THE UNIVERSITY OF NOTTINGHAM

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

MEng Industrially Orientated Project

Moment generating function based optical receiver modelling

AUTHOR	Kevin Neale
SUPERVISOR	Dr Andy Phillips
MODERATOR	Dr John Walker
DATE	May 2008

ABSTRACT

This project focused primary on the mathematical modelling of the noise experienced by an optical receiver. The two sources of noise considered in this project were thermal and shot noise and their contributions were modelled via various methods including Gaussian approximations and Chernoff bounds. The optical receivers considered were pin and avalanche photodiode detectors in conjunction with optical amplifiers. The advantages and disadvantages of both devices were evaluated with the primary focus of the project being on the accuracy of the various methods used for the evaluation. Extension work was also carried out to incorporate the effects of ISI into the project along with brief consideration of typically encountered front end circuitry. A brief consideration on how the above methods may also apply to Quantum Cryptographic key distribution was also considered and has been included in the appendix.

TABLE OF CONTENTS

1. INTRODUCTION

In a typical communication system data is represented by a signal that can take the discrete values of either a one or zero. In an optical system this digital data can be represented by sending pulses of two different intensities one indicating a one bit and the other a zero bit. A common scheme of data modulation is to actually not transmit any power at all for a zero bit and some intensity for a one bit. This is known as an On off Key (OOK) modulation scheme and initially seems like the most sensible scheme to use. An inherent problem with OOK is that it is not spectrally efficient as the fast change in amplitude of the signal to and from zero power requires excessive bandwidth [1]. With either of the schemes a threshold limit needs to be determined that we distinguish between a one and zero bit.

The receiver can be seen as ideal photon counter [2] which is used to detect the presence of light within a bit interval. If no light is detected then it is assumes that a 0 bit was transmitted and if some light is detected then a 1 bit was transmitted. This is known as a 'direct detection' and will not be error free as the photons that make up the received light follow a Poisson arrival process. This leads into the possibility of a rate of erroneous detection of bits or a Bit Error Rate and this limits the overall performance of a communication link. Design specifications for optical links normally require a certain level of performance e.g. a certain BER rate. The desired BER is normally between 10^{-9} and 10^{-12} which means respectively for every 10^{9} or 10^{12} bits received, one is detected incorrectly.

In the case of an ideal receiver, the stream of incident photons can be considered arriving with a power 'P' and at a mean rate of 'P/hf_c', where 'f_c' is the frequency of the photons and 'h' is planks constant (6.63×10^{-34}). When a 0 bit is transmitted then the receiver will not make any errors. However when a 1 bit is transmitted, the photons may not be detected during the bit interval. Therefore by inclusion of the bit period 'T_b', the probability that 'n' photons are detected during a bit interval can be expressed as:

$$\exp\left(\frac{-PT_b}{hf_c}\right)\frac{\left(\frac{PT_b}{hf_c}\right)^n}{n!}$$

The probability of not receiving any photons is e^{-PT_b/hf_c} and then assuming the transmission of a 0 or a 1 bit is equally likely, the BER rate becomes:

$$BER = \frac{1}{2}\exp\left(-\frac{PT_b}{hf_c}\right)$$

A final step can then be taken by letting $M = PT_b/hf_c$ where 'M' is equal to the mean number of photons received during a 1 bit's transmission. Now the expression for the BER becomes:

$$BER = \frac{1}{2}\exp(-M) \qquad (1)$$

Solving in for a BER of 10^{-12}, results in M = 27. Therefore 27 photons on average would be needed for a one bit. This provides a way of determine the required sensitivity of the receiver. This is the error rate of the receiver and is known as the Quantum limit. This is only the case for a simple ideal receiver, as sources of noise also have an effect on the BER.

1.1. Receiver Stages

In an optical link or network a photodiode is used to detect the received power and the photocurrent generated within the device represents the transmitted data. A major problem is that the signal is subjected to a variety of impairments during its transmission. Figure 1.1 shows the various stages of a receiver that may introduce additional noise. At stage A, the signal may have undergone attenuation as it has been transmitted over a length of fibre. At stage B an optical amplifier maybe used and can introduce ASE noise which will be discussed further on. At stage C physical noise effects such as thermal and shot noise can lead to further distortion of the signal.

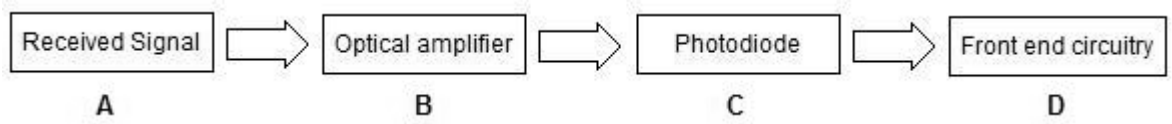

Figure 1.1 - Stages of a optical receiver channel

At stage D the receiver circuitry and possible post receiver amplification can introduce additional noise components. Figure 1.2 illustrate a theoretical signal that might be produced in an optical device with D indicating the mean current that would typically be set as a threshold for determine the received bit. I_0 and I_1 represent the average current that might be received for a one and zero bit and a degree of variance from the mean may be commonly experienced. This variance is due to the addition of the noise and if large enough can lead to the incorrect decoding of received bits.

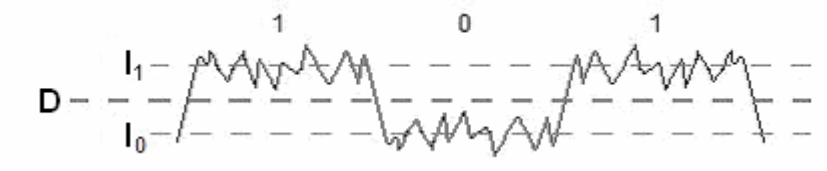

Figure 1.2 - Received photocurrent

The project focused on the current mathematical methods used to evaluate systems and how they might compare to the real BERs. In terms of the receivers, standalone PIN and APD photodiodes were initially considered.

1.2. Sources of noise

The two sources of noise that were considered for this project were thermal (Johnson) and shot noise. The thermal noise generated within a PIN diode is a result of the intrinsically generated electron hole pairs. These thermally generated EHPs increase proportionately with temperature and are the major contributor to the noise within a PIN photodiode. These thermally generated EHPs will be present regardless of whether or not a voltage is present.

For example the electrons present in the resistor have a random thermal energy ‘kT’ with ‘k’ is Boltzmann’s constant (1.38×10^{-23} J/K) and ‘T’ is the temperature (K). Therefore there will be a randomly fluctuating voltage (or current) that can be measured across the circuit containing it. The power associated with this voltage is proportionate to the bandwidth of system [1]. The instantaneous value of the thermal noise follows a Gaussian distribution and it is this property that can be utilised to model the noise.

Typically the thermal noise value is expressed in relation to its power spectral density i.e. the units of A^2 / Hz, thus indicating the current's variance per hertz of bandwidth. In this project this value is referred to as the thermal noise value 'Sn'. The variance due to the thermal noise can be expressed as follows:

$$\sigma_{therm}^2 = SnB_e \quad \text{where } B_e = 0.75 * Bit\_rate \tag{1}$$

'B_e' stands for the electrical bandwidth of the receiver and its value is chosen in accordance with the bit rate of the signal. The value for 'B_e' is chosen to achieve the best noise performance but to still allow the largest bandwidth possible. The value can range from the full bit rate to half of the bit rate, but within the project the value of ¾ was used.

Shot noise produces variations in the average current that flows in a device. This is because the arrival of the charge carriers that make up the current, follow a Poisson arrival process. The variance due to the shot noise can also be expressed in a similar form:

$$\sigma_{shot}^2 = 2qIB_e \tag{2}$$

Unlike thermal noise, the shot noise is dependent on the current within the device and is considerably important when the current in question is very small e.g. the photocurrent generated within a photodiode. Another type of noise known as flicker noise or 1/f noise is a noise contribution that can be found at low frequencies [3]. This noise is due to carrier trapping and detrapping effects and was not considered for this project.

1.3. Optical detection

In order for optical detection to occur, a time varying optical signal needs to be converted to an electronic signal. Figure 1.3 show a pn junction illuminated by a light source in a reverse bias configuration. Optical absorption occurs when a photon with energy greater or equal to the band gap of the material provides sufficient energy to create an electron-hole pair. This generated electron-hole pair then diffuses to the junction and is swept across to the depletion region within the device resulting in an movement of charge per second e.g. Cs^{-1} or a photocurrent. Ideally the speed of the detector and ultimately the bandwidth is determined by how fast the generated EHPs are swept across by the junction voltage. Therefore ideally most EHPs should be generated within the depletion region to eliminate the time need for the diffusion process. The width of 'W' is major design consideration and will determine the characteristics of the device. If the region is large, then most of the EHPs will be created within the depletion region producing a greater sensitivity. If the region is too large, then it takes longer for the EHPs to be swept across causing a reduction in bandwidth [3].

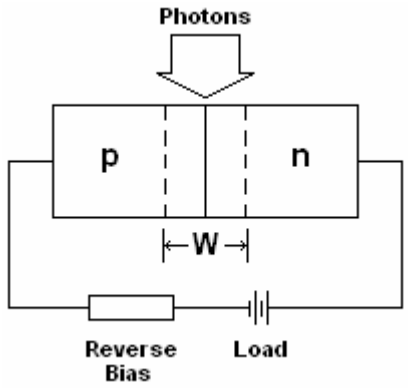

Figure 1.3 – PN junction diode

A PIN device allows control of the depletion region by the introduction of an intrinsic region but as a result means that the device is highly affected by thermal noise. APD devices have an advantage in that they can compensate for the thermal noise by providing gain for the current. This is achieved by the avalanche process and will be discussed further on. An inherent problem with the gain process is that it causes an increase in the contribution from shot noise.

Another important characteristic of an optical receiver is the signal to noise ratio that it can offer (see figure 1.4). This will obviously be different for both types of device and will vary according to the incident power and for the amount of gain that the device can provide.

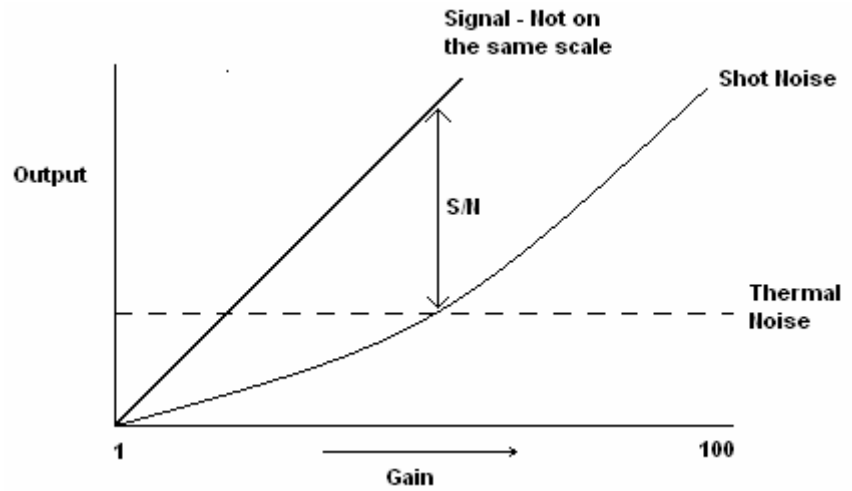

Figure 1.4 – Typical APD Signal to noise characteristics

2. GAUSSIAN APPROXIMATION

The Gaussian approximation involves considering the generated photocurrent within the receiver as a mean current with some variance about that mean. For example a zero bit would produce a mean current relating to the signal with a variance that is caused by the noise. The currents for both a zero and one bit can be modelled as Gaussian probability distributions that centre on the mean value of the current.

Figure 2.1 illustrates this concept and also shows a value indicated by 'Decision threshold or D' which is used as decision threshold to distinguish between a one and zero. It can be clearly seen that the noise variance could cause the threshold to be triggered and result in an incorrectly decoded bit. Therefore the BER of a system is related to the area associated with the overlapping of these probability density functions [2].

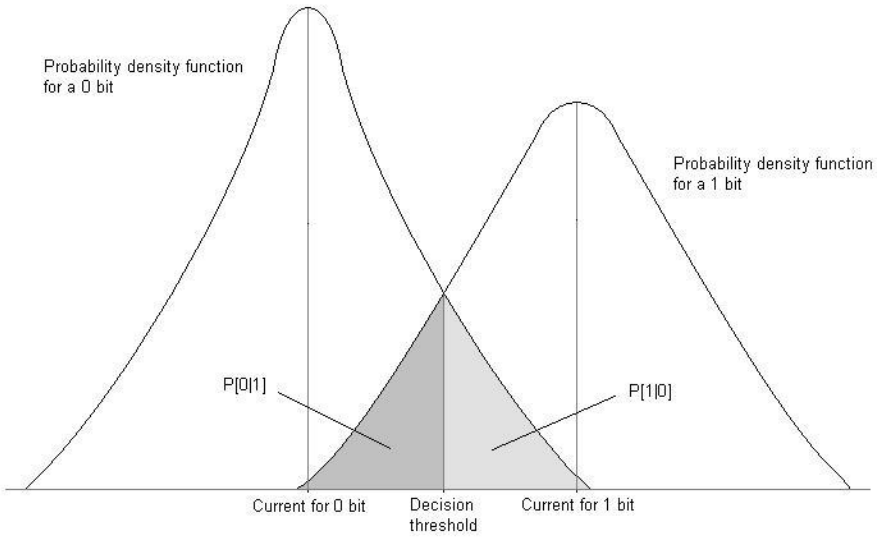

Figure 2.1 – Photocurrent PDF for a one and zero bit

Equation 3 shows an expression for calculating a BER value as the sum of the probability of incorrectly detecting a one bit when a zero bit was transmitted e.g. P(1|0) and the detecting a zero bit when a one bit was transmitted e.g. P(0|1). This also assumes that the transmission of either of the bits is equally probably.

$$\text{BER} = P(1\,|\,0)P(0) + P(0\,|\,1)P(1)$$

$$\text{Simplifies to} \quad BER = \frac{1}{2}\left[P(1\,|\,0) + P(0\,|\,1)\right] \tag{3}$$

Using the basic BER formula the evaluation of these overlapping sections $P(0\,|\,1)$ and $P(1\,|\,0)$ can be accomplished with using the error function. The complementary error function or erfc simply allows the evaluation of the tail end of a Gaussian probability distribution. Therefore with use of the complementary error function a new algebraic expression can be derived to evaluate the areas:

$$\text{BER} = \frac{1}{2}\left[\frac{1}{2}\text{erfc}\left(\frac{\text{I}_1 - \text{D}}{\sqrt{2}\sigma_1}\right) + \frac{1}{2}\text{erfc}\left(\frac{\text{D} - \text{I}_0}{\sqrt{2}\sigma_0}\right)\right] \tag{4}$$

It is this result that forms the basis of the Gaussian approximation and the next necessary steps were to find expression for the mean currents, the noise variance and the decision threshold. The Gaussian probability density function can be calculated via equation 5 [4] and a distribution with a mean of zero and a standard deviation of 1 is shown in figure 2.2. The distribution is centred on its mean and the width of the curve is determined by the standard deviation from the mean value.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right) \tag{5}$$

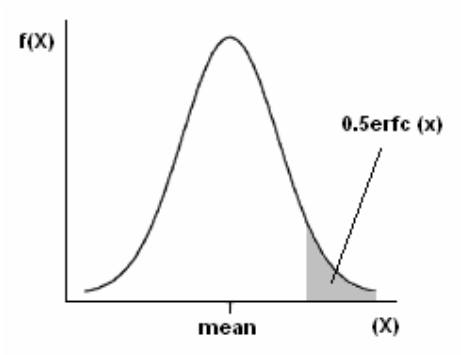

Figure 2.2 – Gaussian Curve and complementary error function [ref]

The thermal noise can be modelled very closely as a Gaussian process. The shot noise initially can be ignored in the case of the pin diode as its contribution is very small in comparison to the thermal noise. A Gaussian approximation can also be taken of the shot noise allowing a crude evaluation of its contribution to the total noise variance. Like any other normalised probability density function the area under the normalised Gaussian PDF or curve is equal to one. The error function (erf) and the complementary error function (erfc) can be used to calculate the area under these curves. As mentioned previously the erfc can be used to evaluate the tail end of the Gaussian curve and the erf can be used for the rest e.g.

$$\mathrm{erf}(x) + \mathrm{erfc}(x) = 1 \tag{6}$$

The erfc was therefore useful in this case as it was the area of the tail ends of the currents PDFs that were of interest. An expression for the erfc is shown in equation 7.

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_{z}^{\infty} e^{-t^2}\, dt \tag{7}$$

Equation 8 now shows that the tail end of a Gaussian distribution can be evaluated which is exactly what was required.

$$\frac{1}{2}\mathrm{erfc}\frac{I_D - I_o}{\sigma_0\sqrt{2}} = \frac{1}{\sqrt{\pi}} \int_{\left(\frac{I_D - I_0}{\sigma_0\sqrt{2}}\right)}^{\infty} e^{-t^2}\, dt = \frac{1}{\sqrt{2\pi\sigma_0}} \int_{I_D}^{\infty} e^{-\left(\frac{x - I_0}{\sigma_0\sqrt{2}}\right)^2} dx \tag{8}$$

The same procedure can be repeated for the other current's PDF.

2.1. Pin Diode

The first attempt at modelling an optical receiver was focused on a pin photodiode subjected to an amount of incident optical power. A pin photodiode was chosen at first as the additional gain factor associated with an APD is at unity in this case e.g. M = 1, making it slightly simpler.

The PIN photodiode's structure allows a photocurrent to flow by the generation of electron-hole pairs, within the intrinsic region of the device. The electron hole pair is generated via the incident optical radiation, and the rate at which these pairs are produced depends on the quantum efficiency of the device. The relationship that governs the electron-hole pair generation is:

$$\mathrm{EHP}_{\mathrm{rate}} = \frac{\eta}{\mathrm{hf}} \mathrm{P}_{\mathrm{opt}}(\mathrm{t}) \quad \textbf{(9)}$$

With P_{opt} being the incident optical power, h = planks constant, f = frequency of the incident light and η being the quantum efficiency of the device. The quantum efficiency of the device detects how effectively the device can convert the incident photons into electron-hole pairs and is usually quoted as a percentage.

Typically the incident power is specified in dBm e.g. -18dBm which is the decibel expression of power in milliwatts. E.g. -18 dBm ≡ 0.0158489 milliwatts or 0.000015848 watts.

The average received incident light P_{opt} typically takes one of two distinct intensities representing either a one or a zero bit. As mentioned earlier the data modulation will follow a particular scheme and the simple NR scheme was used for the initially basic Gaussian approximation. A ratio of the two intensities is normally specified e.g. in decibels to indicate how the two powers compare to each other. This ratio (r) is know as the extinction ratio and is governed via equations 10 and 11.

$$\frac{2r}{r+1} \times P_{opt} = P_1 = \text{the power in a one bit} \quad \textbf{(10)}$$

$$\frac{2}{r+1} \times P_{opt} = P_0 = \text{the power in a zero bit} \quad \textbf{(11)}$$

Figure 2.3 features a scheme with an extinction ratio of 10. The values for P_0 and P_1 are expressions that can be used in the Gaussian approximation [2].

