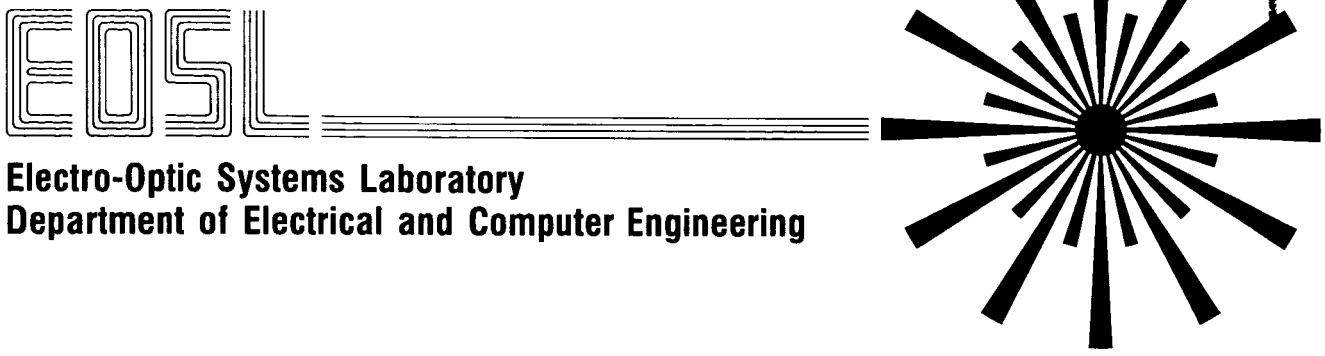


Electro-Optic Systems Laboratory
Department of Electrical and Computer Engineering

Comparison of Direct and Heterodyne Detection Optical Intersatellite Communication Links

C. C. Chen • C. S. Gardner

EOSL No. 87-002 • March 1987

University of Illinois at Urbana-Champaign

UIUL-ENG-87-2548

COMPARISON OF DIRECTION AND HETERODYNE DETECTION OPTICAL INTERSATELLITE COMMUNICATION LINKS

by

C. C. Chen
C. S. Gardner

EOSL No. 87-002

Technical Report
March 1987

Supported by
Contract No. NASA NSG-5049

NATIONAL AERONAUTICS & SPACE ADMINISTRATION
Goddard Space Flight Center
Greenbelt, Maryland 20771

ELECTRO-OPTIC SYSTEMS LABORATORY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
COLLEGE OF ENGINEERING
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

Abstract

The performance of direct and heterodyne detection optical intersatellite communication links are evaluated and compared. It is shown that the performance of optical links is very sensitive to the pointing and tracking errors at the transmitter and receiver. In the presence of random pointing and tracking errors, optimal antenna gains exist that will minimize the required transmitter power. In addition to limiting the antenna gains, random pointing and tracking errors also impose a power penalty in the link budget. This power penalty is between 1.6 to 3 dB for a direct detection QPPM link, and 3 to 5 dB for a heterodyne QFSK system. For heterodyne systems, the carrier phase noise presents another major factor of performance degradation that must be considered. In contrast, the loss due to synchronization error is small.

The link budgets for direct and heterodyne detection systems are evaluated. It is shown that, for systems with large pointing and tracking errors, the link budget is dominated by the spatial tracking error, and the direct detection system shows a superior performance because it is less sensitive to the spatial tracking error. On the other hand, for systems with small pointing and tracking jitters, the antenna gains are in general limited by the launch cost, and suboptimal antenna gains are often used in practice. In which case, the heterodyne system has a slightly higher power margin because of the higher receiver sensitivity.

TABLE OF CONTENTS

1. INTRODUCTION

Intersatellite links (ISLs) between satellites in low earth orbits (LEOs) and the geosynchronous orbit (GEO) are essential for future space missions [1],[2]. An effective ISL can reduce the dependence on ground tracking and data relay stations, increase the flexibility in choosing satellite orbits, and improve the data rate. Recently, with the advances in semiconductor laser and detector technologies, there has been considerable interest in developing an optical ISL as an alternative to microwave systems. Compared to the state-of-the-art microwave link, the optical system offers higher bandwidth and lower power requirements. In addition, the shorter optical wavelength implies smaller antenna diameters and hence a smaller size and weight of the communication system. The combination of these advantages makes the optical system highly attractive for ISL implementations.

Optical links can be implemented using either direct or heterodyne detection receivers [1]-[8]. In a direct detection receiver [3], the received optical intensity is detected without extensive front-end optical processing. The direct detection system is structurally simple and easy to implement. The technology needed to implement the direct detection receiver has been studied for land-based optical fiber links [9],[10] and, more recently, for free space communication channels [4], [5], [11]-[19]. For the free space optical link, pulse-position modulation (PPM) is known to be the optimal modulation scheme [11], [12]. The performance of the PPM channel has been analyzed extensively [3],[4], [11]-[14], and progresses have been toward implementing a PPM channel capable of communicating at 2.5 bits per detected photon [17]-[20]. In order to achieve the detection sensitivity needed for the ISL, avalanche photodiodes (APDs) [21], [22] with sufficiently high gain and low noise are required. The performance of APD-based direct detection channels has been studied for both fiber and free-space communication links [4], [5], [23]-[25].

Unlike the direct detection receiver which detects only the intensity fluctuation, the heterodyne receiver detects both amplitude and phase of the signal by first mixing the the received optical signal with a strong local oscillator (LO) output, then detecting the combined optical signal. This heterodyne process effectively shifts the spectrum of the incoming signal from the optical frequency down to an intermediate frequency (IF) where it can be easily processed. Because of the large IF signal strength, detector thermal noise can often be ignored in the heterodyne receiver. Furthermore, because the frequency filtering is carried out at the IF frequency

where a much narrower bandwidth can be used, noncoherent background radiation can be better rejected by the heterodyne receiver. As a result, the heterodyne detection receiver offers a near quantum limited detection sensitivity [3], [26], [27]. Receiver sensitivity can be further improved by using frequency and phase modulations. In the limit of large LO strength, the LO shot noise at the output of the heterodyne receiver can be approximated by a white Gaussian noise. Therefore, the analysis of the optical heterodyne system is similar to that of an additive white Gaussian noise (AWGN) channel [28]-[31]. The application of heterodyne technology in fiber-optic links has been studied [32]-[37], and the use of heterodyne receivers in ISL applications has also been proposed [6]-[8]. These studies suggest that the heterodyne system offers a 5-10 dB advantage in the detection sensitivity over the comparable direct detection PPM channel [6]-[8]. However, the heterodyne receiver is more complicated than the direct detection receiver. Furthermore, the performance of the heterodyne receiver is very sensitive to the carrier phase noise which is prominent in semiconductor lasers [38]-[43]. The impact of the laser phase noise on the performance of the heterodyne system has been studied by various authors [44]-[53]. Because of the difficulties in analyzing the statistics of the phase noise, however, most of these studies offer only an approximation to the system error rate.

Both direct and heterodyne detection systems have been proposed for ISL implementations. The purposes of this study are to compare the performance of these systems under a realistic operating environment and to analyze the various design trade-offs.

Before a comparison can be carried out between the direct and heterodyne detection ISLs, performance of individual systems must be carefully studied. In Chapter 2, components of an optical channel are discussed individually and the signal and background powers detected by the receiver are derived as functions of link parameters. Given the receiver SNR and the modulation format, the performance of both direct detection and heterodyne receivers will be studied in Chapters 3 and 4, respectively. For direct detection systems, it is shown that the PPM system offers a significant performance advantage over the on-off-keying (OOK) system [3], [11], [12]. For heterodyne systems, the sensitivity towards carrier phase noise suggests that the noncoherent frequency-shift keying (NCFSK) scheme should be used for the heterodyne ISL [6]-[8].

In addition to the receiver SNR and the modulation scheme, the performance of the ISL is very sensitive to the spatial and temporal tracking errors between the two satellites. Both direct detection and heterodyne systems are susceptible to tracking errors in the mutual line-of-sight (LOS). The implementation of the spatial

tracking and acquisition system has been studied by various groups [54]-[59]. The impact of spatial tracking error on the performance of direct detection ISLs has been analyzed in terms of the probability of burst errors [57], [60]-[63]. The use of burst error probability to analyze the design of optical ISLs is simple, straightforward, and in most cases provides adequate design rules. However, it is of interest to derive an optimal system design which maximizes the overall system performance. Such an analysis is carried out in Chapter 5. It is shown that, in the absence of spatial tracking error, the performance of the ISL improves with increasing transmitter and receiver antenna gains. When a random pointing error is present in the ISL, however, an optimal transmitter antenna gain exists which minimizes the required signal power. When pointing errors in each of the two gimbal axes can be modeled as identically distributed, zero-mean, Gaussian random variables, the product of the optimal transmitter antenna gain and the mean square pointing jitters is shown to be a constant which is independent of the RMS pointing error. Similarly, for heterodyne systems, the presence of random tracking error implies that an optimal receiver antenna gain exists which optimizes the receiver SNR. The product of this optimal receiver antenna gain and the mean square tracking error is again a constant which is independent of the RMS tracking error. In addition to limiting the antenna gains, random pointing and tracking errors also impose an additional power penalty in the system link budget. This power penalty is shown to be between 1.5 to 3 dB for a direct detection QPPM system, and between 3.8 to 5 dB for a heterodyne QFSK system.

