THE DEVELOPMENT OF A HARDWARE RANDOM NUMBER GENERATOR FOR GAMMA-RAY ASTRONOMY

R. C. BOTHA B.Sc.

Dissertation submitted in partial fulfilment of the requirements for the degree Master of Science in Physics at the North-West University

Supervisor:	Prof. O. C. de Jager
Assistant Supervisor:	Dr. C. J. Reinecke

January 2005
Potchefstroom Campus

Acknowledgements

- Many thanks to my supervisor, Prof. De Jager, for his guidance and patience. Many illuminating discussions of various fields covered by physics is greatly appreciated and equipped me with the correct mental attitude towards further studies in this field.
- Thanks are also extended towards the staff of the School of Physics at the North-West University, Potchefstroom Campus, for all the extra support they provided and their help in smoothing out administrative processes.
- My gratitude towards the Hartebeesthoek Radio Astronomy Observatory for providing me with ample time and resources to complete my last year of study while employed there.
- Especially to my family, for their continuous support and assistance during all my studies thus far, I express the greatest appreciation.
- *The financial assistance of the Department of Labour (DoL) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DoL.*

Abstract

Pulsars, as rotating magnetised neutron stars got much attention during the last 40 years since their discovery. Observations revealed them to be gamma-ray emitters with energies continuing up to the sub 100 GeV region. Better observation of this upper energy cut-off region will serve to enhance our theoretical understanding of pulsars and neutron stars.

The H-test has been used the most extensively in the latest periodicity searches, whereas other tests have limited applications and are unsuited for pulsar searches. If the probability distribution of a test statistic is not accurately known, it is possible that, after searching through many trials, a probability for uniformity can be given, which is much smaller than the real value, possibly leading to false detections. The problem with the H-test is that one *must* obtain the distribution by simulation and cannot do so analytically.

For such simulations, random numbers are needed and are usually obtained by utilising so-called pseudo-random number generators, which are not truly random. This immediately renders such generators as useless for the simulation of the distribution of the H-test. Alternatively there exists hardware random number generators, but such devices, apart from always being slow, are also expensive, large and most still don't exhibit the true random nature required.

This was the motivation behind the development of a hardware random number generator which provides truly random U(0,1) numbers at very high speed and at low cost. The development of and results obtained by such a generator are discussed. The device delivered statistically truly random numbers and was already used in a small simulation of the H-test distribution.

KEY WORDS: H-test, pulsar searches, periodicity searches, gamma-ray astronomy, random number generator, true randomness.

Opsomming

DIE ONTWIKKELING VAN ‘N HARDEWARE KANSGETAL GENERATOR VIR GAMMA-STRAAL ASTRONOMIE

Gedurende die laaste 40 jaar sedert hul ontdekking het pulsare, as roterende gemagnetiseerde neutronsterre, baie aandag gekry. Hulle is observeer as gamma-stralers met energië tot in die sub 100 GeV gebied. ‘n Meer indiepte observering van hierdie boonste energie afsnit gebied sal ons teoretiese kennis van pulsare en neutronsterre verbeter.

Hedendaags word die H-toets uitgebreid gebruik in soektogte na periodisiteite, waar ander toetse beperkte toepassing het en ongeskik is vir pulsarsoektogte. As die waarskynlikheidsverdeling van ‘n toetsstatistiek nie akkuraat bekend is nie is dit moontlik dat, na deur baie toetse gesoek is, ‘n waarskynlikheid vir uniformiteit gegee kan word wat heelwat kleiner as die ware waarde is en kan lei tot valse deteksies. Die problem met die H-toets is dat die distribusie met simulasies verkry *moet* word aangesien dit nie analities moontlik is nie.

Vir sulke simulasies word kansgetalle benodig en wat gewoonlik verkry word deur van skyn kansgetal generators gebruik te maak, wat sagteware tegnieke gebruik. Wanneer heelwat ($>10^4$) sulke getalle gebruik word, duik probleme gewoonlik op aangesien die generators nie waarlik stogasties is nie. Sulke generators is dus nie geskik vir simulasie van die distribusie van die H-toets tot op ‘n beter as die huidig bekende vlak van 10^{-8} nie. As alternatief bestaan daar hardeware kansgetal generators maar sulke toestelle, bo en behalwe dat hul altyd stadig is, is ook duur, groot en meeste toon steeds nie die ware stogastiese eienskap nie.

Dit het as motivering gedien vir die ontwikkeling van ‘n hardeware kansgetal generator wat ware ewekansige U(0,1) getalle teen ‘n baie hoë spoed en lae koste lewer. Die ontwikkeling daarvan en die resultate verkry vanuit so ‘n generator word bespreek. Die toestel het statistiese ‘n ware stogastiese uitset gelewer, wat reeds in ‘n klein simulasie van die H-toets se distribusie gebruik is.

SLEUTELWOORDE: H-toets, pulsarsoektogte, periodisiteitssoektog, gamma-straal astronomie, kansgetal generator, ware kansgetalle.

Contents

Chapter 1

Introduction

Gamma-ray astronomy had a spectacular development in the last few years. This started with the CGRO EGRET telescope (Fichtel et al. 1983) operational from 1991 to 1999, which revealed the very high energy universe much richer than expected. Of the EGRET catalogue of 271 sources, 6 are galactic pulsars. At about the same time, ground-based gamma-ray astronomy came into being with instruments like Whipple (Krennrich et al. 1997) and CANGAROO (Kifune et al. 1997), which progressed to instruments such as VERITAS (Weekes et al. 1997), MAGIC Telescope (Petri & The MAGIC Telescope Group 1999) and H.E.S.S. (Krawczynsky & H.E.S.S. Collaboration 1999). Recently, very high energy gamma-rays resulting from particle acceleration in the shell of a supernova remnant was discovered with H.E.S.S. (Aharonian et al. 2004). All these instruments work on the principle that gamma radiation entering the atmosphere initiates particle showers which emit Cerenkov light. This Cerenkov light is then detected. A newer space gamma-ray telescope called GLAST is underway (Wood, Michelson & The GLAST Collaboration 1995) which should discover many gamma-ray pulsars.

Pulsars were named so because of the observed regular, high stability pulses of emission radiated in a certain direction. Pulsars, as rotating magnetised neutron stars, have received much attention during the last 40 years since their discovery by Hewish et al. (1968). Observations revealed them to also be gamma-ray emitters with energies continuing up to the sub-100 GeV region, depending on the theoretical model used. Better observation of this upper energy cut-off region will serve to enhance our theoretical understanding of pulsars and neutron stars. For the studies to become more generalised, one must include the data in this cut-off region from even more than the currently 1400+ known pulsars to see if it fits the predictions of certain theoretical models. Also, the pulsar flux in the energy cut-off region is quite low on reaching the Earth. Even though attempts are made to maximise the collection area of detectors and minimise the acceptance of background events, further improvement of searches for pulsed emission can be achieved if optimised search techniques are employed.

One way to achieve this is by implementing statistical tests for uniformity, of which the H-test (De Jager et al. 1986) has been used the most extensively in the latest periodicity searches. De Jager et al. (1989) discussed the merits of a number of statistical tests for uniformity for pulsar detections: the H-test was found to be powerful against a wide variety of pulsar pulse profiles and has been used to detect a number of X-ray and gamma-ray pulsars (Hessels et al. 2004; Kaspi et al. 2000; Chang & Ho 1997). Other tests like the Z_m^2 test (Buccheri et al. 1983) have limited applications and are unsuited for pulsar searches. This applies specifically because some pulsars may be radio quiet (like Geminga) so that one does a blind search in γ-rays. To effectively employ the H-test to the high levels of accuracy required, the distribution of this test must be known very accurately. Unfortunately the problem with the H-test is that it is impossible to obtain an analytical distribution; one must obtain the distribution by simulation. Currently the distribution is known up to an accuracy of $\sim 10^{-8}$, but even for modern instruments, this is not accurate enough. This comes as a result of the large collection areas and high sensitivities, resulting in many more events and the signal still being hidden in the noise. To be able to detect fainter pulsars, the H-test must be accurate up to a level of 10^{-10} or better. Simulation up to this level is therefore required, i.e. $> 10^{10}$ random numbers are needed to obtain a better distribution of the H-test than previously known.

If the probability distribution of a test statistic is not accurately known, it is possible that, after searching through many trials, a probability for uniformity can be given, which is much smaller than the real value, possibly leading to false detections such as discussed by De Jager et al. (1988) and finally proved by Nel et al. (1993). Therefore a false detection may be claimed- even after the proper number of statistical trials have been taken into account. A proper evaluation of the H-test is due, given the fact that it is already widely in use as discussed above.

For simulations, random numbers are usually obtained by utilising pseudo-random number generators, which employ software techniques. On levels of $> 10^8$ numbers certain problems with these generators, which are due to them not being truly random, start to occur. This immediately renders such generators as useless for the simulation of the distribution of the H-test as required here. Alternatively, there exist hardware random number generators, but such devices, apart from always being slow, are also expensive, large and most still don't exhibit the true random nature required. This motivated the

Unit for Space Physics at the North-West University to develop and implement a hardware random number generator which has as features:

1. Truly random number output of U(0,1) distribution
2. Very high speed
3. Relatively low cost

The first prototype of this device was used in the simulation of the distribution of the H-test. The device was also awarded a patent during 2004 for the unique implementation and plans to further the commercial development of the device were successfully implemented thus far. Random numbers have a wide variety of applications, from the scientific need for simulations, to the security of data which uses random numbers to encrypt data, to the modelling of financial market changes and long-term effects by employing Monte Carlo techniques. Potential applications of such a device are therefore widespread and it provides an important spin-off from the research done.

This dissertation is split into two main sections: chapter 2 on gamma-ray astrophysics and chapter 3 on random number generation for astrophysical purposes. In chapter 2 we progress from a general introduction to gamma-ray astrophysics towards pulsars and a basic model of how the radiation is created. After considering the emission spectra, the Atmospheric Cerenkov Technique is discussed as detection method of the interaction of high energy radiation with the Earth's atmosphere. A further discussion of the H.E.S.S. ground-based detector of such Cerenkov light follows after which it is placed in the framework of worldwide past, present and future gamma-ray astronomy ventures.

As discussed, the H-test is the method employed for pulsar searches and since it has its origins in statistics we start off in chapter 3 with a short overview of the relevant statistical theory. A general framework for tests of uniformity is discussed after which the H-test is discussed as one specific case. From the nature of the test it is then clear that no analytical distribution exists for the test. Therefore we move on to the problem of generating the random numbers in an appropriate fashion for the simulation of the distribution of the H-test. This is achieved by first considering problems surrounding most random number generators, after which the development towards a device with the required properties follows. The Quantum Bit Extractor is the first step in such a direction and a discussion of the employed statistical tests for randomness follow. Problems with the basic Quantum Bit Extractor pushed us towards a more in-depth statistical analysis as to where the problems originate and how to compensate and correct for such problems. A complete discussion of this process towards the final hardware implementation of a device providing statistically truly random numbers forms the heart of this dissertation. The applicability of such a device covers wide fields in various sectors and a general overview of this follows. Since the initial motivation for such a device originated from the need to simulate the H-test's distribution, a short discussion of the obtained distribution is given.

The code used to simulate the H-test distribution under null hypothesis in given in Appendix I in the Borland Delphi 5 programming language. Appendix II contains the patent description of the device. Appendix III contains a list of acronyms and abbreviations used in this text.

Chapter 2

Gamma-Ray Astrophysics

2.1. Gamma-Ray Astrophysics in general

In 1911 Victor Hess established the existence of a mysterious radiation of *"extremely high penetrating power"* entering the atmosphere from space, which started the eventual development of current very high energy (VHE) TeV gamma (γ)-ray astronomy. This radiation was called "*cosmic rays*", and it was only realized at a later stage that cosmic rays consisted mostly of particles. Forty-two years later the first ground-based detection of the Cerenkov radiation, associated with γ-rays (see section 2.3), was made by Galbraith & Jelley (1953).

Primary cosmic rays consist mainly of nuclei, some electrons, positrons, neutrinos and γ-rays, all with energies ranging from 10^8eV - 10^{20}eV. Charged cosmic rays are deflected by the interstellar and intergalactic magnetic fields. Thus detection of such a particle gives no indication as to where or by what specific mechanism they originated, or the distance travelled. The arrival times of the charged component of cosmic radiation can be treated as random events with a uniform distribution in both space and time. Therefore the sources of most of the cosmic rays are still unknown and the solution to this is one of the major goals of VHE astronomy. There are however good theoretical reasons to believe that shell-type supernova remnants (SNRs) should be VHE γ-ray sources because they are also thought to be major sources of galactic cosmic rays (Völk 1997).

Theoretically, high-energy γ-rays are usually the result of particle acceleration in collective processes involving wave and particle interactions. This interaction of charged particles and waves is equivalent to the reflection of charged particles on moving magnetic mirrors and the particle energy distribution is therefore not in thermal equilibrium. To observe cosmic rays is therefore to observe the non-thermal universe. Gamma-rays produced by charged particles accelerated in magnetic fields are, under certain conditions, unaffected by magnetic fields and then move along uncurved paths, allowing us to detect where interactions with cosmic rays occur. Thus to study γ-rays also serves towards studying cosmic rays and the regions where electromagnetic interaction with cosmic rays occur.

2.2. Pulsars as Gamma-Ray sources

A group of Cambridge astronomers headed by Anthony Hewish detected astronomical objects having pulsed radio emission in 1967 (Hewish et al. 1968). It was a significant event for subsequent astrophysical research and Hewish was awarded a Nobel prize in 1974 for the discovery. At the time of the discovery Pacini (1967) had already published a preliminary model of a simple magnetic dipole rotator capable of converting neutron star rotational energy into electromagnetic radiation, with his work supported by theories of Hoyle, Narlikar & Wheeler (1964), Tsuruta and Cameron (1966), Woltjer (1964) and Wheeler (1966).

The identification of pulsars with neutron stars was not immediately obvious to astrophysicists, but Gold (1968) argued that the observed pulsars were in fact rotating neutron stars with surface magnetic fields of $\sim 10^{12}$G. Shortly thereafter the slowdown of the Crab pulsar was discovered and he showed that the implied energy loss was approximately the same as the energy required to power the Crab nebula (Gold 1969). The success of Gold's model led to the acceptance of the rotating magnetized neutron star as the basis for all subsequent pulsar models.

Further observations of pulsars revealed them to be γ-ray emitters up to the sub-100 GeV region, the theoretical upper limit dependant on the pulsar magnetic field strength at the emitting regions (Figure 2.1) (Thompson 2000). Figure 2.2 shows the typical emission patterns of three well-studied pulsars. The double peak is a typical feature of observed pulsar emission and must therefore be explained by the theoretical models.

Associations between SNRs and neutron stars have also traditionally been identified with the detection of radio pulsars. Three such objects, the Crab, Vela and PSR B1509-58, have been included in a review by Helfand & Becker (1984). The associations between SNRs and neutron stars are often dubious because of a lack of supporting evidence of association, rather than evidence against association.

2.2.1. Pulsar models in general

The key observational facts of pulsars may be summarised from the ATNF Pulsar Catalogue (2004) and Shapiro & Theukolsky (1983) as follows:

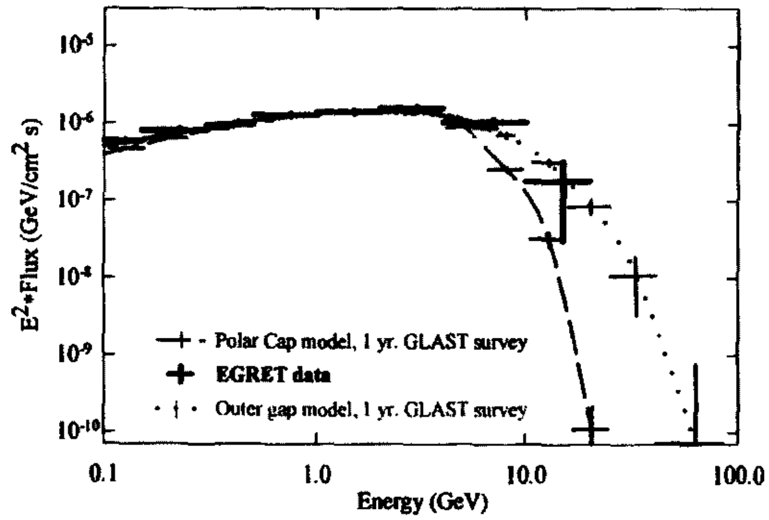


Figure 2.1. High-energy spectrum of the Vela pulsar showing a spectral turnover near 10GeV. The Polar Cap and Outer Gap models are discussed in section 2.2.1.

1. they have periods in the range 1.56 ms to 11.76 s,
2. the periods increase very slowly, except for occasional ‘glitches’,
3. pulsars are remarkable clocks, with some periods measured to 13 significant digits.

Only rotating neutron star models can explain all of the above observed features. Unfortunately, no single model developed for the pulse emission mechanism explains all of the observed features in the pulse profile. Nonetheless, the simple rotating-dipole model illustrates how pulses of high regularity may arise (Bowers & Deeming 1984), thus it will be the basis for the models discussed here.

Consider a rotating magnetic neutron star with mass $\sim 1M_\odot$ and radius $\sim$ 10 km. Assuming that generally the rotation and magnetic axis of the pulsar makes an angle φ with each other, the speed of the magnetic field lines at a distance r from the rotation axis is expected to be $v = \omega \times \mathbf{r}$. The speed of the field lines can be greater than the speed of light because magnetic field lines do not exist physically. If particles are coupled to the magnetic field by some mechanism they will co-rotate with the field and then the restriction $|v| = |\omega \times \mathbf{r}| = \omega r \sin\theta \leq c$ applies, since the tangential velocity of the particles may not exceed the speed of light. This defines the light cylinder which is the furthest point from the rotation axis of the star where the matter can co-rotate with the magnetic field and it has a radius $r_L = c/\omega$ from the rotation axis (Bowers & Deeming 1984).

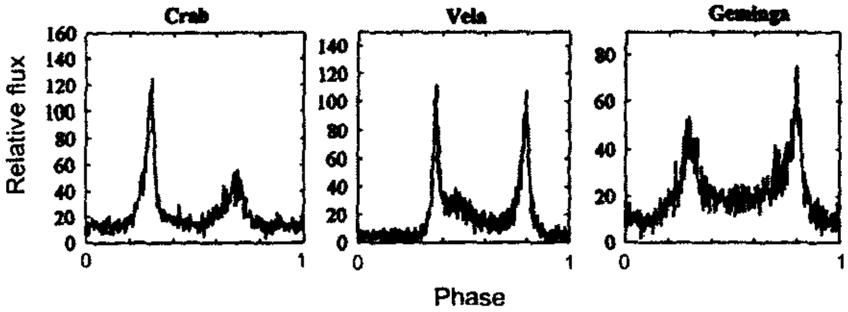


Figure 2.2. Full pulse profiles as from EGRET observations of pulsars for E > 100MeV (Kanbach 1997). The relevant spin periods are ~33 ms (Crab), ~89 ms (Vela) and ~237 ms (Geminga).

Using this approach, there exist both open and closed field lines: the closed field lines are those which do not cross the light cylinder, whereas those field lines, which in the absence of rotation would have closed at larger distances, penetrate the light cylinder and become open field lines. Charged particles will co-rotate with these field lines; for the closed field lines the particles are stationary relative to the field lines and for the open field lines the particles can move along them (Figure 2.3).

For simplicity we assume an aligned rotator (i.e. $\varphi = 0$)[1], following the case considered in Bowers & Deeming (1984), since this allows us to easily demonstrate emission mechanisms and estimate order-of-magnitude results: the characteristics of the magnetosphere must be similar to those of the non-aligned rotator (Mestel 1971).

We define the polar cap to be the region $\theta < \theta_p$ (using spherical coordinates) on the surface of the star, such that all the field lines crossing the light cylinder originate from this region. For a dipole field, $(\sin^2\theta)/r$ is a constant, thus for the angle θ_p defining the polar cap region and the last field line just touching the light cylinder (see Figure 2.3), we have

$$\frac{\sin^2\theta_p}{R} = \frac{\sin^2\frac{\pi}{2}}{r_L} = \frac{\omega}{c}$$

with R the radius of the star and r_L the radius of the light cylinder. From this the radius of the polar cap is

$$R_p \simeq R\sin\theta_p = R\left(\frac{\omega R}{c}\right)^{1/2} \tag{2.1}$$

Charged particles leaving the star's surface through this polar cap region can escape to infinity along the open field lines.

[1] This aligned rotator model however, cannot produce the observed radiation since it has been shown that an inclined rotator is essential (Michel 1997).

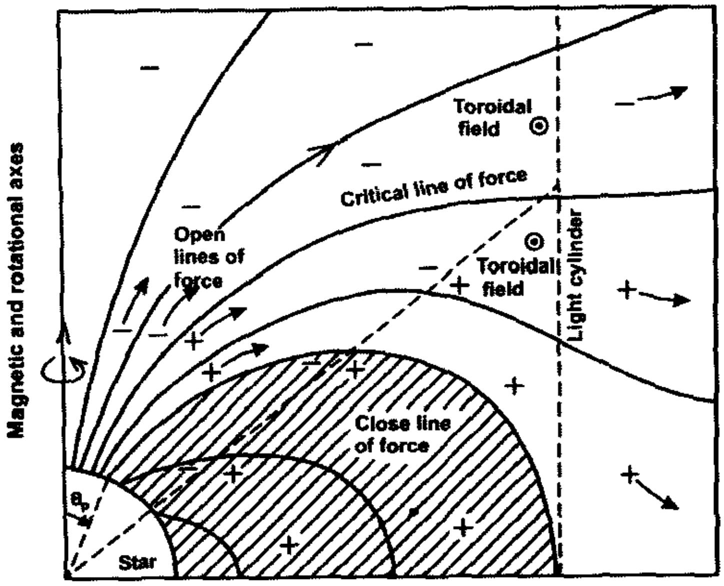


Figure 2.3. Magnetospere of an aligned pulsar showing the open and closed field lines. Charges of different signs are the result of the magnetosphere attempting to obtain a state of minimum energy.

To retrieve the equations for the electromagnetic fields in- and outside a pulsar we consider a frame at rest relative to a star with conductivity σ_*. Ohm's law relates the current density $\mathbf{j}_*$ and electric field $\mathbf{E}_*$ by

$$\mathbf{j}_* = \sigma_* \mathbf{E}_*$$

The reference frame in which $\mathbf{E}_*$ is measured moves relative to a frame at rest with respect to the Galaxy with velocity $|\upsilon| = |\omega \times \mathbf{r}|$. Performing a Galilean transformation from the star's rest frame to the Galaxy's rest frame yields

$$\mathbf{E}_* = \mathbf{E} + \frac{1}{c}(\upsilon \times \mathbf{B})$$

with $\mathbf{E}$ and $\mathbf{B}$ measured in the Galaxy's rest frame. The electrical conductivity σ_* of matter inside a neutron star is extremely large ($\sim \infty$) because of the degenerate state of the matter, therefore one can treat the pulsar as a rotating magnetized perfect conductor. Since $\mathbf{j}_*$ must be finite inside a conductor, $\mathbf{E}_*$ must be zero from Ohm's law and thus

$$\mathbf{E} = -\frac{1}{c}(\upsilon \times \mathbf{B}) = \frac{1}{c} \times \mathbf{B}(\omega \times \mathbf{r}), \tag{2.2}$$

which is the equation we must solve for inside a neutron star.

From the boundary conditions in electrodynamics, the component of the electric field parallel to the surface of the conducting star is continuous across the surface, thus the component of the electric field parallel to the star's surface, just outside the star, is zero. Since magnetic fields are also present on the surface of the star, the electric field is contained in the component perpendicular to the surface at the surface, because we deal here with a non-static magnetic field. For typical magnetic field strengths of 10^{12}G, we can make an order of magnitude estimate for the electric field strength outside the surface from (2.2) as

$$E \sim \frac{B\omega r}{c} = \frac{2\pi BR}{cP} \approx 10^{12} V.cm^{-1}$$

with P the pulsar rotation period.

The gravitational force F_g on the particles at the surface is much weaker than the force F_E on the particles due to the electric field:

$$\frac{F_E}{F_g} = \frac{qB\omega r}{c}\left(\frac{GMm}{R^2}\right)^{-1} \approx 10^9 \text{ to } 10^{12}$$

depending on the charged particles under consideration, with M the stellar mass and m the particle mass. Thus charged particles can be drawn out of the crust into the surrounding magnetosphere (Mészáros 1992; Lyne & Graham-Smith 1990). Charge flowing into the magnetosphere produces currents that induce additional electromagnetic fields which modify the structure of the magnetosphere. A self-consistent solution to the problem has not yet been obtained. We do however know that the induced electric field wants to arrange the configuration outside the star in such a way that the lowest possible energy configuration is reached, resulting in a zero net force on particles. The magnetosphere is highly conducting along but not perpendicular to the magnetic field lines, helping with the flow of charge in such a way as to reach equilibrium. This condition in the magnetosphere is similar to the high conductivity of the stellar interior and the magnetosphere seems to be an extension of the solid interior. In both regions, then, the induced electric field is cancelled by a static field, so that (2.2) is valid for outside the star as well. Note that this holds for the region of co-rotation of the matter and magnetic field lines, i.e. the region of the closed field lines. For the open field lines no static configuration is possible, there is a net outflow of charge and the electric field has a non-zero component parallel to the magnetic field, which also serves to accelerate the charges along these field lines.

