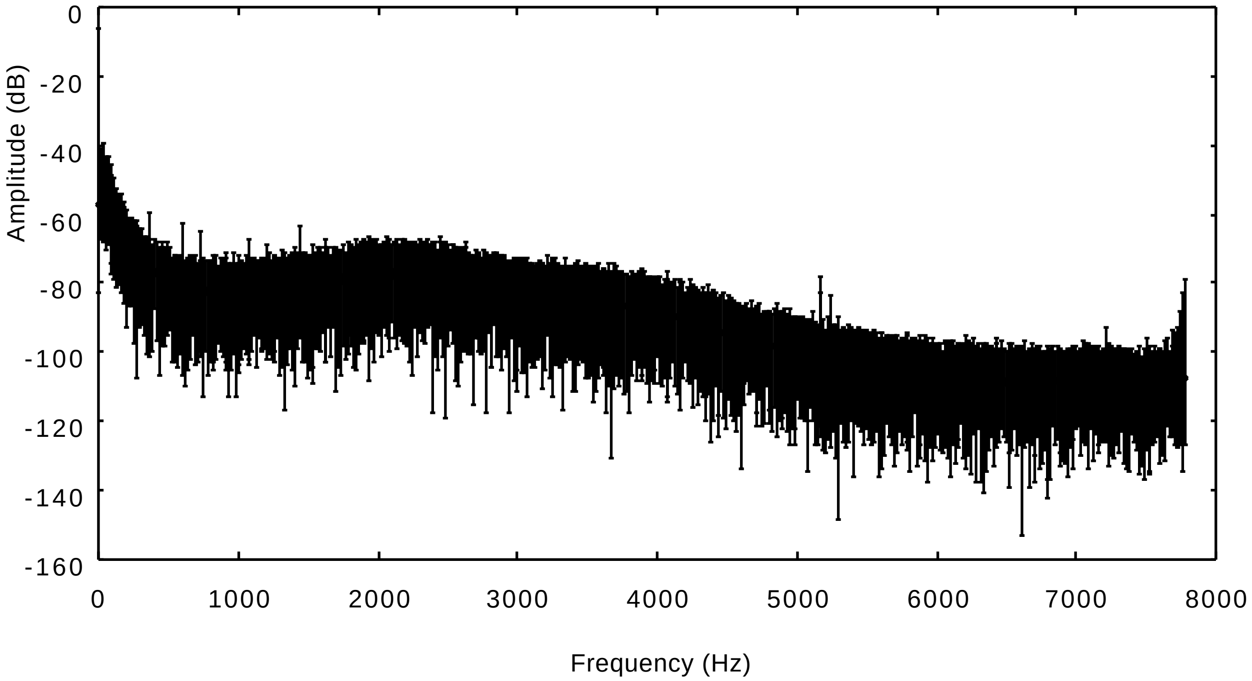


Figure 3: Amplitude vs. frequency at count = 257

Each point in figure 3 is the average of 128 frequency bins computed from a 128 kilosample FFT with the error bars marking the minimum and maximum values within those 128. The noise level is fairly obvious. The bump at 2 kHz is visible but not exceptional.

Distribution of Samples

A fundamental assumption of this study is that particle emission is a random stationary process. This implies a homogeneity of first-order statistics (mean, variance) and second-order statistics (covariance) of the distribution of random variates (rv's). In order to draw inferences from empirical data generated by the radioactive decay process it becomes necessary to identify, with a high level of confidence, an underlying theoretical distribution. To this end we view the generated data as random numbers which are a realization of random variates from a known distribution.

It has long been known that particle emission from radioactive decay occurs as a Poisson process [FAH,REA] i.e. emissions follow the Poisson distribution with probability density function (pdf)

$$\Pr(X = x) = \frac{((e^{-\lambda})\lambda^{x})}{x!} \qquad (x = 0, 1, 2, \ldots)$$

where $\lambda > 0$ is the expectation or mean and x the observed random variable. If we let $\lambda = \nu t$ then ν may be interpreted as the mean rate of decay per unit time. Our sample of Am241 has a mean decay rate of 3.7×10^4 per second. In theory, a sampling interval of say 27 microseconds would correspond to a mean number of events $\lambda \sim 1$ per interval.