Receiver timing synchronization is another major factor of performance degradation. For direct detection PPM systems, early-late gate [3], [28], [31] and phase-locked loop (PLL) [28]-[31], [64]-[68] timing synchronizers can be used to synchronize the receiver timing. It can be shown that, with nonlinear preprocessing, PLLs can be used to achieve effective symbol synchronization of the direct detection PPM receiver [66]. For heterodyne systems, techniques for synchronizing the RF channel can be applied without extensive modification. The impact of symbol synchronization on the performance of direct and heterodyne detection ISLs is studied in Chapter 6. It is seen that, with a narrow tracking loop bandwidth, receiver synchronization can be implemented with only a small power penalty.

Finally, Chapter 7 summarizes the findings of previous chapters and analyzes the various design tradeoffs. It is shown that, for systems with large (>0.5 μrad) RMS pointing and tracking jitters, spatial tracking errors are the dominating loss factor in the link budget. As a result, the direct detection PPM system

has a higher power margin because it is less sensitive to the spatial tracking error. On the other hand, for systems with small (<0.2 μrad) pointing and tracking jitters, losses due to spatial tracking errors are small and the heterodyne system actually shows a superior performance because of the higher receiver sensitivity.

2. THE OPTICAL LINK

Like most communication links, the performance of optical links depends critically on achieving high signal-to-noise ratio (SNR) at the receiver. The higher the receiver SNR, the lower the error rate. Unlike the RF receiver, however, the output of the optical detector is contaminated by the shot noise in addition to the thermal noise. Given the received signal power, the presence of the shot noise imposes a limit on the maximum SNR that can be achieved by the detector. This shot noise limited SNR dictates a lower bound on the performance of the optical link. In this chapter, we are interested in evaluating the detector SNR as a function of link parameters. In addition, expressions for the signal and background power received by the receiver will be derived and evaluated.

2.1 Optical Detectors and Noise

Detection of optical radiation differs from conventional RF field detection in that the signal intensity, not the amplitude, is detected. Detectors commonly used in optical communication include photomultiplier tubes (PMT), *pin* photodiodes, and avalanche photodiodes (APD). *Pin* diodes and APDs are solid state devices, while PMTs are vacuum photoemissive devices.

A typical PMT consists of a photocathode, followed by a series of dynodes and an anode. An incoming photon event is detected by the generation of a photoelectron at the photocathode. This photoelectron is accelerated toward the first dynode in the tube by the applied field. The impact of this accelerated electron on the dynode creates secondary electrons that are then accelerated by the external field toward the next dynode. This process is repeated until photoelectrons hit the anode. However, by this time the original photoelectron has been amplified considerably so that a very high current gain is associated with a single photon event. Gains of 10^5 to 10^6 are typical for PMTs. When operating at such high gains, the detector thermal noise and gain noise usually can be ignored. Consequently, PMTs are essentially quantum limited devices. However, because of the time it takes the electrons to traverse the dynode chain, PMTs have an inherent bandwidth limitation. Typical bandwidths for PMTs are around 20-200MHz. Such bandwidths are considerably lower than the desired information bandwidth. Furthermore, the lack of suitable materials for photocathodes at the high visible and near IR wavelengths is a significant disadvantage. The detector quantum efficiency, which is the ratio between the number of detected and incoming photons, is typically on the order of 10-20% at visible

wavelengths, and drops off rapidly in the near infrared region [69], [70]. In addition to being slow and inefficient, PMTs are bulky, and require a high power supply voltage ($\approx 1\text{–}3KV$) and extensive thermal cooling.

In contrast, solid state detectors have an inherently higher bandwidth because of their smaller size. In addition, solid state devices are more reliable and exhibit higher quantum efficiency. Typical quantum efficiencies for solid state detectors range from 20 to 90% throughout the visible and near IR ranges [69]-[72]. Solid state detectors commonly used include the *pin* diode and the APD. *Pin* diodes are simpler to operate and easier to fabricate than APDs [21]. Under a normal operating condition, the *p-n* junction is reverse biased so that in the absence of light only a small leakage current will flow through the diode. When a photon enters the depletion region, it is absorbed and the excess energy generates an electron-hole pair. Both electron and hole are rapidly drawn toward the opposite electrodes where a current pulse is generated. The generation of electron-hole pairs depends on the detailed interaction between the radiation field and the detector material. For detection of multimode radiation, photoelectron generation can be modeled as a Poisson arrival process with arrival rate $\lambda(t)$, which can be related to the total incident optical power P_R by [3], [74]

$$\lambda(t) = \frac{\eta P_R(t)}{h\nu} \quad , \tag{2.1}$$

where η is the detector quantum efficiency, and $h\nu$ is the energy of the photon. The output of the photodetector, which can be modeled as the superposition of detector response to the primary photocurrent, can therefore be written as [3], [74]

$$i(t) = \sum_{\{\tau_j\}} eh(t-\tau_j) + i_d + i_{th} \quad , \tag{2.2}$$

where the summation is extended over the photon arrival times $\{\tau_j\}$, e is the electron charge, $h(t)$ is the detector impulse response, i_d is the detector dark current, and i_{th} is the detector thermal noise current. Because there is no internal gain mechanism, the bandwidth of the *pin* diode is limited only by the time it takes the electron and hole to travel through the depletion layer. Bandwidths of 10GHz can be easily achieved for InGaAs *pin* diodes [21] and bandwidths as high as 20 GHz have been reported [73].

The main disadvantage of *pin* diodes in optical ISLs is their low gain. When operating at extremely low signal levels, thermal noise current generated by the detector front-end can be more significant than the signal

current. This thermal noise limited situation can severely degrade the signal-to-noise ratio of the detector and increase the system error rate. Consequently, *pin* diodes are typically used in heterodyne systems where the large local oscillator (LO) strength ensures that the SNR is not limited by the detector thermal noise. For applications where the incident optical power on the detector is weak , such as in a direct detection ISL receiver, the APDs with their internal gain are generally preferred.

Avalanche photodiodes are similar to *pin* diodes in that they operate under reverse bias [21], [22], [71], [72]. Unlike *pin* diodes, however, APDs operate with high internal electric fields so that when carriers are generated, they are accelerated to high velocities. These high velocity carriers can subsequently ionize other electron-hole pairs which in turn generate additional carriers. This impact ionization mechanism gives the APD an internal gain so that the device can operate above the detector thermal noise limit. Typical gains of APDs range from 100 to 300 and, because the ionization process is random, gains associated with different photon events are independent. By assuming that each primary electron-hole pair is associated with a random gain G, the output of the APD can be modeled as [3], [75]

$$i(t) = \sum_{\{\tau_j\}} eG_j h(t-\tau_j) + i_d + i_{th} \quad , \tag{2.3}$$

where G_j is the gain associated with the j^{th} photoevent.

Because of the large gain (100-300) associated with each photon event, detector thermal noise can often be ignored at the APD output. The principal noises at the APD output are, therefore, the detector shot noise and noise associated with the avalanche gain mechanism. The avalanche noise is lowest in devices where the avalanche process is generated primarily by the carriers of higher mobility, because although a large number of secondary carriers can be generated by the primary electron-hole pair, the higher carrier mobility implies that carriers will spend a shorter time within the avalanche region and, consequently, a smaller number of secondary carriers will induce further ionization. The ratio between the hole and electron ionization coefficients is known as the effective ionization coefficient ratio, k_{eff}. For a well-designed APD, k_{eff} should be much smaller than 1. The value of k_{eff} as low as 0.007 has been reported for selected commercial devices.

The exact gain distribution of the APD is difficult to characterize [75]-[77]. However, for most practical applications, it is sufficient to characterize the APD by its excess noise factor F, defined as

$$F = \frac{<G_j^2>}{<G_j>^2} \ , \tag{2.4}$$

where the angle brackets $<G_j>$ denote the ensemble average. Experimentally, the excess noise factor is found to be closely approximated by [72]

$$F = G^x \ , \tag{2.5}$$

where G is the average gain of the detector, and x is some exponent which for Si APD is between 0.2 and 0.5[1]. The excess noise factor F can also be theoretically related to k_{eff} by [72]

$$F = k_{\text{eff}}G + (1-k_{\text{eff}})(2-\frac{1}{G}) \ . \tag{2.6}$$

The response time of the APD also depends on k_{eff}. After the initial avalanche, the signal persists until no further avalanche is detected. In the best case where k_{eff}=0, the detector response time is limited only by the time it takes for a single electron to travel through the avalanche region. Typical bandwidths of APD range from 100 MHz to 4 GHz for visible wavelengths. At wavelengths longer than 1 μm, the difficulties of obtaining a low bandgap material with small tunnel leakage have limited the use of APDs. However, recent progresses in manufacturing the separate-absorption-and-multiplication-region APD (SAM-APD) have demonstrated a bandwidth of 5-7GHz with a gain-bandwidth product exceeding 50-GHz [79]. At even longer wavelengths ($\simeq 10\mu m$), multiple-quantum well (MQW) APDs [80] can be used to detect the signal effectively.