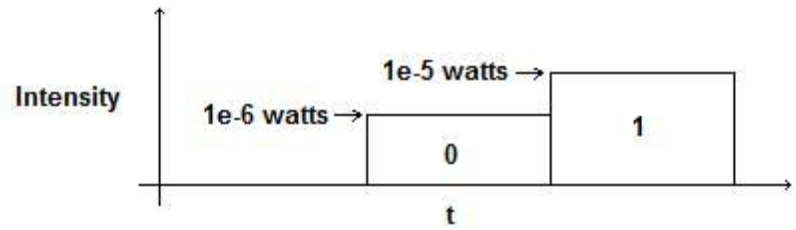

Figure 2.3 – Power present in a zero and one pulse following with an extinction ratio of 10

The responsively of the device dictates how well the photodiode can convert received optical power to a photo current i.e. it's units are Amps/Watts. The equation for responsivity is obviously related to mean creation rate of EHPs multiplied by the value for a single charge carrier e.g. an electron ($q = 1.6 \times 10^{-19}$ C).

$$R = \frac{\eta q}{hf} \quad \textbf{(12)}$$

The current that is produced in this device is related by equation 12 and represents the mean current.

$$I = RP_{opt} \quad \textbf{(13)}$$

In the project the mean currents for zero and one bit were denoted by I_0 and I_1.

Equations were then coded for the thermal and shot noise and for the initial Gaussian approximation were denoted separately as σ^2_{therm} and σ^2_{shot}.

The value for the thermal noise variance σ^2_{therm} is governed by equation 1 as indicated in section 1.1 and the shot noise variance can also be modelled via equation 2. 'B_{rate}' refers to the bit rate of the system and initially was set at 2.5 Gb/s.

The variances that the two noises produce are additive and the total noise contribution can be expressed as: Total noise variance = $SnB_e + 2qIB_e$ e.g.

$$\sigma^2_{total} = \sigma^2 therm + \sigma^2 shot \quad \textbf{(14)}$$

The total variance applied to each current with σ_0 being the total noise for a zero bit and σ_1 being the total noise for a one bit

The equations necessary to take a Gaussian approximation for a PIN diode had been determined and could now be substituted into equation 4. The value for D (the threshold) needed to be optimised for the receiver and ideally will be positioned halfway between the value for a one and zero bit.

An expression can be derived for D for a simple evaluation of the BER and assumes the decision threshold should be exactly halfway between I_0 and I_1 e.g.

$$I_D = \frac{I_1 + I_0}{2} \quad \textbf{(15)}$$

By substitution of equation 15 into equation 4 the expression shown in equation 16 can be derived. This expression can be used to determine the noise variance σ for a PIN diode. This is because in the case of a PIN, the thermal noise dominates and means that the shot noise can initially be ignored in a basic evaluation e.g. $\sigma_0 = \sigma_1 = \sigma_{therm}$.

$$BER = \frac{1}{2} erfc\left(\frac{Q}{\sqrt{2}}\right) \quad \textbf{(16)}$$

$$Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0} \tag{17}$$

The value Q can be determined for a know BER e.g. Q = 7.034 for a BER of 10^{-12} and it can be used to determine the thermal noise value can then be used for further BER calculations. The thermal noise value needed to achieve a BER of 10^{-12} for -18 dBm of received optical power was evaluated:

Converting P_{opt} to watts = -18 dBm = 0.000015848 watts

Using Equation and P_0 = 2.8816e-06 and P_1 = 2.8816e-05

For simplicity initially $R = 1$ A/W although this can actually be greater than 1

$I_0 = RP_0$ = 2.8816e-06 and $I_1 = RP_1$ = 2.8816e-05

Using $Q = \frac{I_1 - I_0}{\sigma_1 + \sigma_0}$ with $\sigma_0 = \sigma_1 = \sigma_{therm} = \frac{I_1 - I_0}{2Q}$

Finally equation gives Sn = $\frac{\sigma^2}{Be}$ *giving the result Sn = 1.812565e-21 A^2/Hz to 6 d.p.*

An alternative algebraic expression for D was derived by assuming that it is a Binary Symmetric Channel and that there will be an equal chance for a one being detected when a zero was transmitted and a zero being detected when a one was transmitted

$$P(1 \mid 0) = P(0 \mid 1)$$

With the use of equation 4 this is equal to:

$$\frac{I_D - I_0}{\sigma_0\sqrt{2}} = \frac{I_1 - I_D}{\sigma_1\sqrt{2}}$$

$$\sigma_1 I_D + \sigma_0 I_D = \sigma_0 I_1 + \sigma_1 I_0 \qquad \Rightarrow \qquad I_D = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_1 + \sigma_0} \tag{18}$$

This expression can then be used in conjunction with the noise contributions for a one and zero bit and substituted back into equation 4. The result value will then produce a BER value for a particular power e.g. -18dBm. Thus if the sensitivity is know the thermal noise can be deduced.

In order to produce a BER curve, the BER values need to be calculated for a range of powers e.g. -30 to -16 dBm. This was achieved in Matlab by simply specifying the variable 'Pav' as a range instead of one particular value i.e. -30:1:-16 will increment the value in steps of 1 from -30 to -16. Initially the presented problems as the division and multiplication values carried out in the BER calculations normally work with one specific value and generated incorrect results at the plotting stage. This was rectified by using the bitwise operator in order to allow single values to be used

within each calculation within the entire range of powers. A BER was then able to be plotted producing the following results.

In the particular case of the PIN photodiode the thermally generated noise will be the major contributor and will override the impact of the shot noise.

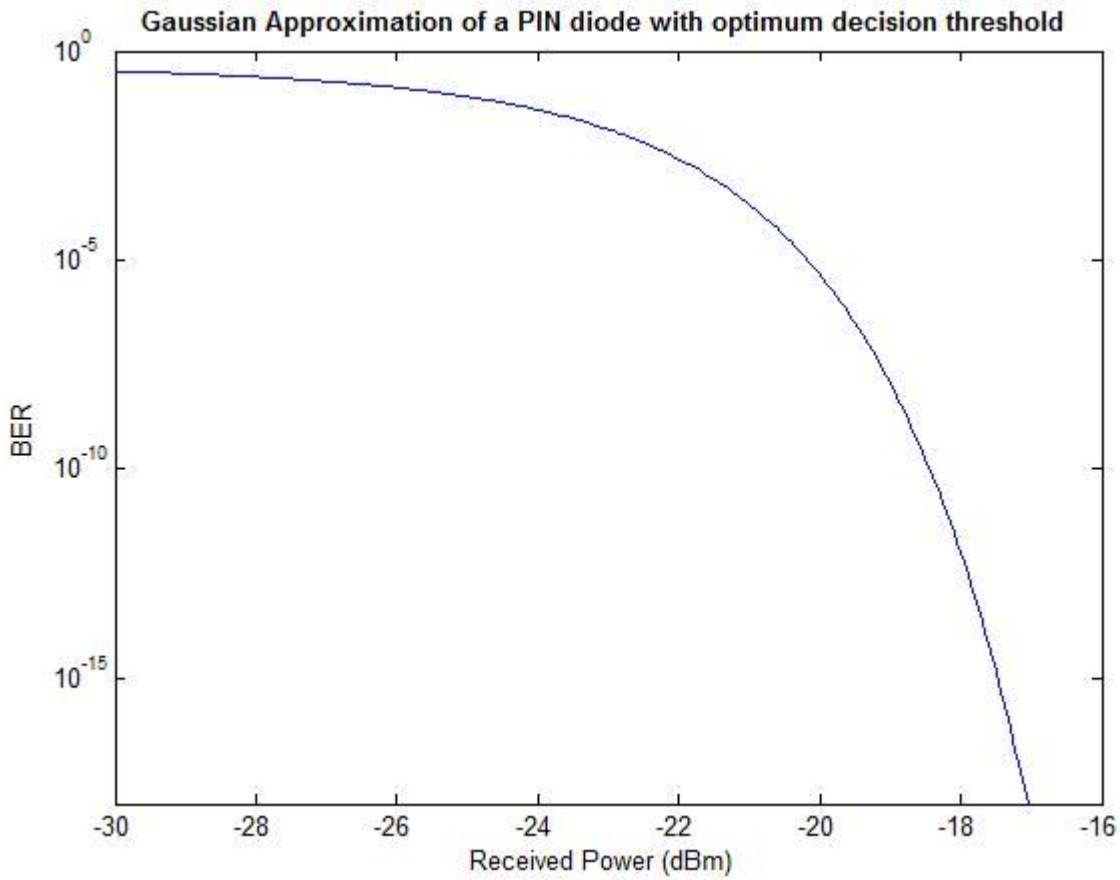

Figure 2.4 – Guassian approximation of a PIN diode

It can be seen from the BER curve that the typical BER which is expected for the system simulated would have a BER of 10^{-12} when the incident power is of -18dBm. The model therefore produced an accurate model of the PIN photodiode.

An important point to note is that the optimum decision threshold shown in equation 18 is based on the assumption that it was a binary symmetrical channel [11]. The actual statistics of the situation may imply that the noise variances vary in magnitude with respect to each other e.g. $\sigma_1 > \sigma_0$. Therefore the derived decision threshold in this case is not necessary the optimum. An improved method can be used to obtain the optimum decision threshold by minimising the BER e.g.

$$\frac{\partial BER}{\partial i_D} = 0 \text{ and then solving for } i_D \text{ e.g.}$$

$$i_D = \frac{1}{\sigma_0^2 - \sigma_1^2}\left(i_1\sigma_0^2 - i_0\sigma_1^2 - \sigma_0\sigma_1\sqrt{(i_0 - i_1)^2 - 2(\sigma_0^2 - \sigma_1^2)\ln\left(\frac{\sigma_1}{\sigma_0}\right)}\right) \quad \textbf{(19)}$$

When $\sigma_1 \approx \sigma_0$, the previous optimal threshold is obtained. However when, the variances differ in comparison to each other then the optimum threshold moves away from the solution for the BSC. The difference in magnitude between the variances may be due to signal dependent noise such as that from the excess noise caused by avalanche gain in an APD which will be discussed later.

2.2. Avalanche Photodiode Detector

The second device that was considered was an Avalanche Photodiode Detector. The APD exploits the 'avalanche' effect in order to produce gain for a generated signal. The 'avalanche effect' occurs when the device is subjected to a high voltage. When an electron-hole pair is generated via the incident light, the generated electrons are accelerated via the large potential difference. This results in an increased velocity of the electrons which can strike or knock out other electrons to take part in the process. These newly freed electrons in turn knock out more free electrons, and these in turn produce more. This is formally known as impact ionisation. This process allows the amplification of a relatively small photocurrent generated by the incident light to a larger more noticeable level.

The gain factor of an APD is denoted as M (also stated as E{g}) and this is the mean value of excess number of electrons that are generated as a result of the avalanche process. The mean multiplication factor produces 'M' times as much current than it would for a typical diode and this was modelled via the following set of equations for each bit.

$$I_0 = MRP_0 \quad \text{and} \quad I_1 = MRP_1 \tag{20 \& 21}$$

The thermal noise makes exactly the same contribution as it did for the pin diode's Gaussian approximation. Therefore the thermal noise is signal independent and was modelled with the same equations as for pin.

The shot noise in the case of the APD is governed by equation 22 and its dependency on the signal current can clearly be seen. The excess noise factor M is associated with the avalanche multiplication process (Equation 23).

$$\sigma_{shot}^2 = 2qRM^2F(M)P_{opt}Be \tag{22}$$

$$\text{Where} \quad F(m) = kM + \left(2 - \frac{1}{M}\right)(1-k) \quad k = \text{rate of ionisation} \tag{23}$$

The Gaussian approximation was obtained by replacing the pin diode's noise equations with that of the newly derived ones shown above. Resultant plots of the Gaussian approximation are shown in figure 2.5. This again was obtained via the use of the optimum threshold. It also shows how an increase in the ionization rate affects the approximate BER. The impact ionization that takes place is subjected to some variation and therefore introduces excess noise. This variation is less significant when the ionisation is due to only one type of charge carrier as more fluctuations occur when both types of carriers are participating e.g. electrons and holes [3]. The rate 'k' is equal to zero for an ideal device and in practically devices can be made that are very close to this e.g. a

Silicon diode. With a Silicon APD the ionisation due to electrons is greater and therefore can produce more gain with a smaller amount of noise. Silicon APDs are rarely used in practical links as Silicon is transparent at the required wavelengths (e.g. 1.33 ~ 1.55 μm). More commonly InGaAs devices are used but the ionisation rate within these compounds for both types of charge carrier are fairly similar. Thus additional noise is encountered and 'k' is closer to 0.7.

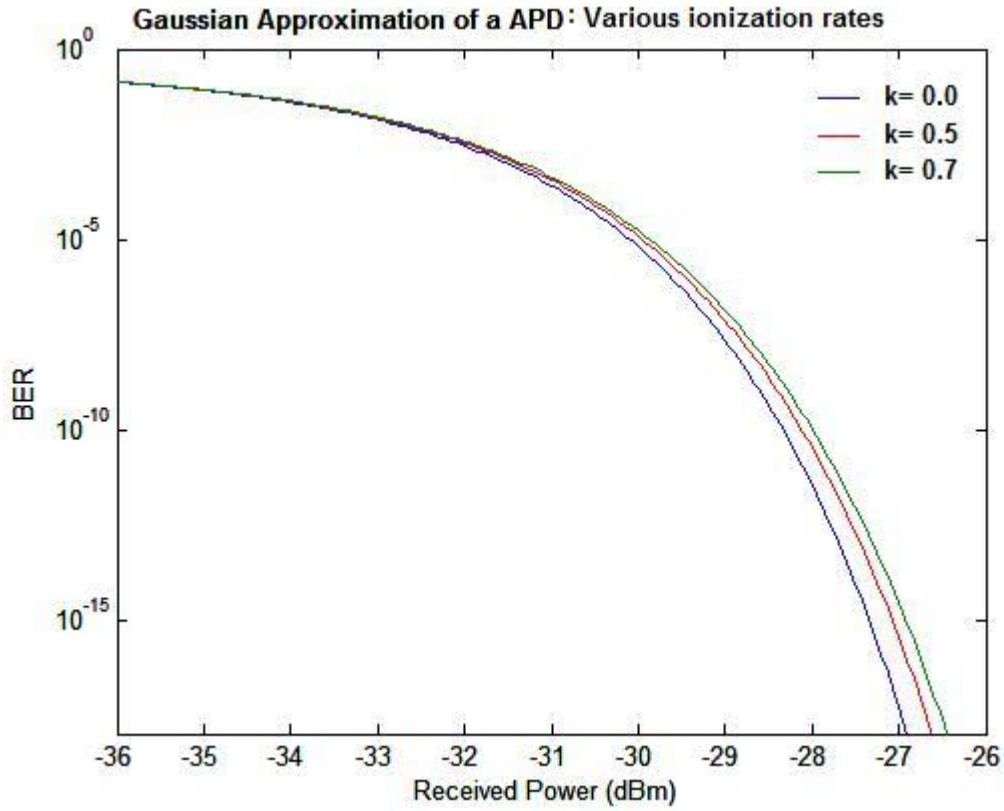

Figure 2.5 – Various ionisation rates for an APD with a gain of 10.

3. MOMENT GENERATING FUNCTION BASED MODEL

The second approach in producing a mathematical model of the detection process was based around the idea of using moment generating functions. Moment generating functions [5&6] can be used to generate the moments of a probably distribution allowing you to model an event that is based on probability such as electron hole pair generation. They are a complete description of a random variable's behaviour (as also is a pdf) and are useful when the pdf is unknown or otherwise difficult to use. The usefulness of the moments is that the first moment gives the mean value and in conjunction with the second moment can be used to determine the variance [7]. The moment generating function allows the generation of these moments directly via the use of a probability density or mass function.

MGFs can be used with both continuous and discrete probability distributions. For example, the expectation of a continuous distribution can be seen to be a long term average or a mean value:

$$E(X) = \int xp(x)dx \tag{24}$$

The variance can then be calculated from this:

$$V(X) = E[(X-\mu)^2] \text{ or } E[(X^2)] - (E[(X)])^2 \tag{25}$$

The expectation of a discrete distribution can also be calculated e.g. of a function f(X), with p_r being the probability of X = r:

$$E(f(X)) = \sum_{r=0}^{\infty} p_r f(r) \tag{26}$$

This result can be used to show that the expectance or mean of the distribution:

$$E(X) = \sum_{r=0}^{\infty} p_r r \tag{27}$$

Therefore again the variance can then be calculated via equation 25.

The MGF can be used to derive these various moments of the distribution:

$$\text{The r}^{\text{th}} \text{ moment of a distribution is } M_x(s) = \sum_{r=0}^{\infty} \frac{E(X^r)s^r}{r!} \tag{28}$$

It can then be then shown that:

$$M_x(s) = E(e^{sX}) \tag{29}$$

This means that the moment generating function can be using to model a random variable by introducing a second variable 's'.

The above simply highlights how the moments contain the statistical information that describes the probability distribution. An important note is that although the first and seconds moments can be used to calculate the mean and variance, it is the contribution from all of the moments that determine the probability density. A full derivation explanation of the MGF and an example is included in the appendix of this report (Appendix A).

Another property is that the various moments can also by derived by differentiation of the MGF [8] e.g.

Taking the first derivative of the MGF gives the mean of the distribution:

$$E\{X(t)\} = \left.\frac{\partial M_{X(t)}(s)}{\partial s}\right|_{s=0} \quad (30)$$

Taking the second derivative of the MGF gives the variance of the distribution:

$$V\{X(t)\} = E\{X(t)\} = E\{X^2(t)\} - [E\{X(t)\}]^2 \quad (31)$$

To summarise the moment generating functions can provide the ability to model a probability distribution such as that encountered in the detection process e.g. with the expectance as the mean current and the shot noise as the variance.

3.1. Pin model

The initial steps in creating a MGF based model were to create MGFs for the mean current and the shot noise and then a separate MGF for the thermal noise for both a zero and one bit.

The steps for the derivation [6,9] have been omitted in this document (see Appendix B) as it was deemed necessary to simply state the final form:

$$M_{X(t)}(s) = \exp\left\{\int_{-\infty}^{\infty} \frac{\eta}{h\nu}(e^{sqh(t-\tau)} - 1)P.d\tau\right\} \quad (32)$$

The individual currents were made by introducing the quantum efficiency of the device 'η', the energy of the photons 'hf' and the specific terms for the relevant powers i.e. when detecting a one bit P_1 and P_0 for a zero bit. This would follow a rectangular pulsed Non Return to Zero scheme.

In a NRZ modulation scheme the value of the pulse does not return to a zero signal level between subsequent pulses which is the characteristic of a Return to Zero (RZ) modulation scheme as seen in figure 3.1.

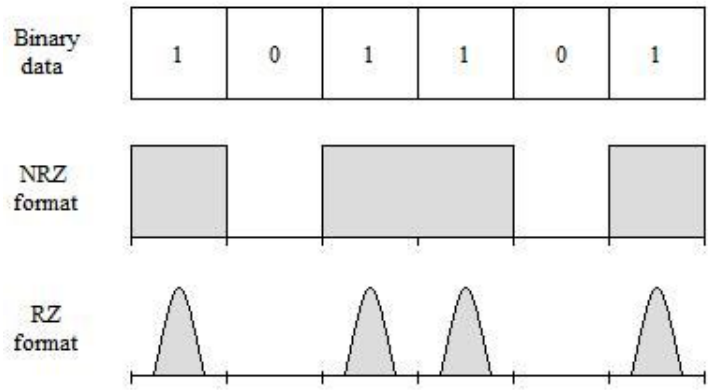

Figure 3.1 – Modulation formats for a digital transmission [2]

This meant that the pulse that was being considered would take the form of that illustrated in figure 3.2 with T_b being the symbol period. The pulse initially considered was centred about zero as this made the evaluation simpler.