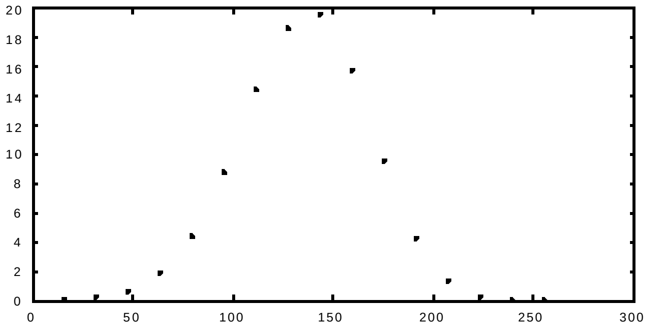
However, as previously pointed out, observable ionization events occur at a mean rate of approximately 2.685×10^3 per second which corresponds to 1 event per 372.4 microsecond interval. It is clear therefore that using a sampling interval of 27 microseconds, we would measure approximately 1 event only in roughly 14 intervals. Given that observations were in fact collected more frequently than this we must conclude that, in addition to the ionizations themselves, we were measuring some residual energy which persists between ionizations.

As a result, we would expect to observe not a Poisson process, but random deviates which follow a different exponential distribution, for example the Gaussian (normal) which has pdf:

$$f(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

where μ is the mean and σ^2 is the variance.

Using 8-bit integers, the observations were tabulated in frequency histograms to assist in identifying a parent distribution. Figure 4 is an example and as shown, the observed frequency distribution has an exponential form. Similar histograms from additional samples provided convincing evidence for having generated random numbers consistent with theory. Further, the empirical distribution closely resembles a skewed normal.



```
interval=27usec
frequency=1.85kHz
sample size=1,000,000
mean=127.21
variance=1021.89
skewness=-0.182
kurtosis=3.015
maximum=247
minimum=0
```

Figure 4. Frequency distribution of samples for count=108

The normal distribution, being quite tractable in linear transformations, is commonly used to approximate other distributions (JKP,NLJ). The next step therefore was to convert normally distributed data to a realization of independent and identically distributed (iid) random variates of the uniform, or rectangular distribution.

Linear Transformations

Many examples exist in the literature of transforming rv's from one distribution to another [FAH,JKP,NLJ]. In order to preserve their random properties the operations must be linear. The simplest method as it turns out is also very effective. This is the process of local averaging, or aggregation [NLJ] as described by Vanmarcke [EV] and which, for a continuous parameter stationary process X(t) with mean μ and variance σ^2, a family of "moving average" processes $X_T(t)$ may be obtained as follows:

$$X_T(t) = \frac{I_T(t)}{T}$$

where

$$I_T(t) = \int_{(t-T/2)}^{(t+T/2)} X(u)du$$

and T denotes the averaging time. The principal results remain valid when X(t) is sampled discretely as in this study and $I_T(t)$ used in place of $X_T(t)$. The mean μ is not affected by the averaging process, however the mean by integration is $\mu_{I_T} = T\mu$. The variance of $X_T(t)$ is expressed as $(\sigma_T)^2 = \gamma(T)\sigma^2$ where $\gamma(T)$ is the variance function of X(t) [EV] which measures the reduction of the point variance σ^2 under local averaging. Since the random functions $X_T(t)$ and $I_T(t)$ differ only by a factor of T and their respective variances by T^2, the variance of $I_T = (T^2)(\sigma_T)^2 = (T^2)(\sigma^2)\gamma(T)$. Simply put, by choosing local integration over local averaging we expect to avoid excessive smoothing of the data which would result from a large reduction in variance. The overall result is an increase in the disorder of the system. With repeated local aggregation disorder tends toward chaos [EV], a desirable outcome in generating random numbers.

In addition, by using the results of Mardia [KVM, p. 51], the ionization measurements can be viewed as directional data from a distribution on the circle. Then by including a moving total in the local integration process it quickly converges to a uniform distribution as follows:

$$y_i = \left(S = S + \sum_{j=1}^{k} x_j \right) \quad \mathrm{mod}\ 2^n$$

where y_i is the next uniform random number on $(0,2^n)$, S is initially set to 0, x_j is the next observation, k is the integration (summation) interval and n is the bit length of the y_i. The reader is also referred to the work of Petrov [VVP] on the convergence of sums to the uniform distribution. With k set to 11 and n set to 8, x_j was derived from the raw A/D input by first subtracting 128. The output y_i was derived from the lower 8 bits of S and transmitted to a computer with storage capability. It was found that, by using the parameters shown above, data could be transmitted at approximately the same rate as it was collected and still pass all the tests described below.