Given the incident optical power, the mean and variance of the photodetector output can be given by Campbell's theorem [73]:

$$E[i(t)] = eG\lambda(t)*h(t) + Gi_b + i_s \ , \tag{2.7}$$

$$Var[i(t)\,|\,\lambda(t)] = <G^2>e^2\lambda(t)*h^2(t) + (<G^2>i_b+i_s)*h^2(t) + \sigma_{th}^2 \ . \tag{2.8}$$

In writing Eqs.(2.7) and (2.8), it has been assumed that the detector dark current consists of a gain dependent part i_b and a gain independent portion i_s. From Eq.(2.8), it is seen that the variance of APD output current consists of three terms. The first term on the right-handed side of Eq.(2.8) is due to the signal shot and the avalanche noise, the second term is due to the detector dark current, and the third term is the contribution of the thermal noise. The thermal noise term in Eq.(2.8) can be written as

[1]For $10 \le G \le 100$, $x=0.3$ is a good approximation.

$$\sigma_{th}^2 = \frac{2k_B T_{eq} B}{R_L} \quad , \tag{2.9}$$

where k_B is the Boltzmann's constant, R_L is the detector resistance, T_{eq} is the equivalent noise temperature of the detector, and B is the (two sided) detector bandwidth (Hertz) which can be related to the detector impulse response $h(t)$ by

$$B = \int_{-\infty}^{\infty} h^2(t)dt \quad . \tag{2.10}$$

Because the detector bandwidth is usually much higher than that of the signal bandwidth, the convolutions in Eqs. (2.7) and (2.8) can be approximated by

$$\begin{aligned} &\lambda(t) * h(t) \simeq \lambda(t) \quad , \\ &\lambda(t) * h^2(t) \simeq \lambda(t) \cdot \int h^2(t)dt = \lambda(t)B \quad . \end{aligned} \tag{2.11}$$

The detector SNR, which can be defined as the ratio between signal and noise power, is therefore [3]

$$SNR_{det} = \frac{(eG\eta P_R / h\nu)^2}{[G^2 F(\frac{e^2 \eta P_R}{h\nu} + i_b) + i_s + \frac{2k_B T}{R_L}]B} \quad , \tag{2.12}$$

where we have substituted the mean square of G by the product of the excess noise factor F and G^2.

2.2 Optical Sources

Ordinary incoherent sources, because of the large beam divergence, are not suitable for long-range optical links. The output of lasers, on the other hand, is both spatially and temporally coherent. At large distances, the use of lasers permits higher signal power to be focused on the detector surface.

Among the possible contenders for ISL laser transmitters are the CO_2 lasers [81], [82], solid state Nd:YAG lasers [83]-[85], and semiconductor lasers based on GaAs-AlGaAs [5], [13], [20]. The He-Ne laser, although well developed, is not suitable for the ISL application because of its low efficiency (0.01%) and small available power output.

Early efforts in developing optical ISL lasers concentrated on the CO_2 laser operating at 10.6 μm [81] [82]. Researchers have demonstrated CO_2 lasers capable of delivering several watts at 10% efficiency and a lifetime expectancy exceeding 50,000 hours [82]. Because of the lack of proper detectors at 10.6 μm wavelength, heterodyne receivers must be used in CO_2 systems. Electro-optic modulation techniques have

been developed to modulate the laser at high data rates. Recent results have shown that CdTe modulators capable of PSK modulation at bandwidths as high as 5 GHz can be manufactured [82]. The modulator produces 1.3 W of averaged modulated sideband power with 6.2 W of input power [82]. However, CO_2 systems are no longer being seriously considered because of the large diffraction loss associated with the long wavelength and the complexity of heterodyne receiver needed to detect the long-wavelength radiation. In addition, the use of a gas laser operating in space poses additional problems for both reliability and system design.

An alternative to the CO_2 laser is the solid state Nd:YAG laser operating at 1064 nm and frequency doubled, 532 nm, wavelengths [2], [83]-[85]. Direct detection laser communication links based on Nd:YAG lasers can be easily constructed because of high detector sensitivities at these wavelengths. In addition, mode-locked Nd:YAG lasers can be constructed with high modulation rates. One serious disadvantage of the Nd:YAG laser is the required external pump source. This results in considerably lower efficiencies than that for the CO_2 laser and lower overall power output. In addition, to compensate for the lower output power, a narrower transmitter beamwidth must be used which in turn demands a more stringent requirement in tracking and pointing of the telescopes. Pumping sources for Nd:YAG lasers include lamp pumping with low pressure potassium rubidium (KRb) arc lamps, solar pumping, and diode pumping with GaAs lasers [2], [83], [85]. Flash lamp pumping has been demonstrated to have a 0.1% efficiency which produces a laser power output of 270 mW with 250 W of electrical input power [2], [83]. Solar pumping has also been demonstrated and the results indicate that over 400 mW of average power can be produced with a 60 cm solar collector [2], [83]. For satellite applications, diode pumping is by far the most attractive technique for pumping Nd:YAG lasers. By doping the GaAs injection laser diode, the device can be made to lase at 810 nm which results in an excellent match to the 810 nm Nd:YAG absorption band [2], [83], [85]. A power efficiency of 0.6% has been demonstrated with only 40 W of supply power needed to generate 270 mW of output power [2], [83]. Recently, by using a tightly focused end-pump geometry, pumping efficiency as high as 8% had been achieved which resulted in 80 mW CW power output with only 1 W of electrical power into a single semiconductor laser array pump [85]. In addition to the required external pump source, Nd:YAG lasers also suffer from the requirement of an external modulator which can severely attenuate the signal power and consume more prime electric power.

With the improvements in device efficiency and reliability in recent years, there has been considerable interest in developing a semiconductor laser based optical ISL as an alternative to Nd:YAG or CO_2 systems. Semiconductor devices have high power efficiency and offer room temperature operation without extensive cooling requirements. However, smaller output power is available from semiconductor devices. The sensitivity of lasing frequency with respect to junction temperature and injection current density implies that precise control over temperature and current must be maintained. Furthermore, the variation of index of refraction in the lasing cavity with varying injection current density can result in a large carrier phase fluctuation which can seriously deteriorate the performance of the heterodyne channel. Direct detection systems employing GaAs laser diodes are insensitive to the phase fluctuations, and have been thoroughly investigated for land-based fiber optics links [10], [86]. For a spaceborne, direct detection optical communication system, laser power on the order of a hundred milliwatts is necessary. Commercial laser diodes are currently available with output power up to 30 mW and product lifetime of approximately 20,000 hours. These devices generally emit in a single spatial mode but provide very little control over the lasing frequency, which depends on the the drive current and temperature [86]. Consequently, such devices are suitable only for direct detection system implementations. Higher laser power (several hundred milliwatts in CW operation) have been reported in laboratory works [87]-[90]. The modulation bandwidth of the laser diode is another important factor in choosing a laser diode. High speed (>10 GHz) [91] operation of the diode laser has been reported. However, it should be noted that laser diodes designed for high speed operation generally have a small output power. In the near future, it is expected that, with improved fabrication technologies such as molecular beam epitaxy (MBE) and metal-organic chemical vapor deposition (MOCVD), better uniformity can be achieved for large structures. Multiple quantum-well (MQW) lasers [92], [93] grown with these techniques should be more efficient than the currently available devices. A uniform fabrication process will also improve device yield and reduce the production cost in addition to offering improved performance. With the inclusion of these technologies, it is expected that single laser diodes with 60-100 mW CW output power level can be reliably manufactured within the next few years.

With the output power of a single laser diode limited, further improvement in the transmitter laser power can be achieved using a phase-locked diode array, or by noncoherently combining the power from several different wavelength sources. The phase-locked laser diode array is currently at its infancy. Lasing power as

high as several watts has been achieved for laboratory devices [94], [95]. However, devices that are presently available suffer from poor spatial beam quality as the lasing mode in the diode array is dominated by the higher order supermodes that are characterized by the double-lobed intensity pattern at the far-field. Several techniques can be used to suppress the higher order supermodes and obtain a single-lobed far-field pattern. One such technique is to injection lock the laser diode array with a coherent injection source. This method has been successfully demonstrated to generate 315 mW of output power using a 20-element GaAlAs gain-guided coupled-strip array with 3.9 mW of incident power [96]. Another technique is to tailor the gain or stripe spacing of the array structure in order to suppress the higher order supermodes [97]. Diffraction coupling [98] and holographic grating [99] can also be used to channel the output of individual lasers into one central far-field pattern. In addition to the linear stripe array, the Y-junction array [100], [101] has also been proposed in which the lowest order mode is selected by the interference between lasing stripes.

Higher laser power can also be obtained by noncoherently combining the output power from a number of laser diodes. Several techniques can be used to combine the laser power. Figure 2.1 illustrates the dichroic-polarizer combiner with 8 laser diodes [102], [103], [104]. These diodes are organized at 4 different wavelengths and different polarizations. A narrow band dielectric polarizer can be designed such that it is highly reflective to one polarization state but appears highly transmissive to the other. Optical signals with different polarizations can be combined using such a polarizer. Dichroic elements that act as optical edge filters can then be used to combine output signals from different wavelengths. Note that the efficiency of the power combining technique depends critically on the property of the optical surfaces. When combining output from a large number of laser diodes, the losses at the surfaces may offset the gain in the power, making the combining system highly inefficient.

In addition to the dichroic combiner, optical power combining at different wavelengths can also be achieved using gratings [102], [103], [105] or waveguide multiplexing techniques [102], [103]. A grating rhomb combiners [105] has been developed with which power combination of more than 100 laser diodes appears feasible. However, precision alignment and frequency control are required for the grating combining technique. For the waveguide multiplexed system, the large insertion and coupling losses of the waveguides can severely reduce the power that can be effectively combined.