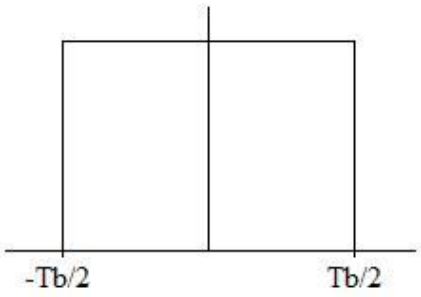

Figure 3.2 – Pulse format initially used for the models

Obviously this would not be practical in a real communication system as it will be discussed later but the limits for the pulse meant that the required integrals now became:

$$M_{I_1}(s) = \exp\left\{\int_{-T_b/2}^{T_b/2} \frac{\eta}{h\nu}(e^{sqh(-\tau)} - 1)P_1 d\tau\right\} \quad \textbf{(33)}$$

$$M_{I_0}(s) = \exp\left\{\int_{-T_b/2}^{T_b/2} \frac{\eta}{h\nu}(e^{sqh(-\tau)} - 1)P_0 d\tau\right\} \quad \textbf{(34)}$$

In order to calculate the moment generating function a numerical integration routine needed to be created. There were various methods available to evaluate this integral numerically including the mid-ordinate, Simpson and trapezoid rule. The trapezoid rule [4] can provide a fairly accurate result with regards to a numerical integration as long as the number of steps taken is quite large.

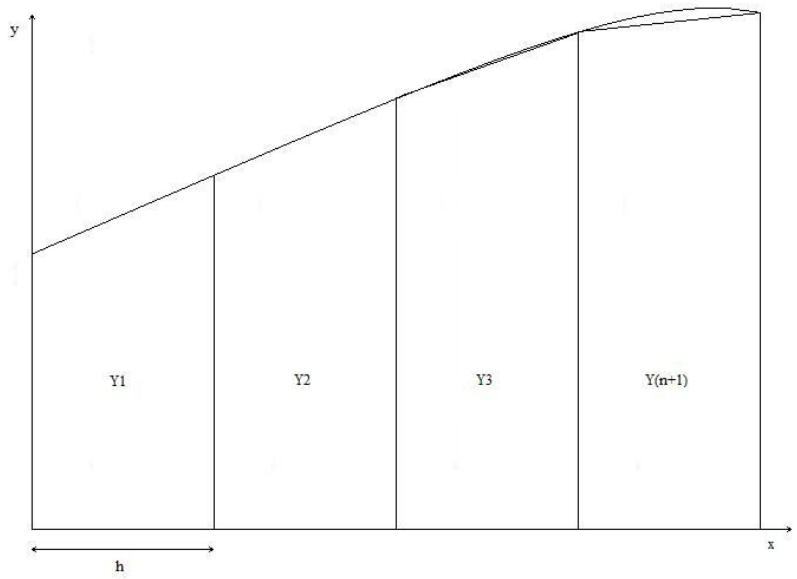

Figure 3.3 – Trapezoidal evaluation of an integral [4]

The basic process for numerically evaluating an integral via the trapezoid rule involved first coding the function to be evaluated. Obviously the integral needs to be evaluated over a range e.g. .from a to b, so in order to perform this computation, the range is divided by a value 'n' thus giving discrete range of values for evaluation e.g. dx = b-a/n. This value of 'dx' and is known as the step size and a smaller value leads to an increase in the accuracy of the evaluated integral. The basic equation used to implement the integration routine can be seen in equation 35.

$$\int_a^b f(x)dx = (dx)\left\{\frac{1}{2}\left(First\_ordinate + Last\_ordinate\right) + \left(Sum\_of\_remaining\_ordinates\right)\right\} \quad (35)$$

The coded routine utilised simple for loops with conditional statements, a specified step size and user defined limits. The values of 'x' to be evaluated need to be increased by a value of dx for each subsequent interval.

The integration routine used for the project was tested comprehensively for a range of known integrals and evaluation to a high degree of accuracy. A very simple example of the routine providing the correct results can be seen by completing the integration of equation analytically and evaluating between the specified limits.

$\int_0^2 2xdx \quad \rightarrow \quad [x^2]_0^2 = 4$ Result from Trapezoid integration routine = 4.0000

$\int_0^{\frac{\pi}{8}} \cos(x)dx \quad \rightarrow \quad [\sin(x)]_0^{\frac{\pi}{8}} = 0.3827$ Result from Trapezoid integration routine = 0.3827

In the case of using a NRZ scheme and choosing an appropriate value of 's' the pulse are rectangular and therefore the impulse response h_r can be set equal to 1. The result of the integration is therefore the area of a rectangle.

Now that a suitable integration routine had been produced the next step simply required raising the result to the power of e. The whole process also needed to be completed for the zero bit's current MGF and these were both coded as functional routines. They were coded in a way so that they could be called as functions given a parameter e.g. from Matlab command line. mgf(s) and mfg1(s). This resulted in the files mgf.m and mgf1.m.

3.2. Thermal noise

The MGF for the thermal noise is known as it is for a Gaussian random variable. Thus the resultant MGF used for calculation can be seen in equation 36.

$$M_{therm}(s) = \exp\left\{\frac{\sigma_{th}^2 \cdot s^2}{2}\right\} \quad (36)$$

This was simply coded as a script file and could be called as a function with a parameter in the same way as the other MGF's e.g. mgftherm(s).

3.3. The Chernoff Bound

The Chernoff bound [5] allows the evaluation up the upper bound of a event based on probability and can be used to express the probability of the value of a random variable *(X)* being greater than some value *(γ)* e.g.

$$P(X > \gamma) \leq e^{(-s\gamma)} M_x(s) \quad when\ s > 0$$

This can be used to determine the probability of the value of I_0 being larger than the decision threshold I_D e.g. an incorrect detection event. This can therefore be applied directly to the first photocurrent I_0:

$$P_{CB}(1|0) = \exp(-s_0 i_D) M_{I_0}(s_0) \qquad s_0 > 0 \quad (37)$$

The same can then be carried out for the second photocurrent I_1

$$P_{CB}(0|1) = \exp(-s_1 i_D) M_{I_1}(s_1) \qquad s_1 < 0 \quad (38)$$

Following from equation 3

$$BER_{CB} = \frac{1}{2}\left[P_{CB}(1|0) + P_{CB}(0|1)\right] \quad (39)$$

Then the use of one variable 's' instead of 's_0' and 's_1'makes the evaluation easier at the expense of a very slight loosening of the bound results in:

$$BER_{CB} = \frac{1}{2}\left[\exp(-si_D)M_{I_0}(s) + \exp(si_D)M_{I_1}(-s)\right] \quad \textbf{(40)}$$

In order to find a minimum for the BER function, partial differentiation of the function of was necessary. Taking a partial derivative allowed i_D to be expressed as a function of s.

$$\frac{\partial BER_{CB}}{\partial i_D} = \frac{1}{2}\left[-s\exp(-si_D)M_{I_0}(s) + s\exp(si_D)M_{I_1}(-s)\right] = 0 \quad \textbf{(41)}$$

Solving by analytical means, results in an expression for i_D that can be evaluated in the original BER function in order to numerically determine a minimum point.

$$\exp(si_D)M_{I_1}(-s) = \exp(-si_D)M_{I_0}(s)$$

$$\exp(2si_D) = \frac{M_{I_0}(s)}{M_{I_1}(-s)}$$

$$i_D = \frac{\ln\left(\frac{M_{I_0}(s)}{M_{I_1}(-s)}\right)}{2s} \quad \textbf{(42)}$$

Substituting equation 42 into 41 and introducing the thermal noise MGF, results in the final equation for the Chernoff Bound.

$$BER_{CB} = M_{therm}(s).\sqrt{M_{I_1}(-s).M_{I_0}(s)} \qquad \text{with } s > 0 \quad \textbf{(43)}$$

Figure 3.4 shows the space associated with the variable's' for a bound for a particular indecent power e.g. -18 dBm. It results in a uni-modal distribution with a minimum value. This minimum value needed to be found in order to find the lowest value for the sensitivity possible. This function was coded into a Matlab script file berVAL.m

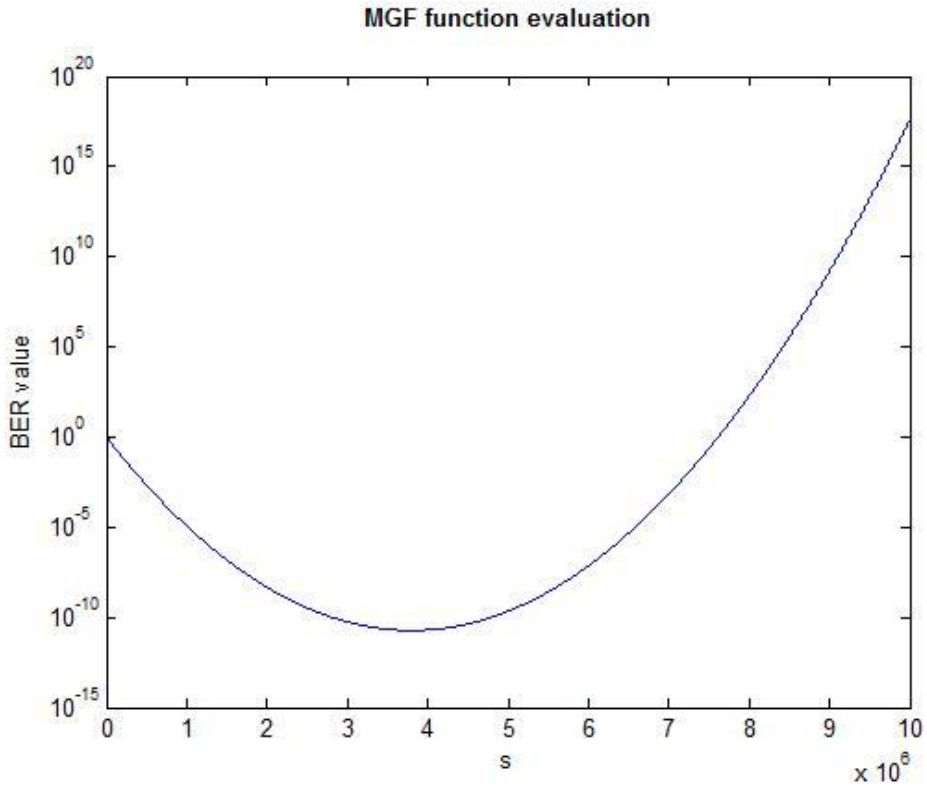

Figure 3.4 – Bound evaluation for Pin device at -18 dBm.

3.4. Golden Section Search

In order to evaluate the Chernoff bound a method needed to be found in order to find the global minimum for the variable 's'. When the function is on one side of the minimum, the value of the function should decrease as the minimum is approached. When the minimum point is reached the value of the function will then start to increase again as you move away from the minimum point. This approach is sufficient for most situations but requires you to specify the starting point near to the minimum and if this is not possible requires you to complete the search over the entire space of the variable. This was undesirable as the minimum could typically take a value between 0 and $1x10^7$. A more suitable approach is to use a minimum finding search algorithm known as the golden section search [10], which reduces the space that has to be searched by use of the golden ratio.

The actual Golden section search algorithm works by reducing the range for the function for which a minimum is known to exist in. This is a more efficient method and determines the location of the minimum in fewer evaluations and iterations. This has a significant impact when the range of the function in question is large. Although this is more computation intensive approach the algorithm determines the minimum in a smaller amount of iterations thus a trade off is reached.

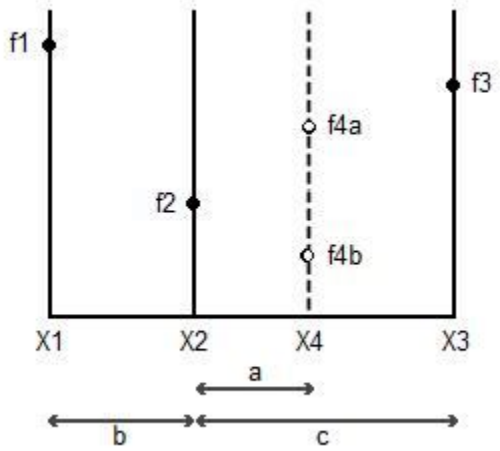

Figure 3.5 – Golden Section Search algorithm [ref]

The basic algorithm initially evaluates the current function between three random points and with the assumption that the function is continuous. It can now be established that if f2 is less than f1 and f3, that the minimum is guaranteed to exist between f3 and f1 or x3 and x2. A new point of evaluation is chosen (f4) which should be selected so that it is positioned within the larger interval namely between f3 and f2 (x3 and x2). Now if the f4 produces a larger value e.g. f4a then it is clear that the minimum lies between f1 and f4a and then f1, f2 and f4 (x1, x2, x4) can then be used for further evaluation a new set of points to begin the process again. If the f4 turns out to produce a smaller result e.g. f4b, then the minimum is going to be located between f3 and f2 (between points x2, x4, x3) and (x2,x3,x4) can then be used for the next evaluation. The reduction of the range utilises the golden ratio in order to ensure spacing used between the values remains in the same proportion to ensure that the minimum is captured. The choice of the initial starting points has a great effect on how quickly the algorithm converges. In order to guarantee the specified starting and end points contain the minimum, evaluation over the range of 's' needs to be carried out. This is discussed in the next section and is which also shows how this process can be optimised.

This search routine was coded a separate Matlab function called gss.m and was designed to be supplied with three parameters (f,a,b,N). 'f' is this case is the function, 'a' and 'b' is the range for evaluation and 'N' the number of iterations for the algorithm. Testing of this routine was initially carried out by using various quadratic functions with known minimums as they are unimodal in nature. The number of iterations needed for the models was found to be 100 as this produced accurate results with the evaluation time being relatively short.

3.5. Dynamic limits

Using to the golden ratio to determine the functions minimum across the entire variables space is computational inefficient. A more appropriate method was found which determined limits for the golden section search algorithm to be evaluated over. A evaluation of the 's' space for a power of -18dBm is shown in figure 3.6 indicating that the minimum occurs between 10^6 and 10^7 and therefore it makes sense to apply the search strictly between this interval. This was achieved by evaluating the function between two points (powers of 10, starting from 10^0 to 10^1) and comparing for a change of sign. Once this change occurs e.g. f(10^x2) > f(10^x1) then a minimum (turning

point) lies within the values of x2 and x1. As the value of 's' was expected to be very large it made no sense to simple increment by a value and evaluating the functions e.g. f(x+1).

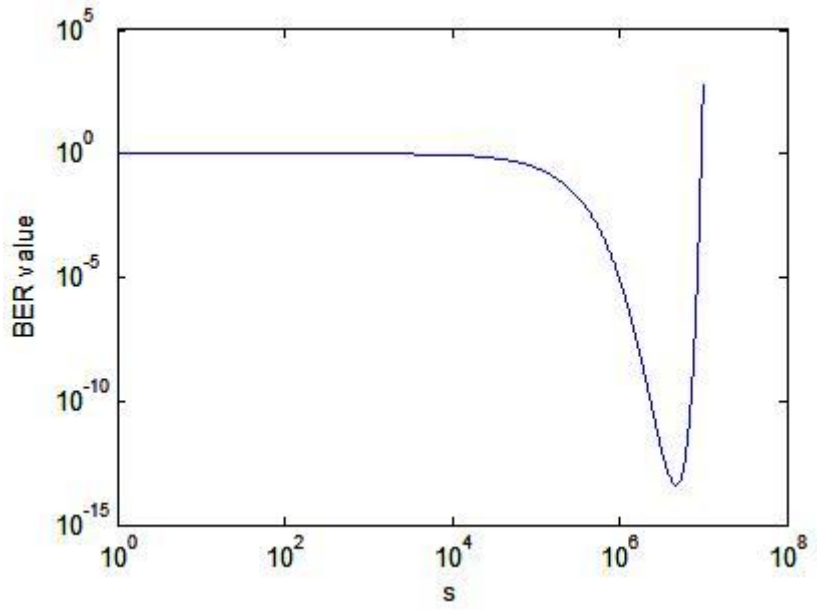

Figure 3.6 – Minimum's location within a range of 's'

Instead the range was increased by a power of 10 each time e.g. 10^0, 10^1, 10^2 etc. This resulted in arriving at a more likely region for the value of 's' more quickly than simple integer incrementation. Once the searching produces the minimum, this indicates the maximum achievable BER value for that particular power.

In order to code the whole process via a series of steps, it required executing the mentioned function files (*.m) with the appropriate parameters as dictated below,

[Specified parameters] - main.m

- *Define incident optical power with a know wavelength*
- *E.g. Popt = -18 dBm and the light is of 1550 nm (mid infra red range)*

[Main Routine] - main.m

- *Use minfind.m to and starting at a value of 's=0' evaluate the other equation using the dynamic limits approach to home in on the minimum and then apply the GSS*

.

[Sub main routine] – Executed for every value of 's'.

- *Convert power to watts*
- *Calculate power P and P for use in each MGF by using equation*
- *Calculate the mgf for a one and a zero bit e.g. mgf and mgf1*
- *Calculate the thermal noise mgf e.g., mgftherm.m*
- *Use berval.m to evaluate the Chernoff bound for that particular power and the current value of 's' and return that value*

[Main Routine] - main.m

- *The value returned via minfind.m is the local minimum for the bound and is used at the BER value for that particular value of incident optical power.*

Continue for a range of powers using a for loop e.g. -30 to -16 dBm.

Used in the above way the script files evaluate the MGFs, find the minimum and make the BER calculations for a range of powers which can then be used to generate a BER curve. Initially an attempt at evaluating the curve provided values that did not seem to be physically sensible according to the values used for the calculations. The issue was found to be that the original MGFs for the currents had not been normalised which means the results were being generated that were greatly out of range. A normalisation was achieved by introduction of a factor of $1/T_b$ e.g.$1/2.5x10^9 = 4x10^{-10}$ for the impulse response h_r instead of simply 1. After this normalisation was completed the calculated values were deemed sensible.

With this amendment the result of main.m being executed can be seen in figure 3.7 which shows how the MGF based approximation compares to that of a Gaussian approximation. The range of powers used was -24 to -16 dBm.