For the aligned rotator, $\omega = \omega \hat{e}_z$ using cylindrical coordinates. Assuming that inside the star the magnetic field is

$$\mathbf{B}_{in} = B_0 \hat{e}_z \tag{2.3}$$

and outside we have a dipole field

$$\mathbf{B}_{out} = \frac{B_0}{2}\left(\frac{R}{r}\right)^3 (2\cos\theta\hat{e}_r + \sin\theta\hat{e}_\theta) \tag{2.4}$$

we have for the electric field inside the star from using (2.3) in (2.2) that

$$\mathbf{E}_{in} = -\frac{B_0\omega r}{c}\sin\theta\,(\sin\theta\hat{e}_r + \cos\theta\hat{e}_\theta) \tag{2.5}$$

Note that $\mathbf{E}.\mathbf{B} = 0$ inside the neutron star, as expected from a conductor. Using $\mathbf{E}_{in}$ and the boundary conditions at the surface, the electric field outside the neutron star is

$$\mathbf{E}_{out} = -\frac{B_0\omega R^5}{cr^4}\left(\frac{3\cos^2\theta - 1}{2}\hat{e}_r + \cos\theta\sin\theta\hat{e}_\theta\right) \tag{2.6}$$

Using (2.4) and (2.6) the component of $\mathbf{E}$ parallel to $\mathbf{B}$ outside the pulsar is

$$E_\parallel = \frac{\mathbf{E}_{out}.\mathbf{B}_{out}}{B_{out}} = -\frac{2\omega R}{3c}\left(\frac{R}{r}\right)^4 B_0\cos\theta \tag{2.7}$$

This is valid for the region of the open field lines where current flow occurs. For the closed field line region of the plasma we have a force-free charge distribution because the charge cannot escape and reaches a static configuration.

The magnetic field line which intersects the light cylinder at right angles[2] is called the critical field line (see Figure 2.3) and is assumed to leave the star's surface at polar angle θ_c. This critical field line is at the same electrostatic potential as the interstellar medium at the star's surface. For $\theta < \theta_c$ the field lines at the surface are at lower potential than the surrounding medium. Therefore electrons stream out along these magnetic field lines which pass through the polar cap of radius $r_c = R\sin\theta_c$. For $\theta > \theta_c$ the electrostatic potential exceeds the interstellar value and positive ions stream out along the field lines lying in the annular region $\theta_c < \theta < \theta_p$. The value of θ_c is fixed by the requirement that the net current flow through the polar cap must be zero. Therefore a negative current flows out along the the poles ($\theta \leqslant \theta_c$) and an equal positive current (protons and ions)

[2] Mathematically, $B_z(r_L) = 0$

flows out in the annular sheath $\theta_c < \theta < \theta_p$. Each current distribution induces a magnetic field that is toroidal about the magnetic axis. Near the light cylinder the toroidal field bends backwards as it passes the light cylinder and trails the co-rotating magnetosphere. The maximum energy an electron can obtain from the dynamo potential difference over the polar cap region, for an acceleration distance of order of the polar cap radius, is

$$\xi \sim eEr_c \leq 3 \times 10^{-2} \frac{r_c B}{P} \text{ eV} \sim 3 \times 10^{17} \text{eV}.$$

The total particle loss rate is

$$\dot{N}_p \sim \frac{\omega^2 B_0 R^3}{ec} \simeq 2 \times 10^{19} \frac{B}{P^2} \text{ s}^{-1}$$

according Mészáros (1992). For pulsars with a Goldreich-Julian charge density streaming out at relativistic speeds which utilizes the full potential difference across the polar cap, the particle energy per second produced is

$$L_p = e \bigtriangleup \Phi_{maks} \dot{N}_p = \frac{\omega^4 R^6 B_0^2}{4c^3}$$

with $\Delta\Phi_{maks}$ the maximum accelerating potential according Goldreich & Julian (1969). Bowers & Deeming (1984) have estimated the total electromagnetic energy radiated from the pulsar as

$$\frac{dE}{dt} \approx \left(\frac{\omega R}{c}\right)^3 \frac{B_0^2 R^3 \omega}{4\pi^2},$$

which is equal to the energy loss rate from a dipole in vacuum within an order of magnitude. Thus the total particle luminosity is proportional to the magnetic dipole radiation power:

$$L_p \propto \frac{dE}{dt}.$$

Taking general relativistic (GR) effects into account, the primary electron luminosity according to Venter (2004) is

$$L_{GR} = \frac{3}{4}\left(1 - \frac{1}{\eta^3}\right) \kappa \left(1 - \kappa\right) \frac{dE}{dt}$$

with η and κ dimensionless constants. This has the same functional form as with the non-GR case. A more detailed analysis of the aligned rotator can be found in Fitzpatrick

& Mestel (1998). The simplified model presented here must however be modified in at least two ways to obtain a reasonable model of pulsars:

1. it must be generalized to non-aligned rotation and magnetic axes,
2. a self-consistent model including the induced fields must be obtained and solved.

Great advances have been made with recent models explaining nearly all of the observational details, but no fully working model has as yet been developed (Arons 1996). Two popular models for pulsar emission are the polar cap and outer gap[3] models. The first places the source of emission immediately above a magnetic pole; the other places it far out in the outer magnetosphere, close to the light cylinder.

The polar cap (PC) model assumes that radiation is emitted primarily from the region of the field lines which delineate the polar cap, by charges being accelerated from the star's surface along these field lines.

The outer gap (OG) models for γ-ray pulsars assume the existence of a vacuum gap in the outer magnetosphere between the last open field line and the null charge surface ($\mathbf{\Omega}.\mathbf{B} = 0$) in charge-separated magnetospheres. These gaps arise because charges escaping through the light cylinder along open field lines above the null charge surface cannot be replenished from below.

If a pulsed flux is detected from young pulsars at TeV energies, polar cap models will be obsolete. If however, upper limits to pulsed flux above 100 GeV continue to decrease, outer gap models will be terminally constrained (Harding & De Jager 1997).

Roughly six or more γ-ray pulsars were observed by the CGRO/EGRET instrument during its mission between 1991 and 1997. Several hard-spectrum unidentified EGRET sources were also observed and are thought to be γ-ray sources for which the EGRET statistics are too small to resolve periodicity (Grenier 2001). With the H.E.S.S. telescope (see par. 2.4) we hope to observe several of these objects.

2.2.2. Radiation mechanisms

In both the PC and OG theories the location and direction of emission generated is mainly determined by the dipolar magnetic field. The high-energy radiation observed from pulsars is very broad-band, which is typical of synchrotron and curvature radiation. The radio regime is narrow-band, which is typical of coherent radio mechanisms (Lyne & Graham-Smith 1990). Primary accelerated particles at high enough energies give rise to

[3] This is also known as the relativistic beaming or vacuum gap model.

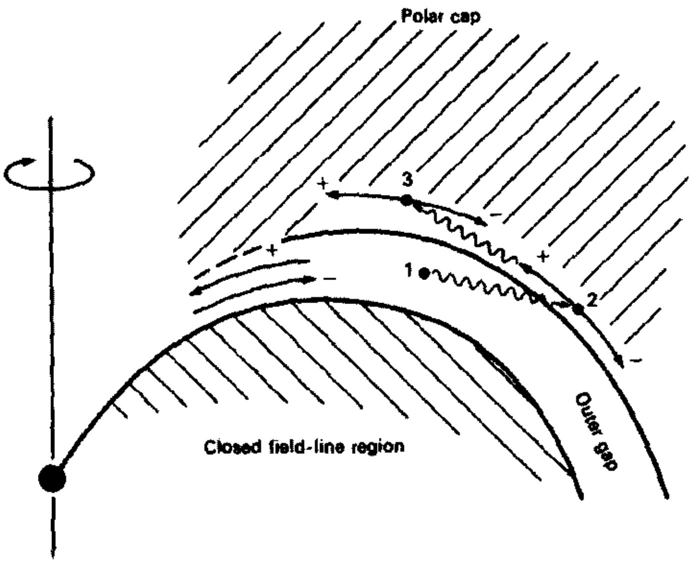


Figure 2.4. The cascade process in the outer magnetosperic gap. Electrons and positrons (-/+) accelerated in the gap emit γ-rays ($\sim\sim$), which, in turn, create e^+e^- pairs, the process moving progressively into the polar cap region.

e^+e^- cascades producing high-energy radiation in the strong magnetic field via the synchrotron or inverse Compton mechanisms. The contribution to pulsed γ-ray emissions from the inverse Compton mechanism is much less than that of curvature radiation (Harding & Muslimov 1997, Harding 2001).

In the pulsar magnetosphere the motion of the particles is in general a combination of both gyration about the field lines and streaming along them. The gyration of the particles about the field lines cause synchrotron radiation and the streaming of the particles along the field lines causes curvature radiation. Both the gyration and streaming cause a loss of energy for the particles. Due to the strong electric field, the particles have relativistic velocities, and being constrained to move along the field, the radiation is strongly beamed. This beaming, coupled with the rotation of the star gives rise to a pulsed emission pattern.

In PC models the basis of pulsed emission lies in the synchrotron and curvature radiation from particles as they are accelerated by the electric field (2.7) along the field lines delineating the polar cap. This can create γ-rays of high enough energy to produce e^+e^- pairs, giving rise to a cascading process described below.

In the OG model, thermal X-rays and soft γ-rays from near the neutron star surface interact with primary radiation and produce e^+e^- pairs. This pair production plays a critical role in the production of the high-energy emission: it allows current to flow and particle acceleration to take place in the gap (Harding 2001). The creation of an e^+e^- pair (at point 2 in figure 2.4) results from the interaction of a γ-ray with either the magnetic field or a lower-energy photon. The created particles accelerate along the field lines, reaching energies comparable to the available potential. In this process they radiate γ-rays, either by curvature radiation or by inverse Compton collisions with low-energy photons. These γ-rays can then create further e^+e^- pairs, giving rise to a cascading process. This process is discussed in more detail in section 2.2.3 on page 21.

Some idea of the balance between these processes can be obtained from a paper by (Cheng, Ho & Ruderman 1981). Since both theoretical models have emission associated with synchrotron and curvature radiation as well as inverse Compton mechanisms, we will consider all three physical processes in short.

2.2.2.1. Synchrotron radiation

Synchrotron radiation implies a change in the transverse momentum[4] of the particles. Since synchrotron radiation is relativistic cyclotron radiation we first consider the more simplistic cyclotron radiation process. Cyclotron motion is the circular motion of a charged particle in a magnetic field and is described by the cyclotron formula

$$evB = m\frac{v^2}{R}$$

with m the mass of a particle, charge e moving with non-relativistic speed v around a circle of radius R. The radiation from such a particle is at the Larmor frequency

$$\nu_L = \frac{eB}{mc}$$

For an electron or positron, $\nu_L = 2.8 MHz.gauss^{-1}$. The rate of energy loss through this radiation is

$$\frac{dE}{dt} = -\frac{8\pi^2\nu_L^2\beta^2e^2}{3c} \tag{2.8}$$

with $\beta = v/c$ (Lyne & Graham-Smith 1990). A polar diagram of the radiation is shown in Figure 2.5 a.

Harmonics of the cyclotron frequency are generated when the particle's circular Larmor orbit is distorted by the wave fields $\mathbf{E}_1$ and $\mathbf{B}_1$ which are first order approximations to

[4] with respect to the magnetic field **B**

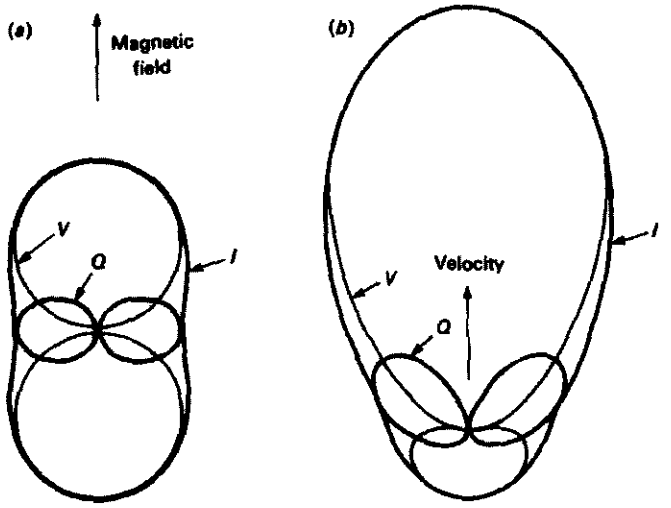


Figure 2.5. The radiation pattern (a) from an electron in a circular orbit perpendicular to the magnetic field, (b) from an electron streaming along the magnetic field. The total power (I), linearly (Q) and circularly (V) polarised are the Stokes parameters (after Lyne & Graham-Smith 1990).

oscillations in the fields. This usually happens when particles have relativistic velocities, reducing the gyration frequency

$$\nu_g = \frac{eB}{2\pi m}$$

below the Larmor-frequency because of the increased mass of the electron, thus distorting the Larmor orbit. Most of the radiated power now lies in the harmonics but the fundamental cyclotron frequency is still emitted with intensity (2.8) but with the actual gyro frequency ν_g substituted for ν_L.

A charged particle with high velocity, i.e. $\Gamma = (1-\beta^2)^{-1/2} \gg 1$, radiates a spectrum of harmonics which extends to frequencies of order $\Gamma^2\nu_L$ (i.e. to $\Gamma^3\nu_g$). When Γ is large this radiation may be regarded as a continuous spectrum. This is then synchrotron radiation. Consider the electric field radiated by a single electron, gyrating perpendicular to the magnetic field, and observed in its plane of orbit. This gives observable pulses, each occurring as the electron travels towards the observer. The relativistic velocity concentrates the field in the forward direction (Figure 2.6).

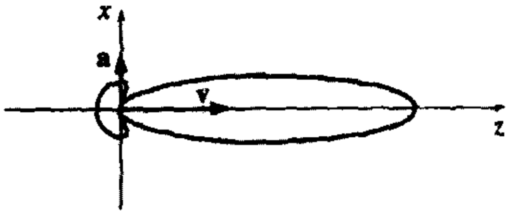


Figure 2.6. Radiation lobes for charges having relativistic velocities. The **v** indicates the direction of propagation and the **a** the direction of acceleration, this specific case referring to synchrotron radiation [28].

2.2.2.2. Curvature radiation

Curvature radiation implies a change in the longitudinal component of the momentum[5] of the particles. In the super-strong magnetic field of a pulsar magnetosphere, an electron may follow the path of a magnetic field line very closely, with pitch angle nearly zero. The magnetic field lines are generally curved due to their dipolar nature, so that the electron will be accelerated transversely and radiate along the tangential direction of the field line. This radiation, which is closely related to synchrotron radiation, is called curvature radiation.

Electrons gain energy by being accelerated by the electric field (2.7) along the magnetic field lines, having radiation lobes as shown in figure 2.5 b. These velocities are highly relativistic, so we have an energy gain rate

$$\left(\frac{dE}{dt}\right)_{gain} \simeq eE_{\|}c \tag{2.9}$$

and radiation lobes similar to that of synchrotron radiation depicted in figure 2.6. An electron with relativistic velocity, constrained to follow a path with radius of curvature ρ_c radiates in a similar way as in a synchrotron process and the theory will not be repeated here. The particle radiates with a typical maximum frequency of

$$\nu_0 = \frac{3c\Gamma^3}{2\rho_c}$$

[5] with respect to the magnetic field **B**

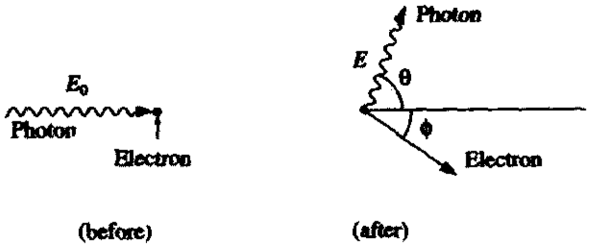


Figure 2.7. Mechanism of the normal Compton scattering. The process in reverse describes inverse Compton scattering. [28]

The rate of curvature radiation energy loss is given by (Mészáros 1992)

$$\left(\frac{dE}{dt}\right)_{loss} = -\frac{2e^2c}{3\rho_c^2}\beta^3\Gamma^4 ergs.s^{-1} \tag{2.10}$$

The total rate change of energy considering only curvature radiation and $E_\|$ acceleration is then, from (2.9) and (2.10),

$$\frac{dE}{dt} = \left(\frac{dE}{dt}\right)_{gain} + \left(\frac{dE}{dt}\right)_{loss} \simeq eE_\|c - K_1\frac{\Gamma^4}{\rho_c^2} \tag{2.11}$$

with K_1 only dependent on the velocity. Thus we have a limited acceleration region because as the particles accelerate away from the surface, $E_\|$ decreases and the second term in (2.11) starts to dominate, making dE/dt negative.

2.2.2.3. Inverse Compton radiation

When a photon of energy E_0 'bounces' off an electron and both the photon and electron travel off at a different energy and angle, the process is called Compton scattering (see figure 2.7). The change of wavelength for Compton scattering is given by

$$\triangle\lambda = \frac{h}{m_e c}(1-\cos\theta)$$

with θ the deflection angle. The maximum change in the wavelength is therefore $\triangle\lambda = 2h/m_e c$ (Griffiths 2003).

The most extreme case of the inverse of the process is that an electron and photon can collide, the photon absorbs all the energy of the electron and a single high-energy photon travels off. For electrons with high relativistic velocities this can be a significant

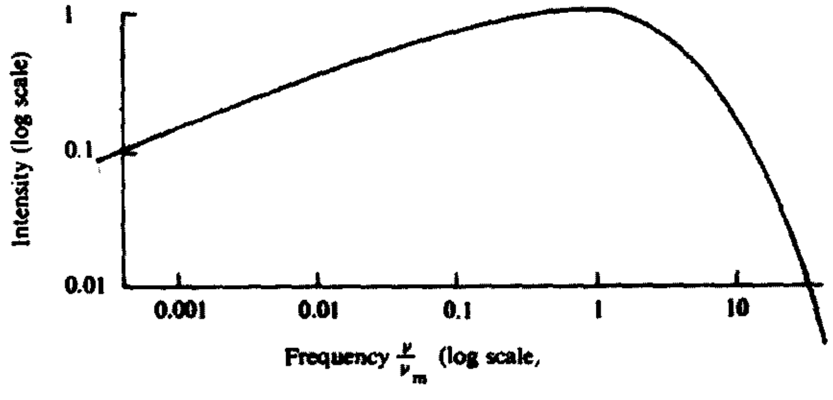


Figure 2.8. Spectrum of synchrotron radiation [52].

contribution to the energy of a photon if all the momentum from such a electron is absorbed by the photon. This process then contributes to the high-energy unpulsed γ-ray flux up to TeV energies, since the photons are not emitted in preferred directions.

2.2.3. Emission spectra

For synchrotron radiation we have a continuous spectrum of harmonics that is emitted for large Γ. For a single charge, most of the radiated power is in the harmonics, but the fundamental cyclotron frequency is still emitted with intensity (2.8) with the actual gyro frequency ν_g substituted for ν_L. For an individual charge, this radiation may be insignificant in comparison with the harmonic radiation, but the coherent radiation from many electrons may be concentrated in the fundamental and the lower harmonics only. We have a frequency $\nu_m = 4.6B(E_{MeV})^2$ where the radiated power is the maximum. Below ν_m the spectrum is a power law proportional to $\nu^{1/3}$ and above ν_m it falls exponentially as $e^{-\nu/\nu_c}$ with ν_c the critical frequency (Ginzburg & Syrovatskii 1969). The radiated synchrotron spectrum is shown in figure 2.8.

Electrons with relativistic velocities constrained to follow a path with radius of curvature ρ_c radiate in a similar way as an electron in a circular orbit with gyro frequency $c/2\pi\rho_c$. As in synchrotron radiation there is a critical frequency given by

$$\nu_c \simeq \frac{c}{2\pi\rho_c}\Gamma^3$$

and the maximum intensity of radiation is at a frequency

$$\nu_m \simeq 10^{-6}\frac{E_e^3}{\rho_c}.$$

This spectrum is of the same form as that for synchrotron radiation, see figure 2.8.

We have so far considered the synchrotron and curvature radiation of a single charged particle. In practice there is an ensemble of charged particles with a range of energies. The radiation from each is concentrated about its critical frequency, so that the resultant spectrum depends on the distribution of critical frequency among the ensemble. If the particle energies are distributed according a power law with index κ so that $N(E) \propto E^{-\kappa}$ then the spectrum also follows a power law $P(\nu) \propto \nu^{-\alpha}$. Here

$$\kappa = \begin{cases} 3\alpha + 1 & \textit{curvature} \\ 2\alpha + 1 & \textit{synchrotron} \end{cases}$$

This applies only if the energy power law extends over a sufficient range of energies. If there is a change of exponent κ in the energy spectrum, it will be reflected in a change of exponent in the radiation spectrum, but the change will be smoothed out over a range of frequencies. A full analysis is given by (Ginzburg & Syrovatskii 1969).

With the radiation from the inverse Compton effect we have the transfer of energy from high-energy electrons to radiation. The radiation from a cloud of high-energy electrons therefore increases the total flux of radiation energy and puts the increased energy in shorter wavelengths. This radiation mechanism does not depend on collisions or on a steady magnetic field, and therefore gives rise to the unpulsed TeV emissions which are observed.

The energy of the γ-rays is

$$E_\gamma = \hbar\omega_0 \propto \left(\frac{E_{max}}{mc^2}\right)^3 \rho_c^{-1}$$

and it can escape the magnetosphere if, with ϕ the angle between the γ-ray and the magnetic field,

$$E_\gamma B \sin\phi < K_2 \tag{2.12}$$

where K_2 is a critical value for photon-photon pair production and the cut-off energy of emitted γ-rays is

$$E_0 = \frac{K_2}{B_\perp} \tag{2.13}$$

with $B_\perp$ the magnetic field strength perpendicular to the direction of the photon. If $E_\gamma B \sin\phi > K_2$, e^+e^- pair production takes place and the secondary electrons can emit a further generation of γ-rays.

The e^+e^- pair creation processes alter the charge density sufficiently to short out the strong accelerating electric field. This happens at a well-defined 'pair formation front' above which the beam coasts, creating more γ-rays whose subsequent cascades radiate a spectrum of γ-rays observable by favourably located observers. Thus, depending on the magnetic field strength, the maximum energy observable is either from the real maximum energy obtainable by the electrons, or by the cut-off at the critical energy value where e^+e^- pair creation starts cascading. This places a very sharp upper limit to the radiation spectrum. Estimates for this upper cut-off energy range from 5 GeV to 100 GeV. Following the outer gap model, it seems that long period pulsars with low magnetic fields will be the best candidates for detection above 20 GeV (Harding 2001). Also, as the pulsar grows older, we expect the multiplicity for pair creation to decrease, with the resulting effect of spectra becoming harder with increasing age.

A generic PC model for the tails of differential spectra is given by

$$\frac{dN}{dE} = k\left(\frac{E}{E_n}\right)^{-m} e^{-(E/E_0)^b} \tag{2.14}$$

according Nel & de Jager (1995). A PC model is assumed here because it gives more conservative upper cut-off energies as the OG model. If b is consistently greater than 1, it would make ground based detections more difficult, since the collection area $A(E)$ increases with energy E and a significant overlap of $A(E)$ and dN/dE would be required for a detection. They assumed $b = 2$ for the most conservative rates. The spectral parameters for pulsars for $E > 1$ GeV which will be used in par. 2.4 for calculation of detection rates are listed in table 2.1.

2.3. The Atmospheric Cerenkov technique

As γ-rays of more than a few GeV enter Earth's atmosphere they produce Cerenkov radiation, which is electromagnetic radiation of 90 to 330 nm, emitted by a beam of high-energy

Table 2.1. Assumed pulsed spectral parameters ($E > 1$GeV) with parameters m and b as defined in (2.14) (De Jager 2002)

Object	k ($\times 10^{-8}$)(/cm^2/s/GeV)	m	E_0(GeV)	b	$F(> 1$ GeV) (/cm^2/s)
Crab	24.0	2.08	30	2	22
Vela	138	1.62	8.0	1.7	148
Geminga	73.0	1.42	5.0	2.2	76
PSR B1951+32	3.80	1.74	40	2	4.9
PSR B1055-52	4.00	1.80	20	2	4.5
PSR B1706-44	20.5	2.10	40	2	20

charged particles passing through a transparent medium at speeds $> c$ for that medium. Cerenkov radiation was discovered by Pavel Cerenkov in 1934 while observing radioactive radiation underwater and in 1958 he shared the Nobel Prize for Physics with Igor Tamm and Ilya Frank for their help in explaining the phenomenon.

The idea to detect Cerenkov light flashes from extensive air showers (EAS) comes from simple physical reasoning. Primary cosmic rays of a high energy entering the atmosphere produce a cascade of the secondary charged particles and γ-ray photons having energy well above the energy threshold of Cerenkov light production. Thus a single high-energy primary particle can produce an EAS of secondary particles distributed over a large area (Weekes 1994; Konopelko 1997). With each conversion in the cascade, the mean energy of each particle or photon halves, giving $\widetilde{E} \simeq 2^{-d/\ell}E_0$ with d the distance travelled into the atmosphere, ℓ the mean free path of the particle in the atmosphere and E_0 the primary particle's energy (Figure 2.9).