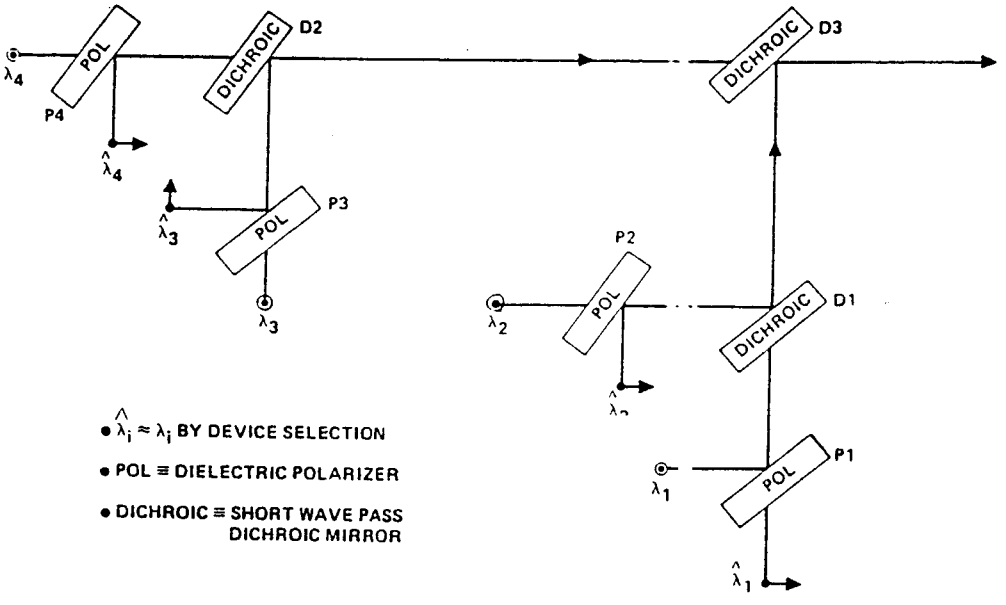


Fig.2.1. The structures of a noncoherent power combiner for laser diodes.

Noncoherent combining of several laser diode outputs can be used to generate sufficient power output for the direct detection ISL. When a heterodyne receiver is used to detect the optical signal, however, additional constraints must be placed on the frequency stability of the transmitter laser. Laser diodes designed for high power output and diode arrays generally have spectral linewidths on the order of 0.01-1 nm or, equivalently, 10^9–10^{11}Hz, and provide little control over the lasing mode. These devices are not suitable for the heterodyne system because the lack of mode stabilization implies a large frequency fluctuation at the optical output. Furthermore, the large linewidth implies a large carrier phase fluctuation which can seriously affect the performance of the decoder. Instead, mode stabilized laser diodes with narrow spectral linewidths must be used. These laser diodes generally employ feedback to improve the frequency stability and reduce the linewidth. Several approaches can be used to achieve stable single mode operations; they are (a) short cavity lasers [106], [107], (b) coupled-cavity lasers [108], [109], (c) external light injection [106],[110], and (d) lasers with grating feedback [106], [111]-[116]. However, these frequency control techniques generally have very low gains such that small signal power is available from the laser diode. Fortunately, over the past few years, steady improvements have been made in the power output and the spectral linewidth of laser diodes. Output power on the order of 10 mW and linewidth on the order of several megahertz have been achieved on laboratory devices [106], [111]-[116]. When coupled with an external cavity, linewidths as narrow as tens of kilohertz have been reported [117]-[119]. Again, it should be noted that lasers designed for narrow spectral linewidth operations generally have small output power. In the future, it is expected that, with an improved fabrication technology, frequency stabilized lasers with averaged power output on the order of 20-30 mW can be manufactured [6]-[8].

The comparison of various candidates for spaceborne laser transmitters is shown in Table 2.1. Note that semiconductor lasers offer smaller device size, lower cooling requirement, and higher reliability. Consequently, for the rest of the discussion we shall assume that semiconductor lasers are used. The output powers of lasers are assumed to be 200 mW for lasers used in direct detection systems, and 25 mW for signal mode, frequency stabilized lasers used in heterodyne systems.

Table 2.1 Comparison of ISL laser transmitters.

System	Advantages	Disadvantages
CO_2	High power output	Large diffraction loss Low detection sensitivity at 10.6 μm wavelength Low reliability
Nd:YAG	High power output High detection sensitivity at 1.06 and 0.53 μm Small diffraction loss	Low pumping efficiency Stringent pointing and tracking requirements
GaAs-AlGaAs	High efficiency High reliability High detection sensitivity Small diffraction loss Easy to modulate Small size	Large frequency and phase fluctuations Low power output Stringent pointing and tracking requirements Large spurious AM and FM

2.3 Transmitter Design

The purpose of the optical transmitter is to focus the optical signal onto the receiver aperture. When a laser is used as the primary optical source, a beam expander is often employed to reduce the diffraction loss. Depending on the cost, size, and weight constraints, a beam expander can be of either refractive or reflective design. Figure 2.2 shows the alternative beam expander configurations. Because of the high cost of low-loss lens materials, reflective beam expanders are used almost exclusively for large telescope designs.

The propagation of the optical signal through a dispersion free media has been studied extensively. The intensity of the far-field signal at an angular distance θ_T from the transmitter line-of-sight (LOS) can be written as

$$I(\theta_T) = \eta_T G_T(\frac{1}{4\pi z^2}) L_T(\theta_T) P_L \quad , \tag{2.13}$$

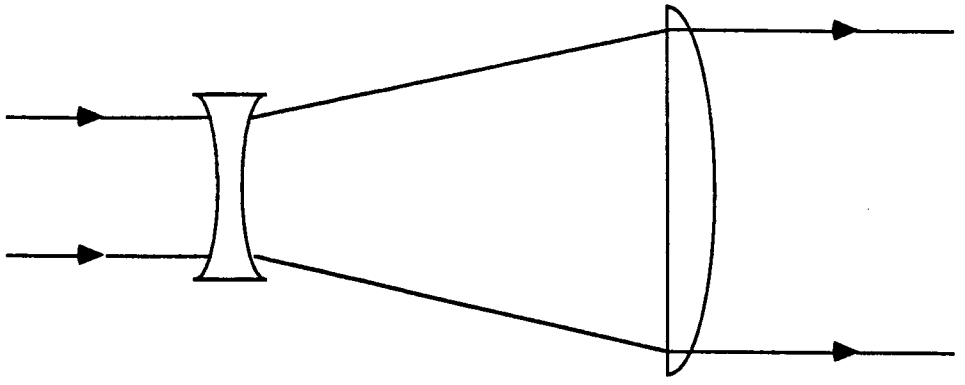
where G_T is the transmitter antenna gain, P_L is the transmitter power, η_T is the transmitter optical efficiency, and $L_T(\theta_T)$ is the pointing loss factor associated with the angular pointing error θ_T. For the case where the signal at the output of the beam expander can be approximated by a uniformly illuminated, circular disk of radius d_T, the transmitter antenna gain and the pointing loss term can be written as [60], [120]

$$G_T = \left[\frac{\pi d_T}{\lambda}\right]^2 , \tag{2.14}$$

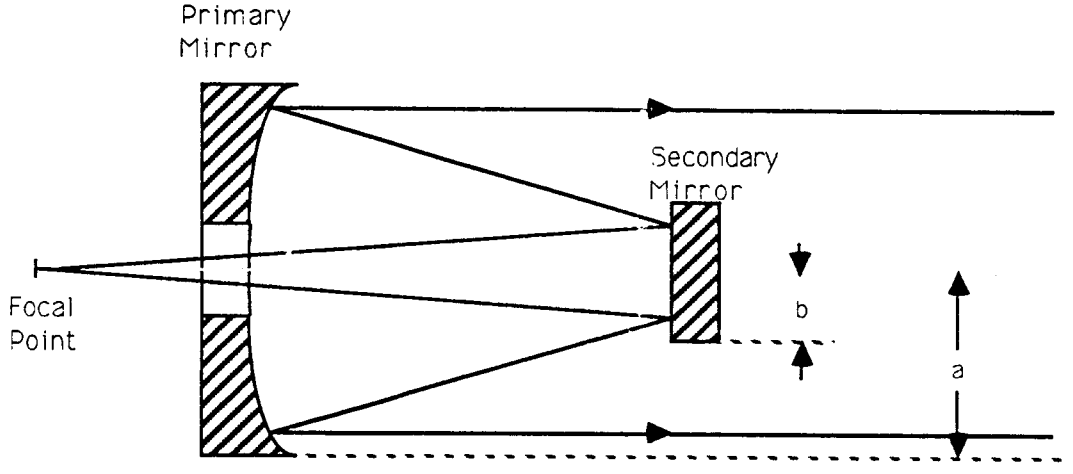
$$L_T(\theta_T) = \left[\frac{2J_1(\pi\theta_T d_T/\lambda)}{\pi\theta_T d_T/\lambda}\right]^2 . \tag{2.15}$$

The far-field intensity distribution of a uniformly illuminated circular aperture is known as the Airy pattern. Figure 2.3 is a plot of the Airy pattern versus the radial angular distance θ_T from the transmitter LOS. It is seen that the far-field intensity drops off rapidly when the angular deviation approaches λ/d_T. In fact, the magnitude of the Airy function is zero when $r/z \approx 1.22\lambda/d_T$.

The transmitter antenna gain and the pointing loss factor in Eqs. (2.14) and (2.15) are calculated for a uniformly illuminated aperture. In reality, the spatial intensity distribution at the output of the beam expander is better approximated by a truncated Gaussian distribution. Furthermore, because of the center obstruction of the Cassegrainian reflector, the gain of an actual transmitter antenna is usually smaller than that given by (2.14). The effect of field distribution and center obstruction on the transmitted signal has been carefully



REFRACTING BEAM EXPANDER



CASSEGRAINIAN TELESCOPE

Fig.2.2. Typical beam expander configurations.