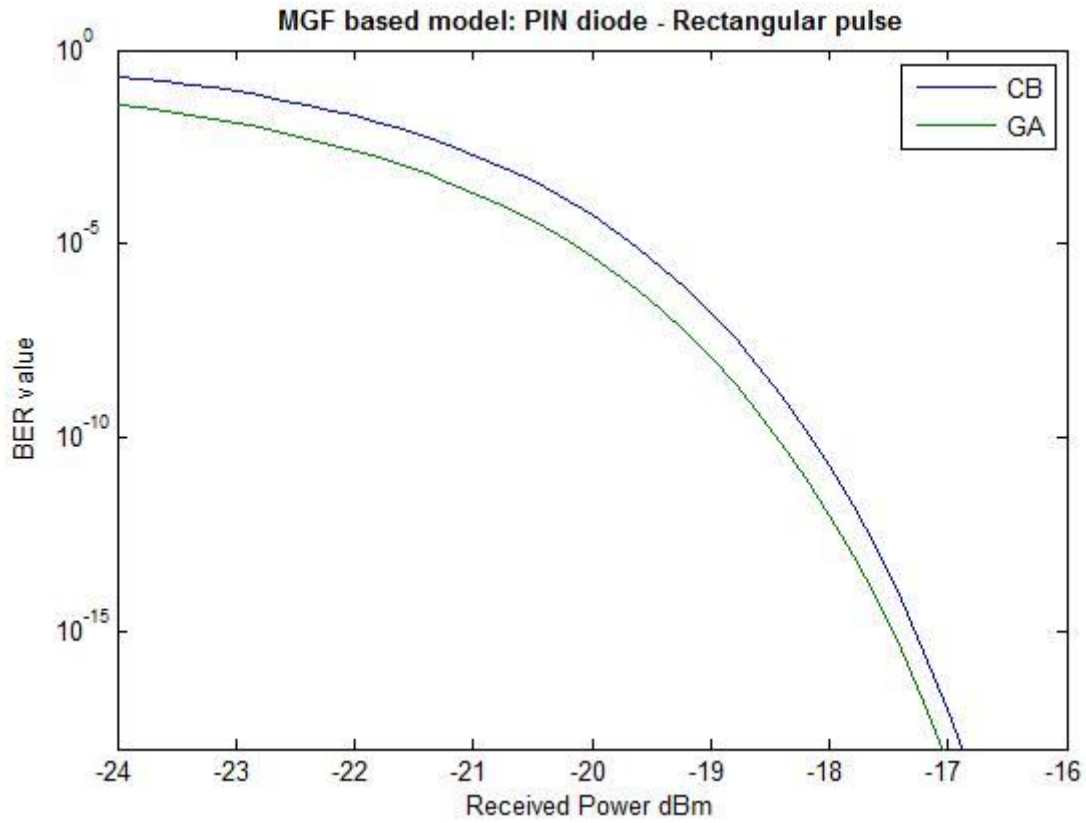

Figure 3.7 – PIN Evaluation- Chernoff bound and Gaussian Approximation

The different functions created for these modules were coded as separate functional elements as this seemed to a logical way to structure the design and reflects the desired industrial style for Matlab based code. A schematic of typical model including the created functions and how they relate to each other can be seen on the next page

Insert Schematic here

3.6. Critique on the golden section search approach

An inherent problem with the implementation of the golden section search (minfind function) is that some situations it can produce erroneous results. These erroneous and subsequently omitted by Matlab when it proceeds in plotting the BER curve. Any omitted results simply result in a break in the curve for he evaluation of the BER calculation for that particular power. After careful examination an explanation for the errors was found by simply evaluating the particular power over its entire 's' space. The role of the 'minfind' function is to detect when the gradient is starting to increase positively as this indicates a turning point (a local minimum) for the function. If the gradient near the minimum changes over a very small amount e.g. 0.00032 – 0.00033 then it can cause the golden section routine to fail [11]. This is because the current approach was to evaluate the function with dynamic levels to 'home in' on a minimum and then apply the GSS over this range. A user defined parameter is the number of iterations that he GSS should take to find the minimum and this produces a result to a specific degree of accuracy e.g. 0.001. If this degree of accuracy is less than the rate of change near the minimum (as specified above) the 'minfind' function will produce a result that is not the true minimum. This was rectified by increasing the number of iterations performed in the GSS.

The validity of the Chernoff bound [9] was calculated by deriving the value of the mean current and shot noise by using the property of MGFs described in on page 16. The various moments can be derived by differentiation of the integral used for the current's MGFs e.g. equation 33 and 34 e.g. at a constant input power -18 dBm.

Table 1 – Mean and variance comparison between the methods

Mean Current E{X(t)}	Shot Noise V{X(t)}
$E\{X(t)\} = \left.\frac{\partial M_{X(t)}(s)}{\partial s}\right\|_{s=0}$	$V\{X(t)\} = E\{X^2(t)\} - [E\{X(t)\}]^2$
Note the value for the responsivity for the Gaussian approximation was set to 1 e.g. $R = \frac{\eta q}{hf} = 1$ A/W	$= \left.\frac{\partial^2 M_{X(t)}(s)}{\partial s^2}\right\|_{s=0} - [R\rho qH(0)]^2$
Therefore the responsivity had to be scaled accordingly for the MGF approach.	$= R\rho q^2 \int_{-\infty}^{\infty} h^2(\tau)d\tau = R\rho q^2 \int_{-\infty}^{\infty} \|H(f)\|^2 df$
$= R\rho q \int_{-\infty}^{\infty} h(\tau)d\tau \quad = R\rho qH(0)$	$V\{X(t)\} = 2qIB_e$ as $B_e = \int_0^{\infty} \|H(f)\|^2 df$

Table 2 – Results for the two methods

Gaussian Approximation		MGF Result	
Current in a zero bit (P_0)	Current in a one bit (P_1)	Current in a zero bit (P_0)	Current in a one bit (P_1)
2.8816e-6	2.8816e-5	2.8848e-6	2.8848e-5
Shot Noise (zero bit)	Shot Noise (one bit)	Shot Noise (zero bit)	Shot Noise (one bit)
1.7290e-15	1.7290e-14	1.7309e-15	1.7309e-14

3.7. APD MGF model

The approaches for creating the MGF model for the APD was very similar to that of the PIN but introduced the gain effect of the device [6]. This was achieved initially by creating a separate MGF namely:

$$M_g(s) = \frac{1}{1 - sM} \text{ With the bound} \qquad s < 1/M \tag{44}$$

The MGF was used to model the gain distribution that would occur due to the avalanche process. Simply integrating the above MGF into the MGF routines for the pin diode and then specifying the additional gain parameter 'M' allowed APD gain approximations to be made.

This led to the MGF for each bit becoming;

$$M_I(s) = \exp\left\{\int_{-T_b/2}^{T_b/2} \frac{\eta}{h\nu}(M_g(sqh(-\tau)) - 1)P\right\}$$

Resulting in

$$M_I(s) = \exp\left\{\int_{-T_b/2}^{T_b/2} \frac{\eta}{h\nu}\left\{\frac{1}{1 - sqh(-\tau)M} - 1\right\}P\right\} \tag{45}$$

The new bound for the MGF can be derived by considering when the input pulse will be at a maximum $h_r(\max)$. Obviously this is constant for a rectangular impulse and causes evaluation of the bound to be:

$$sqh_r(-\tau) < \frac{1}{M} \quad \text{Results in} \quad s < \frac{1}{Mqh_{r(MAX)}} = \frac{1}{1.6\times10^{-19}\cdot 10\cdot 2.5\times10^{9}}$$

This result indicates for a gain value of 10, that the range of 's' is limited by the bound which is equal to 0.25×10^9.

The new MGF for the gain distribution was coded as a function mgfgap.m with 's' as it's parameter. A BER was produced for a gain of 10 and ionisation rate k =0 over a range of incident power (-36 -> -26 dBm). The resultant graph can be seen below along side a Gaussian

approximation. It can be seen again that the MGF model produced a slightly more pessimistic result in comparison. A general characteristic of the APD device is that it can support the same BER at significantly smaller amounts of power e.g. for a gain of 10 (M = 10), sensitivity can be increased by up to 10 dB.

Equation 44 is a basic approximation and assumes a zero ionisation rate k=0. In real terms this would be an ideal avalanche device as this is the best possible condition and is somewhat unrealistic.

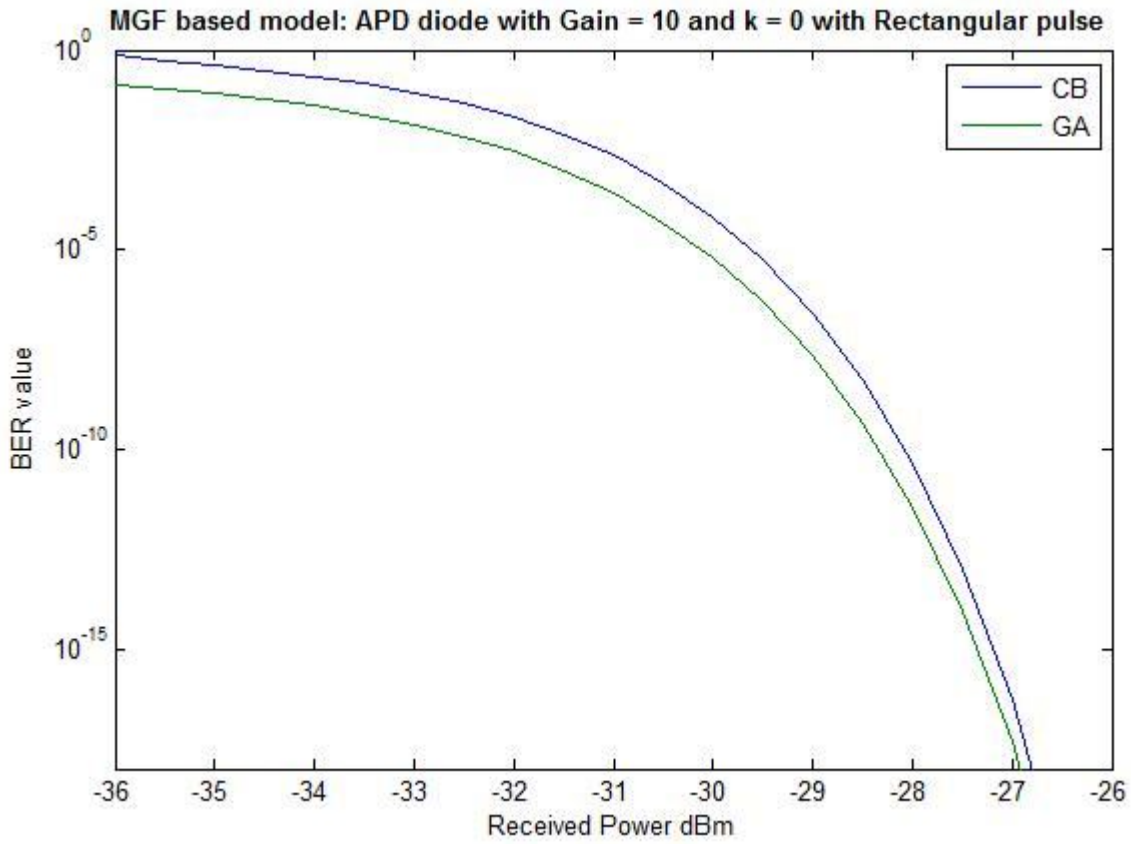

Figure 3.8 – APD evaluation - Chernoff bound and Gaussian Approximation.

As mentioned the APD can offer superior performance over the PIN receiver providing a lower BER over a range of powers as well as equivalent BER at a lower power.

Several gain values were used to evaluate the BER curve for different configurations. Again an obvious result is that a higher gain factor will be able to achieve a BER of 10^{-12} for example at a much lower amount of received power. A clear result is that also as the gain factor increases the gradient of the curve changes more quickly over the range powers.

This is evident from simply looking at figure 3.9 and reading the change in BER for a decrease in 1 dB of received power. For example if we wanted an optical link that would have a performance BER of 10^{-9} for a device with a gain factor of 10 and a change in received power by 1 dB e.g.

at -28 dBm BER $\approx 10^{-10}$ and at -29 dBm the BER $\approx 10^{-6}$
N.B. performance level has already been compromised

This becomes more sever for higher gain values and the decrease of 1 dB for a high gain device can cause significant performance degradation. Therefore a configuration of an APD with a significantly high factor maybe able to achieve target BER at much lower powers but careful control would be needed to make sure that the transmitted power remains at a steady level.

Another modified MGF [6] can be used to model the APD gain taking into the account the excess noise factor and the ionisation rate. This MGF featured in is again an approximation to the gain but provides a more realistic evaluation of APD devices as the previous MGF assumes an APD device with k=0, which would be an ideal device.

$$M_g(s) = \frac{a}{(a-1)^2}[1-\{1-2s(a-1)M\}\}^{1/2}] - \frac{1}{(a-1)}sM + 1 \quad \textbf{(46)}$$

$$a = kM + \left(1 - \frac{1}{M}\right)(1-k) \quad \textbf{(47)}$$

Different gain values were experimented with to see the effect that it has on the BER curve. A clear result is that as the gain value increases a required BER can be achieved at a much lower power level. A side effect is that the range of power diminishes.

The ability to incorporate a more comprehensive gain approximation such as that shown in equation 46, would have allowed MGF based evaluation of the signal to noise ratios that could be expected from a device. For example, a simple analysis is that as the gain factor increases the current increases proportionately e.g. M = 10, results in M × the current.

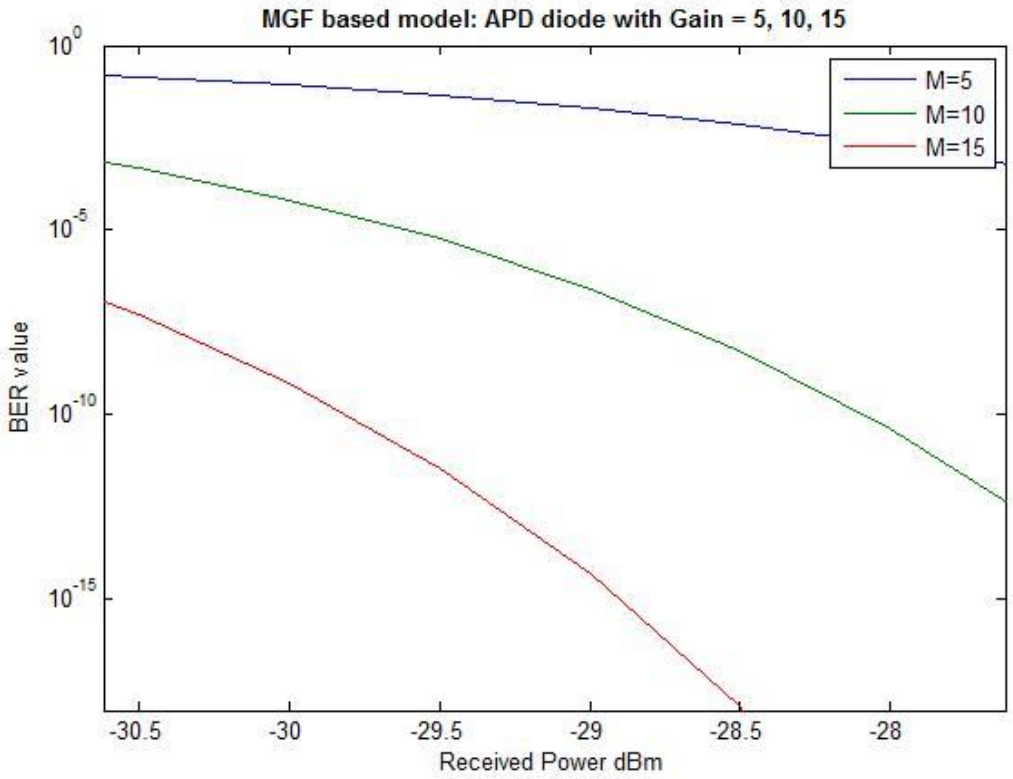

Figure 3.9 – BER value against Received Power – APD (Gain 5,10,15, k=0)

3.8. Modified Chernoff Bound

A slightly tightly bound can be produced by using a modified version of the original Chernoff bound. The modification will not be explored in detail in this report, but the equations necessary were used directly for the evaluation. For a full derivation, see references [6]. This modified Chernoff provides a tighter bound as the derivation specifically takes into account the presence of some Gaussian thermal noise.

$$BER_{MCB} = \frac{M_{th}(s)}{s\sigma_{th}\sqrt{2\pi}} \cdot \sqrt{M_{I_1}(-s).M_{I_0}(s)} \quad (48)$$

The expression shown 48 in can be readily accommodated into the existing models. A typical result produced is shown in figure 3.10, which shows the Chernoff and the modified Chernoff bound produced for a pin photodiode receiver.

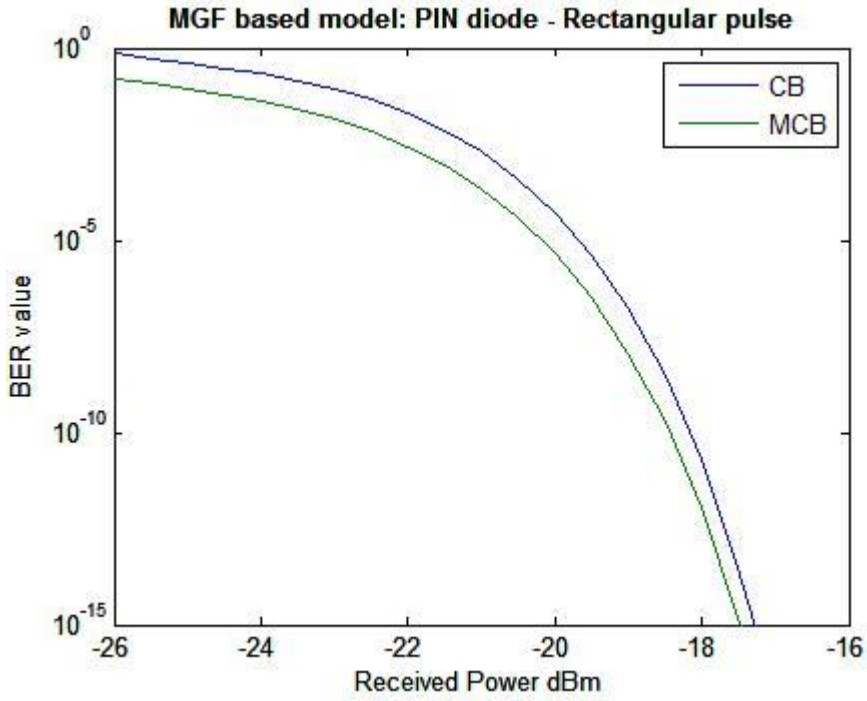

Figure 3.10 – Chernoff bound and modified Chernoff bound

The MCB is an accurate reflection of the physical situation and therefore provides a tighter bound. The saddle point is an additional method that can be used for BER evaluation but was not modelled for this project.

4. OPTICAL AMPLIFICATION

In a practical communication system the received input signal can sometimes suffer from significant attenuation due to transmission over large distances and must undergo an amplification stage. There are a range of optical amplification devices that can be used for this stage e.g. SOA's or Semiconductor Optical Amplifiers and most recently Raman amplifiers, but the most typical are EDFA or rare earth Erbium Doped Fibre Amplifiers. The Erbium Doped fibre amplifier has an amplification window in the 1550 nm wavelength range and this is region where suitable optical gain can be achieved. Therefore an EDFA was a sensible choice for an amplifier that was compatible with the configuration of some existing systems e.g. operating optical wavelength of 1550 nm (C band- Conventional band) [12].

The principle of the EDFA's operation is that optical gain can be achieved by stimulated emission similar to that as within a laser. The received signal in the channel fibre in multiplexed with a light from a pump laser (wavelength band 980 – 1480 nm) which is then connected to a section of Erbium doped fibre. The Erbium doped fibre in connected via splicing it with the channel fibre and then splicing it with another section of convention fibre.