The EAS particles arrive at the Earth's surface in a $\simeq 10^{-8}$s time interval, makeing it possible to measure the Cerenkov light emission within an exposure time of 10 - 30 ns. The amount of night sky light detected for such a short time interval is negligible compared with Cerenkov light flashes from the EAS if the optical reflector used has a sufficient mirror area. To decrease the energy threshold of the detector one can increase the mirror area of the optical detector and use multichannel fast electronics to get more Cerenkov light from the EAS against the night sky background. An effective registration of γ-rays of energy as low as 10 GeV is expected (Konopelko 1997). The low enegry threshold of 10 GeV for ground-based atmospheric Cerenkov Telescopes is complementary to the Compton GRO and EGRET satellites which can measure up to $\sim$20 GeV (Mirzoyan 1997, Weekes 1994, Harding & de Jager 1997).

The Atmospheric Cerenkov Technique (ACT) is unique in astronomy in that the at-

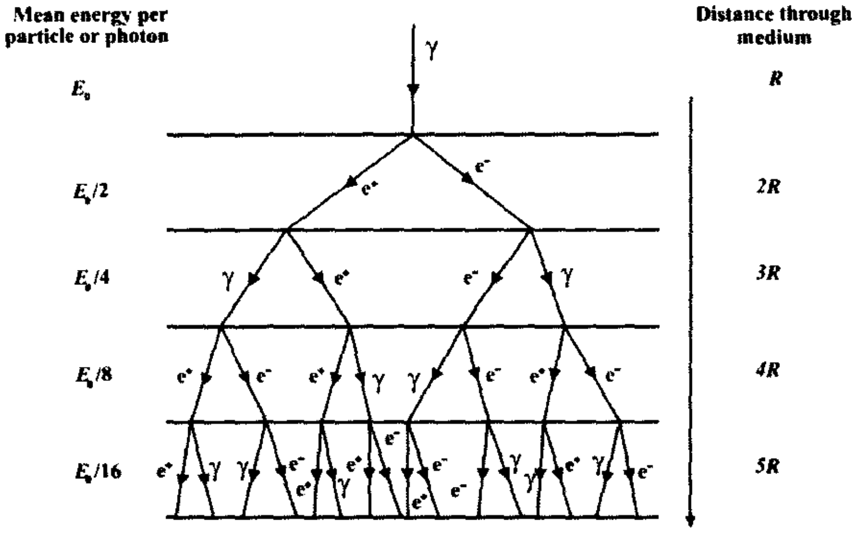


Figure 2.9. Diagram illustrating the basic Cerenkov radiation mechanism

mosphere forms the detection medium. Thus, as well as having to calibrate the telescope, one also needs to know the atmospheric parameters. An extensive analysis of the effects of atmospheric composition on the development of γ-ray cascades has been made by Bernlohr (2000). He concluded that pressure, temperature, ozone, aerosol and water vapour profiles were all significant. Accurate measurement of these parameters is essential to obtain the desired lower detection energy threshold.

The Atmospheric Cerenkov Technique is particularly suited for γ-ray astronomy for a number of reasons, including:

— the inherent angular resolution of the technique is high because the Cerenkov light retains the original direction of the primary photon,

— the light does not spread out appreciably so that the light pool reaching the ground has dimensions of several hundreds of meters, making detection easier,

— the Cerenkov light is a calometric component of the shower and can be used as a good estimator of the primary energy,

— the very short duration of the light pulses is well-matched to fast pulse counting electronics so that the shower can be detected against the night sky.

Initially it was assumed that γ-ray and hadronic showers were identical in a general way but simulations made it clear that, because of the smaller transverse momentum in electromagnetic interactions, the electromagnetic cascade is much more tidy and compact than its

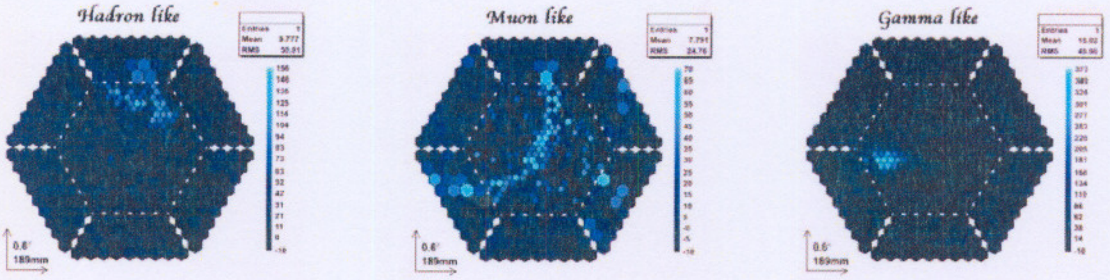


Figure 2.10. Examples of hadron, muon and gamma-like detections. The gamma detection is much more concentrated.

hadronic counterpart (see Figure 2.10). Therefore a drawback of the ACT is the presence of a heavy background from cosmic ray nuclei which produces EAS at a much higher rate than gamma-rays. The eventual implementation of image intensifiers and stereoscopic imaging led to the modern-day detection of γ-rays against a background of hadronic radiation. Stereoscopic imaging relies on the detection of directional anisotropy amongst the arrival directions of cosmic ray air showers, with the assumption being made that the interstellar and interplanetary magnetic fields render the charged component isotropic (Weekes 1996).

2.4. The H.E.S.S. Telescope

H.E.S.S. (High Energy Streoscopic System) is a 3$^{\mathrm{rd}}$ generation ground based Imaging Atmospheric Cerenkov Telescope (IACT) detecting Cerenkov radiation from EASs. The first phase, which consists of four 13-m diameter dishes and cameras with 10-ns detection time, went fully operational on 10 December 2003. It is located in Namibia and is one of the major atmospheric Cerenkov telescopes together with CANGAROO (Kifune et al. 1997), MAGIC (Petri & The MAGIC Telescope Group 1999) and VERITAS (Weekes et al. 1997). The basic goals of the H.E.S.S. group are to study processes in the universe with high energy turnover and to find the origin of cosmic rays. This includes exploring the TeV γ-ray sky and other non-thermal sources, surveying of the galactic centre, all-sky surveying, studying of supernova remnants and extragalactic sources like relativistic jets from black holes. It will also contribute to the theory of active galactic nuclei (The H.E.S.S. Project 2004).

H.E.S.S. has a low threshold energy and high sensitivity, an order of magnitude better than previous instruments (see figure 2.11). It has a field of view of $\sim 5° \times 5°$ degrees with

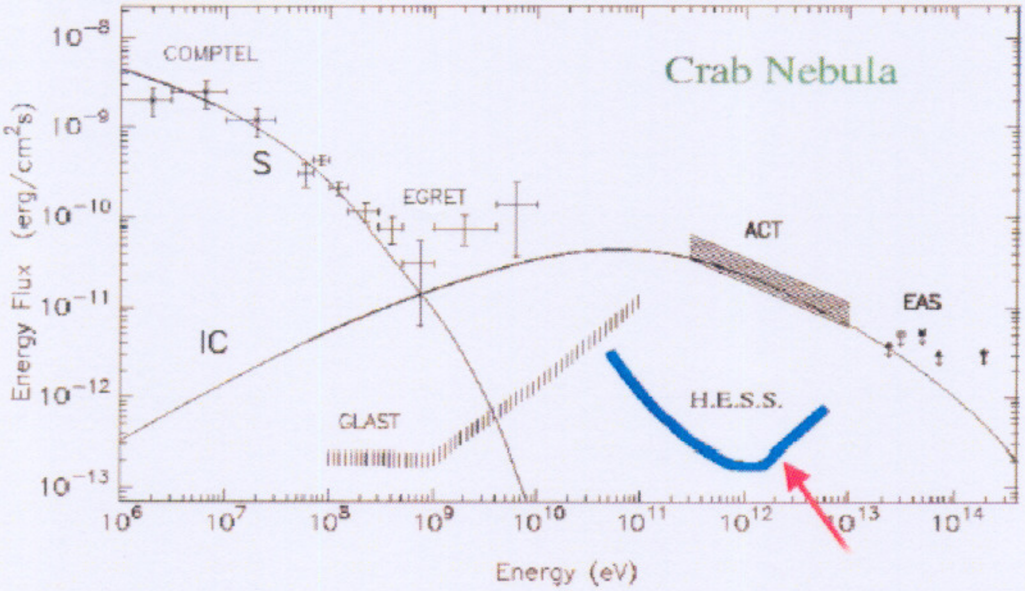


Figure 2.11. Comparison of sensitivity between H.E.S.S. and other past and present projects.

a 0.1 degree angular resolution and an effective area of a few times 10^4m^2. Its detection energies for imaging range from 50 GeV to 100 TeV and the stereo imaging capability results in a significant rejection factor against background events. It can however trigger on events above 10 - 30 GeV and this feature will be exploited for pulsar searches, even though imaging is not possible in this energy range. The stereo imaging capability still allows some background events to be rejected at these energies.

The stereo imaging technique has several advantages (Punch 2002):

— Being able to locate the origin of the shower unambiguously, giving the instrument a good angular resolution.

— Multiple measurements of a shower allows for a good energy resolution and a high level of hadron shower rejection.

— The stereo trigger mechanism helps with the complete rejection of the local muon background and gives the instrument a lower energy threshold.

The relationship between the detection area and γ-ray energy of H.E.S.S. is given in figure 2.12, with the fit to the detection area having functional form (Konopelko 2001)

$$A(E_\gamma) = \frac{a_0 E_\gamma^{a_1}}{1 + (E_\gamma / a_2)^{a_3}}$$

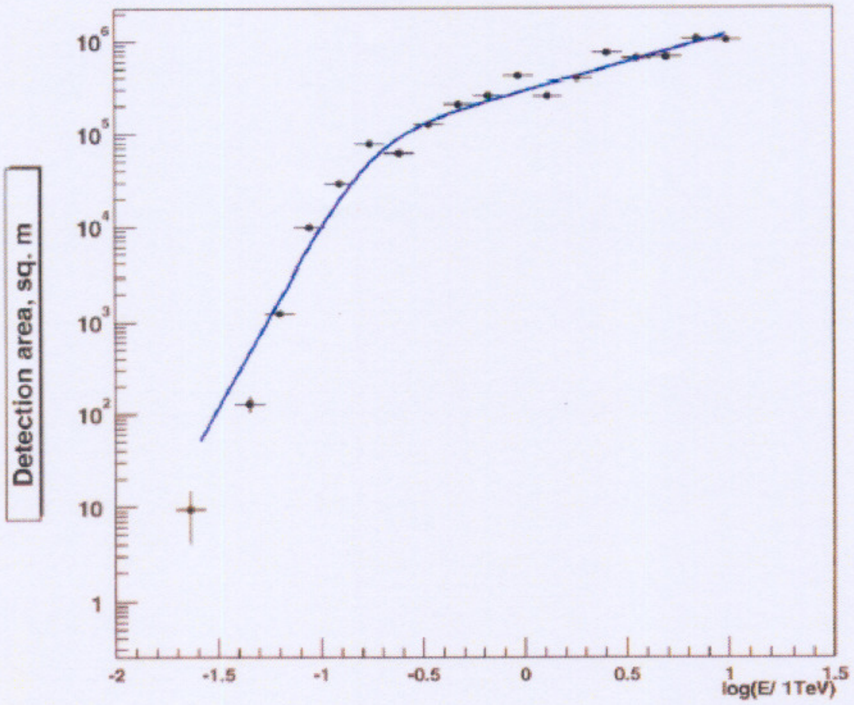


Figure 2.12. The detection area of H.E.S.S. as a function of incident γ-ray energy.

Using the collection area $A(E)$ the expected rate of triggers from pulsed Cerenkov showers is given by

$$R_p = \int A(E) \left(\frac{dN_\gamma}{dE}\right) dE \tag{2.15}$$

Statistics and an additional trigger rate can reduce an incoming cosmic ray background trigger rate R_b from about 1 kHz to about 8 Hz. From these pulsed and background trigger rates, the detection sensitivities have been calculated by (De Jager 2002) for canonical high-field pulsars as shown in table 2.2. It is clear that H.E.S.S. will only be able to detect

Table 2.2. Estimated pulsed rates $R_p$and observation times for H.E.S.S.

Object	R_p (hr^{-1})	T (10-hr days)	E_0 (GeV)
Crab	100	3	30
Vela	8	400	8.0
Geminga	$\ll$1	-	5.0
PSR B1951+32	180	1	40
PSR B1055-52	8	420	20
PSR B1706-44	240	1	40

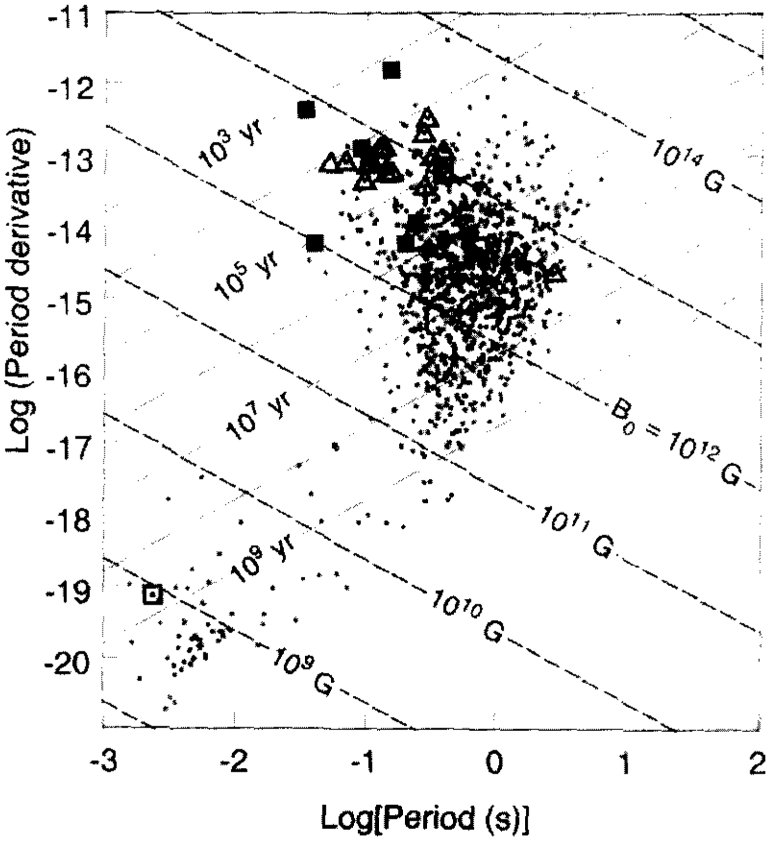


Figure 2.13. Plot of period derivative $\dot{P}$ of radio pulsars (grey dots) from the ATNF Pulsar Catalogue, with confirmed (solid squares) and candidate (open squares) γ-ray pulsars. Dashed diagonal lines indicate constant magnetic field strength.

pulsed emission if $E_0 > 30$ GeV, which is realised at least for PSR B1706-44 and PSR B1951+32.

Pulsars visible in the radio regime have rotation periods P between 1 ms and 10 s, with magnetic field strengths B_0 of 10^8 to 10^{14} G as shown in figure 2.13. To select possible candidates for observation with H.E.S.S., we see from (2.12) that the lower the value of the magnetic field of a given pulsar, the higher the energy of the escaping γ-rays. In figure 2.13 this represents the lower left section of the plot.

Important discoveries has been made with the H.E.S.S. instrument thus far. One important discovery was proof of high-energy particle acceleration in the shell of a supernova remnant (Aharonian 2004), which point to the origin of the galactic cosmic rays. Many exciting projects are also underway, including one to measure the proposed general relativistic frame dragging effects on the emission around a millisecond pulsar (Venter, 2004).

2.5. Other telescopes: past, current and future

H.E.S.S. Phase 2, currently in the planning phase, will expand the system of 4 telescopes to 5 by erecting an additional very large 28-m telescope in the centre of the current array. This will bring down threshold energies, increase sensitivity and make the system even more appropriate for pulsar searches (The H.E.S.S. Project 2004). H.E.S.S. Phase II aims to close the gap to GLAST (The Gamma-ray Large Area Space Telescope), which will be launched during 2006 and has a detection range in the energies from 10 keV to 300 GeV. H.E.S.S. and GLAST have an overlapping energy regime so they can observe the same particle populations. Simultaneous observations will also allow inter-calibration of the devices. The key scientific objectives of the GLAST mission are (Kniffen 2002):

— To understand the mechanisms of particle acceleration in Active Galactic Nuclei, pulsars and SNRs.
— To resolve the gamma-ray sky: unidentified sources and diffuse emission.
— To determine the high-energy behaviour of gamma-ray bursts and transients.
— To probe dark matter and the early universe.

GLAST will serve as a follow-up space telescope to EGRET, which was operational between 1991 and 1999. EGRET had a much smaller field of view, higher angular resolution and detected in the energy range from 30 MeV to 10 GeV. Some major EGRET results include the determination with high certainty that cosmic rays are galactic and measurement of the diffuse, presumably extragalactic, high energy gamma-ray spectrum. In the EGRET sky however, there still remain many unidentified sources.

MAGIC (Lorenz 2004) is currently the largest gamma-ray telescope with the lowest energy threshold. Currently, a second improved clone is being constructed on the same site under the name MAGIC II. Plans are eventually to install devices to give access to even lower gamma-ray energy and are codenamed ECO-1000.

CANGAROO III (Kubo et al. 2004) consisting of an array of 4 10-m IACTs, was completed during March 2004 and has a threshold energy of about 100 GeV. As with H.E.S.S. it is also located in the southern hemisphere to be able to observe the galactic centre.

With just these few examples it is clear that gamma-ray astronomy is a rapidly expanding field. The main goal of the last few years was to close the observation gap of 10 GeV (upper threshold for EGRET) and 300 GeV. New telescopes have already lowered the threshold energy to below 60 GeV. At around 10 to 40 GeV the universe becomes near-transparent and one should be able to see γ-emitting objects as far as a redshift of $>$ 3.

Chapter 3

The development of a hardware random number generator

3.1. Introduction

In astronomy the need quite often arises to identify a periodicity in data dominated by counting statistics. Usual procedures involve superposition (or folding) of the arrival times on some phase interval, normally $[0, 1)$ or $[0, 2\pi)$, using appropriate parameters like the period P and period derivative $\dot{P}$. This interval then represents one full period of rotation and in the absence of any periodicity the folded events will be distributed uniformly. One can then test for any periodicities by applying a test for uniformity on the chosen interval.

In observing cosmic rays, the charged cosmic ray component (i.e. the nuclei) is of an isotropic and incoherent nature and cannot be traced back directly to particular sources. This is thought to be the effect of both the production mechanisms and interstellar and intergalactic scattering mainly by magnetic fields. Therefore the arrival times of the charged cosmic ray component are stochastic, i.e. independent and uniform. The γ-ray component of cosmic rays, however, arrives at Earth in a nearly undisturbed fashion and can be associated directly with certain point sources. The Cerenkov radiation from both the charged and γ-ray components of cosmic rays is detected by IACTs. The γ-rays from pulsars however are of a periodic nature as discussed earlier, but the ratio of these γ-ray fluxes to the charged cosmic ray flux is low and one is forced to approach the problem with proper statistical methods if these γ-rays are to be properly identified. Therefore one has to rely on hypothesis testing to provide an answer to the possible presence of pulsed γ-rays.

The necessity to have a sound statistical basis, which we discuss in the next section, is quite clear. This also comes in handy when simulating arrival times (or events) of Cerenkov showers at an IACT, with goal to test the accuracy and distribution of the H-test (see section 3.6). This calls for the use of truly random numbers, section 3.5.1 explaining in detail what we imply with *truly*. The lack of a generator being able to generate such truly random numbers at the frequency required gave the motivation for the Space Research Unit at the Potchefstroom University to develop such a truly random number generator, described in section 3.5.2. This generator unit in itself does not generate numbers which are totally independent and uniformly distributed, so procedures had to be implemented to account for and correct these problems (see section 3.5.4). Also, one often needs numbers that have a standard normal distribution for simulations, this being discussed in section 3.5.5.

The design and implementation for such a truly random number generator is not motivated by astrophysical needs alone, but has an extensive range of applications. These spin-offs and further possible applications are discussed in section 3.7.

3.2. Statistical basics

A quick overview of important and relevant statistical concepts and laws is given here, forming the basis of the statistical tools needed further on.

A *random variable* is a variable which can take on more than one value, either discreet or continuous, this value not being predictable in advance. The distribution of the variable may well be known, this giving the probability of a given value (or infinitesimal range of values) being obtained. The *probability density function* of a random variable u gives the probability of finding the random variable u' within $\mathrm{d}u$ of a given value u, denoted by $g(u)$ as

$$g(u)\,\mathrm{d}u = P(u < u' < u + \mathrm{d}u)$$

This is normalised in such a way that the integral over all u is 1, implying that the total probability of finding all the possible values is one. The *distribution function* is defined through

$$G(u) = \int_{-\infty}^{u} g(x)\,\mathrm{d}x$$

and is a monotonically increasing function taking on values from zero to one.

Considering two random variables u and v with density $h(u,v)$, u and v are *stochastically independent* if and only if $h(u,v) = p(u)q(v)$. For more than two variables the concept of independence becomes more complicated and all possible pairs, triplets, etc. (i.e. all combinations but not permutations) have to be considered.

The definition of *random sampling*, which is a fundamental point of departure in statistics can now be given:

Let $x_1,\ x_2,\ .\ .\ ,\ x_n$ be some sample drawn. The sample is random if and only if

1. x_i is independent from x_j for all $i \neq j$ and $i,\ j = 1,\ .\ .\ .\ ,\ n$.
2. the probability density function of each x_i is the same, meaning that the x_i's are identically distributed.

The *expectation value* of a function $f(u')$ is defined as the average or mean value of the function:

$$E(f) = \int f(u)g(u)\mathrm{d}u$$

The *variance* of a function or variable is the average of the squared deviation from its expectation:

$$Var(f) = E\left[(f - E(f))^2\right] = \int (f - E(f))^2\,\mathrm{d}u = E\left(f^2\right) - \left[E\left(f\right)\right]^2$$

Expectation is a linear operator but variance is non-linear. Also note that

$$Var(x + y) = Var(x) + Var(y) + 2Cov(x, y)$$

with $Cov(x, y)$ the *covariance* between x and y. For independent variables the covariance between them is zero. Often, instead of the variance, the *standard deviation* is used and is given by the relation

$$\sigma(f) = \sqrt{Var(f)}$$

It can be interpreted as the r.m.s. deviation from the mean.

The *law of large numbers* concerns the behaviour of sums of a large number of random variables. Choosing n numbers u_i randomly with an identical probability density, thus $E(u_i) = \mu$ and using the *average* $\overline{U}_n = \frac{1}{n}\sum_{i=1}^{n} u_i$, we have that

$$\overline{U}_n \rightarrow \mu \text{ as } n \rightarrow \infty$$

An important theorem is the *central limit theorem*, essentially stating that the sum of a large number n of identically distributed and independent random variables will always be normally distributed, provided n is large enough. It does not however specify when n is considered large enough, that has to be induced from convergence properties for the specific case under consideration. Mathematically the central limit theorem can be written as

$$\begin{matrix} lim \\ n \rightarrow \infty \end{matrix} P\left(\frac{\overline{U}_n - \mu}{\sigma/\sqrt{n}} \leq u\right) = \Phi(u) \tag{3.1}$$

with Φ the standard normal distribution, basically stating that it is distributed normally with average μ and standard deviation $\sigma/\sqrt{n}$,

$$\overline{U_n} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

This implies that

$$\frac{U_n - n\mu}{\sigma\sqrt{n}} \sim \mathrm{N}\left(0, 1\right) \qquad n \gg 1.$$

3.3. Tests for Uniformity

Beran (1969) derived a complete class of tests for uniformity on the circle and some of the most useful astrophysical tests are special cases from his class. Important cases are Pearson's well-known χ^2-test with K bins and the $\mathrm{Z_m}^2$-test (Buccheri et al. 1983) which involves the sum of the Fourier powers of the first m harmonics. Both these tests are dependent on a smoothing parameter: the number of bins K or harmonics m. If the periodic shape is unknown a priori it is impossible to make the correct choice for the smoothing parameter. De Jager et al. (1989) and references therein showed that the capability or "power" of these tests to detect specific pulse shapes is strongly dependent on the choice of a smoothing parameter if the signal is weak. Proposed as a solution the H-test. In this section we consider tests for uniformity in general and in the next section we give attention to one special case, the H-test.