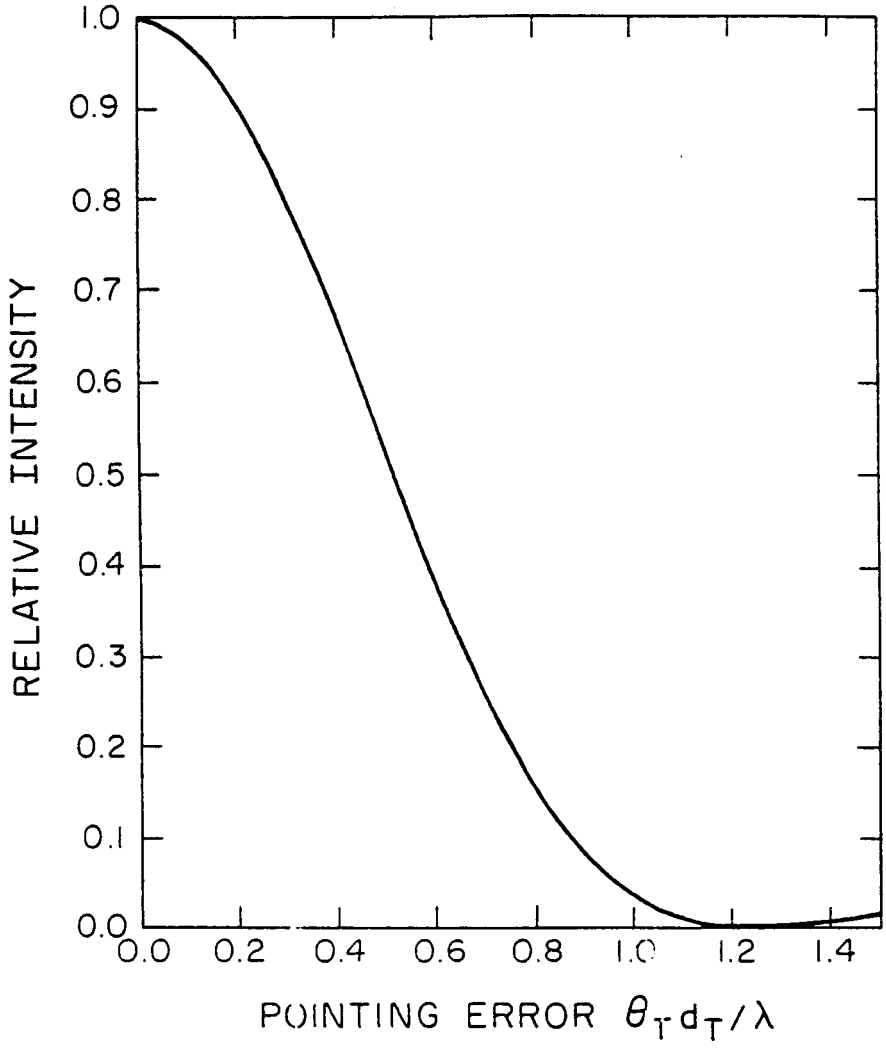


Fig.2.3. Airy pattern of the transmitted signal intensity versus the angular pointing error $\theta_T \lambda / d_T$.

analyzed [120]. It is seen that for a given center obstruction and the transmitter diameter, an optimal RMS width of the Gaussian beam can be found which maximizes the on-axis gain at the far-field. The power penalty associated with the center obstruction is then plotted in Fig. 2.4. It is seen that a 1 dB power penalty is incurred on the system by using a truncated Gaussian beam rather than a uniformly illuminated aperture. The power penalty then increases with the obscuration ratio, which is the ratio between the primary and secondary mirror diameters. When the obscuration ratio increases from 0 to 0.5, over 3.5 dB of power penalty can be expected. In some cases, the spider support for the center reflector can also have a measurable effect on the far field intensity [121]. However, such effects are usually insignificant except when the obscuration ratio is small (<0.2) [121].

In addition to reducing the diffraction loss of the transmitted signal, the transmitter must also maintain the line-of-sight (LOS) tracking of the receiver. Because of the small beam divergence of the optical signal ($\approx$ 10 μrad), care must be taken at the optical transmitter to maintain the receiver within its footprint. Because both transmitter and receiver terminals are in motion for an ISL link, their relative angular position is constantly changing. In order to maintain the receiver within the primary diffraction pattern of the transmitted signal, receiver motion must be acquired and tracked at the transmitter. In practice, this tracking control is accomplished using a separate spatial tracking subsystem. Because of noises in the system, however, some residual errors will remain in the LOS tracking. Such tracking errors can inflict an additional constraint in designing the optical transmitter. This design constraint will be discussed in detail in Chapter 5.

2.4 Receiver Design

At the receiver, the transmitted optical signal is collected and focused onto the photodetector. In general, a large collecting disk, usually of similar design to that of the transmitter telescope, is used to collect the optical radiation. Receiver and transmitter can share the same telescope to reduce the system size, cost, and weight. However, for systems employing the single telescope design, cross talk between the transmitter and receiver resulting from optical leakage must be carefully minimized. Alternatively, optical links may be designed using separate transmitter and receiver telescopes. Such systems offer better transmitter-receiver isolation and allow modular design for the optical component. However, the size and weight of the optical system must be significantly increased to accommodate the additional optical components.

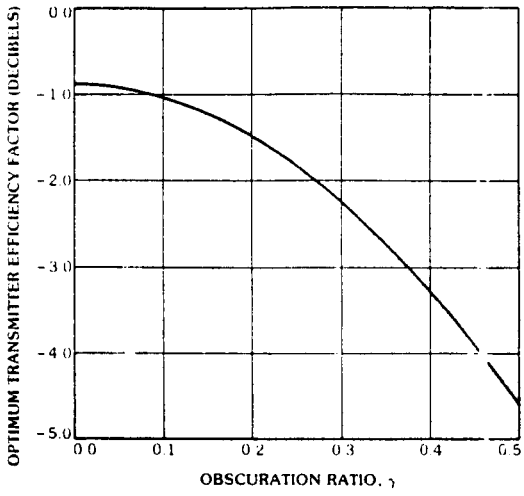


Fig.2.4. Optimal transmitter telescope efficiency relative to an unobscured uniformly illuminated aperture as a function of the obscuration ratio. (From B. J. Klein and J. J. Degnan, "Optical Antenna Gain, 1. Transmitting Antennas," Appl. Opt., 13, Sept. 1974; p.2139.)

Given the transmitted intensity pattern of (2.13), the total signal power collected by the receiver can be related to the transmitter power by [60]

$$P_S = \eta_T \eta_R G_T G_R (\frac{\lambda}{4\pi z})^2 L_T(\theta_T) P_L \ , \tag{2.16}$$

where G_T, G_R are the transmitter and receiver antenna gains, η_T, η_R are the transmitter and receiver optics efficiencies, and $L_T(\theta_T)$ is the loss associated with a angular pointing error of θ_T. The gain of the receiver antenna can be related to the diameter of the receiver aperture and the wavelength by [122]

$$G_R = \left[\frac{\pi d_R}{\lambda} \right]^2 . \tag{2.17}$$

Equation (2.16) was derived assuming that the area of the receiver aperture is small compared to the primary diffraction pattern such that the total power can be approximated by the product of the aperture area and the signal intensity. Furthermore, it was assumed that the receiver aperture is uniformly illuminated without center obstruction. A smaller receiver gain can be expected when using a Cassegrainian telescope [122].

Optical receivers can be divided into two classes: those that detect only the intensity fluctuation, and those that detect both amplitude and phase of the optical radiation [3]. The former type, usually known as the direct detection receiver, employs little optical processing and detects the signal by directly focusing the collected optical signal onto a photodetector. The latter, known as the heterodyne receiver, requires a precise alignment between the received optical signal and the local oscillator (LO) laser, but offers a more optimal detection sensitivity.

Figure 2.5 shows the basic block diagram of a direct detection receiver. The optical signal is collected by the receiver telescope and, after passing through optical and spatial filters, is focused directly onto a photodetector. The optical filter in the receiver allows only radiations near the signal wavelength to pass through. Optical filtering can be accomplished using an interference filter, the passband of which is typically on the order of tens of angstroms. The spatial filtering in the receiver is achieved using a field stop iris, and is necessary to reduce the receiver field-of-view (FOV). Because no use is made of the spatial and temporal coherence of the transmitted signal, direct detection receivers are also known as noncoherent receivers.