Tests for Randomness

When testing numbers for randomness, the primary objective is to determine how well the samples emulate random deviates from a specific theoretical distribution i.e., is the empirical data in fact a realization of the theoretical distribution? To this end we must determine whether the data resemble iid rv’s. For independence there must be no local or long range correlation within a sample and no more than a first-order (or one-step) Markovian dependency between values. Identically distributed rv’s are ergodic in that all information about the probability distribution (and its statistics) can be obtained from a single realization of the distribution [EV].

Numerous tests exist for determining the random properties of empirical data. For our purpose we classify the more important of these into two groups:

1. Traditional. This includes the runs test [DEK,ND,REA], the spectral (or lattice) test [RRC,DEK], the chi-square test for uniformity [LO,REA], the Kolmogorov-Smirnov (KS) deviation test [RVM,GMJ,REA] and the discrete Fourier transform (DFT) [AR, NR, OEB, SMK].

2. Modern. Diaphony, which is a weighted spectral test suitable for all RNG's [PH1,PH2], the serial n-dimensional digit test [HL] and Diehard [GM] which is a group of fifteen powerful tests built into one computer program.

Tests were selected for use from the two groups above according to two basic requirements:

1. The need to test raw data from the RNG during the construction stage and also to screen data prior to generating very large samples for final acceptance. This was carried out with runs, chi-square, KS and digit tests. DFT using the fast Fourier transform (FFT) was used for signal testing and also to compute autocovariance and autocorrelation estimates.

2. Final acceptance was contingent upon passing all fifteen Diehard tests. Diaphony, being a theoretical test like its predecessor the spectral test, served as a double-check. Although Diaphony may not yet be in wide use it was found to be very accurate in assessing long range statistics i.e. testing non-overlapping s-tuples for correlation in s-dimensional space.

Analysis of Results

Although uniformity was achieved relatively early, it was found that local and long-range correlations persisted in later samples (from counts 160-200, defined below). This was disclosed by several of the screening tests and confirmed by Diehard. A summary of these results appears in Table 1.

The microcontroller collected the data from a fixed timebase of four megahertz. An internal 16 bit register was set to a specific value that forced the A/D converter to grab the next sample of analog signal. The value "count" in Table 1 is the value placed in the compare register, divide by 4 x 10^6 to compute sample time.

Following the change to count 257 a 40% reduction in variance of raw data was achieved. As seen in the following statistics the distribution remained approximately normal with an increase in kurtosis of about 0.06 i.e. narrower. This was done to ensure no saturation of the signal at the 0 or 255 limits. In addition, for transformed data, the maximum coefficient of correlation had dropped to 0.0549 and the Diaphony normalized mean had converged to 1.000 for dimension 2 (optimal), 0.998 and 0.992 for dimensions 3 and 4 respectively.

Table 1 provides a summary of distribution statistic corresponding to changes in experiment parameters. Beginning with count 233, samples passed all Diehard tests, however only results for the most important, the Monkey tests (GM), are shown. (Monkey tests include Bitstream, OPSO, OQSO and DNA.)

Scaling up

The methods described in this paper can easily be scaled up to higher data rates. One method described in the early 1970's [HFM] is to have many devices in parallel. We can also increase the solid angle of detection by constructing the ionization chamber with a cylindrical geometry rather than flat.

Combining these two ideas, it is clear that multiple grids in the ion chamber will give multiple data sets. To help increase immunity to outside noise the ion chamber can be encased in a sealed chamber which has a different atmosphere than STP. One experiment which should be investigated is the signal level as a function of gas type and pressure. If the ion chamber has more ionized plasma than neutral gas there may be more cross talk between grids, but there may also be larger signals at much higher frequencies.

Finally, the activity can easily be increased. By using more than one isotope and decay mode, the noise level can be increased several million times. A spent fuel pellet from a nuclear reactor is only 1 cm^3 but generates several curies even after 10 years of cooling [GGE]. Clearly more research is needed, but data rates in excess of 100 megabytes per second are possible.