Consider a set of arrival times $t_i (i = 1, \ . \ . \ . \ , n)$. Assume firstly that the frequency parameters of the source are known so that folding of the t_i's according to these parameters gives the phases $\theta_i \in [0, \ 2\pi)$. Testing for the presence of a periodic signal is accomplished by making the null hypothesis H_0, which is characterized by a test for uniformity, and testing for a rejection of this. The null hypothesis may, in this case, be written as

$$H_0 : f\left(\theta\right) = \frac{1}{2\pi} \text{ with } \theta \in [0, \ 2\pi) \tag{3.2}$$

In the presence of a known periodic signal denoted by a source function $f_s\left(\theta\right)$ giving the relative radiation intensity, (3.2) will differ and may be written under the alternative hypothesis as

$$H_A : f(\theta) = pf_s(\theta) + \frac{1-p}{2\pi}\ ,\ p \in [0,\ 1) \tag{3.3}$$

with $p \in [0, 1)$ the strength of the periodic signal. To determine the presence of a periodic signal in thc data one tests thc hypothesis $H_0 : p = 0$ against $H_A : p > 0$ and to do this we measure the deviation between $f(\theta)$ and the uniform density $1/2\pi$. A good measure is given by the functional

$$\psi(f) = \int_0^{2\pi} \left(f(\theta) - \frac{1}{2\pi} \right)^2 d\theta \tag{3.4}$$

H_0 should then be rejected when $\psi(f)$ is too large.

A problem in this is that $\psi(f)$ is unknown since $f(\theta)$ is unknown and therefore one has to estimate the functional. This is usually done in a 'crude' manner by replacing f with a consistent estimator $\hat{f}_h$, these esimators being characterised by some smoothing parameter h. Statistical literature gives an extensive discussion of various choices of $\hat{f}_h$ and their properties under the heading "density cstimators". De Jager ct al. (1986) also discussed their implementation in perodic analysis.

For a histogram-like dataset one sets $h = K$ for the corresponding number of bins. In this case Eq. (3.4) results within constants to the well-known χ^2 statistic of Pearson, this being asymptotically distributed to χ^2 with $K - 1$ degrees of freedom under H_0:

$$\chi^2_{K-1} = 2\pi n\psi\left(\hat{f}_K\right) = \sum_{j=1}^{K} \frac{\left(X_j - \frac{n}{K}\right)^2}{\frac{n}{K}} \tag{3.5}$$

with X_j the number of events in the the j'th bin. This statistic is however dependent on the choice of bin positions, unfortunately making it variant under rotations.

Specifying $\hat{f}_m$ as the Fourier Series Estimator (FSE) with m harmonics, we have from De Jager et al. (1986) that

$$\hat{f}_m(\theta) = \frac{1}{2\pi}\left[1 + 2\sum_{k=1}^{m} \left(\bar{c}_k \cos(k\theta) + \bar{s}_k \sin(k\theta)\right)\right] \tag{3.6}$$

where the empirical trigonometric moments are given by

$$\bar{c}_k = \frac{1}{n}\sum_{i=1}^{n} \cos(k\theta_i) \text{ and } \bar{s}_k = \frac{1}{n}\sum_{i=1}^{n} \sin(k\theta_i) \tag{3.7}$$

with both $\hat{c_k}$ and $\hat{s_k}$ distributed $N\left(0,\ 1/\left(2n\right)^{1/2}\right)$ for uniform θ_i's $\left(U\left(0,2\pi\right)\right)$ according Mardia (1972). For this case Eq. (3.4) reduces to the Z_m^2 statistics, introduced by Buccheri et al. (1983), within constants:

$$Z_m^2 = 2\pi n\psi\left(\hat{f_m}\right) = 2n\sum_{k=1}^{m}\left(\hat{c_k^2}+\hat{s_k^2}\right) \tag{3.8}$$

This is rotationally invariant and therefore has an advantage over Pearsons χ^2-test. It is a potentially good test in the sense that the number of harmonics m can be adjusted to detect both narrow (large m) and broad (small m) pulse shapes. The problem here remains that $f_s\left(\theta\right)$ is not known a priori, however, and m cannot be chosen before inspection of the data has been done, thus we still have a problem with the smoothing parameter.

3.4. The H-Test for uniformity

Most tests for periodicity suffer from the problem that they are only powerful enough to detect some kinds of periodic shapes in the case of weak signals. This causes one to, after inspection of the data, subjectively select specific tests for the identification of weak periodic signals, causing a possible false identification of sources. The H-test (De Jager 1987, De Jager et al. 1989) is a consistent and powerful test for uniformity against most periodic shapes encountered in astronomy and is especially useful if no a priori information about the periodic shape is available. It is independent of a subjectively chosen smoothing parameter and is simple to calculate, requiring order 40n calculations, with n the number of recorded events.

Z_m^2 statistics form the basis of the H-test with m chosen as a specific function of the data θ_i using Hart's rule (Hart 1985). From Hart's rule one obtain a value for m, say M, so that it minimizes an estimator of the mean integrated squared-error (MISE) between the FSE, eq. (3.6), and the true unknown periodic shape $f\left(\theta\right)$. This is written as

$$MISE(m) = E\int_0^{2\pi}\left[\hat{f}_m\left(\theta\right) - f\left(\theta\right)\right]^2 d\theta \tag{3.9}$$

effectively giving MISE(M) $\leq$ MISE(m), $m = 1, 2, ...$ For practical purposes one usually only searches through the first 20 harmonics because the main signal component is expected to be contained in it. Also neglecting terms of the order $1/n$ arising from Hart's rule, we can define the H-test as

$$H = \underset{1 \leq m \leq 20}{max} \left(Z_m^2 - 4m + 4\right) = Z_M^2 - 4M + 4 \tag{3.10}$$

and it only assumes positive values. This is now our test for uniformity and is independent of any subjectively chosen smoothing parameter. The H-test is useful in general cases, except where more than three peaks are anticipated in the periodic shape, in which case the Z_M^2-test is preferable. The number M obtained in Eq. (3.10) is a good estimate of the true optimal number of harmonics. This smoothing parameter is chosen in an objective and automatic way without requiring any knowledge a priori of $f_s(\theta)$ and it is nearly the most powerful test for most possible light curve shapes (De Jager et al. 1989). It is important to realise that, because the H-Test scans through 20 harmonics, it actually comes down to the same as using 20 independent tests. This fact makes it powerful for blind pulsar searches.

The probability distribution of H under uniformity is used to determine the significance of a detection. Unfortunately this distribution cannot be obtained analytically and must be determined from simulations. These simulations use a set of uniform random numbers, ~U(0, 2π) for the unit circle, to obtain the distribution of H.

De Jager et al. (1989) considered both the small n and large n cases. Nowadays it is not necessary to consider small n values since collection areas and integration times have increased to the extent that only large n values are obtained ($n \gg 100$). The probability distribution for large n-values was obtained as

$$\begin{array}{ll} \text{Prob}\,(H > h) = ae^{-bh} & \text{if } 0 < h \leq 23 \\ \text{Prob}\,(H > h) = ce^{-dh+eh^2} & \text{if } 23 < h \leq 50 \end{array} \tag{3.11}$$

with the constants of the fit given by $a = 0.9999755$, $b = 0.39802$, $c = 1.210597$, $d = 0.45901$ and $e = 0.0022900$. For h-values larger than 50 they could not give a reliable parametric equation, because they simulated only 10^8 values for H but they also mentioned that Prob($H > 50$) $\simeq 4 \times 10^{-8}$. For uniformly distributed events both the mean and standard deviation are 2.51.

The efficiency of any statistical test to reject the null hypothesis, given it has been falsely assumed, is quantified by its power. The power of a test as seen from a periodic point of view, is the probability that a periodic source will be identified above a given detection threshold for a given sample of size n. The detection threshold is the probability

α to falsely reject H_0. The choice of α is very subjective and depends on the situation but usually $0.001 \leq \alpha \leq 0.05$. The H-test stands out as a fairly powerful test compared to other tests for uniformity, for both broad and narrow periodic peaks. If the number of peaks increases above three however, it is best to use the Z^2-test, this being independent of the duty cycles involved. Very few physical processes call for this approach however.

Numerous examples exist of where the H-test has been employed in periodicity searches (Hesssels et al. 2004; Kaspi et al. 2000; Chang & Ho 1997). It has been applied for pulsar searches in EGRET data and will also be applied for this cause on MAGIC and H.E.S.S. data. It will also serve as one of the periodicity search algorithms on GLAST (Kniffen 2002).

3.5. Random number generation for simulations

As mentioned in section 3.4, random numbers are needed to simulate the distribution function of the H-test. This is by far not a unique case, as there are numerous fields in physics alone that require simulations based on random numbers. One practical example having relation to cosmic ray research is the simulation of the Cerenkov shower itself. The time of existence (and therefore the distance travelled) for any entity in the shower before the next stage can be modelled on a statistical basis including for example the collisional cross section.

Having fast access to truly random numbers is an essential component in effective modelling, especially if one needs accurate information. The uses of a random number generator (RNG) stretches from applications in everyday homes, industrial applications to the frontiers of research and development. In this section we consider the problems surrounding current RNGs, discuss a solution and its implementation as well as further uses for a truly RNG.

3.5.1. Problems surrounding random number generators

Recall in section 3.2 the definition of a random sample. It states that *the elements of the sample must all be independent and the distribution of each element must be identical.* Even the slightest deviation from these conditions will render the sample as non-random, but not necessarily useless. This has been exploited for many years by the pseudo-random number generators built into PC's and software, and we must therefore distinguish clearly between truly random and pseudo-random sequences. It is common to confuse the randomness properties of a sequence with its distribution. A truly random sequence may have any distribution whereas a perfectly uniformly distributed sequence may not at all be random. We now consider two general types of pseudo-random generators, namely the *mixed congruental* and *r250* methods, for the demonstration of these general properties.

The most commonly used type of random number generator is the *mixed congruental* generator discussed in James (1980). A sequence of random numbers is obtained by using

$$r_i = (ar_{i-1} + b) \mod m$$

with a, b and m constants and the MOD operation giving the remainder after division with m. A computer is however a deterministic machine with no room for statistical fluctuations or the like. Using this method it generates random sequences by starting with a seed r_0, chosen by some or thc other mechanism simulating near-random behaviour. Thereafter it performs a series of simple operations on it, the result forming the next number in the sequence. This is all well if only a few 'random' numbers are needed, but when a large number of them are needed (typically $> 10^4$), the situation changes. In the sequence of random numbers produced, one may obtain one of the numbers previously occurring in the sequence, having as effect that the same sequence of numbers follows, one then being stuck in an infinite loop of the same numbers repeating in exactly the same order. This is due to the deterministic nature of computers and cannot be undone, this effect violating both the conditions that have been set for a sequence of numbers to be truly random. The maximum period for this type of generator is generally of length $m/4$ according James (1980). The numbers generated are unfortunately also sparse and not useful when high resolution is needed. Another drawback is the regularity in numbers produced in a d-dimensional space, known as the Marsaglia effect, which he described in his classical paper *Random numbers fall mainly in the planes,* Marsaglia (1968), which finally brought some genuine understanding into the art of pseudo-random number generation. He showed that if successive d-tuples produced by the multiplicative congruental method are taken as coordinates of points in d-dimensional space, all the points will lie on a certain finite number of parallel hyperplanes, depending on the bit length of integer arithmetic on thc machine. This was onc of thc typical problems with IBM's RANDU function in the 360 series. Some values are given in table 3.1, from which it is clear that problems will most certainly arise during simulations using a large number of these pseudo-random numbers. The *Dieter-Ahrens solution* was proposed by Dieter & Ahrens (1971) but it still used the

Table 3.1. Maximum number of hyperplanes = $(d!2^t)^{1/d}$

Number of bits	d=3	d=4	d=6	d=10
16	73	35	19	13
32	2953	566	120	41
36	7442	1133	191	54
48	119086	9065	766	126
60	1905376	72520	3064	290

same basic method of producing random sequences, thus only lessening the extent of the problem by increasing the number of hyperplanes.

Another popular method is the *r250 algorithm*, described by Kirkpatrick and Stoll (1981) and named so because of the 250-element array used. Consider

$$a_k = (c_1 a_{k-1} + c_2 a_{k-2} + \cdots + c_{p-1} a_{k-p+1} + a_{k-p}) \bmod 2$$

with the a's generated bits and the c's coefficients, with values of either 1 or 0. Thus having a set of bits that have been previously generated, one can multiply them by the coefficients and add them together to create a new pseudo-random bit after the mod operation. The maximum period is achieved by choosing the polynomial primitive. This is usually done by setting only coefficients 103 and 250 as 1, and we simply have a simple XOR-operation between these two bits. As the new bit is generated the sequence shifts on one place and the same operation is now applied to the bits now residing in the same positions. If one needs a 16-bit number, then 16 of these sequences are used in parallel. This method is an order of magnitude faster than the mixed congruental method. Problems with this method mostly occur when the different sequences of 250 bits are linearly dependant and regularities arise. This usually results because the seed bits (250 of them for each parallel sequence) are chosen in the same way as with the mixed congruental method by some or the other mechanism simulating near-random behaviour, this introducing some extent of linear dependence.

The effect of pseudo-randomness can usually be detected by using the law of large numbers at levels of $> 10^8$ numbers (somtimes even at levels of 10^4 numbers). At such high levels patterns start to emerge in both the numbers produced and in the sequence they are generated in. Therefore we need to use a *truly RNG* for simulations requiring more than 10^8 random numbers, especially if one wants to study distribution tails or if high resolution is required. A sequence of truly random numbers is *totally unpredictable* and therefore *irreproducible* and can only be generated by using some random physical process such as radioactive decay, thermal noise, cosmic ray arrival times, etc. which has a high entropy. In a wide sense entropy can be interpreted as the measure of disorder; the higher the entropy the greater the disorder. In practice however there has been great difficulty in constructing such high entropy truly RNGs which are fast, small and inexpensive enough, at the same time being accurate, producing uniform (or Gaussian) numbers.

In 1978 Frigerio and Clark used a radioactive source and high-resolution counter for fixed time intervals of 20ms to extract truly random numbers. Whenever the count was odd they recorded a zero bit, and when even, a one bit, all stored on magnetic tape. Corrections were made for the bias and the apparatus yielded about 6000 31-bit truly random numbers per hour. This is however one of the rare cases of truly random number extractions and speed, size and cost still posed a problem.

3.5.2. The Quantum Bit Extractor

The Unit for Space Physics at the North-West University in South-Africa developed a truly random number extractor (not generator, since that will point to some deterministic process, again indicating a lack of true randomness) specifically to help in determining the distribution of the H-test to a much higher accuracy than previously known. The general importance of such a device has been recognised beforehand as well, resulting in an even higher level of motivation and input for the project.

Goals for the final Quantum Bit Extractor (QBE) implementation are:

1. It should be a small and compact device that can fit onto any portable implementation
2. to provide random numbers of required bit-length
3. according a specified distribution
4. at an appropriately high speed.

Research and design up to the level before implementation into a single-chip device has been completed and required results have been achieved. A patent was awarded for the QBE under *A hardware generator for uniform and Gaussian deviates employing analogue and digital correction circuits*, which is contained in Appendix I. Only a short discussion of the basic elements is given here.

The basic design is shown in figure 3.1. The Noise Element can be a resistor or avalanche process which is known to generate true randomlike signals, given the physics involved. This has to be chosen carefully because electromagnetic interference, temperature and

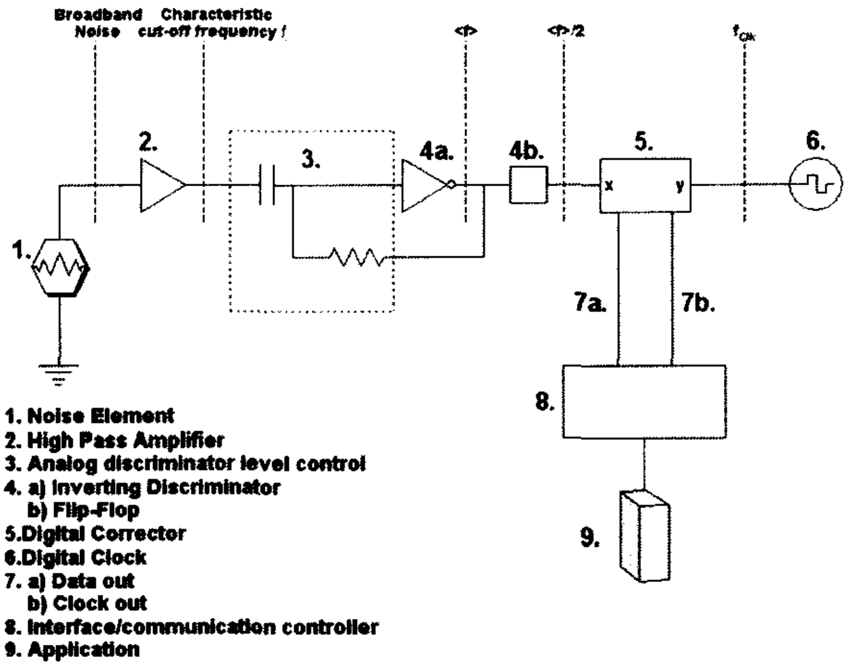


Figure 3.1. Basic QBE diagram.

humidity fluctuations, etc. can have a measurable influence on the output of such a device, which we want to be as stable as possible for the expected parameter ranges. From this device the high-frequency noise is amplified using a high-pass amplifier so that it is detectable by any discriminating technique such as a Schmidt trigger, comparator or discriminator. Feedback loops can regulate parameter variations so that some extent of unbiasing (i.e. production of a near-equal number of "0" and "1" bits) occurs on an analogue level. The flip-flop output then serves as the input to the digital corrector section as a Digital Input Sequence (DIS) shown in figure 3.2, the digital corrector discussed in section 3.5.4.

Output from the noise element is some continuous function, sampled in one way or another to produce a digital output sequence as input to the digital corrector. This digital output sequence therefore has some correlation between the bits sampled at a fixed rate because of the continuous nature of the source signal obtained from the noise element. Also, a certain level of biasing remain because the analogue debiasing can not always be perfect due to parameter variations. The digital output stream is sampled at discreet times, regulated by a clock (6 in figure 3.1) with frequency lower than that of the maximum frequency f_0 of the broadband noise. This bits resulting from the sampled stream therefore

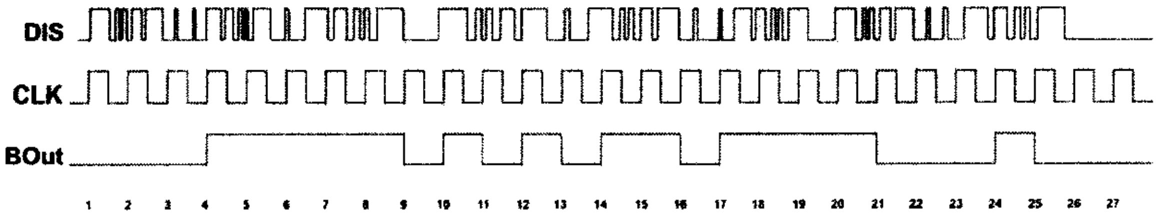


Figure 3.2. Logic of Digital Input Sequence (DIS) as the digital corrector receives it from the flip-flop. The clock is shown as CLK, which is nr. 6 in figure 3.1, and BOut is the Bitstream Output as the digital corrector then effectively receives it.

may still have some level of bias and probably even correlation, thus giving motivation for implementation of a digital corrector.

3.5.3. Employed tests for randomness

Before coming to the actual corrections that have to be made to the raw data to achieve both independence and uniformity on decimal level, we give a brief overview of the statistical tests used to verify randomness on certain levels. Many tools are available to test randomness and the method devised by Knuth is often used. Knuth suggested that statistical analysis such as distribution of values and auto correlation of different sets of values could be made. The DIEHARD test suite developed by Marsaglia (1995) is popular and regarded as the *de facto* test suite. The Federal Information Processing Standard (FIPS 140-2) document developed by National Institute of Standards and Technology (NIST) recommended the Monobit Test, Poker Test and Runs Test to test random number series (NIST 2004). All new standard RNGs need to pass through these test suites to prove themselves as generators of 'true' random numbers. The results of a battery of tests must be interpreted carefully. As Marsaglia states:

> By all means, do not, as a statistician might, think that a $p < .025$ or $p > .975$ means that the RNG has "failed the test at the .05 level." Such p's happen among the hundreds that DIEHARD produces, even with good RNG's.

If the null hypothesis is true, and if the statistical test is exact, we expect $|p| < 0.05$ 5% of the time.

We will not submit the developed Truly RNG to the DIEHARD and NIST test suites at this developmental stage but only submit it to some basic tests of randomness. The runs test and Chantilley test are variations of the same basic test for randomness in a random binary sequence. After converting a binary sequence to decimal numbers we also need a test for uniformity on decimal level to verify that the numbers used satisfy the necessary conditions for randomness. Pearson's Chi^2 (χ^2) test comes in handy for this purpose.

3.5.3.1. The runs test

A nonparametric test for randomness is provided by the theory of runs by Bradley (1968). A run is defined as a set of identical symbols contained between two different symbols (or no symbol, at the beginning or end of a sequence). Consider the sequence

$$a\,a \mid b\,b\,b \mid a \mid b\,b \mid a\,a\,a\,a\,a \mid b\,b\,b \mid a\,a\,a\,a \mid .$$

Proceeding from left to right the first run, indicated by a vertical bar |, consists of two a's; similarly the second run consists of three b's, the third run of one a, etc. There are seven runs in all ($V = 7$) and the number of a's are $N_a = 12$ and $N_b = 8$.

It is clear that some relationship exists between randomness and the number of runs. The sequence

$$a \mid b \mid a \mid b \mid a \mid b \mid a \mid b \mid a \mid b \mid$$

has a cyclic pattern which has very low probability of being random: in this case we have too many runs V. On the other hand, for the sequence

$$a\,a\,a\,a\,a\,a \mid b\,b\,b\,b\,b\,b\,b\,b \mid a\,a\,a\,a\,a \mid b\,b\,b\,b\,b\,b \mid$$

we have a trend pattern of clustering and there are too few runs V, and we would not consider the sequence to be random.

To quantify the idea of randomness, we consider sequences with a total of $N = N_a + N_b$ symbols and the collection of all these sequences provides us with a sampling distribution. In this way we are led to the sampling distribution of the statistic V. This sampling distribution has a mean and variance, from Shaum & Spiegel (1961), given by

$$\mu_V = \frac{2N_aN_b}{N_a + N_b} + 1 \qquad \sigma_V^2 = \frac{2N_aN_b\,(2N_aN_b - N_a - N_b)}{(N_a + N_b)^2\,(N_a + N_b - 1)}. \tag{3.12}$$

If both N_a and N_b are at least 8 then the sampling distribution of V is very nearly a normal distribution. Thus

$$z = \frac{V - \mu_V}{\sigma_V} \tag{3.13}$$

Table 3.2. Distribution of Runs for a 1'048'555 bit stream

Runs	Expected	Data	Runs	Expected	Data
1	262144	263231	11	256	265
2	131072	130657	12	128	131
3	65536	65496	13	64	64
4	32768	32783	14	32	22
5	16384	16373	15	16	13
6	8192	7982	16	8	4
7	4096	4222	17	4	5
8	2048	2081	18	2	1
9	1024	1048	19	1	0
10	512	501	>19	0	0

is very nearly a standard normal distribution (mean 0 and variance 1) and can easily be tested for using the standard tables.

3.5.3.2. The Distribution of Runs test

The distribution of runs test of Hawthorne (2002) is a versatile test for accessing important characteristics of streams of different kinds. In evaluating the merits of an apparent random bit stream there is a tendency to be suspicious of long runs of 0's or 1's. Most are surprised to learn that in a random stream of 1 048 555 bits, the expectation is that 127 of the runs will be in excess of 12!

Say we have a binary stream starting with a 1 bit. Under conditions of randomness and uniformity, the probability that the next bit will be a 1 is 0.5, in which event the run will have accumulated to 2. The probability that the next two bits will be 1 is 0.25 and so on. In general,

$$\frac{P\left(run = N\right)}{P\left(run = N+1\right)} = 2.$$

For this we need a sample stream of not less than $N = 2^{P+1} - P - 2$ terms with P a positive integer. Thus for $P = 19$, $N = 1\,048\,555$. The expected distribution of run lengths in such a stream is shown in table 3.2 together with some test data. This test also gives conclusion as to whether a stream is satisfactory for use in an encryption algorithm.

3.5.3.3. Uniformity test on decimal level

The test that will be used here is Pearson's well known χ^2 test as described in Rice (1995), with binning employed where necessary. A goodness of fit may be assessed informally by comparing the observed O and expected counts E which appear to agree quite

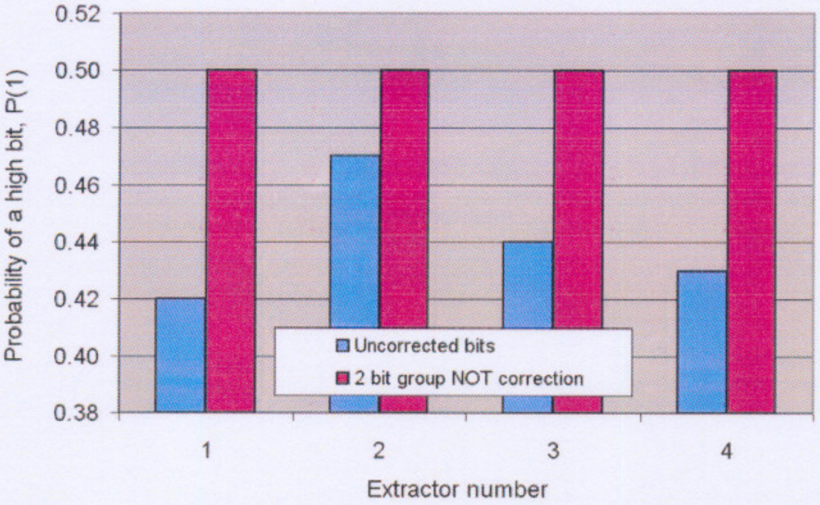


Figure 3.3. The biasing plot of the 4 extractors before and after the NOT correction, for each extractor producing 10^8 bits.