In the absence of the background radiation, the signal-to-noise ratio (SNR) of the direct detection receiver is simply the SNR of the detector given by Eq. (2.12). When the background radiation is detected in

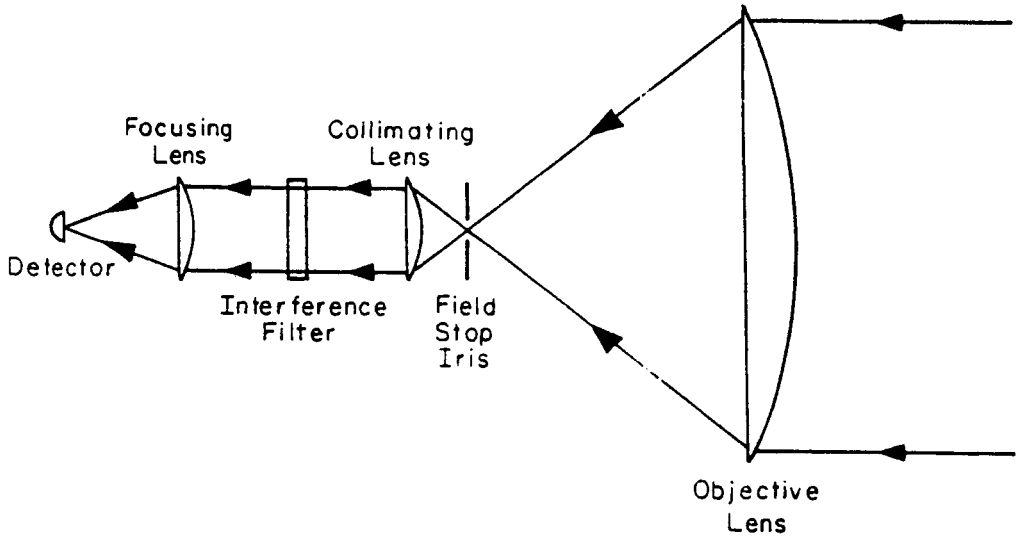


Fig.2.5. Block diagram of a direct detection receiver.

addition to the signal, the receiver SNR is given by

$$SNR_{DD} = \frac{(\eta G P_S / h\nu)^2}{\{G^2 F [(\frac{\eta}{h\nu})(P_S+P_B)+\frac{i_b}{e}]+\frac{i_s}{e}+\frac{2k_B T}{R_L}\}B} \quad . \tag{2.18}$$

The numerator in (2.16) is the signal power due to a constant (dc) signal power P_S, and the denominator is the sum of shot noise and thermal noise power admitted by the receiver of bandwidth B. When the signal power is much higher than the thermal and background noise powers, and when the detector dark current can be ignored compared to the signal current, the signal shot noise is the dominant noise source. In this case the detector is said to be shot noise limited and the receiver SNR is given by

$$\lim_{P_S \to \infty} SNR_{DD} = \frac{\eta P_S}{F h\nu B} \quad . \tag{2.19}$$

The principal advantage of the direct detection receiver is its simplicity. When operating under a restricted environment, the simplicity in design usually leads to an improved reliability and a longer mean-time-to-failure (MTTF). However, when operating under low signal power and high background levels, the SNR of the direct detection receiver can degrade appreciably because of the large thermal and background noise. Performance can be improved by increasing signal power or by using high gain photomultipliers. However, both methods are impractical for a spaceborne system because of limitations on the available output power and the lack of suitable photomultipliers at the optical wavelengths of interest. The effect of thermal noise on the detector sensitivity can be reduced by using the avalanche photodetector. Alternatively, the detection sensitivity can be improved by using the heterodyne receiver.

Figure 2.6 depicts the structure of a typical heterodyne receiver. The incoming signal is spatially mixed with a strong local oscillator (LO) beam. The combined optical signal is then focused onto a photodetector. Optical field mixing at the detector surface can be described in terms of the diffraction patterns at the focal plane or, equivalently, in terms of the back-projected signal and LO field distributions over the receiver aperture plane. If we let $\vec{E}_S$ and $\vec{E}_{LO}$ denote the electric field vector of the signal and LO field at the receiver aperture plane, respectively, then the total power incident on the detector surface is

$$P_R = \frac{1}{z_0}\int_{A_R} |\vec{E}_S+\vec{E}_{LO}|^2 d^2r \quad , \tag{2.20}$$

where z_0 is the impedance of the media surrounding the photodetector, and the integral is extended over the

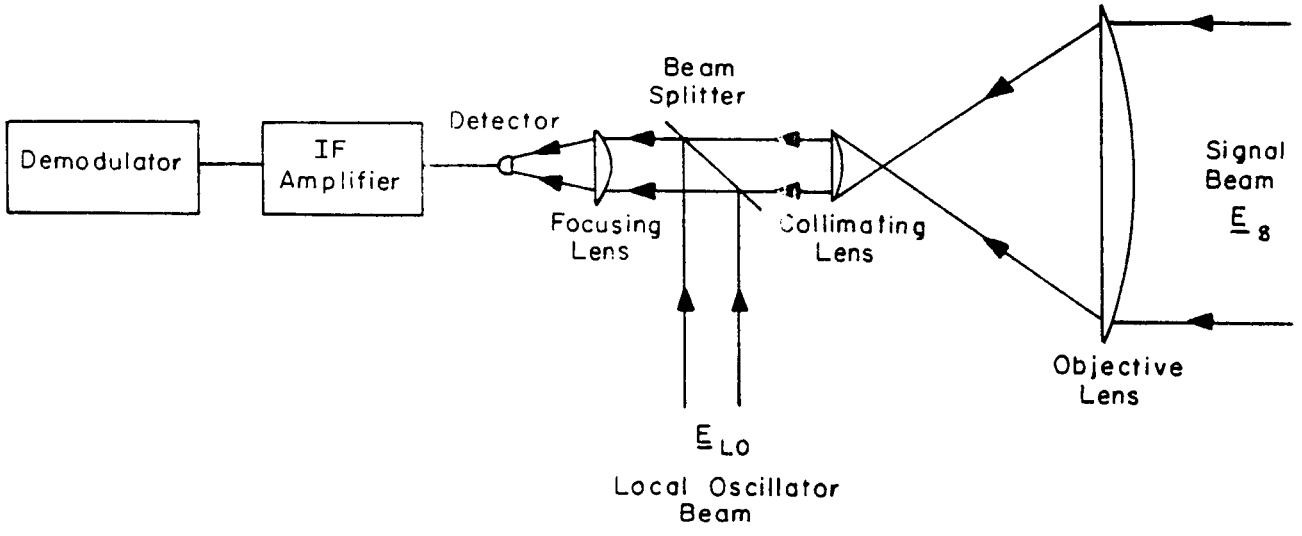


Fig.2.6. Block diagram of a typical heterodyne receiver.

receiver aperture area A_R. The electric field components $\vec{E}_S$ and $\vec{E}_{LO}$ can be written in the form

$$\vec{E}_S(\vec{r},t) = \vec{A}_S(\vec{r})\cos(2\pi f_S t + \phi_S) \ ,$$
$$\vec{E}_{LO}(\vec{r},t) = \vec{A}_{LO}(\vec{r})\cos(2\pi f_{LO} t + \phi_{LO}) \ , \tag{2.21}$$

where $\vec{A}_S(\vec{r})$ and $\vec{A}_{LO}(\vec{r})$ are the vector amplitudes, f_S and f_{LO} are the optical frequencies, and ϕ_S and ϕ_{LO} are the phase of the signal and LO, respectively. The amplitudes and phases of both signal and LO can depend upon time and position. When the optical frequencies f_S and f_{LO} are closely spaced, the interference of the optical signals creates a beat frequency at $f_{IF}=f_S-f_{LO}$. By substituting Eq.(2.21) into (2.20) and integrating over a time period that is much longer than the optical period, but short compared to the beat period, the total optical power incident on the detector surface can be written as

$$P_R = P_S + P_{LO} + P_{IF} \ , \tag{2.22}$$

where P_S and P_{LO} are the total power of the signal and LO, respectively, and P_{IF} is the IF signal power which is due to the interference between the signal and LO:

$$P_{IF}(t) = \frac{1}{z_0}\int_{A_R} \vec{A}_S(\vec{r})\cdot\vec{A}_{LO}(\vec{r})\cos(2\pi f_{IF} t + \phi_S - \phi_{LO}) d^2r \ . \tag{2.23}$$

The integral in Eq.(2.23) is very sensitive to the (spatial) mode mismatch between the signal and LO. The magnitude of $P_{IF}(t)$ is largest when the signal and LO fields are spatially matched over the entire receiver aperture. For intersatellite links, the distance between the receiver and the transmitter is usually so large that the received signal can be approximated by a plane wave on the receiver aperture. Then for a uniform LO field, the integral in Eq.(2.23) will depend only on the relative polarization and propagation directions of signal and LO. By substituting the expressions for plane waves into (2.23) and integrating over the aperture area, the IF signal power can be written as

$$P_{IF}(t) = 2\sqrt{P_S P_{LO}\eta_{het}}\cos(2\pi f_{IF} t + \phi_S - \phi_{LO}) \tag{2.24}$$

where

$$P_S = \frac{1}{2z_0}\int_{A_R} |\vec{A}_S(\vec{r})|^2 d^2r \ ,$$
$$P_{LO} = \frac{1}{2z_0}\int_{A_R} |\vec{A}_{LO}(\vec{r})|^2 d^2r \ , \tag{2.25}$$

are the signal and LO powers, and η_{het} is the heterodyne efficiency which is defined as

$$\eta_{het} = \left[\frac{\frac{1}{z_0}\int_{A_R} \vec{A}_S(\vec{r}) \cdot \vec{A}_{LO}(\vec{r}) d^2r}{\sqrt{P_S P_{LO}}} \right]^2 . \tag{2.26}$$

When both the signal and LO are plane waves that cover the entire receiver aperture, the heterodyne efficiency can be evaluated to be:

$$\eta_{het} = \cos^2\psi_p L_R(\theta_R) , \tag{2.27}$$

where ψ_p and θ_R are the angular differences between the polarization direction and the propagation axes of the signal and LO, respectively. The factor $L_R(\theta_R)$ is the loss factor due to the spatial tracking error between the signal and LO. When both signal and LO are plane waves extending over the receiver aperture of diameter d_R, $L_R(\theta_R)$ can be given by

$$L_R(\theta_R) = \left[\frac{2J_1(\pi\theta_R d_R/\lambda)}{\pi\theta_R d_R/\lambda} \right]^2 . \tag{2.28}$$