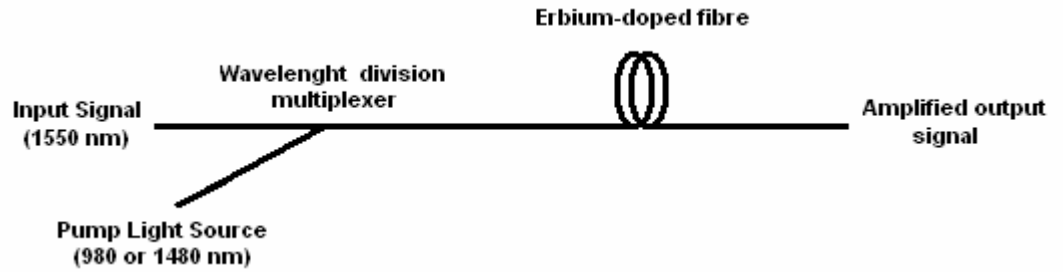

Figure 4.1 – Typical EDFA configuration [12]

The Erbium, doped fibre section is conventional fibre that has been doped with erbium ions (Er^{3+}) during manufacturing. This doping allows energy to be transferred from the light produced via the pump to the received optical signal. The incident photons provided by the pump laser can raise the energy of the dopant ions to a higher energy level where they can then decay back down producing a photon with the same energy, phase, direction and polarisation as the lasers. The ultimate aim is to produce a population inversion of a significant amount of dopant ions to a higher level so that a number of additional photons are produced thus achieving gain [12]. Not all of the ions decay as required and some decay spontaneously giving off their energy as photons (with random frequencies and phase etc) or by other non-radiative methods. Therefore the amplification stage inherently introduces additional noise into the received signal. This additional noise is referred to as ASE or Amplified Spontaneous Emission noise as a proportion of the photons produced by the spontaneous emission are guided by the optical fibre when they then interact with other dopant ions. These photons then undergo their own amplification process and contribute to the signal that arrives at the receiver. The ASE noise has an associated power spectral density and for a single polarisation state can be expressed via equation 49.

$$N_0 = (n_{sp}(G-1)hv) \quad \textbf{(49)}$$

With the 'n_{sp}' being the spontaneous emission noise parameter, 'G' being equal to the gain of the device and 'hv' related to the photon energy. The values used for the model were n_{SP} = 1.78, Gain = 27 dB or 501 watts. The term optical bandwidth refers to a filter that would normally be present in a practical system which has the effect of a bandpass filter on the light received from the fibre. This can be used to limit the effect of the spontaneous-spontaneous beat noise by filtering the output from the EDFA with the optical filter. The major contributor of additional noise would then be the signal-spontaneous noise.

The spontaneous emission parameter 'n_{sp}' is related to the level of the population inversion that is achieved by the amplifier. A complete population e.g. all of the atoms in the excited energy state would have an 'n_{sp}'of 1. Typically a complete inversion may not be obtained and the value of n_{sp} can take slightly larger values e.g. 2 – 5 [3]. The noise figure for the amplifier can be calculated via:

$$NF_{dB} = 10\log_{10}(2n_{sp}) \quad \textbf{(50)}$$

This means that the theoretical minimum noise figure would be 3dB e.g. n_{sp} = 1. The NF for the amplifier used results in 5.51 dB.

Other noises

Another type of noise known as the signal spontaneous noise occurs as a result of interaction between the optical signal and the ASE noise. The ASE noise contains a range of frequencies related to spontaneous decays of atoms and the components have different polarisation. For example if a polarization filter is placed in between the EDFA and the receiver only the signal and the ASE with the same polarisation will come through. The signal and the ASE noise will beat together when the polarisations are matched and this leads to a mixing effect which is detected by the receiver. The frequency of this beat noise is the difference between the frequency of the signal and the frequency of the ASE noise with the same polarization (figure 4.2). A property of the ASE noise is that orthogonal modes do not interfere with each other and are said to be independent.

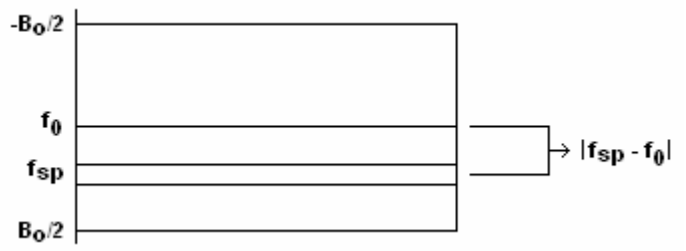

Figure 4.2 – ASE beat noise component [9]

A final type of noise that arises from the ASE is called the Spontaneous-Spontaneous noise. This is a result of the ASE noise beating with itself again causing a mixing effect which is experienced by the receiver.

Practically a polarization filter would not be used and there would be two modes of polarization taking place e.g. within a single mode fibre. As mentioned previously the Signal-spontaneous noise would not be affected by this as the orthogonal modes would not have an effect on each

other. The spontaneous-spontaneous noise on the other hand becomes twice as large as each mode produces ASE noise which in term beats with itself.

Therefore the total noise produced via the amplification process has four components being the thermal and shot noise (in addition), the signal-spontaneous emission beat noise and the spontaneous-spontaneous emission beat noise. The total noise variance can be expressed as:

$$\sigma^2 = \sigma^2_{thermal} + \sigma^2_{shot} + \sigma^2_{s\_sp} + \sigma^2_{sp+sp} \quad (51)$$

There are Gaussian approximations for these processes that can be used to model these additional noise components but were only reviewed for the project and not practically implemented.

4.1. The model

To model the effects of the optical amplifier a modified MGF was used and a complete derivation is available in the [13].

$$M_Y(s) = \frac{\exp\left\{\int_{-\infty}^{\infty}\left[\frac{RG(e^{sqh_r(t-\tau)}-1)h_p(\tau)}{1-RN_0(e^{sqh_r(t-\tau)}-1)}\right]d\tau\right\}}{\exp\left\{\int_{-\infty}^{\infty} Bo.\ln\left[1-RN_0(e^{sqh_r(t-\tau)}-1)\right]d\tau\right\}} \quad (52)$$

The term $h_p(\tau)$ represents the power within the optical pulse that is present at the input of the EDFA. The term B_o or the optical bandwidth refers to a filter that would normally be present in a practical system. This would be placed in between the EDFA and the receiver. This limits the effect of the spontaneous-spontaneous beat noise by filtering the ASE output from the EDFA. The optical bandwidth can be minimised to achieve designed suppression up to a limit as it must not unduly distort the signal. The major contributor of additional noise would now be from the signal-spontaneous noise. The value used for this model was B_o=1.15 nm @ 1550 nm (but this was included as 2*B_o as typically there would be two modes of polarisation.

The implementation of this new MGF required modification of the PIN model's MGF and required two numerical integration routines to evaluate the integrals. The additional parameters mentioned above in equation were also set. Theoretically for a pin with a noise free optical amplifier, a gain 27 dB should produce a BER of 10^{-12} at approximately (-18 – 27 = -45 dBm). Taking into the account of the impairments, the value is a few decibels less. The inclusion of this additional model for the optical amplifier and the pin allow modelling of more complex systems in order to compare its effectiveness to a high gain APD. A typical result of the model can be seen in figure 4.3 showing a BER of 10^{-12} being achievable with a very small amount of optical power present at the EDFA's input.

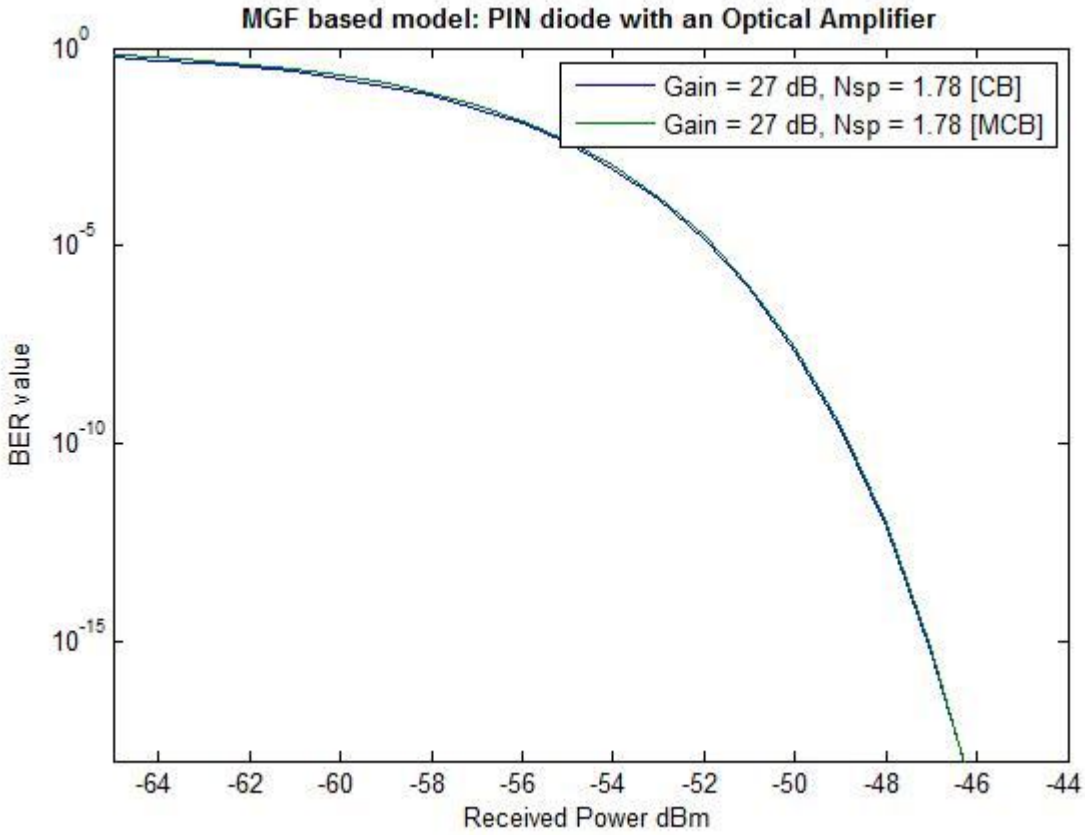

Figure 4.3 – Chernoff and Modified Chernoff bound evaluation for an Optical amplifier

5. INTERSYMBOL INTERFERENCE

The models created for this project have so far assumed a rectangular input pulse following a NRZ format. A major issue with trying to send an actual rectangular pulse via communication systems is that it is composed of a large number of frequencies. This presents a very real performance constraint as most communications links allow the sending of multiple signals by dividing the total available bandwidth into a number of separate channels. Therefore there will be a limited amount of bandwidth per channel meaning that the optimal usage of it will be a major design consideration. The effects of dispersion experienced by an optical fibre cause the transmitted pulse to become elongated in the time domain leading to an increase in the pulse width [14]. The issue concerning the transmitted pulse is how the power of each symbol (1 or 0 bit) might interfere with its neighbours. This is formally known as Inter Symbol Interference or ISI.

A solution to this problem is to either lower the data rate of the link to allow for detection of the elongated pulse or to limit the effects of the spreading with pulse shaping. Pulse shaping can be carried out by use of a raised cosine filter [15]. Figure 5.1 shows the form of a raised cosine filter and how its shape is affected by different values of β .

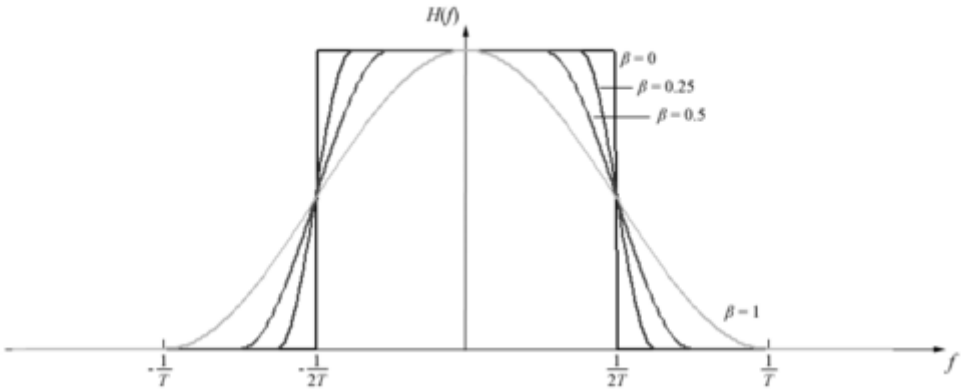

Figure 5.1 – Raised Cosine output pulse Filter – Frequency representation [ref]

In order to limit the effects of the ISI the desired type of filter can be given a value for β which can minimise the ISI to various degrees. A 100% raised cosine filter β=1 was used for the project used as ensure zero ISI at the sampling instant.

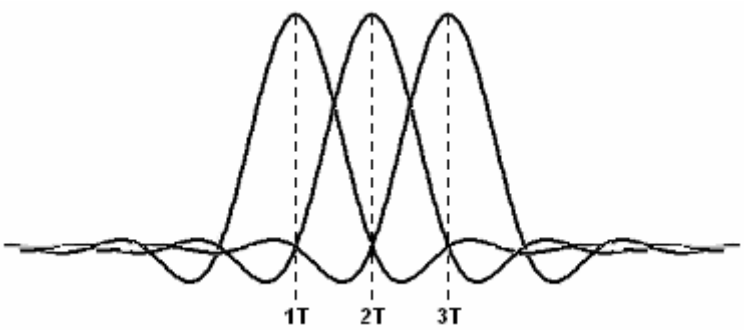

Figure 5.2 – 3 Consecutive pulses demonstrating a Zero ISI digital transmission

This meant that the power from any neighbouring pulses would be zero at this point. The method to achieve the pulse shaping was to supply the rectangular input pulses (NRZ scheme) to a filter which would produce a raised cosine output.

The transforms function necessary to achieve this transformation required calculation of the Fourier transforms. It was deemed easier to calculate the response in the frequency domain as multiplication was easier to deduce mathematically than convolution in the time domain. The input pulse in this case was presumed to be an idealised rectangular pulse modulated via a NRZ scheme and meaning the Fourier transform would be:

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft}\, dt \quad = \quad H_{in}(f) = \int_{-\frac{T_b}{2}}^{\frac{T_b}{2}} e^{-j2\pi ft}\, dt$$

$$\left[\frac{e^{jx} - e^{-jx}}{-j2\pi f} \right]_{-\frac{T_2}{2}}^{\frac{T_b}{2}} \quad = \quad \frac{1}{\pi f}\left[\frac{e^{jx} - e^{-jx}}{-j2} \right]$$

Then using the identity: $\sin(x) = \dfrac{e^{jx} - e^{-jx}}{2j}$

$$H_{in}(f) = \frac{1}{\pi f}\sin(\pi f T_b) \tag{58}$$

Then the normalised Fourier transform for the desired output took the form of equation with β as 1 and inserting the appropriate limits:

$$H_{out}(f) = \frac{T}{2}\left[1 - \sin\left(\frac{\pi f T}{\beta} - \frac{\pi}{2\beta}\right)\right] \quad = \quad \frac{T_b}{2}\left[1 - \sin\left(\pi f T_b - \frac{\pi}{2}\right)\right]$$

Resulting in an expression for the impulse response in the frequency domain:

$$H_r(f) = \frac{H_{out}(f)}{H_{in}(f)} \quad = \quad H_r(f) = \frac{\frac{T_b}{2}[1 - \sin(\pi f T_b - \frac{\pi}{2})]}{\frac{1}{\pi f}\sin(\pi f T_b)}$$

Then finally taking the inverse Fourier transform to transform it into the time domain.

$$h_r(t) = \int_{-\infty}^{\infty} H_r(f) exp(j2\pi ft) df$$

$$h_r(t) = \int_{-\frac{1}{T_b}}^{\frac{1}{T_b}} \frac{\frac{T_b}{2}[1-(\sin(\pi fT_b - \pi/2)]}{\frac{1}{\pi f}\sin(\pi fT_b)} \exp(j2\pi ft) df$$

*Using symmetry argument the factor of jSin2πft can be ignored

$$h_r(t) = 2\int_{0}^{\frac{1}{T_b}} \frac{\frac{T_b}{2}[1-(\sin(\pi fT_b - \pi/2)]}{\frac{1}{\pi f}\sin(\pi fT_b)} \cos(2\pi ft) df \qquad \textbf{(E.q. 56)}$$

This integral was then in a form that could be used to represent the impulse response of a filter that gives a raised cosine output for a rectangular input. Using another numerical integration routine, the impulse response was coded as function called 'hresp.m' then could be used by calling it in Matlab with an argument of 't' e.g. hresp(t) where 't' would be the time for the pulse e.g. tau. This response was tested to ensure that it would produce the required target output pulse given a rectangular impulse. It was tested by coding the integral in MathCAD and by specifying an input pulse. Figure 5.3 shows the result of the output.

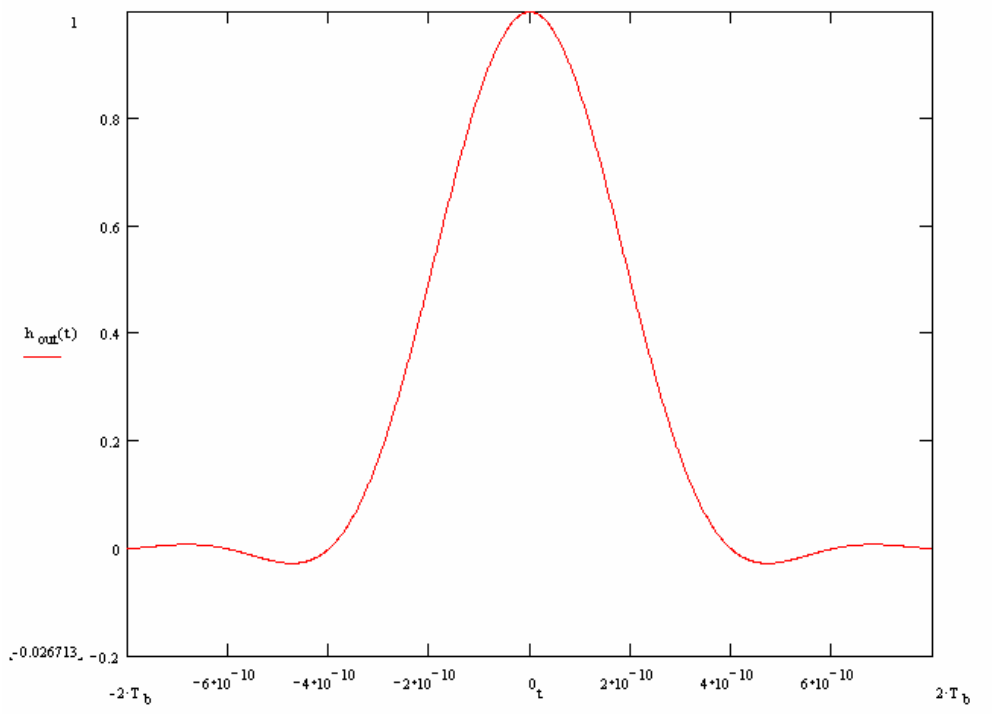

Figure 5.3 – h_{out} – Impulse response for the Raised Cosine Filter

The reason for doing this was that a raised cosine output would guarantee zero ISI at the sampling point. As mentioned previously, this would be advantageous as it can be seen in figure 5.4 is that ISI can lead to an increased amount of incorrect detection of bits leading to poorer system performance [16].

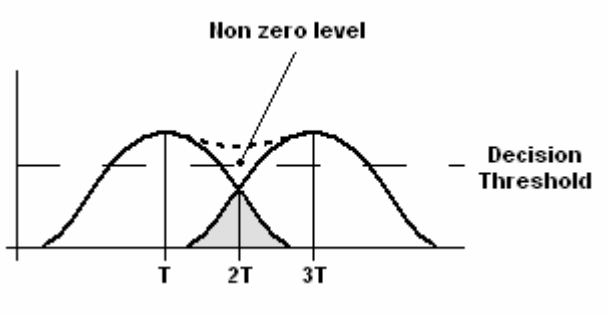

Fig 5.4 – ISI and its effect on the BER [16]

An attempt was made at incorporating a bit sequence into the models by using modified MGFs. For example consider a sequencer of bits that can be transmitted via various pulse of optical power e.g.

$$P_{in}(t) = \sum_{k=-\infty}^{+\infty} a_k . h_p(t - kT). \tag{56}$$

The MGF can be conditioned to represent the sequence by replacing $h_p(t)$ with $P_{in}(t)$.

$$M_Y(s \mid \{a_k\}) = \frac{1}{D(s)} \exp \int_{-\infty}^{\infty} \frac{F(s)}{N(s)} . \sum_{k=-\infty}^{+\infty} a_k . h_p(t - kT) . d\tau \tag{56}$$

With the expression of F(s), N(s) and D(s) equal to the expression for the optical amplifier (see Equation 52).

$$F(s) = GR(e^{sqh_r(t-\tau)} - 1) \qquad N(s) = 1 - RN_0(e^{sqh_r(t-\tau)} - 1)$$

$$D(s) = \exp\left\{ \int_{-\infty}^{\infty} B_0 . \ln[1 - RN_0(e^{sqh_r(t-\tau)} - 1)] d\tau \right\}$$

This was partially coded for this project but never fully completed. Completion of this would have allowed observation of the effect that ISI can have on the systems BER and the improvement once the pulse shaping filter is used.