3.5.4.1. Initial Trials

Initially it was thought that only the biasing on binary level posed a problem and had to be corrected for. Of the 4 QBE units under test, the probability of obtaining a high bit (a "1") was between 0.42 and 0.47. Ideally we want this at 0.500... , within certain fluctuations acceptable from a statistical point of view. This has been achieved by pairing the bits in groups of two bits and applying the NOT operation to both bits in every second group, i.e.

$$x_1x_2\ x_3x_4\ x_5x_6...\ \ 00\ 10\ 11\ 10\ 01\ 10\ 00\ 10$$

became

$$x_1x_2\ x_3^{'}x_4^{'}\ x_5x_6...\ \ 00\ 01\ 11\ 01\ 01\ 01\ 00\ 01.$$

This NOT correction worked to some extent, apparently removing the biasing and giving

$$P(0) \simeq P(1) \simeq \frac{1}{2},$$

figure 3.3 giving a representation of before and after. After this procedure the resulting sequence even satisfied the runs test for randomness on binary level, this shown for different NOT correction group sizes in figure 3.4. Note that only the 2 bit group sizes in the NOT correction were successful. The problems with this approach became clear when these binary sequences were used to obtain decimal numbers of any bit length, even when some

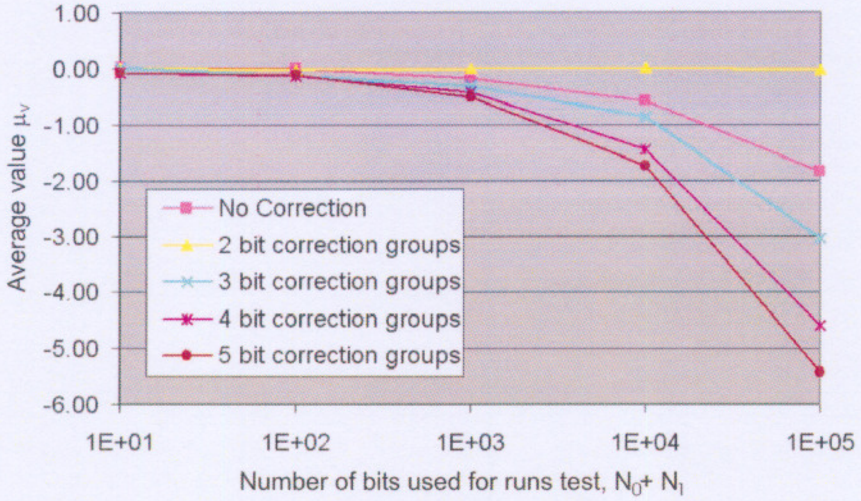


Figure 3.4. Graph showing change of average z-value obtained from the runs test (see section 3.5.3.1), for different number of bits used in the test. Note that only the 2 bit NOT correction groups satisfy statistical requirements by using the law of large numbers.

bits were discarded between adjacent conversion sequences (intermediate bits). In all cases the decimal numbers obtained showed some pattern, as some numbers were more probable to be obtained than others. This was outside of the fluctuations allowed by the statistical approach for uniformity. Figure 3.5 gives a good representation of this problem. The fault with this approach was a violation of one of our important conditions for randomness: the distribution of all elements had to be identical. This was not the case after the NOT operation has been applied since this correction made $0.53 < P(1) < 0.58$ for those inverted bits, while $0.42 < P(1) < 0.47$ still held for the other bits in the sequence, thus the distribution of the elements in the sequence were not identical. Possible cross-talk between extractors can pose problems and should be considered later on during the development of the product.

3.5.4.2. Markov Correction for independence

We follow a similar path as discussed by Peres (1992). Assume we have a sequence $x_1x_2...x_n...$ generated by a stationary two-state (like 0 and 1) Markov process. We also assume that only adjacent bits have some dependence,

$$P(x_i \mid x_{i-2}) = P(x_i) \neq P(x_i \mid x_{i-1}) \quad \forall\, i = 3, 4, ... \tag{3.14}$$

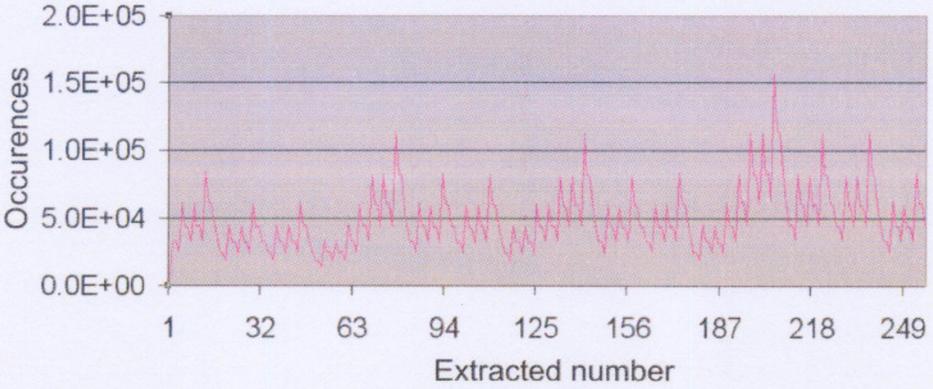


Figure 3.5. Graph showing a measure of uniformity on the decimal level for 2-bit NOT correction groups.

In the simple case of tossing a coin this implies that picking up a coin that shows heads may affect the next toss's outcome. Specifically, in binary language we can write this as

$$\begin{aligned} P\left(x_i = 0 \mid x_{i-1} = 0\right) = \alpha_0 \neq \alpha_1 = P\left(x_i = 0 \mid x_{i-1} = 1\right) \\ P\left(x_i = 1 \mid x_{i-1} = 0\right) = \beta_0 \neq \beta_1 = P\left(x_i = 1 \mid x_{i-1} = 1\right). \end{aligned}$$

This then takes into account the previous bit extracted but no other information. Since we know that one of the basic properties of density functions is that the sum of the probabilities of all possible outcomes must be one, we can use this normalisation condition:

$$\begin{aligned} 1 &= \textstyle\sum_{\text{all } i} p\left(x_i\right) \\ &= p(0) + p(1) \\ &= p\left(x_i = 0 \mid x_{i-1} = 0\right) + p\left(x_i = 1 \mid x_{i-1} = 0\right) \\ &= \alpha_0 + \beta_0 \end{aligned}$$

We have used the Markov property as in equation 3.14 between lines 2 and 3 for trigger state 0. The proof for trigger state 1 follows the same route except that condition $x_{i-1} = 1$ is used. Because, for all four QBE units we had $P(0) > P(1)$, we wish to use this property to obtain the most possible output bits which are independent. Several mappings are given, of which the most efficient then is obtained by using the 0 trigger state:

$$00 \rightarrow 0 \qquad 01 \rightarrow 1 \qquad 10 \rightarrow \emptyset \qquad 11 \rightarrow \emptyset, \tag{3.15}$$

$\emptyset$ implying no output and the resulting independent output stream is denoted by $y_1y_2...y_n....$ Mapping (3.15) produces 0 and 1 outputs for y_k in the 0-state of the process (which occurs more frequently). The output digits of this mapping are then independent but not necessarily equiprobable. The efficiency of mapping (3.15) is $\leq \frac{1}{2}$ but $\geq \frac{1}{4}$ if the more probable trigger state is selected.

It must be noted however that the problem could have been solved to a high degree of approximation by the Von Neumann method (section 3.5.4.3) itself when pairing up non-adjacent items, but those say 10 items apart, say for ex. x_1x_{11} and $x_{21}x_{31}$ and so forth. This however wastes 90% of the data, resulting in a very inefficient procedure which only gives approximate results and will therefore not be considered further on.

3.5.4.3. Von Neumann Correction for uniformity

Von Neumann (1951) described a procedure for generating an output sequence $z_1z_2...z_n...$ of statistically *independent* and *equiprobable* binary digits from an input sequence $y_1y_2...y_n...$ of *independent* bits. For the input stream we have for all n, $P(x_n = 1) = p$ and $P(x_n = 0) = 1 - p = q, \quad 0 < p < 1$. Say possible mappings for each of the pairs y_1y_2, y_3y_4, ... are

$$\begin{array}{cccc} 00 \rightarrow \emptyset & 01 \rightarrow 0 & 10 \rightarrow 1 & 11 \rightarrow \emptyset \\ 00 \rightarrow \emptyset & 01 \rightarrow 1 & 10 \rightarrow 0 & 11 \rightarrow \emptyset \end{array} \qquad (3.16)$$

with $\emptyset$ again representing no output, to produce the output sequence $z_1z_2...z_n...$ of independent and unbiased bits. If the output bit has a probability P' we have from the mapping that

$$P'(1) = pq = qp = P'(0)$$

which proves that the output bits are equiprobable.

The efficiency of this procedure is the expected number of output digits per input digit. For each input pair the probability of generating an output digit z is $2pq$, so the efficiency is pq, which is $\frac{1}{4}$ at $p = q = \frac{1}{2}$ and less elsewhere. The mapping (3.16) is independent of the value of p, the output 0's and 1's are statistically independent and equiprobable (thus uniform) for any $p \in (0, 1)$, but the efficiency depends on p.

3.5.4.4. Results of corrections

The input data $x_1x_2...x_n...$, which were correlated and identically distributed, were fed to the digital corrector section. This started with a Markov correction procedure to deliver a stream of identically distributed but independent output bits $y_1y_2...y_n....$ This was again well. This idea is quantified by Pearson's chi-squared statistic:

$$\chi^2 = \sum_{all\ cells} \frac{(O_i - E_i)^2}{E_i}$$

and it has degrees of freedom df = number of cells - number of indep. params. fitted - 1. As a rule of thumb, the fit is good if the value of χ^2 is about the same (within 10%) as the degrees of freedom. The closer the values are to each other, the higher the probability of a good fit.

This test can easily be employed for a uniformity test on decimal level. For a 12-bit binary number, we have 4096 possible decimal integer outcomes (0 to 4095) and the number of degrees of freedom is 4095. We obtain our χ^2 value by taking the total number of decimal outcomes used and divide that by 4096 to obtain the expected number of occurrences of each number, E_i. We then use the observed number of occurrences of each decimal outcome O_i in the formula for all possible outcomes. This then gives a fast and easy way to statistically analyse the uniformity of the resulting decimal sequence.

3.5.4. Corrections for independence and uniformity

At first it was assumed that only a correction of the biasing (and therefore uniformity) in the bits had to be applied. This has been tried in several ways which, after proving to be unsuccessful, was approached from a theoretical point of view and found to be erroneous. Further theoretical investigation into the statistical approach to unbiasing led to implementation of a process described by Von Neumann (1951) to unbias an independent and identically distributed sequence of binary numbers. This unbiasing alone did not prove effective since problems with uniformity were encountered once these unbiased bits were converted to decimal numbers. After much thought and experimentation it was found that there existed a dependence (and therefore a correlation) between adjacent bits in the output sequence, because of the continuous source signal from which they were extracted, thus violating the condition of independence. Assuming that we had a two-state Markov process, certain steps could be taken to remove this correlation between adjacent bits, according Samuelson (1968). The final process thus involves first correcting for the correlation between adjacent bits assuming a two-state Markov process and then unbiasing the resulting binary sequence using a Von Neumann process to obtain a truly independent, unbiased and random sequence of binary bits. This is then in an appropriate format for conversion to uniform decimal numbers of the desired bit length.

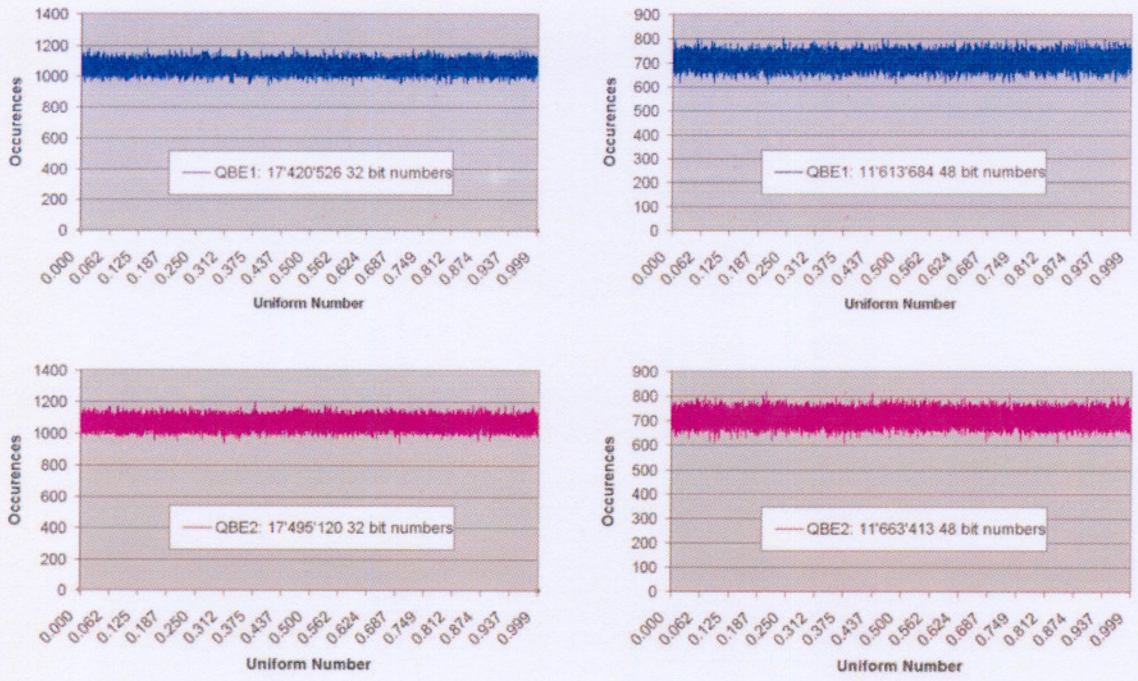


Figure 3.6. The uniformity of 32- and 48-bit numbers from QBE units 1 & 2, with the bits implying the number of bits used from the random binary stream and converted to obtain each decimal number. 16383 bins have been employed to plot the data and to obtain χ^2 statistics.

used as the input for the Von Neumann correction procedure that delivered as output a stream of independent and unbiased bits, $z_1 z_2 ... z_n ...$. The resulting output stream was used to obtain decimal numbers of the required bit length. The results were tested statistically and shown to be successful, as depicted in both the graph of uniformity (figure 3.6) and the χ^2 test (table 3.3). Therefore we had successfully corrected the input stream to the digital corrector to an output stream of independent and unbiased bits which were truly random in nature and appropriate for conversion to decimal numbers, which also proved to be independent and uniform. This is the base format in which we need the decimal random numbers to be of practical use. From these numbers, other distributions like the Gaussian distribution (see section 3.5.5) can quite easily be obtained, leaving the end-user free to choose the required distribution, the QBE hardware and its software drivers taking care of the rest.

For the simulation of the H-test distribution discussed in section 3.6, a dataset of

Table 3.3. Table of expected and obtained χ^2 values from QBE units 1 & 2 for the uniformity test on the decimal level (some shown graphically in figure 3.6). Note that for the 32- and 48-bit numbers binning of size 16384 has been employed.

Decimal Number	QBE1 χ^2	QBE2 χ^2	$E(\chi^2)$
10 bit	1029	1085	1023
11 bit	2061	2095	2047
12 bit	4104	3981	4095
32 bit	16471	16858	16383
48 bit	16288	16316	16383

Table 3.4. Table showing the χ^2 values for the Distribution of Runs (DoR) test and uniformity (U) test on decimal level, for the dataset employed in the simulation of the H-test distribution.

	Generator 1	Generator 2	Generator 3	Generator 4	Expected
DoR Average	21.0	17.8	19.4	18.5	19
DoR Stddev.	11.8	6.0	6.7	6.2	
U Average	997.3	1001.8	998.5	1001.2	1023
U Stddev.	44.5	41.7	44.2	43.7	

$4,361 \times 10^8$ truly random 32-bit numbers were used. Four generator units were used to obtain this data and each unit produced 249 datasets. On the binary level the Distribution of Runs test (section 3.5.3.2) was employed to verify randomness and on the decimal level the χ^2 test for uniformity was employed (section 3.5.3.3). Table 3.4 shows the values obtained, confirming that the generator produces truly random numbers.

3.5.4.5. Hardware implementation of correction processes

The abovementioned correction procedures were first extensively tested using software which proved to be a long and tedious process, especially with the large volumes of data that required processing. This provided motivation for the implementating a hardware solution by means of digital electronics. Since for both the Markov and Von Neumann correction processes groups of 2 adjacent bits are compared for a possible output bit, comparison can be activated by a divide-by-two counter in hardware. Bits are read into a shift register (2 bit length) and as soon as the shift register has been filled the divide-by-two counter activates comparison for a possible output to the next level. The digital clock (6) in figure 3.1 drives the whole process of acquiring data and correcting it.

The full circuit is given in figure 3.7. The Digital Input Stream (DIS) and Clock (CLK) are inputs to the first section taking care of the Markov correction for independence. This section consists of 3 D-type latches and a single 2-input AND gate. The rightmost latch in

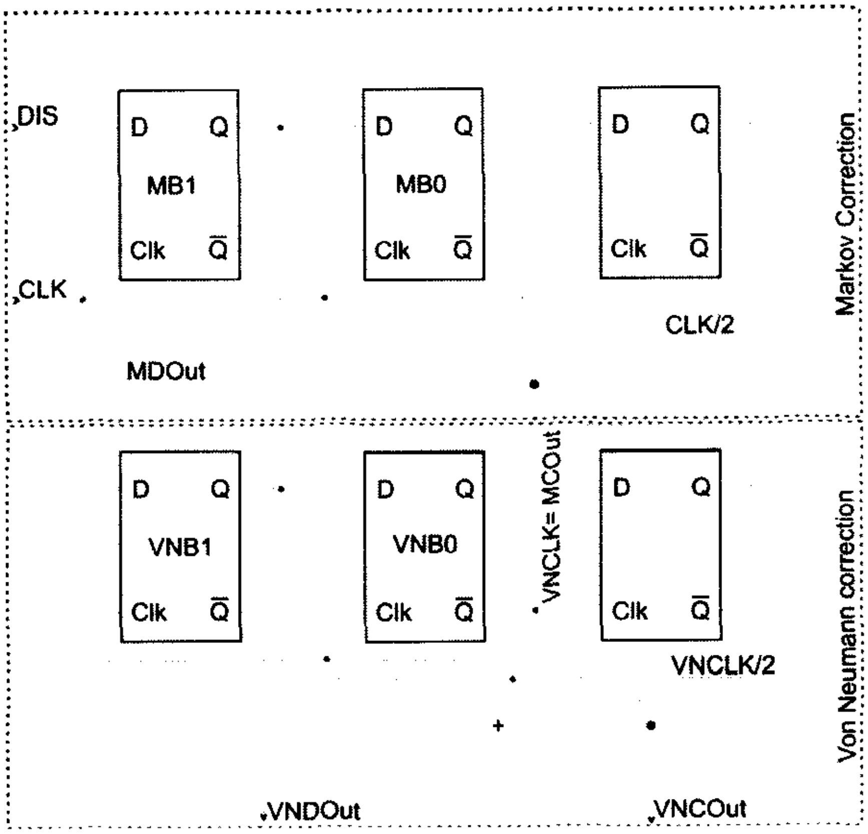


Figure 3.7. Schematic of logic of the Digital Corrector. Symbols are defined in the text.

the diagram is connected as a divide-by-two clock, that activates possible output on every second pulse of CLK, i.e. it divides the DIS in groups of two bits for comparison. The other two D-type latches MB0 and MB1 are connected in a shift register setup to store two adjacent bits of the DIS and the AND gate is connected for a trigger 0 event (the first bit of a group from the DIS must be 0), as depicted in mapping (3.15). The output from the AND gate (VNCLK) drives the Von Neumann corrector section to indicate a valid bit on MDOut, which according the mapping should be the last of the two bits entering the shift register. VNCLK is not periodic anymore but gives a pulse each time a valid output bit from the Markov corrector can be stored.

The Von Neumann corrector works in essentially the same way with a divide-by-two clock and shift register formed by VNB0 and VNB1. At also compares bits in groups of two bits, the XOR gate taking care of the condition of different bits as depicted in mapping (3.16) with the output taken as the inverse of the last bit to enter the shift register (making the output the same state as the first bit to enter). VNCOut gives the output clock pulses to indicate when a new valid output bit can be read from data output VNDOut, and is also not of a periodic nature. This data output stream then consists of independent and unbiased bits which are ready to be used for uniform, truly random numbers.

The Interface / Communication controller stores these data in buffers and gives the appropriate signal to the application once it is ready with at least a specified number of bits (section 8 in figure 3.1). This is the usually input to a PC using USB, PCI, Ethernet or Firewire interfaces. A software driver takes care of the formatting of the numbers according to the user preset selection and provides access to the data in the proper way.

3.5.5. Hardware Gaussian Number generation

Recall that the Central Limit Theorem (section 3.2) gives an estimate of the distribution of a sum of a large number n of independent but identically distributed random numbers as a normal distribution. This is independent of the distribution of the random variables, provided they have finite expectations and variances and n is 'large enough' (see section 3.2). This gives us the means to easily construct a Gaussian number generator using random numbers with any distribution whatsoever, simply by using the sums of random numbers.

We denote the sum of n uniform $U(0,1)$ random numbers as R_n, so that R_1 will be a random number distributed uniformly (between zero and one). Then R_2 would be distributed with a density function which is a triangle with possible outcomes from 0 to 2, the most probable outcome as 1. This kind of distribution is familiar to gamblers using dice, where the outcome is the sum of two numbers uniformly distributed between one and six. The extreme values of the sum (2 and 12) are the least likely, where the middle value 7 is the most likely. R_3 already has a continuous-like curve which is beginning to look like the well-known bell-shaped Gaussian curve. After R_6 the distribution is almost indistinguishable from a true Gaussian by eye, except for the extreme tails which are of course of finite length whereas true Gaussian tails go to infinity in both directions. Figure 3.8 gives a graphical representation of this. Because of the importance of the tails in the simulation of the H-test, care must be taken to ensure that the approximation is close enough.

Since the expectation and variance of the uniform U(0,1) distribution are $\frac{1}{2}$ and $\frac{1}{12}$ respectively, we have

$$E(R_n) = \frac{n}{2} \qquad Var(R_n) = \frac{n}{12}. \tag{3.17}$$

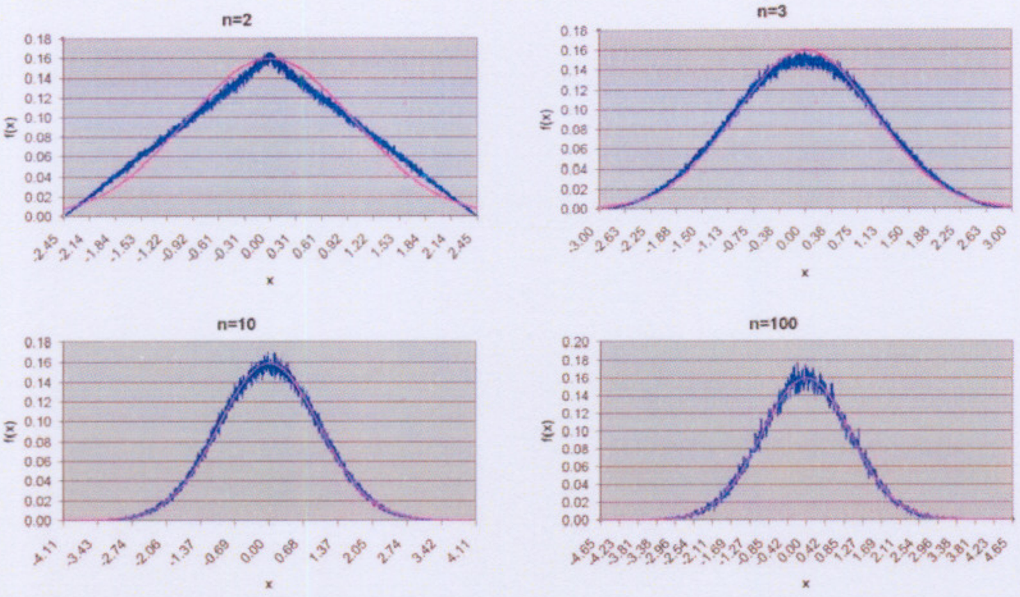


Figure 3.8. Normalised distributions of R_n in blue with n as denoted for each graph, compared to the standard Gaussian distribution N(0,1) in purple. In each case R_n was calculated using the same set of 8 672 418 32-bit uniformly distributed truly random numbers.

Usually we want a standard Gaussian distribution $N(0,1)$, therefore we take

$$\frac{x-\mu}{\sigma} \equiv \frac{R_n - n/2}{\sqrt{n/12}} \to N(0,1).$$

This also makes it more appropriate to use specifically uniform numbers in the hardware process since these parameters are known analytically beforehand, whereas with other distributions, these parameters would need to be calculated each time.