Statistics	27 µsec	40 µsec	50 µsec	58 µsec	63 µsec	64 µsec	67 µsec
count	108	160	200	233	251	257	267
Freq. (kHz)	1.845					3.525	
Mean	127.205					127.111	
Variance	1021.89					613.444	
Skewness	-0.182					-0.183	
Kurtosis	3.015					3.075	
Maximum	247.0					223.0	
Minimum	0.0					0.0	
Correlation	0.069	0.068	0.067	0.063	0.056	0.055	0.069
Diaphony							
dim 2	1.180	1.044	0.989	0.994	0.942	1.000	0.955
dim 3	1.131	1.026	0.989	0.993	0.998	0.998	1.019
dim 4	1.010	0.993	0.998	0.997	1.003	0.992	1.004
Diehard Monkey test							
lower p	0.000	0.000	0.000	0.025	0.026	0.022	0.002
upper p	1.000	1.000	1.000	0.999	0.996	0.993	0.994
Uniformity							
lower p	0.000	0.002	0.538	0.133	0.025	0.164	0.157
upper p	0.933	0.997	0.977	0.719	0.860	0.963	0.890

Table 1: Summary of test results corresponding to parameter changes

References

[AR] A. Ralston and H.S. Wilf, “Mathematical methods for digital computers” volume 1, John Wiley & Sons, Inc., 1967

[CHV1] C.H. Vincent, “The generation of truly random binary numbers”, Jour. Phys. E, Scientific Inst., V3, 1970, p594-598

[CHV2] -, “Precautions for accuracy in the generation of truly random binary numbers”, Jour. Phys. E, Scientific Inst. V4, 1971, p825-828

[DEK] D.E. Knuth, “The art of computer programming volume 2 Semi-Numerical Algorithms”, Addison-Wesley, 1981

[EV] Erik Vanmarcke, “Random fields: analysis and synthesis”, MIT Press, 1988

[FAH] Frank A. Haight, “Handbook of the Poisson distribution”, John Wiley & Sons, Inc., 1967

[FFC] F.F. Chen, “Introduction to Plasma Physics”, Plenum Press, 1976, chap. 4

[GGE] G.G. Eicholz, “Environmental Aspects of Nuclear Power”, Ann Arbor Science, 1976

[GM] George Marsaglia, “Monkey tests for random number generators”, Computers & Mathematics with Applications, 9, 1-10, 1993

[GMJ] G.M. Jenkins and D.G. Watts, “Spectral analysis and its applications”, Holden-Day, 1968

[HFM] H.F. Murry, “A General Approach for Generating Natural Random Variables”, IEEE Trans. Computers, Dec. 1970, p1210-1213

[HL] Hannes Leeb, “On the digit test”, ACPC Technical Report Series Nr. ACPC/TR 95-4, August 1995

[KVM] K.V. Mardia, "Statistics of directional data", Academic Press, 1972.

[JKP] Jagdish K. Patel and Campbell B. Read, “Handbook of the normal distribution”, Marcel Dekker, Inc., 1982

[LO] L. Ott, “Introduction to statistical methods and data analysis”, PWS-Kent, 1988

[ND] N. Draper and H. Smith, “Applied regression analysis”, John Wiley & Sons Inc.,1966

[NLJ] Norman L. Johnson, Samuel Kotz and N. Balakrishnan, “Continuous univariate distributions” volume 1, John Wiley & Sons, Inc., 1994

[NR] “Numerical Recipes in C: the Art of Scientific Computing”, Cambridge University, 1992

[OEB] O. E. Brigham, “The fast Fourier transform”, Prentice-Hall, 1974

[PH1] Peter Hellekalek, “Correlations between pseudorandom numbers: theory and numerical practice”, ACPC Technical Report Series Nr. ACPC/TR 95-4, August 1995

[PH2] -, “General discrepancy estimates V: Diaphony and the spectral test”, preprint, Institute of Mathematics, University of Salzburg, Austria, 1995

[REA] Research and Education Association, “The statistics problem solver”, 1994

[RRC] R.R. Coveyou and R.D. MacPherson, “Fourier analysis of uniform random number generators”, JACM 14, 1967

[RVM] R. von Mises, “Mathematical theory of probability and statistics”, Academic Press, 1964

[SMK] S.M. Kay, “Modern spectral estimation”, Prentice-Hall, 1988

[VVP] V.V. Petrov, "Convergence of independent random variables", Springer-Verlag, 1975.