6. INDUSTRIAL RELEVANCE

6.1. Network Design

The main design considerations for optical networks include the physical size of the network, the transmitted and received power and the quality and efficiency of electrical circuits and components. It is common in industry for required BERs for optical links to target values of 10^{-12} to 10^{-15}. The real BER that a system will support is based on numerous factors such as the transmitter and receiver characteristics, the length of the link and the encountered level of interference. The methods used for BER evaluation in this project produce theoretical estimations to certain degrees of accuracy and can produce results that a better or worse than the actual real BER. Figure 6.1 shows how the upper bound generated by the Chernoff bound might compare to the real curve. For a certain situation a Gaussian approximation might produce a curve that lies to the right of the upper bound and therefore the Chernoff bound provides a more realistic evaluation of the performance. In another situation a Gaussian approximation might produce a bound that is to the left of the Chernoff bound and therefore may or may not produce a more realistic evaluation.

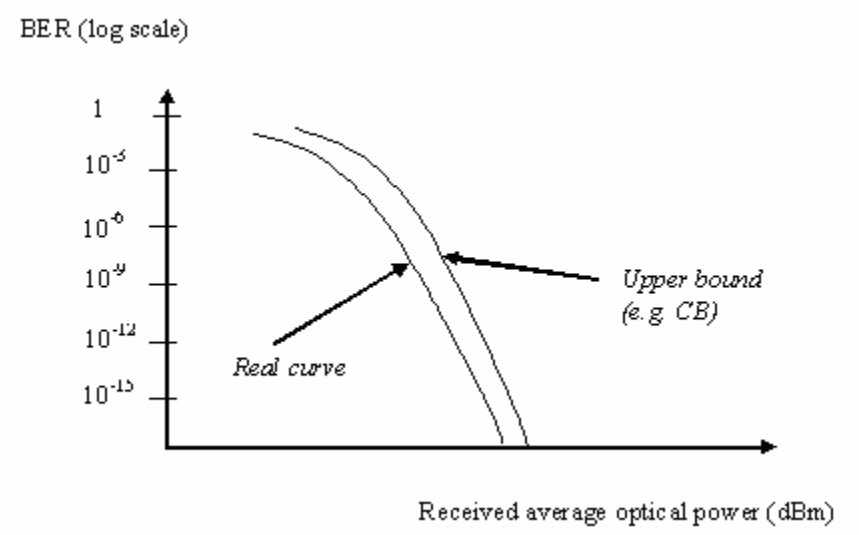

Figure 6.1 – Bound comparison for system evaluation [9]

Being able to accurately model a proposed link and determine very closely the actual required power for maintaining a specific BER is therefore a major advantage as the a little power will be wasted trying to maintain a pessimistic or loose BER. Therefore it is important for accurate methods to be developed to accurately evaluate the theoretical BER of a proposed system. The reasoning behind this is that it is inefficient and costly to be transmitting more power than necessary for a given system.

Although these models can provide performance estimations the fail to take into account practical issues that may be experienced. Typically a power budget maybe proposed for a link which takes into account the power losses associated with typical components such as losses experienced by the laser, component degradation etc. Also in reality, as the received power from a transmission increases in intensity it can have the effect of causing electrical amplifiers to saturate. This can in

fact impede the performance thus lowering the achieving BER back towards a lower rate as shown in figure 6.2.

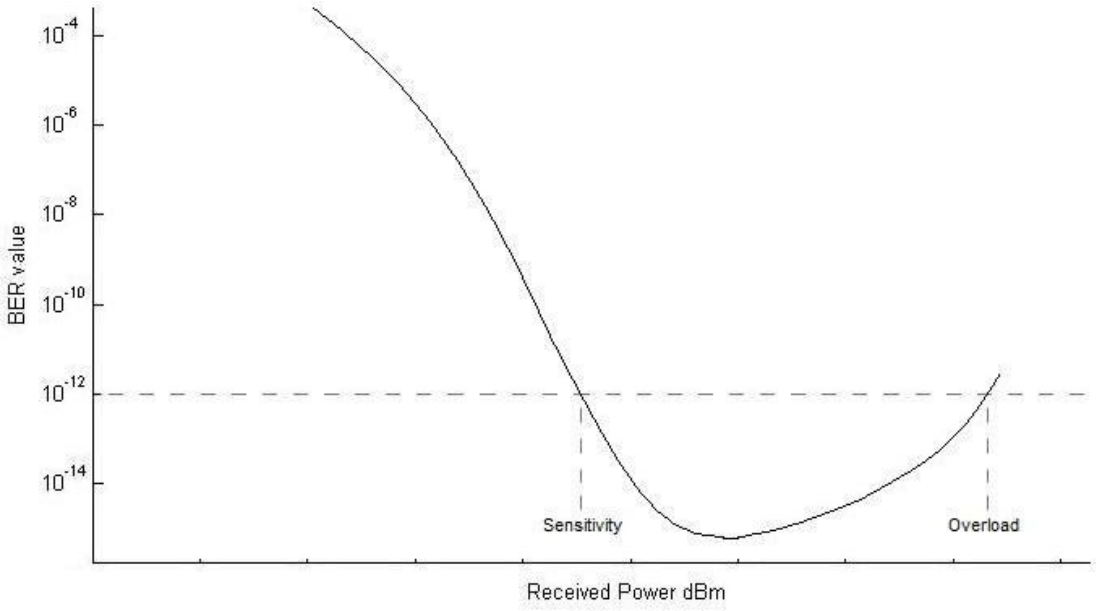

Figure 6.2 – Practical BER due to amplifier saturation [9]

An example of this is that simple evaluations of performance such as taking a Gaussian approximation produce continuously improving BER values for increasing amounts of received power and may not be truly reflect the situation that actually may occur. Therefore from a design perspective it might be seen initially that providing more power for detection will greatly improve the BER of the system where actually this may not be the case. This in turn could lead to links being designed with transmitted power levels that are more than necessary producing more costly and inefficient designs.

Additionally, the comprehensive modelling of devices and typically encountered optical components provides the designer with the ability to determine whether or not similar performance can be achieved with slightly cheaper or less complex components. For example APDs are currently expensive in comparison to PIN diodes and require careful control of a large reverse bias voltage for the device to operate in the avalanche region. A pin diode and an EDFA might be a more suitable option in this particular case achieving similar or better performance.

7. CONCLUSION AND FINDINGS

The Gaussian approximation is a useful estimate for at determining the variances of the two noises and how they compare to each other. Figure 7.1 shows how the shot and thermal noise values compare to each other in terms of their magnitude for a pin photodiode. As the shot noise level is relatively low in comparison to the thermal noise for a range of powers, it follows that the in the case of the pin, the thermal noise dominates. In the case of the APD, the thermal noise initially dominates for low gain values until a particular value of M is reached where the contribution of the shot noise becomes more prominent e.g. when M is greater than 7 in this particular case See fig 7.2. As the gain increases the shot noise increase by a factor of F(M) which is a fairly obvious result (dictated by equation 23). This is because more gain results in more current and thus more shot noise. Practically the gain could be limited via Automatic Gain Control circuitry as the received optical signal can vary. When the received power is close to the minimum sensitivity of the APD configuration then the AGC is not necessary and the gain should be kept at a value that minimises the noise contribution. As the received power increases then the AGC should be able to compensate for this and reduce the gain to avoid overloading the receiver.

A conclusion expressed in section 3.6 is that an APD can offer increased receiver sensitivity in comparison to pins in certain situations. However these devices can offer advantages over each other in certain situations. Further analysis can be used to perform comparisons of the devices and the effects that changing system parameters will have on them. Figure 7.3 shows how the systems bit rate was varied between three values and the resultant BER curves generated via a Gaussian approximation. The plots show as the Bit rate for a system is increased then the achievable BER for a specific power decreases. For example the incident power needed for the PIN was at a bit rate of 2.5 Gb/s and. In practical terms an increase in Bit rate will lead to an increase in the thermal noise and shot noise and the effective bandwidth thus causing an increase of erroneous detections. The relationship between the two noises and the bit rate is evident from the equations used in this project to model their contributions. Table 3 shows the sensitivities of the different receivers needed to achieve a BER of 10^{-12}.

Table 3 – Sensitivities for various Bit rates obtained via the models

	2.5 Gb/s	5 Gb/s	10 Gb/s
PIN	-18 dBm	-16.5 dBm	-15 dBm
APD	-27.5 dBm	-25.8 dBm	-24 dBm

The results for similar analysis performed for an APD are shown in figure 7.4. It can be seen from the graph that the Bit rate has a similar effect to that of the PIN in that an increase leads to a decrease in the supportable BER. The increase in bit rates will again lead to an increase in the thermal noise but this is compensated for by the Gain factor of the APD. The shot noise is also affected by the increase in the bit rate. The results for the APD suggest that an increase in bit rate has less of an effect on the sensitivity and produces less of a deviation from a required BER. This suggests that an APD may be more suitable for higher bit rate systems and practically they are used for exactly this situation. APDs are actually the preferred type of optical receiver for high speed networks in industry being able to support transfer rates up to 12.5 Gb/s. In the case of device sensitivity and the maximum usable bit rate a trade off needs to be reached. The APD for

example can provide additional gain but this has a direct effect on the maximum supportable bit rate. This is due to the fact that the gain produced via the avalanche process actually takes a certain amount of time to build up. The time for this process means that the response time of the device is affected by this i.e. the bandwidth is dependent on it. This results in a gain bandwidth product which is typical of any common amplifier e.g. at low gain the device's bandwidth is fairly constant, but at high gain the bandwidth decrease proportionally with increasing gain.

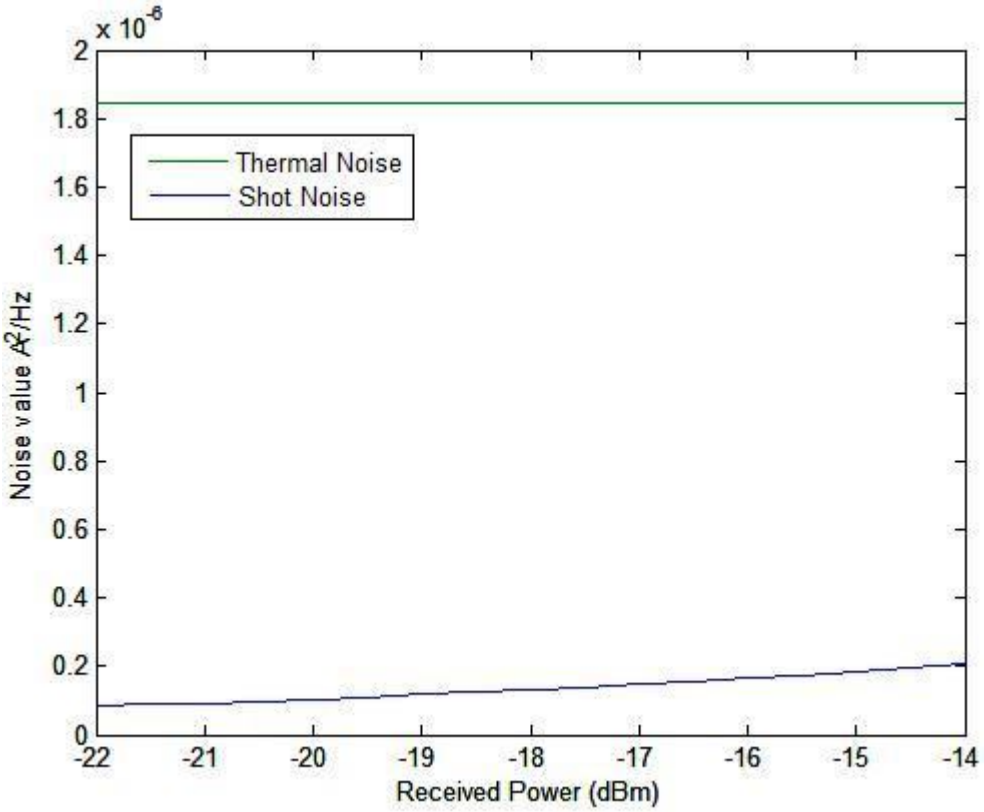

Figure 7.1 – Shot and thermal noise comparison for a PIN diode

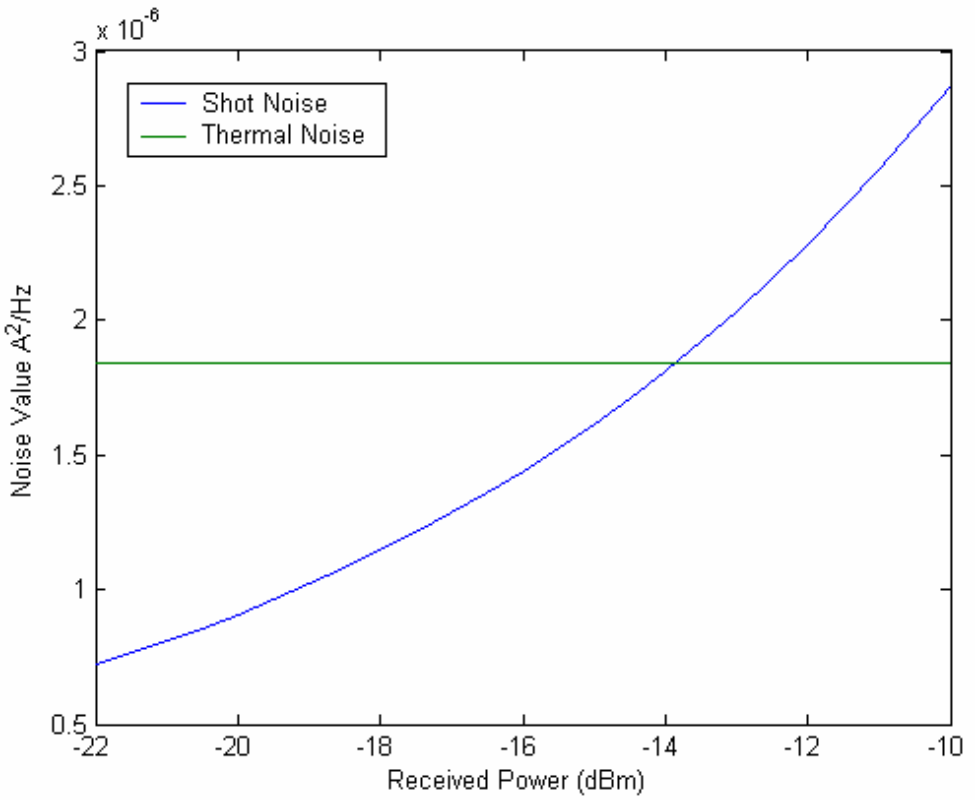

Figure 7.2 – Shot and thermal noise comparison for a APD (Gain=10 , k=0.7)

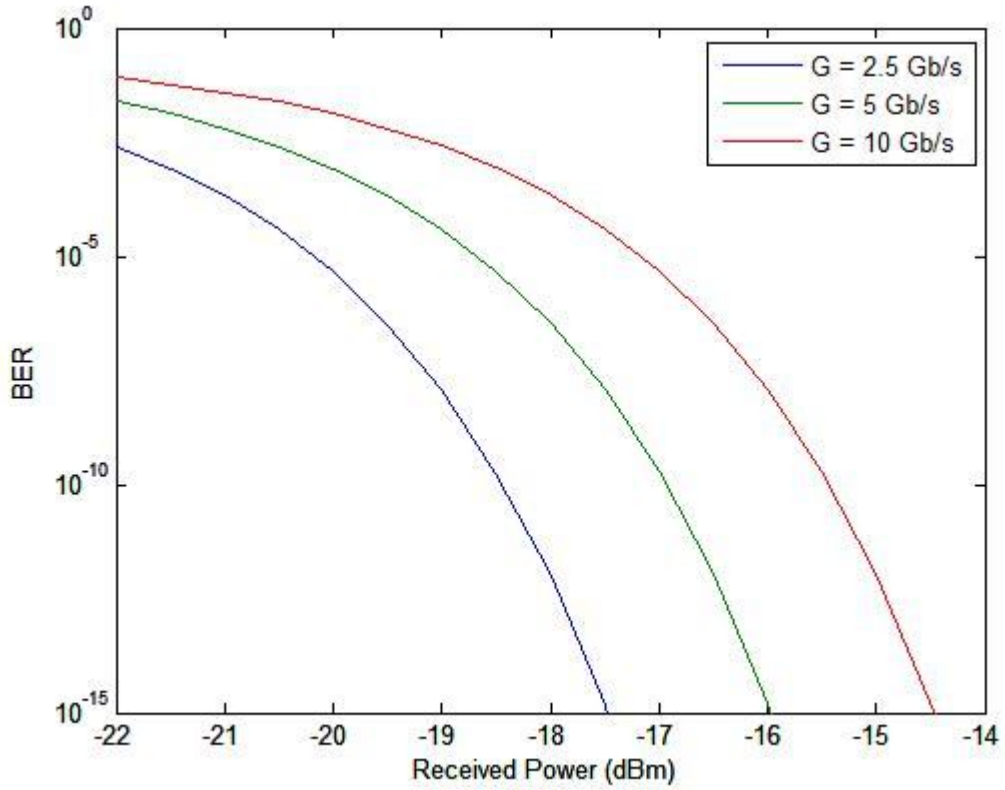

Figure 7.3 – Bit error rate versus receiver power (PIN diode)

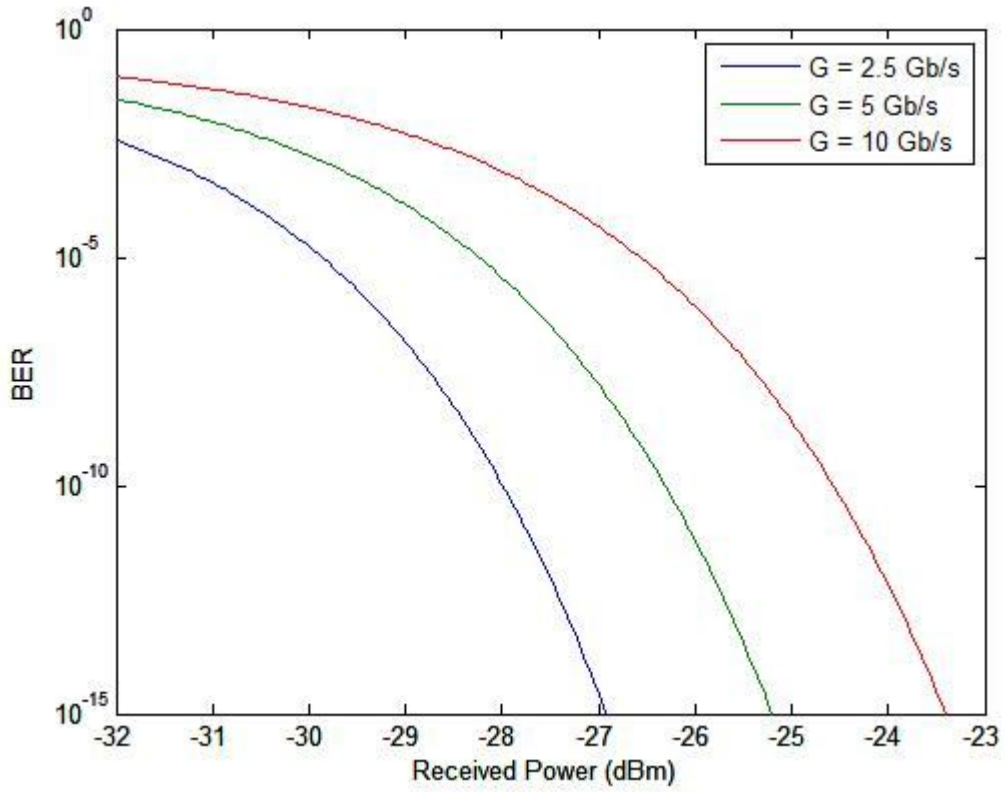

Figure 7.4 – Bit error rate versus received power (APD - Gain=10, k=0.7)