During tests, it was noted that with R_{100} and R_{1000} the main form of the Gaussian was accurately represented (the χ^2 test proved this). Clearly the tails of the distribution of R_n, $n \leq 1000$ will only accurately represent a Gaussian if the number of times R_n is obtained is large enough, since the probability included in the tails is so small. Table 3.5 shows how the possible outcomes for $z \sim N(0,1)$ is influenced by n. It seems that it is better to take $n = 100$ and obtain more z-values in such a way as to accurately represent most of the tails as well, with the boundaries also being large enough since a z-value of ± 7 already has a probability of $\sim 10^{-12}$, well within the accuracy with which we want to simulate the H-test.

Table 3.5. The boundary and obtained values of $z \sim N(0,1)$ for different n values and input set size. Note the improved observed outcomes when the input set size was increased by an order of magnitude.

		Possible Outcomes		Observed Outcomes	
32-bit numbers	n	min	max	min	max
8 672 418	10	-5.48	5.48	-4.30	4.04
8 672 418	100	-17.32	17.32	-4.55	4.22
8 672 418	1000	-54.77	54.77	-3.31	4.15
34 915 646	1000	-54.77	54.77	-3.85	4.46

Implementation of the R_{100} arithmetic in hardware should be straightforward: A binary adder of 101 32-bit numbers, with the starting number equal to $-\frac{n}{2} = -50$ is used after which the result is divided by $\sqrt{\frac{n}{12}} = 17.32....$ The resulting numbers are then distributed $N(0,1)$ as required. This hardware process is considerably faster than the processing via software counterpart on a PC, making it quite useful as a possible addition to the final device.

3.6. Simulation and Results: H_M test distribution

The goal here, and the final goal for this dissertation, is to determine the distribution of the H_M test (as defined by 3.10 on page 38) by utilising the developed hardware truly random number generator. This will serve to demonstrate the applicability of this generator to such tasks as well as build confidence in the obtained output. Later simulations of the H-test up to an accuracy of 10^{-12} will help in identifying a periodic signal in a large dataset (>10 000 events) with a very low signal-to-noise ratio, especially when no previous knowledge of the light curve is available.

To obtain the distribution of

$$H = \underset{1 \leq m \leq 20}{max} \left(Z_m^2 - 4m + 4\right) \tag{3.18}$$

we first need to obtain

$$Z_m^2 = \sum_{k=1}^{m} 2n\overline{R}_k^2 = 2n\sum_{k=1}^{m}\left(\overline{c}_k^2 + \overline{s}_k^2\right) \tag{3.19}$$

with

$$\begin{aligned} \overline{c}_k &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n}\cos\left(k\theta_i\right) \\ \overline{s}_k &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n}\sin\left(k\theta_i\right), \end{aligned} \tag{3.20}$$

$\theta_i \epsilon \, [0, 2\pi)$ and n the number of events. The ϕ_i are the phases of the events modulo the pulse period searched for, and are obtained by folding the event data on the unit circle by using

$$\phi_i = \phi_0 + f\,(t_i - t_0) + \frac{1}{2}\dot{f}\,(t_i - t_0)^2 + \frac{1}{6}\ddot{f}\,(t_i - t_0)^3, \qquad \theta_i = 2\pi\phi_i \tag{3.21}$$

with f the tested for pulse frequency and t_i the time of event i. For pulsar 0833-45 (Vela) for example we have $f = 11.1390092984739\,s^{-1}$, $\dot{f} = -1.55781 \times 10^{-11} s^{-2}$, $\ddot{f} = 8.92 \times 10^{-22} s^{-3}$ and $t_0 = 52775.000000202$ MJD from the ATNF Pulsar Catalogue.

The modus operandi in this section will be to show that using the statistical properties of the cosmic ray background arrival times, the phase folding produces a uniform distribution as expected. Therefore it will not be necessary to simulate this phase folding but one can use the random numbers in the same format as the distribution obtained after phase folding. This will also serve to verify the true randomness of the QBE since it will be clear that the obtained distribution and analytical distribution are within statistical limits of each other. This process will be repeated through the sine and cosine moments distributions, equation (3.20) up to the distribution of the Z_m^2 test, equation (3.19). Once this has been proved to be a χ^2_{2m} distribution, it will only be necessary to use the random numbers from the QBE and transform them to the χ^2_{2m} distribution of the Z_m^2 test, which can then be used in obtaining the distribution of the H-test. Therefore this procedure eliminates the unnecessary use of many random numbers and computing time.

The arrival times of Cerenkov events always include those of the cosmic ray background, which is isotropic and uniform in nature as discussed earlier. These arrival times may however include a pulsed cosmic ray component if observations are made on certain periodic cosmic ray emitters like pulsars. The arrival times t_i of only background Cerenkov events are distributed as a Poisson process with parameter λ as the count rate. The waiting time between these arrival times is then distributed exponentially with parameter λ. For the background Cerenkov events we can take $t_0 = 0$ and

$$t_{i+1} = t_i - \frac{\ln r_i}{\lambda}, \qquad r_i \sim U\,(0, 1) \tag{3.22}$$

as the arrival times of the subsequent events.

For simulation of the distribution of the H_M test we then only need to simulate this cosmic ray background component, since we want to test for the null hypothesis. A rejection of this null hypothesis then indicates the presence of some periodic component in the arrival times. Using the arrival times as obtained from (3.22) in (3.21) to obtain the phases

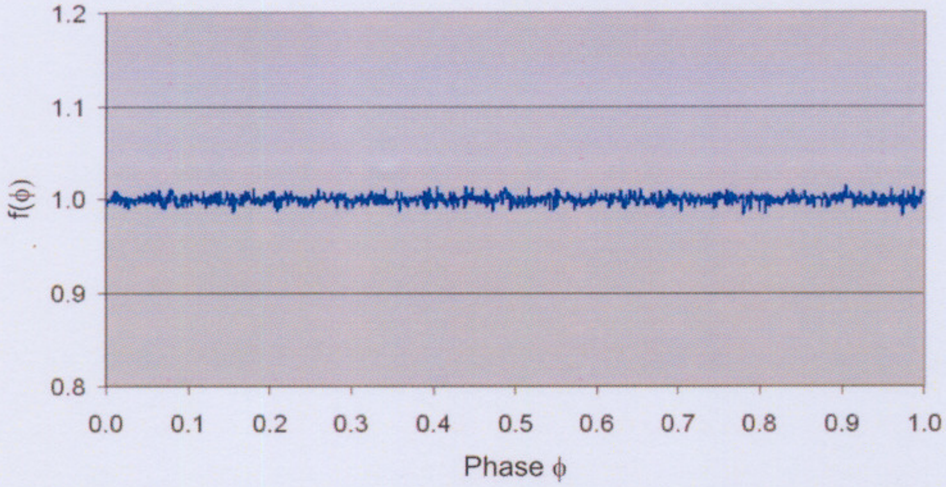


Figure 3.9. Graph of normalised probability $f(\phi)$ using only random arrival times as in eq. (3.22) and folding the data according eq. 3.21 using as parameters those of the Vela pulsar.

according some parameters, we expect the distribution of ϕ to be $\sim U(0,1)$ when no periodicities are present. Figure 3.9 shows that the outcome of the phase folding is again uniform as expected. This clearly indicates that it is not necessary to obtain the phase probabilities by calculation but to rather use the extracted $U(0,1)$ numbers directly and multiply them with 2π to obtain the θ_i's as used in eq. (3.20).

Using $U(0,1)$ numbers in eq. (3.20), we expect that both are distributed independently and asymptotically $N\left(0,\ 1/(2n)^{1/2}\right)$ from Mardia (1972). Note that the use of uniform data result in an independence of the distribution of both $\overline{c}_k$ and $\overline{s}_k$ of the k-parameter. Taking n as 10 and 100 respectively we clearly obtain the expected Gaussian distribution as seen in figure 3.10. Therefore we again need not calculate the $\overline{c}_k$ and $\overline{s}_k$ values of but only need to transform $U(0,1)$ numbers to be distributed $N\left(0,\ 1/(2n)^{1/2}\right)$. These transformed numbers can then be used directly in eq. (3.19).

The Z_m^2-distribution is also distributed as χ^2_{2m}(Bendat & Piersol, 1971), this being easy to simulate since it is by definition the sum of $2m$ values of Z^2 with Z a standard Gaussian random variable. In calculation of eq. (3.19) we need as input numbers distributed as $N\left(0,\ 1/(2n)^{1/2}\right)$. Thus we need to convert the uniform $U(0,1)$ numbers to standard Gaussian numbers and obtain the required Gaussian distribution by using the fact that, if $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$, then $Y \sim N(a\mu + b,\ a^2\sigma^2)$. To obtain the standard Gaussian numbers from $U(0,1)$ numbers, we used a variant of the standard

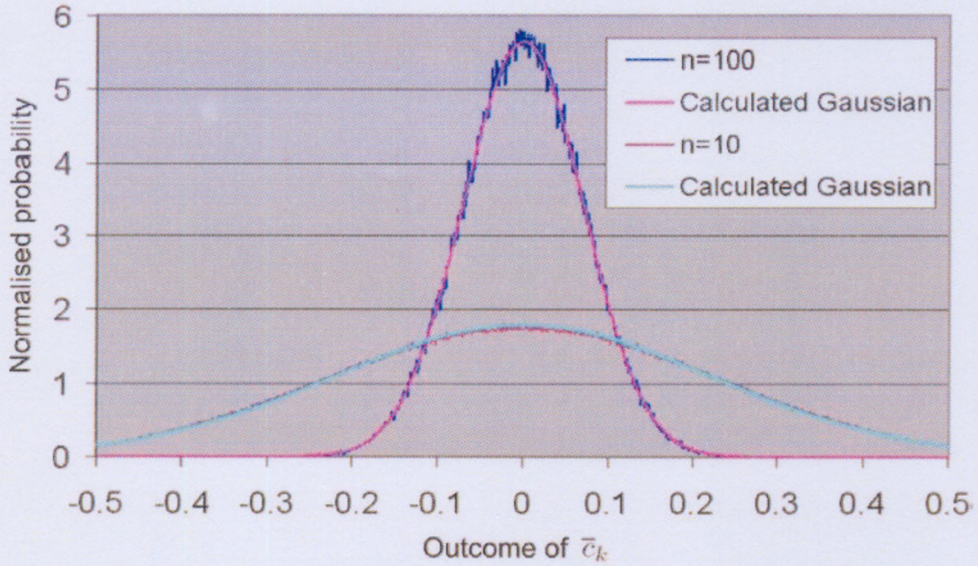


Figure 3.10. Graph showing the obtained and expected distribution of $\overline{c}_k$ for different values of n.

Box-Muller transformation (Gentle 2003). The Box-Muller transformation is also known as the polar method and states that having two uniform numbers U_1 and U_2, one obtains two independent $N(0,1)$ numbers X and Y using

$$X = \sqrt{-2logU_1}\cos(2\pi U_2)$$
$$Y = \sqrt{-2logU_1}\sin(2\pi U_2).$$

Figure 3.11 shows that the distribution obtained using this method agrees extremely well with the analytical standard Gaussian distribution. The variant of this method is called the rejection method and provides a minor increase in speed by eliminating the calculation of sine and cosine functions. The algorithm also uses two uniform numbers U_1 and U_2 to obtain two independent $N(0,1)$ numbers X and Y:

$$V_1 = 2U_1 - 1, \qquad V_2 = 2U_2 - 1,$$

if $w = V_1^2 + V_2^2 < 1$, continue to calculate $k = \sqrt{-2\ln(w)/w}$ to obtain

$$X = kV_1, \qquad Y = kV_2.$$

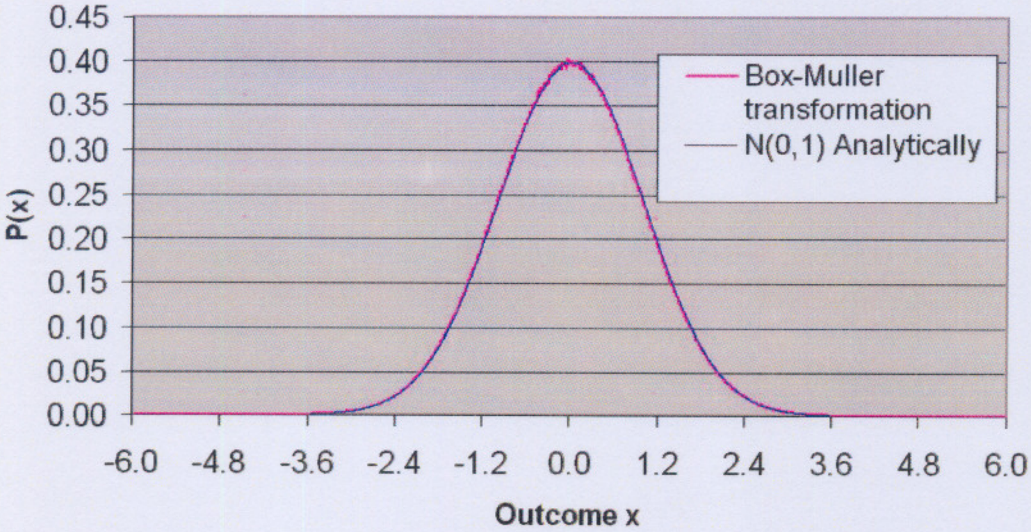


Figure 3.11. Comparison between the analytical standard Gaussian distribution and values obtained using the Box-Muller transformation on uniform random numbers.

The rejection method is generally about 30% faster than the normal Box-Muller transformation.

The analytical density of the χ^2_{2m} distribution involves gamma-functions and is quite difficult to compute for non-integer values. Since the uniform output of the generator has been shown to accurately produce normally distributed numbers, we can use these $N(0,1)$ numbers to simulate the χ^2_{2m} distribution directly, by summing the required number of the squares of such $N(0,1)$ values:

$$\chi^2_{2m} = N_1^2 + N_2^2 + ... + N_{2m}^2 \tag{3.23}$$

Comparing it with the distribution obtained by using $N\left(0,\ 1/(2n)^{1/2}\right)$ numbers in eq. (3.19) one can see the statistical equivalency between the two in figure 3.12.

At this point, all that remains to be done is the simulation of the H-test. This can now easily be obtained by transforming the $U(0,1)$ numbers obtained from the generator to the χ^2_{2m} distribution which is statistically equivalent to the Z^2_m distribution as in eq. (3.23), and use them in eq. (3.18). The code is given in Appendix I and the obtained distribution together with the functional fit obtained by De Jager et al. (1989) is given in figure 3.13. From this figure it is clear that the previous fit for the H-test was unaccurate- it underestimated the uniformity probability for $0 < H < 8$ and overestimated for $H > 9$.

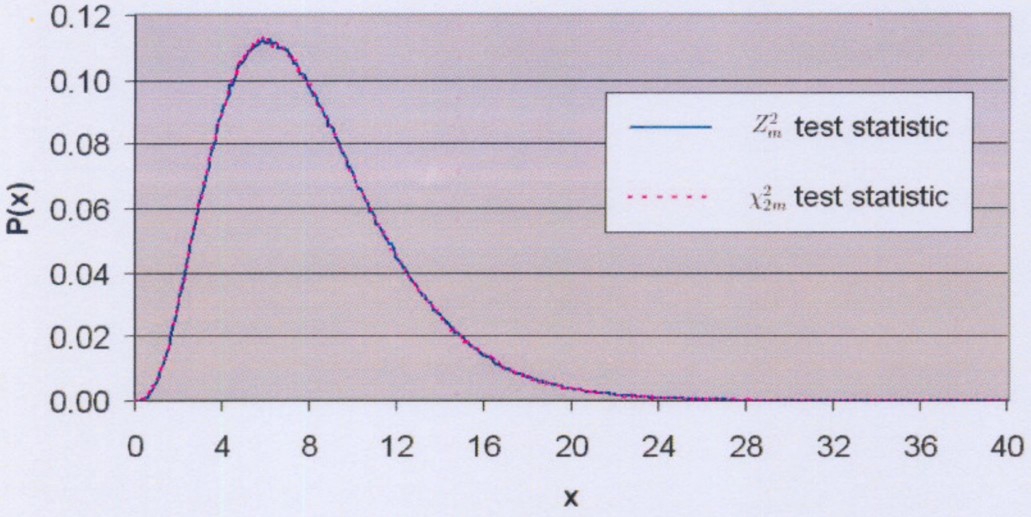


Figure 3.12. Comparison between the distributions obtained for the Z^2_m and χ^2_{2m} statistics.

This simulation utilised $\sim 10^7$ H-values and only contained dense enough data up to $H = 20$. Further simulations of $> 10^{10}$ values of H will be necessary to obtain the tail of the distribution accurately and be able to make a fit to the totality of the H-distribution. It seems like the functional form will be something like

$$y = a\exp(-bh) + cx^3 + dx^2 + ex + x_0.$$

Determination of this within the required accuracy will provide a more powerful tool for periodicity searches.

3.7. Further uses for real random numbers and other generators

Randomness and random numbers have traditionally been used for a variety of purposes. It is used in computer games, scientific experiments and simulations, and in the generation of cryptographic keys. In finanacial markets Monte Carlo simulation is a popular method for pricing financial options and other drivative securities because of recent advances in applying the tool. Thus we can describe randomness on three basic levels: for everyday, scientific and commercial purposes.

Everyday uses include computer applications such as games using dice, shuffled cards etc. Typically, random numbers used in such applications need not be of the truly random

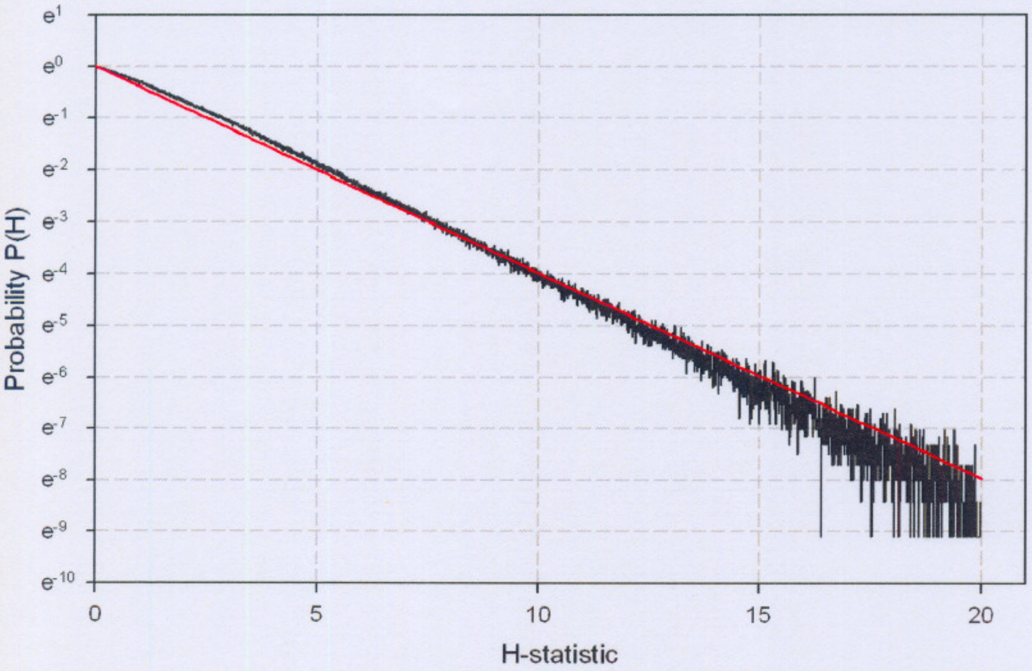


Figure 3.13. The distribution of the H-test as from simulation (black) and the functional fit from De Jager et al. (1989) superimposed (red).

nature as modern pseudo-random generators will suffice in providing 'random enough' numbers. Typically then in such cases there is some level of predictability in the sequence but not at a level that is observable without mathematical and statistical analysis.

For scientific use it may be sometimes convenient that a series of random numbers can be reused, either for several experiments or for debugging purposes. For this purpose pseudo-random numbers are well suited. On the other hand, simulations may need a large number of unpredictable and identically distributed random numbers. For this a truly random number generator is necessary. Usually these truly random numbers are needed at a high rate so as not to slow down any calculations being made with them. This is typically the requirements of the H-test, where the simulation of an accurate distribution is the reason behind the development of the Truly Random Number Extractor.

For encryption purposes, the level of security required determine the type of random number (pseudo or truly) and the encryption algorithm required, together with word length. Cryptographic algorithms come in a variety of flavours. Some are strong (meaning difficult to decipher) but make substantial demands to processing power and key management. Others are weak (meaning easier to decipher) but generally less demanding and therefore better suited for some applications. All strong cryptography requires truly random numbers to generate keys, but how many depends on the encryption scheme. The strongest possible method, One Time Pad encryption, is the most demanding of all; it requires as many random bits as there are bits of information to be encrypted. Weak encryption schemes on the other hand, typically employ pseudo-random generators which are much faster and can be implemented on any PC.

In a modern economy, it is important for firms to be able to select an appropriate level of risk in their transactions. In practice, simulation is often used to price derivative securities. It is also used for estimating sensitivities, risk analysis, and stress testing portfolios. Such simulations are employed using Monte Carlo methods, this being attractive relative to other techniques because it is flexible, easy to implement and modify, and its applicability does not depend on the dimension of the problem. Complex analytical real options models suffer the curse of dimensionality: when several sources of uncertainties and several investment options are considered simultaneously, analytical methods start to fail. Because of this, Monte Carlo methods are used more frequently nowadays since it doesn't suffer from this curse of dimensionality. These methods rely heavily on uniformly distributed random numbers, where it is known that the best pseudo RNG successfully handles up to 623 dimensions. The commonly used LIBOR market model is a practical example of where the dimensionality of the problem easily exceeds 700. The availability of a very fast and truly uniform RNG is therefore of the utmost importance for the continued use of Monte Carlo methods in financial markets.

A few hardware truly random number generators exist on the market. These are rare and usually either costly (since they employ quantum mechanical-like measurements) or don't produce truly random data at a very fast rate (>1Mbps). Two such hardware implementations are the Intel RNG (Intel RNG Whitepaper 1996) and VIA C3's Nehemiah RNG. Intel embedded a RNG into its architecture in 1999. It utilises undriven resistors as source for white noise and a Von Neumann-like correction is applied later on to remove biasing. The VIA C3 Nehemiah RNG (VIA C3 Webpage, 2004) is called the VIA Padlock Data Encryption Engine and is used in several processor ranges. Freewheeling oscillators act as the entropy source and a Von Neumann whitener is used to reduce biases.

Another option is websites providing random data online. *HotBits* (2004) from Fourmilab use sampled radioactive decays which are a really good source of entropy. Another source of entropy could be atmospheric noise as heard on an ill-tuned radio, this being used by *random.org*. The *lavarnd.org* group at Silicon Graphics uses lava lamps to generate random numbers. These are also not suitable for purposes that require large sets of random data at high speeds. A weakness of web-based random numbers for applications calling for strong encryption, is that the random dataset can be intercepted during transmission from the server to the client.

Generally it has been found that there exits a need for a very fast and easily accessible truly random number generator. If a device such as the truly random number generator described here can be fitted into a CPU or IC on the PC motherboard, all users will have effective access to fast and reliable random numbers of any length needed for most applications. This also defines one of the goals for the developed truly RNG, which we hope to realise soon.

Chapter 4

Conclusion

In astrophysics the need quite often arises to search for periodicities in data. Because the H-Test scans through 20 harmonics, it actually comes down to the same as using 20 independent tests, which makes it powerful for blind pulsar searches. Therefore the H-test is one of the most powerful test known for such searches if the number of pulses and pulse shapes are unknown beforehand. In gamma-ray astronomy specifically, it has been applied numerous times for pulsar searches at very high energies and it is being implemented in many experiments which are underway such as GLAST and CANGAROO III. Because of the extremely low signal-to-noise ratio of pulsar gamma-rays, it is very important to know the distribution of the H-test with high accuracy, so that tests against the null hypothesis give significant results.

Previously the distribution of the H-test was obtained by utilising pseudo-random number generators. The problems with these generators originates in the deterministic nature of computing devices. This leads to dependencies between numbers and it also has a sparse, non-uniform distribution. For numbers to be truly random, 2 conditions must be satisfied:

1. Every number must be independent from all the other numbers
2. All numbers must have the same distribution function.

Problems start to occur when many ($> 10^4$) of these numbers are used simultaneously. This can be explained by the Marsaglia effect, which states that that if successive d-tuples produced by the multiplicative congruental method are taken as coordinates of points in a d-dimensional space, all the points will lie on a certain finite number of parallel hyperplanes. This immediately shows the non-random effects of this type of generator through the patterns and the non-uniform distribution it produces. Other more modern software methods usually suffer from the same problems but usually to a lesser extent. On the other hand, hardware generators available are always slow and usually very expensive and not very portable, apart from not always having a uniform output distribution. Therefore most available random number generators are inappropriate for the simulation of the H-test up to the required levels of accuracy.