When the incident wavefronts at the receiver aperture can not be approximated by plane waves, Eq.(2.26) should be used to evaluate η_{het}. In order to obtain a maximum detection sensitivity, both polarization directions and propagation axes of the signal and LO must be carefully matched. Because polarization alignment is relatively easy to accomplish, the principal loss in the heterodyne receiver is due to the tracking error between the propagation axes. By assuming that the polarization directions of the signal and LO are perfectly aligned, the IF power at the output of the receiver can be written as

$$P_{IF}(t) = 2\sqrt{P_S P_{LO} L_R(\theta_R)}\cos(2\pi f_{IF} t + \phi_S - \phi_{LO}) . \tag{2.29}$$

Because the signal power is usually much smaller than the LO power, the output of the photodetector can be modeled as the sum of a large dc current due to P_{LO}, an IF current due to P_{IF}, and the associated shot noise. The dc current can be effectively removed by passing the photodetector output through an IF amplifier-filter or, alternatively, it can be removed by using a dual detector configuration shown in Figure 2.7 [123]. By using one detector at each port of the beam splitter and coherently subtracting the two detector outputs, it is possible to eliminate the large dc contribution from the LO and greatly reduce the noise due to the fluctuation in LO strength. After removing the large dc current due to the LO, the IF signal is given by

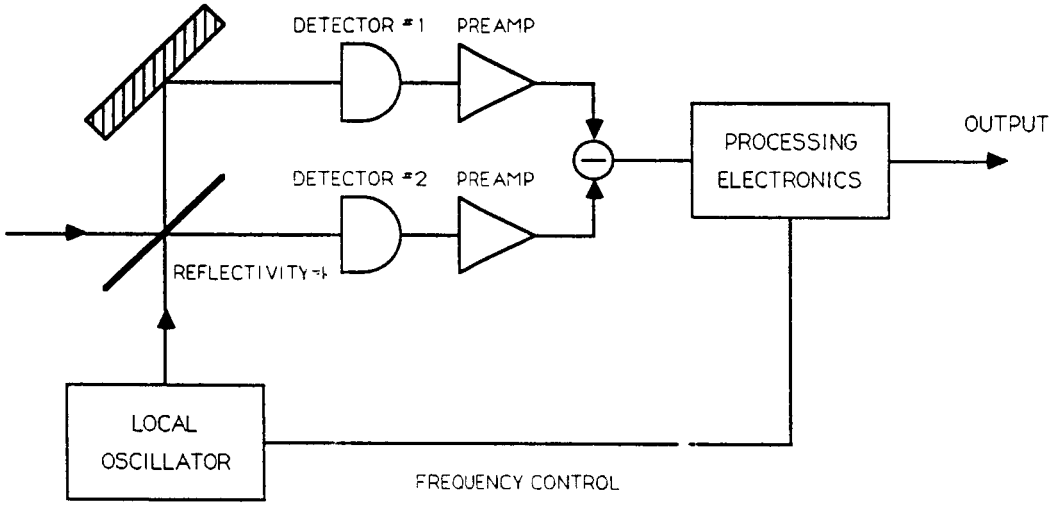


Fig.2.7. Balanced mixer (dual detector) heterodyne receiver.

$$i_{IF}(t) = 2(\frac{\eta}{h\nu})\sqrt{\eta_{het} P_S P_{LO} L_R(\theta_R)}\cos(2\pi f_{IF} t + \phi_{IF}) + n_s(t) + n_{th}(t) \ , \tag{2.30}$$

where ϕ_{IF} is the phase angle of the IF signal, and n_s and n_{th} are the detector shot noise and thermal noise, respectively.

The detector shot noise can be approximated by a white Gaussian noise with power spectral density

$$N_s = \frac{\eta e^2 G^2 F}{h\nu}(P_S + P_{LO} + P_{IF}) \ . \tag{2.31}$$

For a given IF bandwidth B_{IF}, the receiver SNR is therefore

$$SNF_{IF} = \frac{2(\frac{\eta e G}{h\nu})^2 P_S P_{LO} L_R(\theta_R)}{[\frac{\eta e^2 G^2 F}{h\nu}(P_S + P_{LO} + P_{IF}) + \frac{2k_B T}{R_L}]B_{IF}} \ , \tag{2.32}$$

where we have ignored the noise contribution of the detector dark current. In the limit of large LO power, thermal noise and shot noise due to signal and background radiations will be negligible compared to the LO shot noise. In this case the SNR reduces to that of an ideal, quantum limited photodetector [3], [26], [27]:

$$\lim_{P_{LO} \to \infty} SNR_{IF} = \frac{2\eta P_S}{F h\nu B_{IF}} L_R(\theta_R) \ . \tag{2.33}$$

Note that because of the large LO strength, *pin* diode can be used to detect heterodyned signal. The excess noise factor for the *pin* diode is equal to 1. Compared to the SNR of a direct detection receiver given in (2.16) it is seen that, when spatial tracking between the signal and LO is perfectly established, the heterodyne receiver offers a significant improvement in SNR over the direct detection receiver. Background radiation and thermal noise that affect the direct detection receiver can be easily overcome by increasing the LO signal strength.

2.5 Background Radiation

In addition to the transmitted signal power, optical receivers also collect background radiation. The amount of background power collected by the receiver depends on the receiver's field of view (FOV) and the optical bandwidth [124] [125]. In order to reduce the amount of background power admitted by the receiver, it is desirable to have as small a receiver FOV and optical bandwidth as possible. Receiver FOV can be limited using a field stop in the optical path. However, the pointing and tracking problems are more severe with small

FOV. Typical values of the receiver FOV are on the order of 0.1 mrad. Similarly, the receiver optical bandwidth can be limited using an interference filter. For a white background radiation, the amount of power admitted by an interference filter is proportional to its bandwidth. However, the transmittances of typical interference filters decrease with decreasing bandwidth. Consequently, at small optical bandwidths the signal itself may be severely attenuated. For most applications, interference filters with bandwidths of 10-20 Å and transmittances of 50-70% are used.

The amount of background power incident on the receiver can in general be measured in terms of the spectral radiant emittance $W(\lambda)$ which is the radiant power at wavelength λ emitted into a hemisphere per unit area of source in the hemisphere. Assuming a Lambertian source of diameter d_s at a distance z from the receiver, the amount of background power incident on the detector surface can be written as [124], [125]

$$P_B = \frac{\pi \eta_R d_s^2 d_R^2}{16 z^2} \Delta\lambda \, W(\lambda) \ , \tag{2.34}$$

where

η_R is the receiver optical efficiency,

d_S is the source aperture diameter,

d_R is the receiver aperture size, and

$\Delta\lambda$ is the bandwidth of the interference filter.

Equation (2.34) was derived for a background source that is contained within the receiver FOV. For an extended source where the size of the object is larger than the FOV, the background power is given to a good approximation by [124], [125]

$$P_B = \frac{\eta_R d_R^2 \Omega_R}{4} \Delta\lambda \, W(\lambda) \ , \tag{2.35}$$

where Ω_R is the solid angle receiver FOV which can be related to the planar angle FOV by

$$\Omega_R = \frac{\pi \Theta_R^2}{4} . \tag{2.36}$$

For smaller sources such as stars, the background radiation is usually measured in terms of the spectral irradiance $H(\lambda)$ which is the power per unit wavelength incident on a unit area of the receiver. For a given

spectral irradiance $H(\lambda)$, the background power collected by the receiver is given by [124]

$$P_B = \frac{\pi \eta_R d_R^2}{4} \Delta\lambda \, H(\lambda). \tag{2.37}$$

The total background power collected by the receiver is the sum of powers due to earth reflection, the moon, the planets, and the stars. The radiation due to the sun is usually so intense that direct communication with the sun in the background is not achievable using direct detection receivers. Figure 2.8 shows the spectral radiant emittance of the earth versus wavelength. Note that for the wavelength of interest (850 nm for GaAs diode lasers), $W(\lambda)$ is on the order of 10^{-2} W/cm^2-μm. As indicated in the figure, this measurement is carried out when there is no cloud cover in the earth's atmosphere. Radiance from cloud-covered earth is approximately one order of magnitude higher than that given by the figure and represents the strongest background source to be considered. By evaluating the amount of received power using the assumed receiver characteristics shown in Table 2.2, the background power due to earth reflection can be seen to be

$$P_B \approx \begin{cases} 10^{-10} W & \text{with cloud cover,} \\ 10^{-11} W & \text{no cloud.} \end{cases} \qquad \text{(Earth reflection)}$$

Next to the earth reflection, lunar spectral irradiance is the strongest background emission source. Figure 2.9 is a plot of the lunar and planetary spectral irradiance versus wavelength which shows, typically, a spectral irradiance of the moon is approximately 10^{-7} W/cm^2-μm. Because of the close proximity of the moon, the amount of background power gathered due to the moon cannot be evaluated using Eq.(2.37) since it assumes the receiver FOV is greater than the angular span of the moon while actually only a portion of the moon is visible within the receiver FOV. By modeling the lunar surface as a uniformly illuminated Lambertian reflector, the amount of background radiation received due to the moon is given by

$$P_B = \frac{\pi \eta_R \Omega_R d_R^2 z^2}{4 d_m^2} \Delta\lambda H(\lambda) \tag{2.38}$$

where z is the mean distance between the earth and the moon, and d_m is the diameter of the moon. After substituting the lunar irradiance into (2.38), the amount of background power collected due to the moon is given by

$$P_B \approx 5 \times 10^{-12} W \ . \qquad \text{(Moon)}$$

Table 2.2 Typical ISL link paramters.