Simple expressions were derived to help justify the results produced by the models for certain situations. The Q's of the two devices can be compared directly (Appendix C) to one another with the result:

$$\frac{Q_{APD}}{Q_{PIN}} = \frac{\sqrt{\sigma_{th}^2 + 2qRP_1Be} + \sqrt{\sigma_{th}^2 + 2qRP_0Be}}{\sqrt{\frac{\sigma_{th}^2}{M^2} + 2qRF(M)P_1Be} + \sqrt{\frac{\sigma_{th}^2}{M^2} + 2qRF(M)P_0Be}}$$

This is with the assumption that the devices are at the same power and that they have the same thermal noise. Now if the power is increased towards infinity, the top expression tends towards 1 with the bottom expression proportionately increasing in value, leading to an larger excess noise term e.g. F(M). This leads to the conclusion that the larger the amount of the received power the larger the amount of excess noise generate in the APD device which is also dictated by a factor of M. Therefore the PIN offers an advantage when the received power is large.

$$\text{As Pav} \rightarrow \infty$$

$$\frac{Q_{APD}}{Q_{PIN}} = \frac{1}{\sqrt{F(M)}} < 1 \text{ Advantage PIN device}$$

If the power tends towards 0 then the gain factors indicating that the gain from the device will compensate for the thermal noise increasing well as M increases. Therefore this would suggest that the APD offers an advantage over the PIN when the gain factor is greater than one and at smaller powers.

$$\text{As Pav} \rightarrow 0$$

$$\frac{Q_{APD}}{Q_{PIN}} = M > 1 \quad \text{Advantage APD device.}$$

The above analysis is merely a mathematical argument as to why the devices display the characteristics indicated by their BER curves. The main aim of this project was not to compare the device characteristic but to compare and contrast various methods for the BER generation and is carried out in the next section. However the argument above is a useful way of being able to understand how the two noises in these devices play a role and the benefits of using either device in certain situations.

7.1. Method comparison CB, MCB and GA

A method for comparing the three different methods for performance evaluation was carried out by using the methods with the same parameters e.g. responsivity, range of power etc. The resultant curves generated via the methods were then plotted on the same graph to allow calculations of the variance between the curves and to graphically see how the methods compare (Figure 7.5).

It can be seen from that the graph that for the MGF based Chernoff bound produced for the pin diode resulted in a more pessimistic BER curve. The modified Chernoff bound however produced a curve that was very similar in comparison to the Gaussian approximation in this case. The relevant values for three main powers were determined for three BERs including 10^{-9}, 10^{-12} and 10^{-15} (see table 4). These values were chosen as they form a decent range of BER for comparing the curves and they are also commonly quoted or required BERs in industry. The bit rate for the system was set at 2.5 Gb/s.

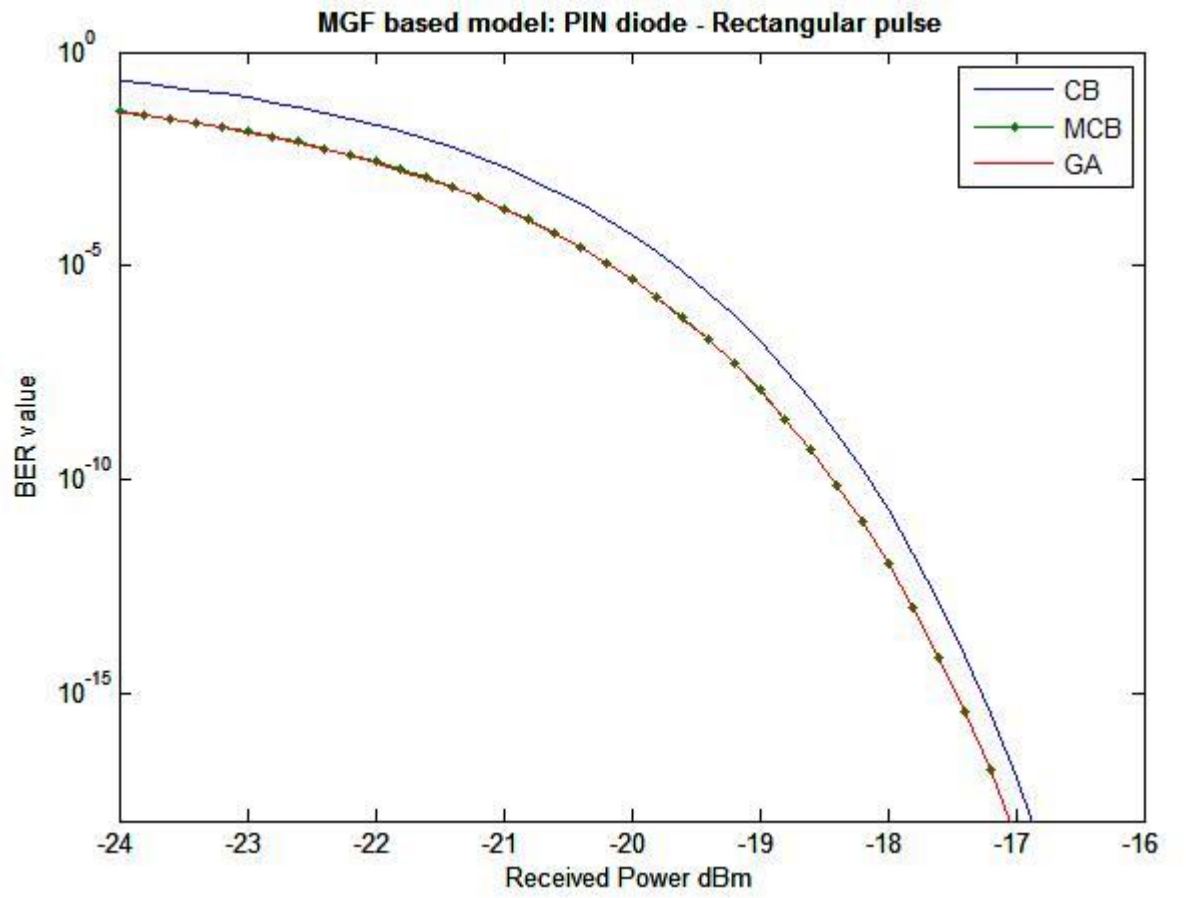

Figure 7.5 – Bound comparison for a PIN

Table 4 - PIN Based receiver – BER comparison

BER	RECEIVED POWER (dBm)			DIFFERENCE FROM GA	
	GA	CB	MCB	CB	MCB
10^{-9}	-18.6891	-18.3816	-18.6893	0.3075	0.0002
10^{-12}	-17.9939	-17.7572	-17.9948	0.2367	0.0009
10^{-15}	-17.4682	-17.2726	-17.4699	0.1956	0.0017

It can be seen from the table that Chernoff bound varied up to 0.3 of a dB in value from the Gaussian approximation.

If a system was designed according to the Gaussian approximation then it would dictate a power level of -17.99 for a BER 10^{-12}. Initially it was seen that a reduction in received power of 0.3 dB could occur in a practical link e.g. laser power output variation, fibre absorption and the BER would then be compromised. Therefore fluctuation form the required BER may be a problem due to a design based on the Gaussian approximation alone. However MCB provide a more definitive bound and in the case of a PIN based receiver, the Gaussian approximation is very similar and therefore can provide a realistic BER evaluation.

The same analysis was carried out from the avalanche photodiode detector with various gain factors as required by the specification of the project. It can be seen that the Gaussian approximation can very somewhat for the APD device. Figure 7.6 shows that the GA is slightly more optimistic than the Modified Chernoff bound suggesting a better level of sensitivity. As the gain is increased the GA tends towards the MCB (see figure 7.7). For high gain values the GA seems to underestimate the receiver sensitivity and indicates a sensitivity that is more pessimistic than the modified Chernoff bound (See figure 7.8). It therefore makes sense that in the case of the APD that the CB or MCB would be a better choice for practical system design.

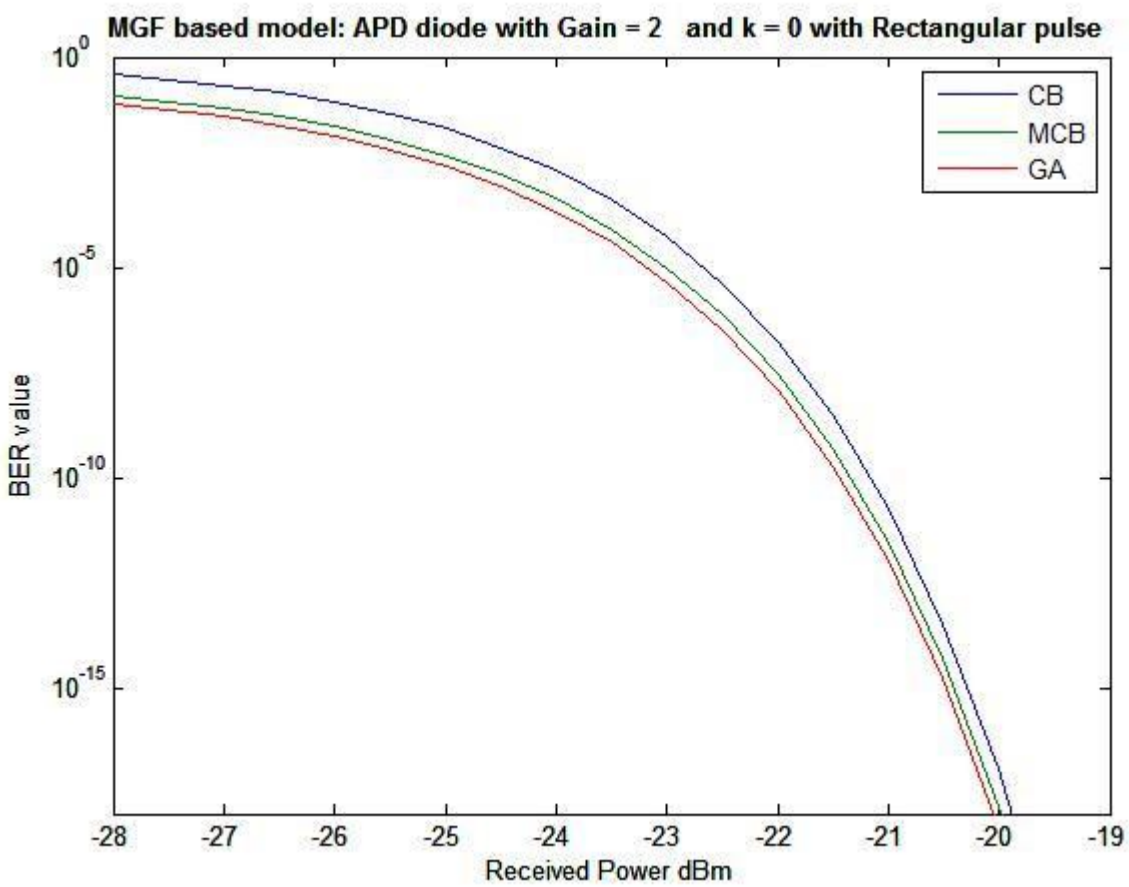

Figure 7.6 – Bound Comparison for an APD

Table 5 – APD Based receiver – BER comparison

BER	RECEIVED POWER (dBm)			DIFFERENCE FROM GA	
	GA	CB	MCB	CB	MCB
10^{-9}	-21.7000	-21.3903	-21.5892	0.3097	0.1108
10^{-12}	-20.9926	-20.7674	-20.9171	0.2252	0.0775
10^{-15}	-20.4665	-20.2807	-20.4034	0/1858	0.0631

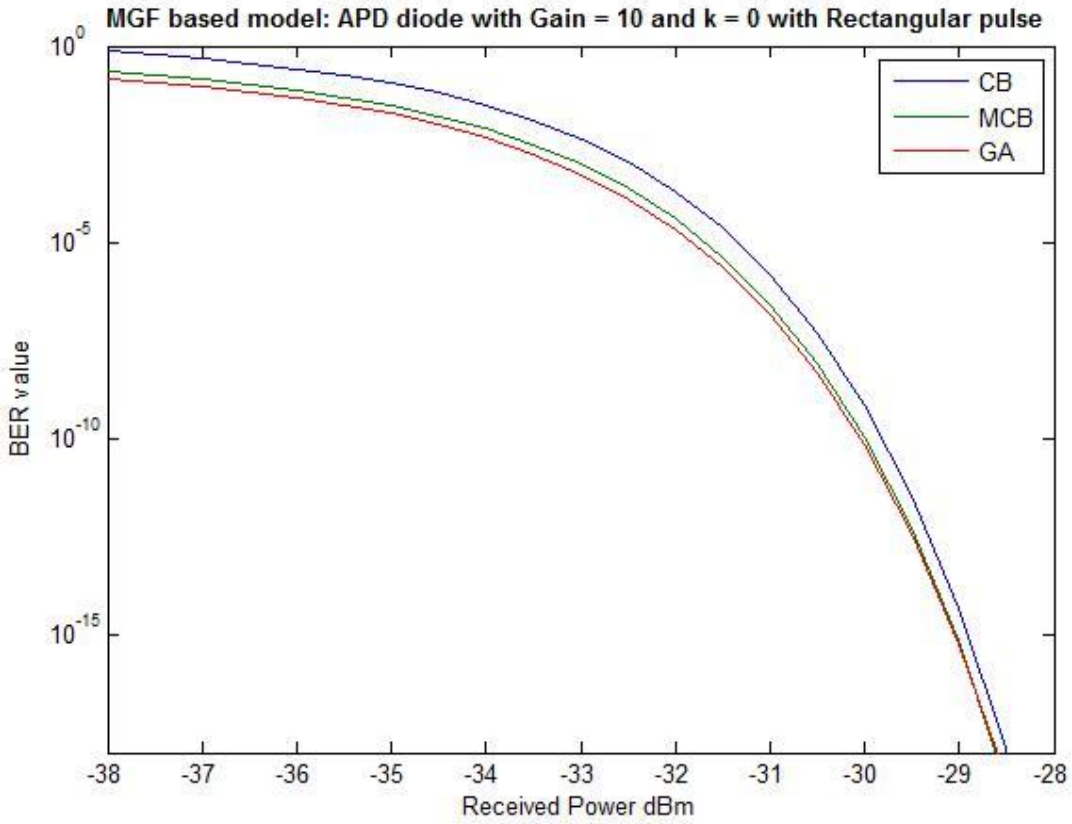

Figure 7.7 – Bound comparison for a APD (Gain=10, k=0)

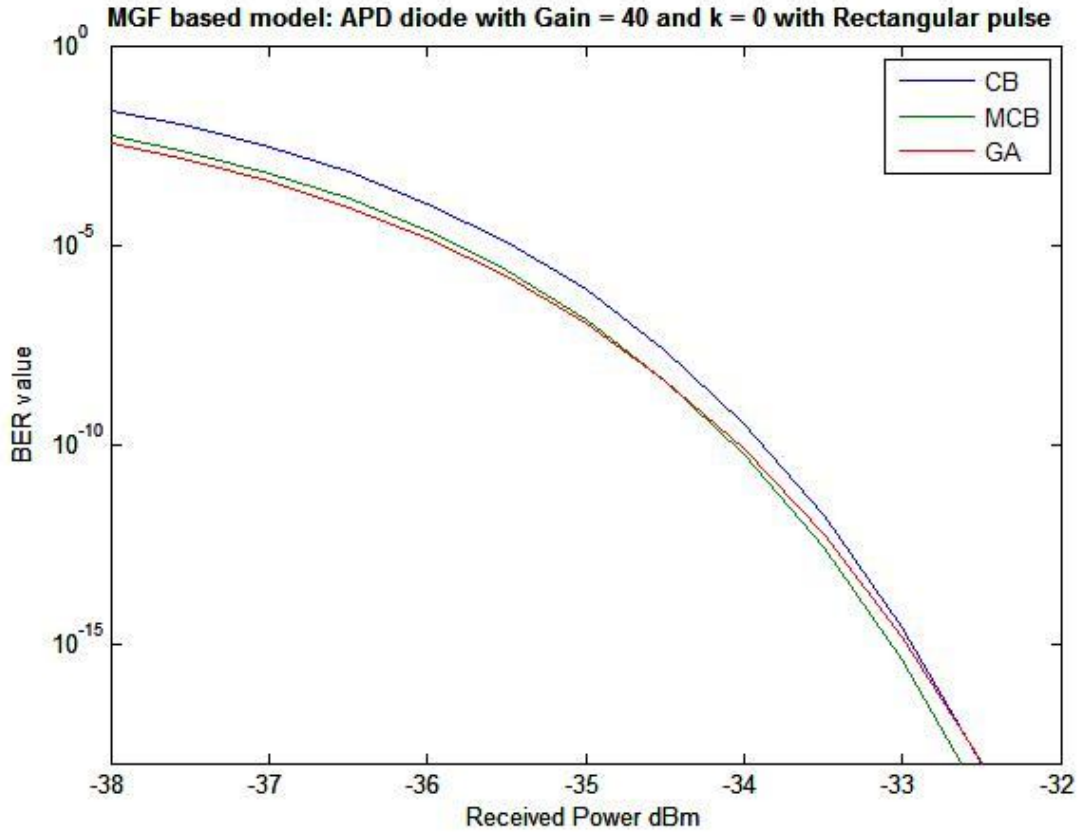

Figure 7.8 – Bound comparison for a APD (Gain=40, k=0)

8. EXTENSION WORK

8.1. Front end circuitry

As mentioned throughout this report, further noise is encountered via the specific front end circuitry that the photodiode is used in conjunction with. Figure 8.1 shows a basic receiver configuration of a photodiode output being fed into a preamplifier.

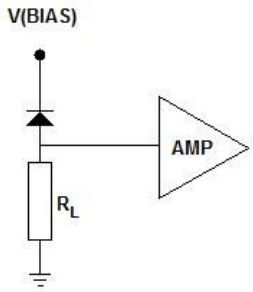

Figure 8.1 – Basic receiver circuit [17]

Typically the preamplifier would then be fed into a post amplifier, an equaliser and then a filter. As the preamplifier is the first stage that is encountered it will be a major contributor to additional noise and has an affect on the final sensitivity of the receiver. Commonly two types of transistor can be used for the amplification process the BJT and FET with both types having different advantages. Different configurations of transistor configurations can be modelled e.g. a FET in a common source configuration and then have their associated input noise currents approximated. The minimum input noise experienced by a FET is shown and results in the expression shown in equation [17].

$$\left\langle i^2 \right\rangle = 4kT\Gamma \frac{(2\pi C_T)^2}{g_m} I_3 B^3 \tag{56}$$

This could be used to provide the user with additional modelling capabilities as the values for equations can be evaluated for real life transistors to determine what noise values they can expect. In section of the values are used for that of a commonly available transistor (high frequency FET 2N4416). The total capacitance is the summation of device specific capacitances such as that of the gate/source etc. The Transconductance g_m and the noise factor Γ are device specific, whilst the term I_3 represents the current which is dictated by the chosen pulse shape. Additional modelling of the thermal and shot noise can be carried out leading to expression for the shunt resistance or leakage current. Therefore the total noise experienced by the device can is the summation of these the two noises.