The development of the Quantum Bit Extractor, with goals:

1. To provide truly random numbers
2. from a device that is in a cheap and portable implementation
3. at very high speed

was hugely successful as comes forth from the fact that a patent was awarded for the device in 2004. It utilises some source of high entropy like thermal noise and, after some sampling and feedback procedures, one ends up with a near-random stream of bits. It is near-random because there still exists a correlation between adjacent bits, which comes from the fact that sampling of the continuous noise yields this correlation factor. This problem is corrected for by assuming a two-state Markov process and that only directly adjacent bits are correlated. The statistics show that the simple procedure to select a common trigger state and to take the next bit as corrected data, yields the required bitstream of totally uncorrelated and identically distributed bits. This resulting bitstream however has the property that $P(0) \neq P(1)$, while we would like to utilise a perfect uniform distribution, $U(0,1)$. To format the data according to this requirement, a procedure by Von Neumann is employed to remove the biasing, which then has as output a stream of truly random and unbiased bits.

Applying these statistical corrections to the raw data as obtained from the source of entropy using software, takes much processing power and therefore makes the process very slow. To obtain the high speeds required, these statistical procedures were implemented using electronic hardware in the form of D-latches and logic gates. This speeds up the process by orders of magnitude. The full hardware implementation has huge applicability in other fields like the financial and data security markets which gives it a great commercialisation and marketing potential.

A small simulation of the distribution of the H-Test was completed, with the results already clearly indicating the necessity for such a generator because of problems with the pseudo-randon number generators used previously. Simulation of $> 10^{10}$ values of the H-test is our next goal and the results will be published once available.

Hopefully all this will further especially pulsar astronomy by helping to tighten the model parameters and determining the correctness of Outer Gap and Polar Cap models. The further applicability of such a device in the financial and data security markets renders this device highly applicable and the obtained positive results agreed perfectly with what was set out to be done.

Appendix I. H-test Code

```pascal
procedure H_Test(Filein,FileOut:String);
Var
   FIn, FOut    : TextFile;
   m, tel       : Integer;
   Uget1, Uget2, v1, v2,w, n1, n2, y, Hmaks, H_stat, chi_sq, a  : Real;
   H_array      : Array[0..6000] of integer;
   InStr        : Array[0..6000] of string;

begin
   for tel := 0 to 6000 do H_array[tel] := 0;
   a := 60/ln(60);             // Constant for scaling
   AssignFile(FIn,FileIn);     // File with 32 bit uniform random numbers
   AssignFile(FOut,FileOut);
   Reset(FIn);

   Hmaks := 0;
   While not EOF(FIn) do
      begin
         For m := 1 to 20 do   // Run through harmonics
         begin
            chi_sq := 0;
            repeat             // Basic transformation to N(0,1) numbers
               If not EOF(Fin) then Readln(Fin,uget1);
               If not EOF(FIn) then Readln(FIn,uget2);
               v1 := 2*uget1 - 1;
               v2 := 2*uget2 - 1;
               w := v1*v1 + v2*v2;
            until ((w > 0) and (w < 1)) or EOF(FIn);
            If not Eof(FIn) then  // If valid data in previous step
               begin   // Continue transformation to N(0,1) numbers
                  y := sqrt(-2*ln(w)/w);
                  n1 := v1*y;
                  n2 := v2*y;
                       // Calculate Z^2 and H statistics
                  chi_sq := chi_sq + n1*n1 + n2*n2;
                  H_stat := chi_sq -4*m + 4;
                  If H_Stat > Hmaks then
                     Hmaks := H_stat;
               end;
         end;          // for m = 1 to 20

         // Use exponential scale to store data
         inc(H_array[round(100*(61-exp((60-Hmaks)/a)))],1);
         Hmaks := 0;
      end;             // EOF(FIn) reached
   CloseFile(FIn);

   try
     Reset(FOut);
   except
      begin            // Create file with H values
         Rewrite(FOut);
         for tel := 100 to 6000 do
             Writeln(FOut,FloatToStr(60-a*ln(61-tel/100)));
         Reset(FOut);
      end
   end;

   for tel := 100 to 6000 do       // Read file lines
         Readln(FOut,inStr[tel]);
   Rewrite(FOut);
   for tel := 100 to 6000 do       // Add new number of occurences of H
      Writeln(FOut,InStr[tel] + '; '+ FloatToStr(H_array[tel]));
   CloseFile(FOut);
end;
```

Appendix II. The RNG Patent

A HARDWARE GENERATOR FOR UNIFORM AND GAUSSIAN DEVIATES EMPLOYING ANALOG AND DIGITAL CORRECTION CIRCUITS

TECHNICAL FIELD

This invention relates to a random number generator (RNG) and more particularly a hardware RNG for generating at an output a train of successive truly random bits of first and second states and/or truly random numbers having a Gaussian distribution.

Both software and hardware random number generators are known in the art. The output signals of the known software generators are not truly random, but pseudo random and these generators are generally slower than hardware generators. Some known hardware RNG's comprise quantum mechanical optical devices, are expensive, bulky and difficult to implement. The outputs of other hardware RNG's do not exhibit statistically truly random behaviour. In this specification the term "truly random" is used to denote a collection of elements wherein the elements are independent from one another and identically distributed.

OBJECT OF THE INVENTION

Accordingly it is an object of the present invention to provide a hardware random number generator and a method of generating random numbers with which the applicant believes the aforementioned disadvantages may at least be alleviated.

SUMMARY OF THE INVENTION

According to the invention there is provided a hardware random number generator (RNG), comprising:

- a source of entropy for providing an input bit stream comprising successive bits of a first state and a second state;
- a first digital corrector comprising a first input and a first output;
- the corrector being configured to provide at the first output from two successive bits in the input bit stream an output bit of a first output bit stream according to a first scheme wherein a first bit of a first state and a second bit of the first state yield an output bit of a third state and wherein a first bit of the first state and a second bit of a second state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the output bit stream at the first output of the corrector are independent from one another.

The first output of the first corrector is preferably connected to a first input of a second digital corrector, the second digital corrector comprising a first output and being configured to provide at the first output from two successive bits at the first input of the second corrector an output bit of a second output bit stream according to a second scheme wherein a first bit of a first state and a second bit of a second state yield an output bit of a third state and wherein a first bit of the second state and a second bit of the first state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the second output bit stream are both independent from one another and unbiased and hence truly random.

The source of entropy may comprise a white noise generating element having an output connected to an analogue high pass amplifier having an output and a cut off frequency of f_0.

The output of the high pass filter may be connected to an input of a voltage level discriminator circuit having an output.

The output of the level discriminator circuit may be connected to a first input of a flip-flop, to generate the input bit stream at an output thereof.

An analogue voltage level compensation circuit may be provided at an input of the voltage level discriminator circuit. The compensation circuit may comprise an RC circuit having an RC time constant which is shorter than $1/f_0$.

The RNG may comprise a Gaussian generator comprising an input connected to either the first output of the first digital corrector or the first output of the second digital corrector, the generator comprising an adder arrangement for generating a sum of j words of i sequential bits each received from the corrector; subtractor means for deriving a difference between the sum and a mean value of the sum; and a divider arrangement for dividing the difference by a standard deviation, thereby to generate at an output of the Gaussian generator a Gaussian deviate.

According to another aspect of the invention there is provided a method of generating a random bit stream comprising the steps of:

- utilizing a source of entropy for providing an input bit stream comprising successive bits of a first state and a second state;
- utilizing a first hardware digital corrector comprising a first input and a first output to provide at the first output from two successive bits in the input bit stream at the first input an output bit of a first output bit stream according to a first scheme wherein a first bit of a first state and a second bit of the first state yield an output bit of a third state and wherein a first bit of the first state and a second bit of a second state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that

successive bits in the output bit stream at the output of the corrector are independent from one another.

The method may include the further step of utilizing a second digital corrector comprising an input and an output in series with the first corrector and wherein the second digital corrector is used to provide at said output from two successive bits at said input of the second corrector an output bit of a second output bit stream according to a second scheme wherein a first bit of a first state and a second bit of a second state yield an output bit of a third state and wherein a first bit of the second state and a second bit of the first state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the second output bit stream are both independent from one another and unbiased.

BRIEF DESCRIPTION OF THE ACCOMPANYING DIAGRAMS

The invention will now further be described, by way of example only, with reference to the accompanying diagrams wherein:

figure 1 is a block diagram of a hardware random number generator (RNG) according to the invention;

figures 2(a), 2(b) and 2(c) are waveforms against time of signals at points 2(a), 2(b) and 2(c) in figure 1;

figure 3 is a logic diagram of one embodiment of a first digital corrector circuit forming part of the RGN in figure 1;

figures 4(a), 4(b), 4(c), 4(d), 4(e) and 4(f) are waveforms against time of signals at points 4(a), 4(b), 4(c), 4(d), 4(e) and 4(f) in figure 3;

figure 5 is a logic diagram of one embodiment of a second digital corrector circuit forming part of the RGN in figure 1;

figures 6(a), 6(b), 6(c), 6(d), 6(e) and 6(f) are waveforms against time at points 6(a), 6(b), 6(c), 6(d), 6(e) and 6(f) in figure 5;

figure 7 is a logic diagram of the first and second digital correctors circuits connected in series; and

figures 8(a), 8(b), 8(c), 8(d), 8(e), 8(f), 8(g), 8(h), 8(i) and 8(j) are waveforms against time of signals at points 8(a),8(b), 8(c), 8(d), 8(e), 8(f), 8(g), 8(h), 8(i) and 8(j) in figure 7; and

figure 9 is a high level block diagram of the RNG according to the invention.

DESCRIPTION OF PREFERRED EMBODIMENTS OF THE INVENTION

A hardware random number generator (RNG) according to the invention is generally designated by the reference numeral 10 in figure 1.

The RNG 10 comprises a source of entropy 12 generating a wide bandwidth analogue output signal 14. The source of entropy may comprise a thermal noise element such as a semiconductor junction, resistor or an avalanche noise generator for generating random or white noise. The element is connected to a high pass amplifier 16 having a cut off frequency f_0 and providing an amplified output signal 18.

The amplifier is connected to an analogue control circuit 20 for controlling a discrimination level for a known inverting discriminator 22. At an output of the discriminator there is provided an analogue signal having an average frequency <f>. The circuit 20 comprises an RC feedback loop 23 to control and correct for temperature and power supply variations which may influence the element 12 and amplifier 22.

The output of the discriminator is connected to flip-flop 24 to yield a digital bit stream 26 shown in figure 2(a). The RC time constant of the circuit 23 is slower than $1/f_0$. In this manner substantially equal numbers of bits of a first state and bits of a second state are generated at the output of the flip-flop. Hence, the bit stream 26 in figure 2(a) comprises a sequence of successive bits having one of the first state and the second state. The states may be a logic high or "1" and a logic low or "0". As illustrated in figure 9, it is known that these bits are typically not independent from one another nor unbiased. Therefore, the bit stream 26 is not truly random. As will hereinafter be described, a hardware digital corrector arrangement 28 comprising a first digital corrector circuit 30 (shown in figures 3, 7 and 9) and a serially connected second digital corrector circuit 32 (shown in figures 5,7 and 9), is used to remove correlation between bits and bias respectively, thereby to provide a truly random bit stream 34. The RNG further comprises a clock generator 36 for generating a clock signal 38 shown in figure 2(b). The RNG 10 further comprises a hardware Gaussian generator 40 which receives the truly random bit stream 34 as an input and generates words with standard Gaussian deviates, mean zero and variance 1, as will hereinafter be described. An output of the generator 40 is connectable to an input of any suitable application 42.

The frequency of the clock signal 38 is typically lower than the average or means frequency of the digital input stream (DIS) 26. In other embodiments a DIS 44 with a lower average frequency than the clock signal may be used. The DIS 26 or 44 is provided as one input to the first digital corrector circuit 30. The other input being the clock signal 38.

The first digital corrector circuit 30 is a circuit for removing inter bit dependence in the DIS 26 or 44 comprising successive bits $s_1,s_2,s_3,\ldots\ldots s_n$. It has been found that a circuit implementing a first scheme according to the following first truth table in respect of two immediately successive bits in the stream 26 or 44 (wherein the first bit is leading the second bit in time) and having one of a first state p and a second state q, would achieve this result:

1stbit, 2ndbit →p,p →c (third state)

1stbit, 2ndbit →p,q →c' (fourth state)

any other— no output

with c and c' being the inverse of the other, and c equal to p or q.

In this first scheme, successive pairs of immediately adjacent bits in stream 26 or 44 are used and no two bits are used more than once.

In one embodiment of the invention, the aforementioned first scheme is implemented by circuit 30 shown in figure 3 and comprising a first flip-flop 45, a second flip-flop 47 and a third flip-flop 49 as shown in figure 3. The circuit 30 has a first input 51, a second input 53, a first output 55 and a first clock output 57. Referring to figures 3 and 4(a) to 4(f), the bit stream 26 is applied to the first input 51 and the clock signal 38 to second input 53. The sequence 44 shown in figure 4(a) appears at the output of flip-flop 45 and at the output of flip-flop 47, one clock cycle later, as indicated at 50 in figure 4(c). A clock signal divided by two is provided at a Q-output of flip-flop 49 and is illustrated at 52 in figure 4(d). The comparison between signals 44 and 50 according to the first truth table is made on the rising edges of signal 52. An output bit stream 56 is generated by the first scheme and in accordance with the aforementioned first truth table at first output 55 of the circuit 30 and is designated 56 in figure 4(f). At the first clock output 57 of the circuit 30, an output clock signal 54 comprising intermittent and not periodic pulses 58.1 to 58.n is generated, successive pulses being associated with and synchronized with successive bits t_1, t_2, $t_3 \ldots t_m$ in the output bit stream 56 of the first corrector circuit. These output bits are substantially independent of one another in that correlation is removed by the first corrector.

In embodiments wherein the clock frequency is higher than the frequency of the DIS 44, the DIS 44 may be used to latch the clock signal. That is, the signals at inputs 51 and 53 are changed around.

The RNG 10 comprises a second digital corrector circuit 32 for correcting bias of aforementioned bits t_1, t_2, t_3....t_m. The second corrector circuit 32 is connected in series with the first corrector circuit 30 as shown in figure 7. The first output 55 of the first circuit 30 is connected to a first input 60 of the second corrector circuit and the first clock output 57 of the first corrector circuit is connected to a second input 62 of the second circuit 32. To illustrate the operation of the second circuit, the second circuit is shown on its own in figure 5.

It has been found that a circuit 32 implementing a second scheme according to a second truth table (shown herebelow) in respect of two immediately successive bits in a DIS 26 or 44 having one of a first state p and a second state q, would substantially remove bias of the bits in the DIS:

1stbit, 2ndbit → p,q → c (third state)

1stbit, 2ndbit → q,p → c' (fourth state)

any other— no output

with c and c' being the inverse of the other, and c equal to p or q.

In one embodiment of the invention, the aforementioned second scheme is implemented by circuit 32 shown in figure 5 and comprising a first flip-flop 64, a second flip-flop 66 and a third flip-flop 68. Apart from the aforementioned first and second inputs 60,62 the circuit 32 has a first output 70 (or a second for the arrangement 28) and a second clock output 72. Referring to figures 5 and 6(a) to 6(f), the bit stream 44 (which for this illustration is the same as that applied to the first corrector hereinbefore) is applied to the first input 60 and the clock signal 38 to second input 62. The sequence 44 appears at the output of flip-flop 66 one clock cycle later than at the Q-output of flip-flop 64 and is designated 74 in figure 6(c). A clock signal divided by two is provided at a Q-output of flip-flop 68 and is illustrated at 76 in figure 6(d). An output bit stream 80 is generated by the circuit 32 in accordance with the second scheme and the aforementioned second truth table at output 70 and is shown in figure 6(f). At the second clock output 72 an output clock signal 78 comprising intermittent and not periodic pulses 82.1 to 82.k is generated, successive pulses being associated with coincides with successive bits i_1, i_2, i_3...i_k in the output bit stream 80 of the second corrector circuit. Bias of bits in the output stream 80 is substantially removed.

As shown in figures 7 and 9 the first and second corrector circuits are connected in series, to remove both correlation between bits and bias respectively and thereby to yield a truly random bit stream 34 in which the bits are independent, unbiased and hence equi-probable or uniformly distributed.

Waveforms for this serial connection of circuits 30 and 32 are shown in figures 8(a) to 8(j). Since the input bit stream 44 is the same as that used in the description of figure 3, figures 8(a) to 8(f) correspond with the waveforms in figures 4(a) to 4(f). However, with the serial connection of circuits 30 and 32 the output signals 54 and 56 are connected to the inputs 60 and 62 respectively of the second corrector circuit 32. The resulting signals at the Q-outputs of the flip-flops 64 and 66 are shown at 56 and 84 in figures 8(f) and 8(g). The clock signal divided by two generated at the Q output of flip-flop 68 is shown at 86 in figure 8(h). The output bit stream of the corrector arrangement 28 is shown at 34 in figure 8(j). A clock signal to indicate successive bits in the output bit stream 34 is illustrated 88 in figure 8(i).

As shown in figure 9, a hardware Gaussian generator 40 is connected to either output 55 or output 70 of the digital corrector arrangement 28.

According to the central limit theorem in statistics, by summing by means of a suitable hardware adder or summing arrangement sufficient random words j results in a sum T which approaches that of a Gaussian deviate with known mean <T> and known standard deviation S_N. The generator 40 is configured to sum j words of i bits in the stream 34 each. The generator is further configured intermittently to compute sums T as aforesaid, to subtract by a suitable hardware arrangement an average <T> of the sums and to divide the difference by a standard deviation S_N utilizing a suitable hardware divider arrangement, to yield Gaussian deviates Z with a zero mean and unity standard deviation.

It is believed that the RGN 10 herein described may be implemented in the form of an integrated chip and housed in a plug-and–play device, comprising a USB memory stick. Once the device is activated, the generator 40 starts adding words until a default number or user specified number is reached, which would result in a Gaussian deviate as hereinbefore described and which would be available for input by the application 42.

CLAIMS

1. According to the invention there is provided a hardware random number generator (RNG), comprising:
    - a source of entropy for providing an input bit stream comprising successive bits of a first state and a second state;
    - a first digital corrector comprising a first input and a first output;
    - the corrector being configured to provide at the first output from two successive bits in the input bit stream an output bit of a first output bit stream according to a first scheme wherein a first bit of a first state and a second bit of the first state yield an output bit of a third state and wherein a first bit of the first state and a second bit of a second state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the output bit stream at the first output of the corrector are independent from one another.

2. An RNG as claimed in claim 1 wherein the first output of the first corrector is connected to a first input of a second digital corrector, the second digital corrector comprising a first output and being configured to provide at the first output from two successive bits at the first input

of the second corrector an output bit of a second output bit stream according to a second scheme wherein a first bit of a first state and a second bit of a second state yield an output bit of a third state and wherein a first bit of the second state and a second bit of the first state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the second output bit stream are both independent from one another and unbiased.

3. An RNG as claimed in claim 1 or claim 2 wherein the source of entropy comprises a white noise generating element having an output connected to an input of an analogue high pass amplifier having an output and a cut off frequency of f_0.

4. An RNG as claimed in claim 3 wherein the output of the high pass filter is connected to an input of a voltage level discriminator circuit having an output.

5. An RNG as claimed in claim 4 wherein the output of the level discriminator circuit is connected to a first input of a flip-flop, to generate the input bit stream at an output thereof.

6. An RNG as claimed in any one of claims 4 and 5 wherein an analogue voltage level compensation circuit is provided at the input of the voltage level discriminator circuit.

7. An RNG as claimed in claim 6 wherein the compensation circuit comprises an RC circuit having an RC time constant which is shorter than $1/f_0$.

8. An RNG as claimed in any one of claims 1 to 7 comprising a Gaussian generator comprising an input connected to the first output of the first digital corrector, the generator comprising an adder arrangement for generating a sum of j words of i sequential bits each received from the corrector arrangement; a subtractor arrangement for deriving a difference between the sum and a mean value of the sum; and a divider arrangement for dividing the difference by a standard deviation, thereby to generate at an output of the Gaussian generator a Gaussian deviate.

9. A method of generating a random bit stream comprising the steps of:

- utilizing a source of entropy for providing an input bit stream comprising successive bits of a first state and a second state;
- utilizing a first hardware digital corrector comprising a first input and a first output to provide at the first output from two

successive bits in the input bit stream an output bit of a first output bit stream according to a first scheme wherein a first bit of a first state and a second bit of the first state yield an output bit of a third state and wherein a first bit of the first state and a second bit of a second state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein the third state is equal to one of the first state and the second state, thereby to ensure that successive bits in the output bit stream at the output of the corrector are independent from one another.

10. A method as claimed in claim 9 wherein a second digital corrector comprising an input and an output is utilized in series with the first corrector, wherein the second corrector is used to provide at the first output from two successive bits at the first input of the second corrector an output bit of a second output bit stream according to a second scheme wherein a first bit of a first state and a second bit of a second state yield an output bit of a third state and wherein a first bit of the second state and a second bit of the first state yield an output bit of a fourth state, wherein the third and fourth states are inverse to one another and wherein he third state is equal to one of the first state and the second state, thereby to ensure that successive bits in

the second output bit stream are both independent from one another and unbiased.

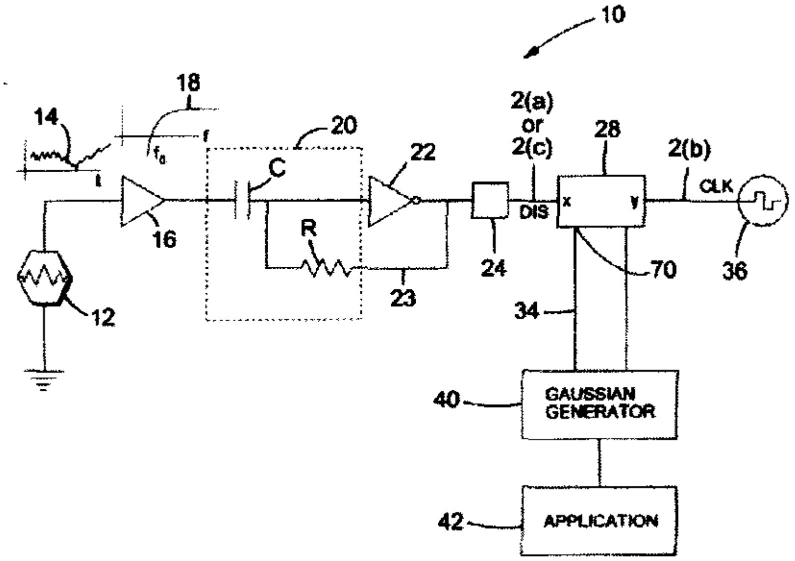

10
18
14
f
f_0
t
20
C
22
2(a)
or
2(c)
28
2(b)
CLK
DIS
x
y
R
16
24
70
36
23
12
34
40
GAUSSIAN
GENERATOR
42
APPLICATION

FIGURE 1

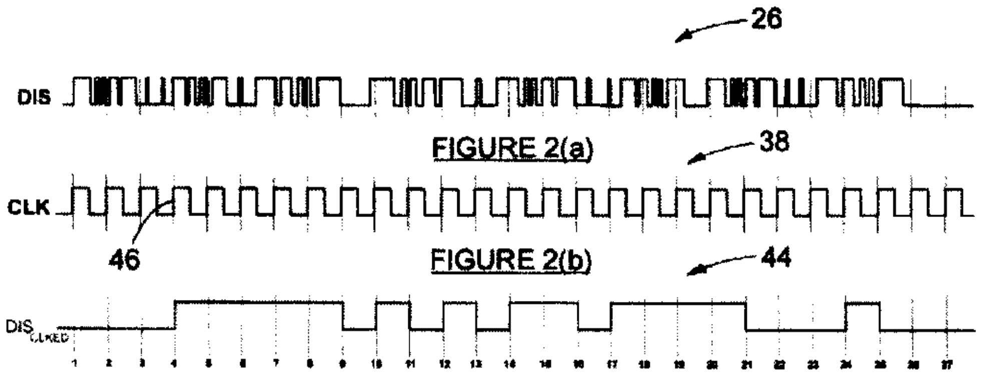


FIGURE 2(a)

FIGURE 2(b)

FIGURE 2(c)

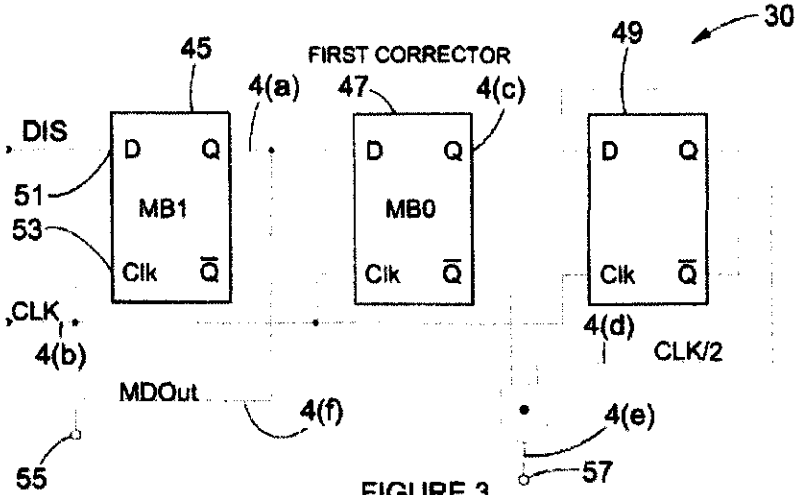

30
45
49
FIRST CORRECTOR
4(a)
47
4(c)
DIS
D
Q
MB1
Clk
Q̄
D
Q
MB0
Clk
Q̄
D
Q
Clk
Q̄
51
53
CLK
4(b)
4(d)
CLK/2
MDOut
4(f)
4(e)
55
57

FIGURE 3

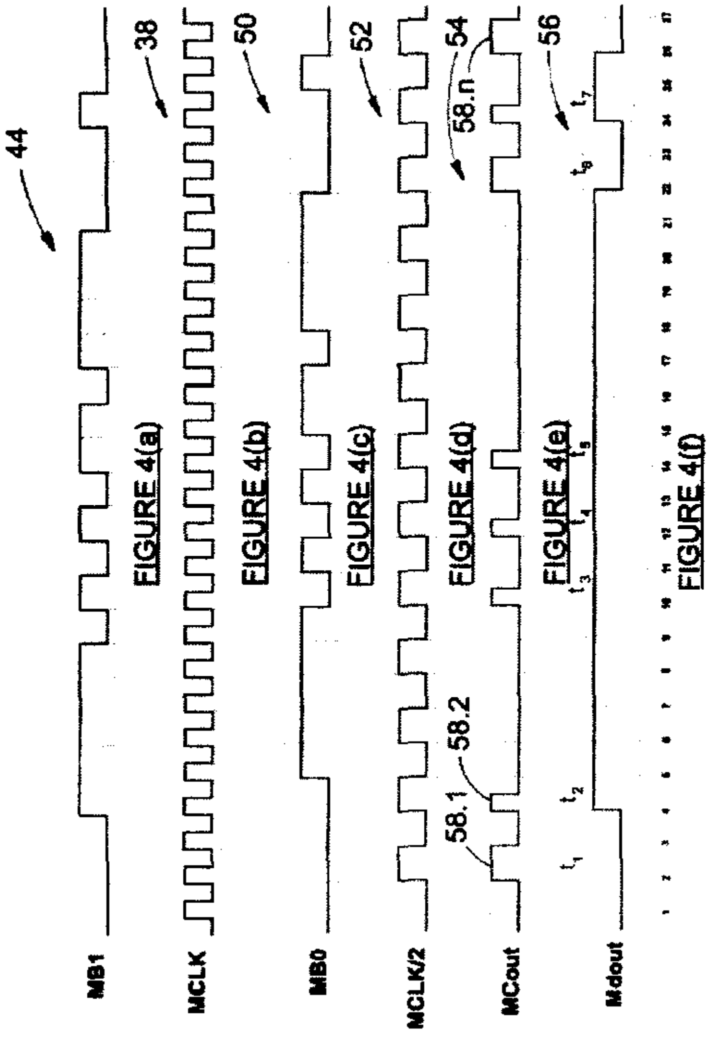


FIGURE 4(a)

FIGURE 4(b)

FIGURE 4(c)

FIGURE 4(d)

FIGURE 4(e)

FIGURE 4(f)

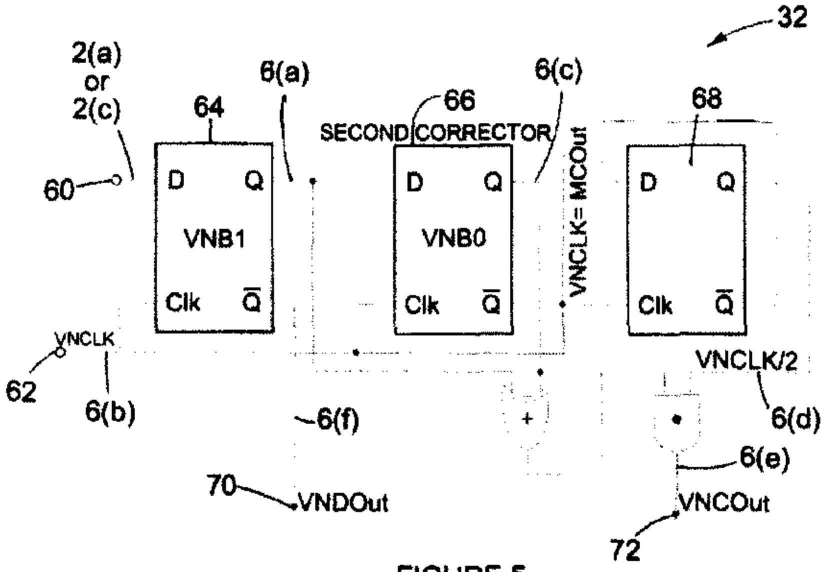

32
2(a)
or
2(c)
64
6(a)
66
6(c)
68
SECOND CORRECTOR
60
D
Q
VNB1
Clk
Q̄
D
Q
VNB0
Clk
Q̄
VNCLK= MCOut
D
Q
Clk
Q̄
VNCLK
62
6(b)
VNCLK/2
6(d)
6(f)
6(e)
70
VNDOut
VNCOut
72

FIGURE 5

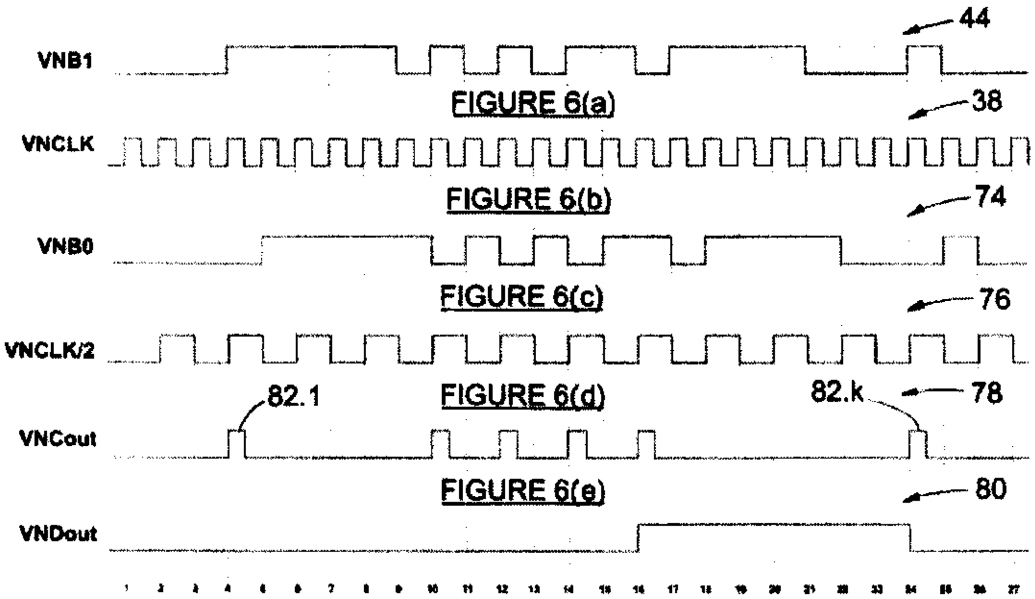


FIGURE 6(a)

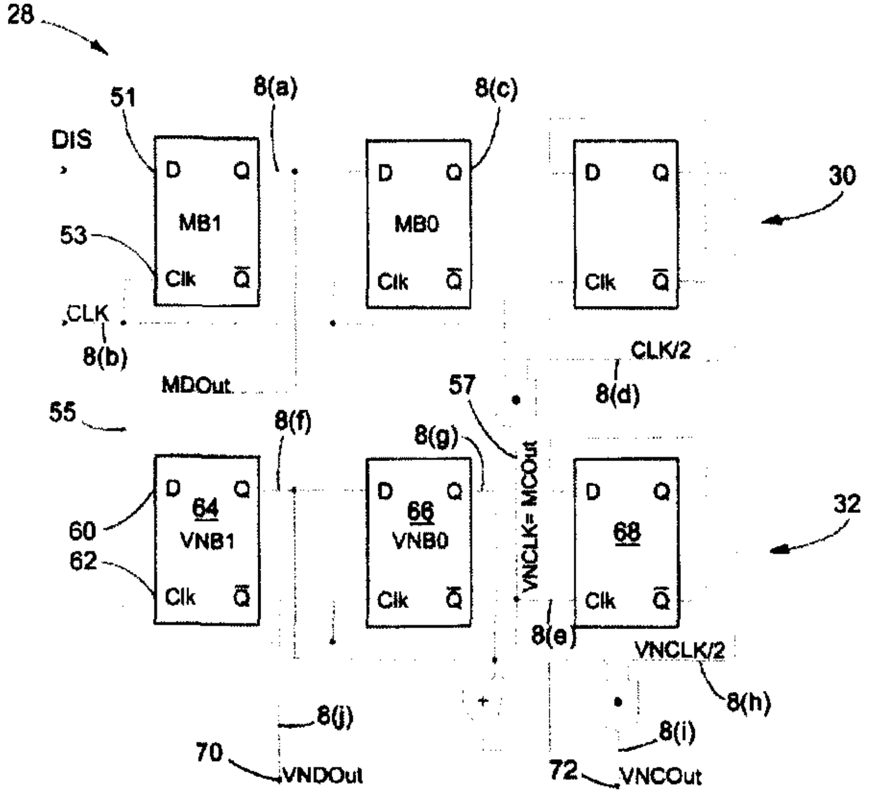
FIGURE 6(b)

FIGURE 6(c)

FIGURE 6(d)

FIGURE 6(e)

FIGURE 6(f)



28
51
8(a)
8(c)
DIS
D
Q
MB1
MB0
Clk
Q̄
30
53
CLK
8(b)
CLK/2
8(d)
MDOut
57
55
8(f)
8(g)
VNCLK= MCOut
64
VNB1
66
VNB0
68
32
60
62
8(e)
VNCLK/2
8(h)
8(i)
8(j)
70
VNDOut
72
VNCOut

FIGURE 7

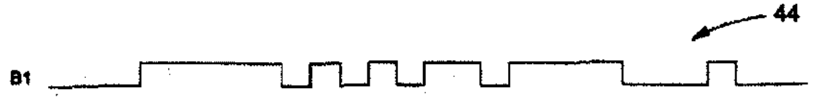


FIGURE 8(a)

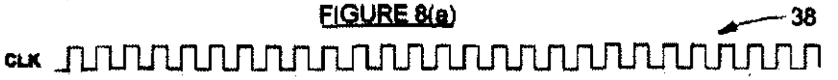


FIGURE 8(b)

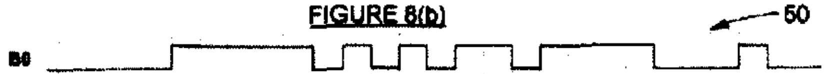


FIGURE 8(c)

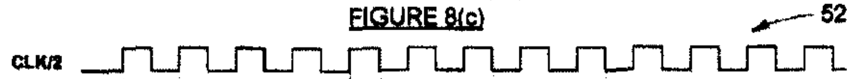


FIGURE 8(d)

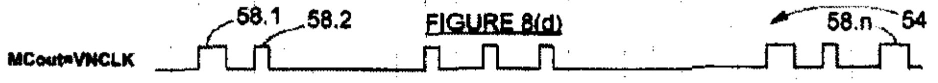


FIGURE 8(e)

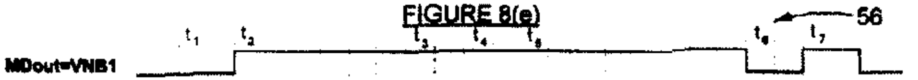


FIGURE 8(f)

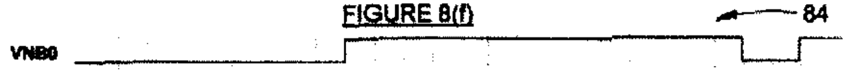


FIGURE 8(g)

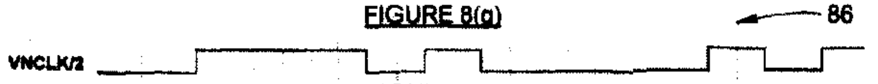


FIGURE 8(h)

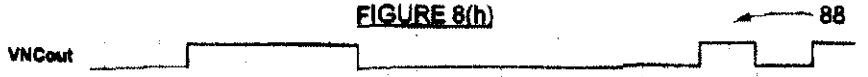


FIGURE 8(i)

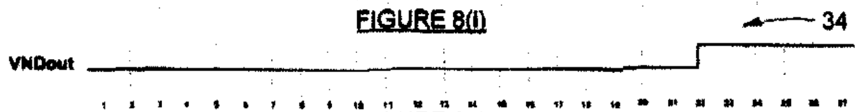


FIGURE 8(j)

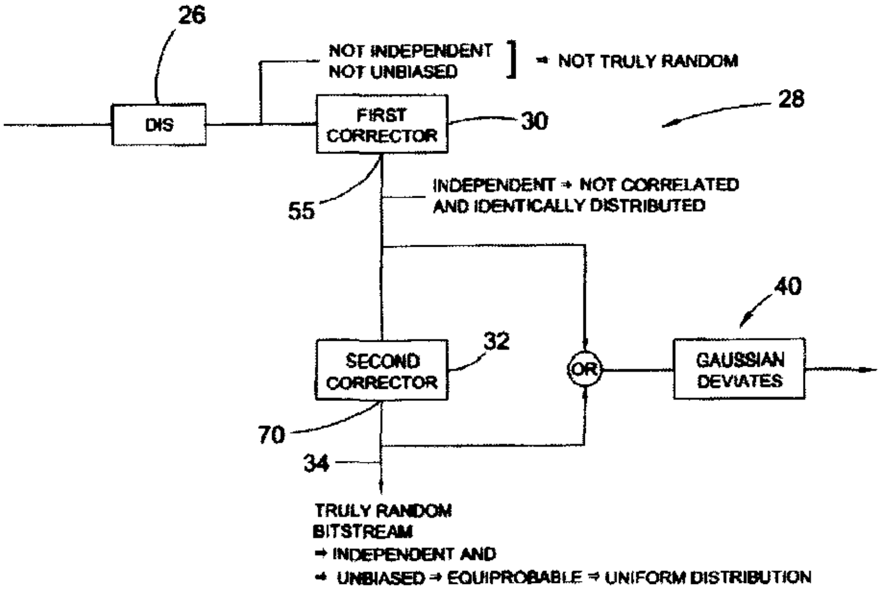

26
NOT INDEPENDENT
NOT UNBIASED
⇒ NOT TRULY RANDOM
28
DIS
FIRST CORRECTOR
30
55
INDEPENDENT ⇒ NOT CORRELATED AND IDENTICALLY DISTRIBUTED
40
SECOND CORRECTOR
32
OR
GAUSSIAN DEVIATES
70
34
TRULY RANDOM BITSTREAM
⇒ INDEPENDENT AND
⇒ UNBIASED ⇒ EQUIPROBABLE ⇒ UNIFORM DISTRIBUTION

FIGURE 9

Appendix III. Acronym and Abbreviation List

Acronym / Abbreviation	*Description*
ACT	Atmospheric Cerenkov Technique
ATNF	Australia Telescope National Facility
CANGAROO	Collaboration of Australia and Nippon for a Gamma Ray Observatory in the Outback
CGRO	Comptom Gamma-Ray Observatory
DIS	Digital Input Sequence
EAS	Extended Air Showers
EGRET	Energetic Gamma Ray Experiment Telescope
FIPS	Federal Information Processing Standard
FSE	Fourier Series Estimator
GLAST	Gamma Ray Large Area Space Telescope
GR	General Relativistic
H.E.S.S.	High Energy Stereoscopic System
IACT	Imaging Atmospheric Cerenkov Telescope
LIBOR	London Interbank Offered Rate
MAGIC	Major Atmospheric Gamma Imaging Cherenkov
MB	Markov Bit
MCOut	Markov Clock Out
MDOut	Markov Data Out
MISE	Mean Integrated Squared-Error
MJD	Modified Julian Date
NIST	National Institute of Standards and Technology
OG	Outer Gap

Acronym / Abbreviation	*Description*
PC	Polar Cap
QBE	Quantum Bit Extractor
SNR	Supernova Remnant
VERITAS	Very Energetic Radiation Imaging Telescope Array System
VHE	Very High Energy
VNB	Von Neumann Bit
VNCLK	Von Neumann Clock
VNCOut	Von Neumann Clock Out
VNDOut	Von Neumann Data Out

Bibliography

[1] Aharonian, F. H. et. al. 2004, Nature, 432, 75

[2] Arons, J. 1996, in TeV Gamma-Ray Astrophysics, ed. H. J. Völk & F. A. Aharonian (Dortrecht: Kluwer Academic Publishers), 235

[3] ATNF Pulsar Catalogue. 2004, www.atnf.csiro.au/research/pulsar/psrcat/

[4] Bendat, J. S. & Piersol, A. G. 1971, Random Data: Analysis and Measurement Procedures (3rd ed.; Wiley)

[5] Beran, R. J. 1969, Annals of Mathematical Statistics, 40, 1196

[6] Bernlohr, K. 2000, APh, 12, 255

[7] Bowers, R. L. & Deeming T. 1984, Astrophysics 1- Stars (Jones and Bartless Publishers International)

[8] Bradley, J. V. 1968, Distribution-Free Statistical Tests (Englewood Cliffs: Prentice Hall)

[9] Buccheri, R. et al. 1983, A&A, 128, 245

[10] Chang, H. K., & Ho, C. 1997, ApJ, 479, L125

[11] Cheng, K. S., Ho, C., & Ruderman, M. 1981, ApJ, 300, 522

[12] De Jager, O. C. 1987, The analysis and interpretation of VHE gamma-ray measurements (Potchefstroom University for CHE)

[13] De Jager, O. C. 2002, BASI 30, 85

[14] De Jager, O.C. et al. 1986, A&A, 170, 187

[15] De Jager, O.C., Raubenheimer, B. C., North, A. R., Nel, H. I., & van Urk, G. et al. 1988, ApJ 326, 831

[16] De Jager, O. C. et al. 1989, A&A, 221, 180

[17] Dieter, U., & Ahrens J. H. 1971, Numer. Math., 17, 101

[18] Epstein, R. I. 1973, ApJ, **183**, 593

[19] Fichtel, C. E. et al. 1983, 18th ICRC Proceedings (Bombay, Tata Institute of Fundamental research), 19

[20] Fitzpatrick R., & Mestel L. 1988, MNRAS, 232, 277

[21] Galbraith W., & Jelley J. V. 1953, Natur, 171, 349

[22] Gentle, J. E. 2003, Random Number Generation and Monte Carlo Methods, (2nd ed.; New York: Springer-Verlag)

[23] Ginzburg, V. L., & Syrovatskii, S. I. 1969, ARA&A, 7, 375

[24] Goldreich, P., & Julian, W. H. 1996, ApJ, 157, 869

[25] Gold, T. 1968, Nature, 218, 731

[26] Gold, T. 1969, Nature, 221, 25

[27] Grenier, I. A. 2001, in AIP Conference Proceedings 558, High energy gamma-ray astronomy, ed. F. A. Aharonian & H. J. Völk (Heidelberg: AIP), 613

[28] Griffiths, D. J. 2003. Introduction to Electrodynamics (3rd ed. Prentice Hall)

[29] Harding, A. K. 2001, in AIP Conference Proceedings 558, High energy gamma-ray astronomy, ed. F. A. Aharonian & H. J. Völk (Heidelberg: AIP), 115

[30] Harding, A. K., & De Jager, O. C. 1997, in Towards a major atmospheric Cerenkov detector-V, ed. O. C. de Jager (Potchefstroom: Wesprint), 64

[31] Harding, A. K., & Muslimov, A. G. 1997, International Conference on Neutron Stars and Pulsars, in Neutron Stars and Pulsars, ed. N. Shibazaki, N. Kawai, S. Shibata, & T. Kifune (Tokyo: Universal Academy Press), 311

[32] Hart, J. D. 1985, JSCS, 21, 95

[33] Hawthorne, W. M. 2002, Distribution of runs test, www.chantilley.com/html/paper5_distrunstest.htm

[34] Helfand, D. J., & Becker, R. H. 1984, Nature, 307, 215

[35] Hessels, J. W. T. et al. 2004, ApJ, 612, 389

[36] Hewish, A. et. al. 1968, Nature, 217, 709

[37] Hotbits. 2004, www.fourmilab.ch/hotbits/

[38] Hoyle, F., Narlikar, J. V., Wheeler, J. A. 1964, Natur, 203, 914

[39] Intel RNG Whitepaper. 1996, http://www.cryptography.com/resources/whitepapers/IntelRNG.pdf

[40] James, F. 1980, PPPh, 43, 73

[41] Kanbach, G. 1997, JENAM-97, 269

[42] Kaspi, V. M. et al. 2000, ApJ, 528, 445

[43] Kifune, T. et al. 1997, in AIP Conference Procedings 410, ed. C.S. Dermer, M. S. Strickman & J. D. Kurfess (Williamsburg: AIP), 1507

[44] Kirkpatrick, S., & Stoll, E. 1981, JCoPh, 40, 517

[45] Kniffen, D. 2002, APS, 2A3.001

[46] Konopelko, A. K. 1997, in Towards a major atmospheric Cerenkov detector- V, ed. O. C. de Jager (Potchefstroom: Wesprint), 208

[47] Konopelko, A. K. 2001, in AIP Conference Proceedings 558, High energy gamma-ray astronomy, ed. F. A. Aharonian & H. J. Völk (Heidelberg: AIP), 569

[48] Krawczynsky, H., & H.E.S.S. Collaboration. 1999, BAAS, 31, 727

[49] Krennrich, F. et al. 1997, in AIP Conference Procedings 410, ed. C.S. Dermer, M. S. Strickman & J. D. Kurfess (Williamsburg: AIP), 563

[50] Kubo, H. et al. 2004, NewAR, 48, 323

[51] Lorenz, E. 2004, NewAR, 48, 339

[52] Lyne, A.G., & Graham-Smith, F. 1990, Pulsar Astronomy, (1st ed.; Cambridge University Press)

[53] Mardia, K. V. 1972, Statistics of directional data, (New York: Academic Press Inc.)

[54] Marsaglia, G. 1968, PNAS, 61, 25

[55] Marsaglia's DIEHARD Test Suite, 1995, http://stat.fsu.edu/pub/diehard/

[56] Mestel, L. 1971, NPhS, 233, 149

[57] Mészáros, P. 1992, High-Energy Radiaion from Magnetized Neutron Stars, (Chigago: The University of Chigaco Press)

[58] Michel, F. C. 1997, International Conference on Neutron Stars and Pulsars, in Neutron Stars and Pulsars, ed. N. Shibazaki, N. Kawai, S. Shibata, & T. Kifune (Tokyo: Universal Academy Press), 263

[59] Mirzoyan, R. 1997, in Towards a major atmospheric Cerenkov detector- V, ed. O. C. de Jager (Potchefstroom: Wesprint), 298

[60] Nel, H. I., & De Jager, O. C. 1995, Ap&SS, 230, 299

[61] Nel, H. I., De Jager, O. C., Raubenheimer, B. C., Brink, C., Meintjies, P. J., & North, A. R. 1993, ApJ, 418, 836

[62] NIST, Cryptographic Standards and Evaluation Programs, 2004, http://csrc.nist.gov/cryptval/

[63] Pacini, F. 1967, Nature, 216, 567

[64] Peres, Y. 1992, Ann. Stat., 20, 590

[65] Petri, D., & The MAGIC Telescope Group. 1999, in Proceedings of BL Lac phenomenon Conference, 249

[66] Punch, M. 2002, on H.E.S.S. First Light Workchop CD

[67] Rice, J. A. 1995, Mathematical Statistics and Data Analysis, (Duxbury: International Thompson Publishing)

[68] Samuelson, P. A. 1968, Journal of the Americal Statistical Association, 63, 1526

[69] Shapiro, S. L., & Teukolsky, S. A. 1983, Black Holes, White Dwarfs and Neutron Stars, (John Wiley & Sons)

[70] Shaum & Spiegel, M. R. 1961, Shaum's Statistics (McGraw Hill Companies)

[71] The H.E.S.S. Project: An Array of Imaging Atmospheric Cerenkov Telescopes, 2004, Official Website, http://www.mpi-hd.mpg.de/hfm/HESS/HESS.html

[72] Thompson, D. J. 2000, in AIP Conference Proceedings 558, High energy gamma-ray astronomy, ed. F. A. Aharonian & H. J. Völk (Heidelberg: AIP), 103

[73] Tsuruta, S., & Cameron, A. G. W. 1966, Nature, 211, 356

[74] Venter, C. 2004, The effect of general relativistic frame dragging on millisecond pulsar visibility for the H.E.S.S. telescope (Potchefstroom University)

[75] VIA C3 Webpage. 2004, http://www.via.com.tw/en/products/processors/c3/

[76] Völk, H. J. 1997, in Towards a major atmospheric Cerenkov detector- V, ed. O. C. de Jager (Potchefstroom: Wesprint), 87

[77] Von Neumann, J. 1951, National Bureau of Standards Applied Mathematics Series, 12, 36

[78] Weekes, T. C. 1994, in TeV Gamma-Ray Astrophysics, ed. H. J. Völk & F. A. Aharonian (Dortrecht: Kluwer Academic Publishers), 1

[79] Weekes, T. C. et al. 1997, Proceedings of the 25th ICRC, ed. M. S. Potgieter, B. C. Raubenheimer, & D. J. van der Walt, (Transvaal, South Africa: Potchefstroom University), 173

[80] Wheeler, J. A. 1966, ARA&A, 4, 393

[81] Woltjer, L. 1964, ApJ, 140, 1309

[82] Wood, K., Michelson, P., & The GLAST Collaboration. 1995, BAAS, 27, 1387