Similarly modelling of a BJT based amplifier can be carried out in a similar way and an expression of the minimum noise is shown in equation.

$$\left\langle i^2 \right\rangle_{circuit,\min} = (8\pi kT)\left(\frac{C_T}{\beta_0^{1/2}}\right)(I_2 I_3)^{1/2} B^2 + 4kT\Gamma r_{b'b}(2\pi)^2 (C_d + C_s^{'})^2 I_3 B^3 \quad \textbf{(57)}$$

These equations were coded in Matlab to show how the input noise current varies with increasing bit rate. The results for both type of transistor can be seen in figure 8.2 and 8.3.

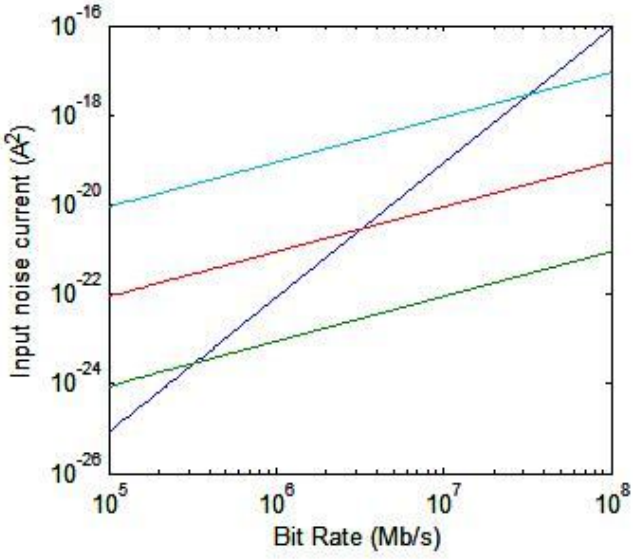

Figure 8.2 – Input noise current versus Bit rate for a FET. g_m = 6mS and Γ = 0.7 with I3 equal 0.03

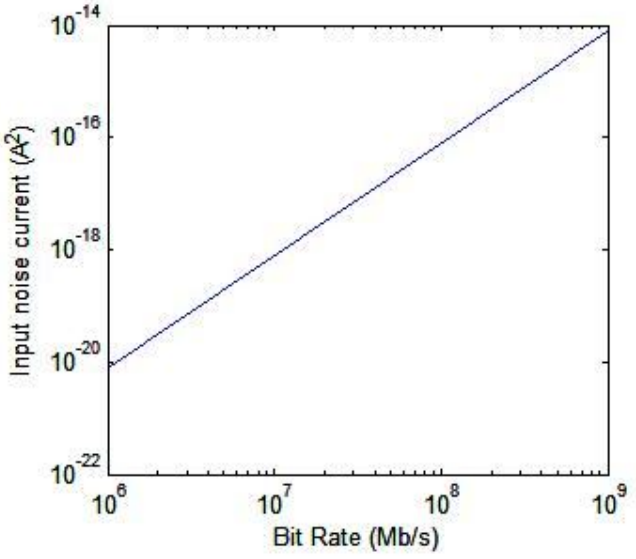

Figure 8.3 – Input noise current versus Bit rate for a BJT. g_m = 6mS and Γ = 0.7 with I3 equal 0.03, β=100, rb'b=100Ω.

Two types of configurations are common within front end receiver designs one being a high impedance amplifier and the other a transimpedance amplifier. The two configurations can also have similar expressions derived for their noise contributions.

Therefore modelling of front end receiver circuitry is an area that has the potentially to be readily incorporated into existing models giving the designer the ability to specify front end configuration with specific transistors parameters.

9. REFERENCES

- [1] Lionel Warnes - Electronic and electrical engineering – Principals and practice – 3rd Edition (2003) - Chapter 24 – Analogue Communications & Chapter 25 – Digital Communications - ISBN – 0-333-99040-4 -

- [2] Rajiv Ramaswami and Kumar N. Sivarajan - Optical networks - A practical perspective (second edition) - (2002) - Chapter 4 Modulation and demodulation – [4a] Section 4.4.6 – Bit Error Rates - Pages 258 – 263, [4b] Section 4.4.1 – 4.4.5 Pages 250 – 258. ISBN: 1-55860-655-6

- [3] Ben G.Streetman – Sanjay Kumar Banerjee - Solid State Electronic Devices (Sixth Edition 2006) – Chapter 4 – Section 4.3.1. – Direct Recombination of Electrons and Holes Page 128 -129, Chapter 8 – Optoelectronic Devices – Section 8.11. to 8.2.2 – Pages, 398 to 414, Section 8.4 .1 Page 422 – 424. - ISBN: 013149726X.

- [4] John Bird - Higher engineering mathematics – Third Edition (2002) – Chapters 44 - 46 - ISBN: 0-7506-4110-X.

- [5] K. W. Cattermole - Mathematical topics in telecommunications – Generating Functions Methods in Probability theory – Pages 1 -25 - ISBN: 1-55860-655-6

- [6] J.J O'Reilly - Mathematical topics in telecommunications – Generating Functions, Bounds and Approximations in Optical Communications – (2002) - pages 119 -133 – ISBN: 1-55860-655-6.

- [7] Charles M. Grinstead, J. Laurie Snell - Introduction to probability – Second Edition (1998) - ISBN – 0-8218 – 0749-8. Also available in electronic format from: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf. Made available by the American Mathematic Society.

- [8] Sheldon M. Ross - Introduction to Probability Models – Sixth Edition – ISBN – 0-12-598470-7. Chapter 2 – Random Variables – Section 2.6 – Moment Generating Functions.

- [9] Dr A. Phillips - H64OCP - Optical communications lecture notes and reference materials – University of Nottingham module.

- [10] Ward Cheney and David Kincaid - Numerical mathematics and computing – Fourth Edition – ISBN – 0534389937 – Chapter 14 – Minimization of multivariate functions – 14.1 – Golden Section Search Algorithm – 555 – 557.

- [11] Richard P. Brent - Algorithms for minimization without derivatives – ISBN – 0-13-022335-2 -] Chapter 5 – An Algorithm with Guaranteed Convergence for finding a minimum of a function of one variable – Pages 61 – 70.

- [12] Ivan P. Kaminow (Editor), Tingye Li (Editor) - Optical Fiber Telecommunications IV-A: Components - ISBN: 0-12-3951720. Chapter 3 – New Materials for optical amplifiers – Pages 130 - 140

- [13] Luís F. B. Ribeiro, José R. F. Da Rocha and João L. Pinto - Journal of Lightwave technology, Vol 13, No. 2, February 1995 – Performance evaluation of EDFA preamplified receivers taking into account Intersymbol interference .

- [14] John Senior - Optical Fiber Communications – Principles and practice – Second Edition – ISBN: 0-13-635426-2, Chapter 3 - Transmission characteristics of optical fibers – [10 a] Section 3.8 Dispersion Pages 102 - 106. [10b] Chapter 8 - Optical detectors – 8.5 – 8.10, Pages 426 – 454.

- [15] Rodger E. Ziemer, Roger L. Peterson - Introduction to Digital Communication – Second Edition – ISBN – 0-13896481-5. Chapter 3 – Basic Digital Communication Systems, Signalling through band limited channels, Section 3.3.2 – Designing for Zero ISI: Nyquist's Pulse shaping criterion, Pages 157 - 160

- [16] Bernard Sklar - Digital Communication – Fundamentals and Applications – Prentice Hall, , ISBN – 0-13-212713-X, Chapter 2 – Formatting and baseband transmission - Sec 2.11 – Intersymbol Interference Pages 98 – 105

- [17] R. G. Smith and S. D. Personick - "Receiver design for optical fiber communication systems" in Semiconductor devices for optical communications, Springer, Verlay, New York, 1980 - Chapter 4.

- [18] C. Gobby, Z. L. Yuan, and A. J. Shields - Quantum key distribution over 122 km of standard telecom fibre – Toshiba Research Quantum cryptography group – http://www.toshiba-europe.com/research/crl/qig/pdfs/scientific-publications/qkd_apl84_2004

- General reading

Amnon Yariv (Author), Pochi Yeh (Author) - Photonics: Optical Electronics in Modern Communication - ISBN – 0-195179463.

APPENDIX A – MOMENTS EXAMPLE

A discrete variable has a sample space consisting of distinctive countable points e.g. x_i which have an associated value of probability p_i. The expectation of this type of distribution can be calculated via equation 31:

$$E(X) = \sum x_i \, p(x_i) \quad \textbf{(1)}$$

A simple discrete example can be illustrated with the various outcomes of a dice e.g.

$$p(x_i) = P(X = x_i) = 1/6$$

I.e. it is the same for all of the faces $x_1, x_2, x_3, x_4, x_5, x_6,$

Therefore $E(X)$ = (1*1/6) + (2*1/6) + (3*1/6) + (4*1/6) + (5*1/6) + (6*1/6) = 3.5

The variance of both types of distributions can be calculated as follow with μ being the mean of expectation:

$$V(X) = E[(X-\mu)^2] \text{ And } V(X) = \sum_{i=1}^{n} p_i (x_i - \mu)^2 \quad \textbf{(2)}$$

In this case the variance is:

$$V(X) = 1/6*(1-3.5)^2 + 1/6*(2-3.5)^2 + 1/6*(3-3.5)^2 + 1/6*(4-3.5)^2 + 1/6*(5-3.5)^2 + 1/6(6-3.5)^2 =$$

2.9167 to 4 d.p

The moment generating function allows the generation of the moments of a distribution by use of its probability density or mass function. The moments that can be generated are often useful statistical indicators such as the mean (the 1st moment) and the variance (derived from the 2nd moment).
For example the moments of a distribution can be calculated via the following expression:

$$\text{The r}^{\text{th}} \text{ moment of a distribution is } \mu_r = \sum_{i=1}^{\infty} (x_i)^r (p_i) \quad \textbf{(3)}$$

Thus for the dice scenario the first moment is given by $\mu_1 = \sum_{i=1}^{\infty} (x_i)(p_i)$ being equal to equation 1.

This result therefore gives the expectance or mean (μ) of the distribution

The second moment is given by $\mu_2 = \sum_{i=1}^{\infty} (x_i)^2 (p_i)$ and this can then be used to calculate the variance e.g. V(X) $= \sigma^2 = = \mu_2 - \mu_1{}^2$

Substituting the appropriate values into the above gives $\mu_1 = 3.5$ and $\mu_2 - \mu_1{}^2 = 2.9176$ to 4 d.p. The above is used to simple highlight how the moments contain the statistical information that describes the probability distribution.

APPENDIX B – PIN DIODE MGF DERIVATION

Firstly it the instantaneous output current from the device was modelled as a random variable

$$X(t) = q\sum_i h(t - t_i) \tag{1}$$

With a generated electron (at t_i) being represented as a impulse $h(t-t_i)$ and with 'q' as the charge of an electron.

By viewing the electrons within a time $\Delta\tau$ centred on a time τ, with generations at times t_i can be seen as a Poisson process with mean value of $\rho\Delta\tau$. Now the contribution each electron makes to the current is given by:

$$M_X(s) = e^{sqh(t-\tau)} \tag{2}$$

This is the contribution that one electron makes to the output current.

Now a second MGF can be determined to represent the count (C) of the photoelectrons generated within $\Delta\tau$. This again is a Poisson distribution and can be represented by:

$$E\left\{e^{sC}\right\} = \sum_{r=0}^{\infty} \frac{e^{-\rho\Delta\tau}(\rho\Delta\tau)^r}{r!} e^{sr} \quad \therefore \text{ Current contribution within } \Delta t \text{ is:}$$

$$X_{\Delta\tau}(t) = Cqh(t-\tau)$$

Leading to the MGF:

$$E\left\{e^{sCqh(t-\tau)}\right\} = \sum_{r=0}^{\infty} \frac{e^{-\rho\Delta\tau}(\rho\Delta\tau)^r}{r!} e^{sqrh(t-\tau)} = \exp(\rho\Delta\tau(e^{sqh(t-\tau)} - 1))$$

$$\text{Achieved using the definition } e^x = \sum_{r=0}^{\infty} \frac{x^r}{r!}$$

Now the final step is involved considering these counts of electrons occurring in discrete intervals denoted by the term j. Ultimately the output current will be an expression involving the summation of these intervals.

$$M_X(s) = E\left\{e^{s\sum_j X_{\Delta\tau_j}(t)}\right\} \quad = \quad \exp\sum_j \rho_j (e^{sqh(t-\tau)})\Delta\tau \tag{3}$$

A final step can then be taken by applying limits decreasing the value of t towards zero, *j* towards infinity and $\rho\Delta\tau$. This means that the summation transforms into an integral with appropriate limits.

$$M_{X(t)}(s) = \exp\left\{\int_{-\infty}^{\infty} \rho(\tau)(e^{sqh(t-\tau)} - 1)d\tau\right\} \qquad j \to \infty, \tau \to 0, \rho_j \to \rho(\tau) \qquad \textbf{(4)}$$

APPENDIX C – NOISE COMPARISON

Simple expressions were derived to help justify the result produced by model for certain situations. For example Q factor for the APD can be expressed as in equation with the relevant power and noise contributions.

$$Q_{APD} = \frac{i_1 - i_0}{\sqrt{\sigma_{th}^2 + \sigma_{1Sh}^2} + \sqrt{\sigma_{th}^2 + \sigma_{0Sh}^2}} = \frac{RM(P_1 - P_0)}{\sqrt{\sigma_{th}^2 + 2qRM^2F(M)P_1Be} + \sqrt{\sigma_{th}^2 + 2qRM^2F(M)P_0Be}}$$

$$= \frac{RM(P_1 - P_0)}{\sqrt{\frac{\sigma_{th}^2}{(RM)^2} + \frac{2q}{R}F(M)P_1Be} + \sqrt{\frac{\sigma_{th}^2}{(RM)^2} + \frac{2q}{R}F(M)P_0Be}}$$

The Q of the APD is equal to Q of that of a PIN when the Gain factor is equal to 1 e.g. M=1

$$= \frac{(P_1 - P_0)}{\sqrt{\frac{\sigma_{th}^2}{R^2} + \frac{2q}{R}P_1Be} + \sqrt{\frac{\sigma_{th}^2}{R^2} + \frac{2q}{R}P_0Be}}$$

Then the two devices can be compared directly to one another via there expressions for their relevant Q's. e.g.

$$\frac{Q_{APD}}{Q_{PIN}} = \frac{\sqrt{\frac{\sigma_{th}^2}{R^2} + \frac{2qRP_1Be}{R}} + \sqrt{\frac{\sigma_{th}^2}{R^2} + \frac{2qRP_0Be}{R}}}{\sqrt{\frac{\sigma_{th}^2}{(RM)^2} + \frac{2qF(M)P_1Be}{R}} + \sqrt{\frac{\sigma_{th}^2}{(RM^{2)}} + \frac{2qF(M)P_0Be}{R}}}$$

$$\frac{Q_{APD}}{Q_{PIN}} = \frac{\sqrt{\sigma_{th}^2 + 2qRP_1Be} + \sqrt{\sigma_{th}^2 + 2qRP_0Be}}{\sqrt{\frac{\sigma_{th}^2}{M^2} + 2qRF(M)P_1Be} + \sqrt{\frac{\sigma_{th}^2}{M^2} + 2qRF(M)P_0Be}}$$

APPENDIX D - QUANTUM CRYPTOGRAPHIC KEY DISTRIBUTION

A current area of research is the practical application of Quantum cryptographic principles to existing networks technologies [18]. One method of implementing quantum cryptography is to encode single photons according to an encryption protocol by changing their polarisation. Successful systems have been developed which can transmit information encrypted by a QC scheme via standard fibre optic networks up to kilometres. The difficult with encoding information via single photons is that they can be absorbed by the fibre or fail to be detected by an optical receiver. The whole system produces successful transmission and detections with various degrees if success. Obviously this has a direct affect on the bandwidth and capacity of the system and means current system performance is dictated by the probability of these transmissions successfully occurring. A typical performance measure for these systems is to similarly quote a QBER or Quantum Bit Error Rate.

The security of this method is based on the concept of the Vernam cipher which provides a mathematically secure cryptographic algorithm. Currently a method is available to implement this system by using alternating polarisation schemes to transmit the keys securely.

For example user A wants to send something to B. The first step would be for A to send a random sequence of photons, polarised with randomly alternated schemes i.e. randomly vertical or horizontal schemes 1 and 2 in figure 1. B could measure the incoming photon with either one of the two detectors/filters and has a half chance of detecting the correct polarisation. A can confirm correct detections with B, but this does not give any information on the polarisation of the photon thus this information is secure from the any eavesdropper. A and B can then use these correctly detected photon's as a one time pad to encrypt future transmission. A then send some test photons which are encrypted with the one time pad, and if the B decrypts them correct the one pad has been not been tampered with and is perfectly secure.

It is very important to understand that any attempt to detect the polarisation of the photons with the incorrect detector will affect the polarised photons that arrive at B. This will reproduce error in the communication when a test is tried with the one time pad. This is related back to the randomness idea with the Vernam cipher that if the probability is equal to 0.5 they will be equal weighting of correct and incorrect interceptions by the eavesdropper, thus voiding the one time pad.

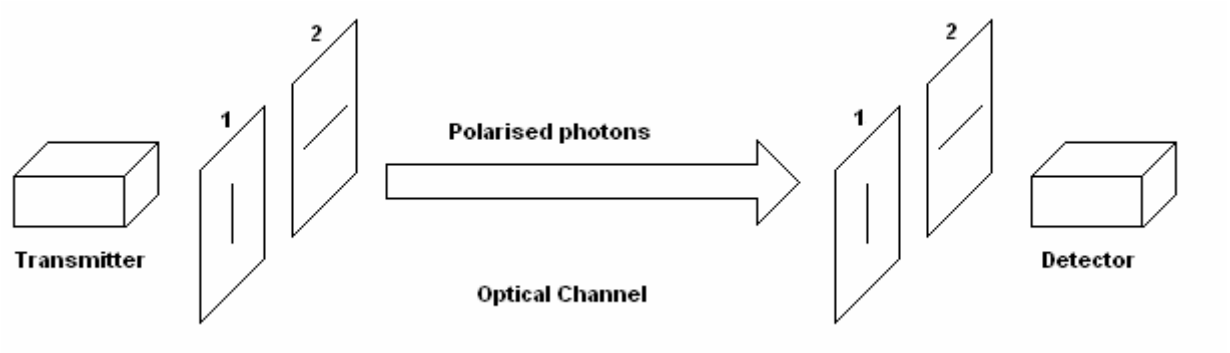

Figure 1 – Theoretical QC system

The ability to model the detection of a single photon via the above system might provide useful performance estimates for working systems. The effects if noise would have a significant impact on the detection of single photons and therefore the methods used for this project may be of some use when trying to determine the probability of single photon detection. This may involve generating probability distributions for the models and this was looked at briefly as possible extension work for this project.