Parameters	Direct Detection System	Heterodyne System
Transmitter Laser		
Power	200 mW	30 mW
Extinction Ratio	5%	-
Modulation Bandwidth	>15 GHz	>15 GHz
Linewidth	-	3 MHz
Optical Antenna		
Transmitter Diameter	10.6 cm†	11.4 cm†
Receiver Diameter	30.0 cm	11.4 cm†
Transmitter and Receiver Optics		
Transmitter Efficiency	50%	50%
Receiver Efficiency	35%	35%
Tracking Split	5%	5%
Receiver Optical Bandwidth	2 nm	-
Receiver Field-of-View	0.1 mrad	-
Optical Detector		
Quantum Efficiency	70%	70%
k_{eff}	0.01	-
Gain	150	1
Excess Noise Factor	3.5	-
Gain Dependent Dark Current	10^{-10}A	-
Gain Independent Dark Current	10^{-8}A	-
Bandwidth	>10 GHz	>10 GHz
Noise Equivalent Temperature	400°K	-
Load Resistance	2KΩ	-

† Optimized for RMS pointing and Tracking errors of 1μrad.

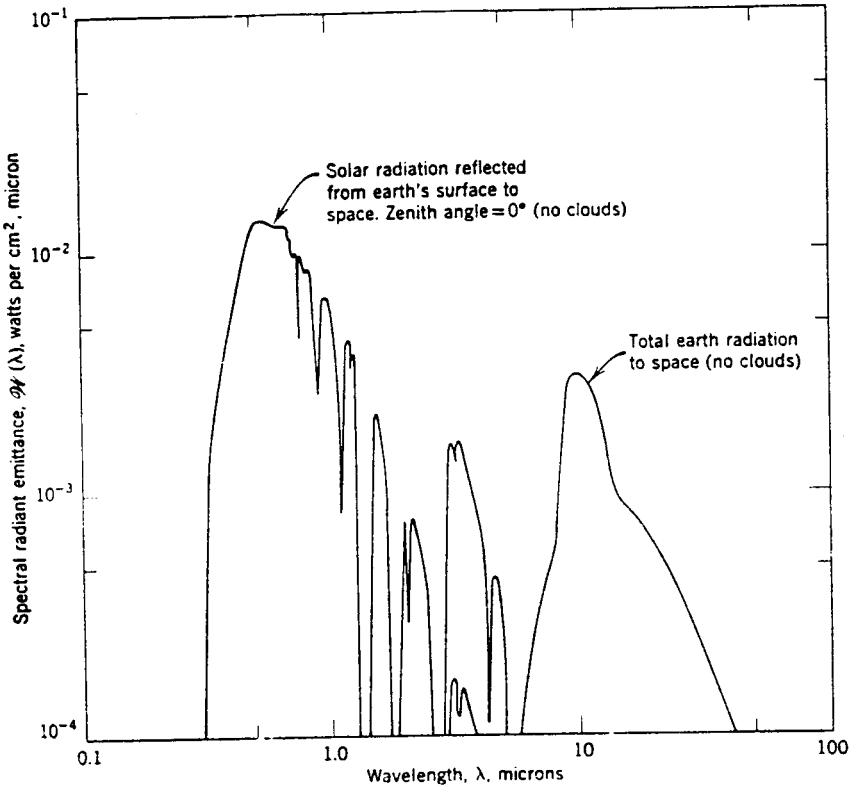


Fig.2.8. Spectral radiant emittance of the earth. (From W. K. Pratt, *Laser Communication Systems*, New York: Wiley, 1968; p. 125.)

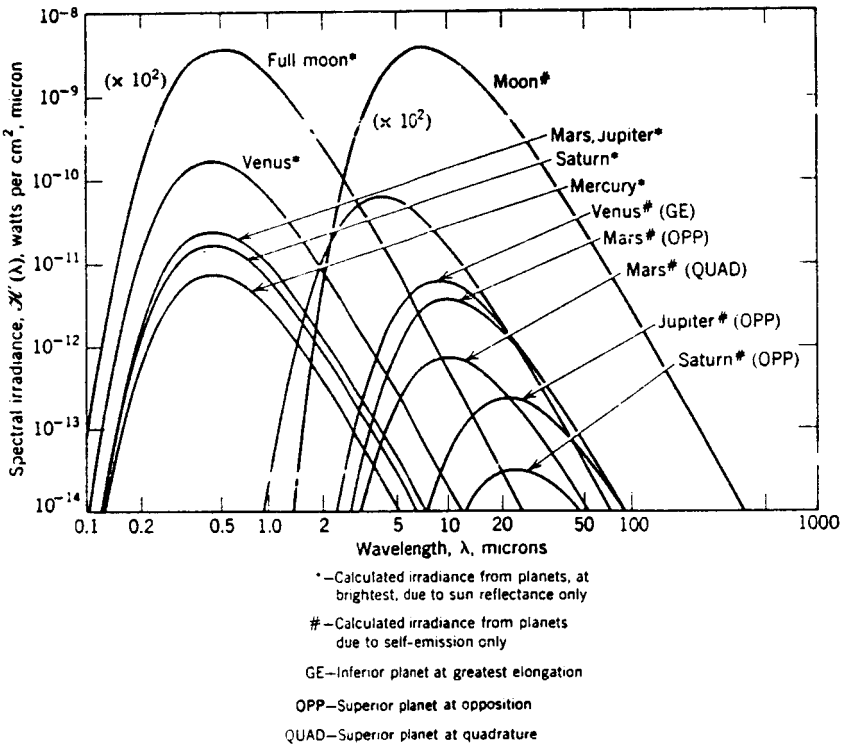


Fig.2.9. Calculated planetary and lunar spectral irradiance outside the atmosphere. (From W. K. Pratt, *Laser Communication Systems*, New York: Wiley, 1968; p. 123.)

Background powers due to planets and stars are many magnitudes smaller than those of the earth reflection and the moon. From Figures 2.9 and 2.10 it is seen that, even under the worst case, the spectral irradiances of the planets and stars are at least three orders of magnitude smaller than the lunar irradiance. This results in a background power of

$$P_B \ll 10^{-13} W \qquad \text{(Stars)}$$

which is considerably lower than the earth and moon contributions.

2.6 Summary

Various components of an optical ISL have been described. It is seen that semiconductor lasers with sufficiently high power and reliability can be used as the primary transmitter for an intersatellite link. The signal and background power collected by the ISL receiver have been derived as functions of the transmitter and the receiver parameters. Both direct and heterodyne receivers can be used to implement the optical link. The direct detection receiver offers simplicity in design and a potentially higher reliability, but is more sensitive to the thermal and background noises. The sensitivity with respect to the thermal noise can be improved at the cost of increasing gain fluctuation by using an APD. On the other hand, the heterodyne receiver promises a much improved detection sensitivity and a better background noise rejection. However, the higher complexity of the receiver implies lower reliability and longer development time. For the design of the practical ISL, it is the probability of bit error (PBE) that is of ultimate interest. Given the signal and background powers at the receiver, the PBE for systems employing the direct and heterodyne receivers will be analyzed in the next two chapters.

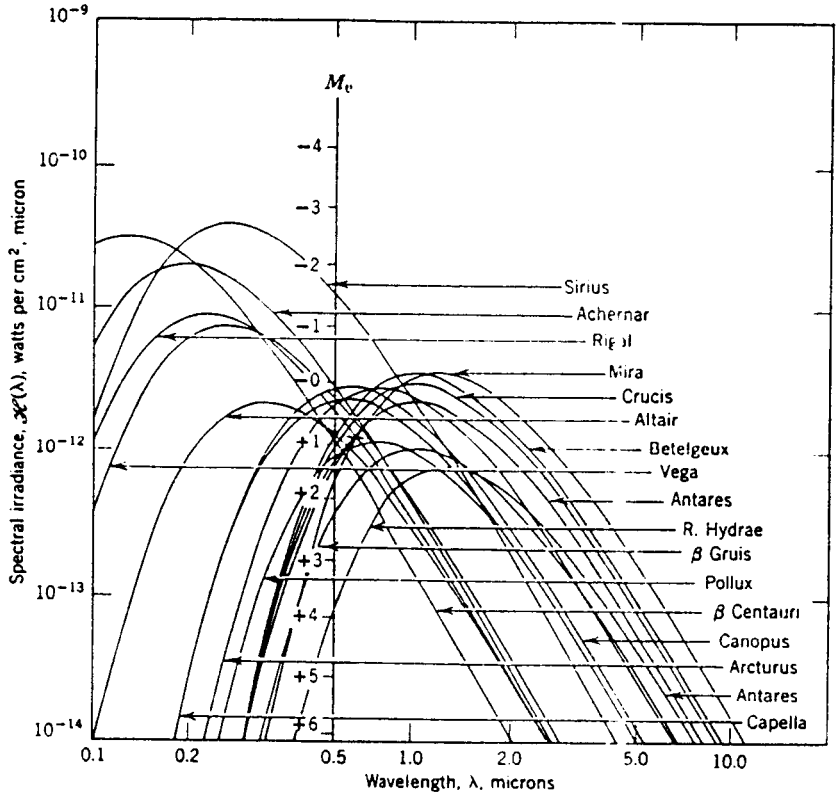


Fig.2.10. Spectral irradiance of brightest stars outside the atmosphere. (From W. K. Pratt, *Laser Communication Systems*, New York: Wiley, 1968; p. 126.)

3. DIRECT DETECTION SYSTEMS

In the previous chapter the signal and background power collected by the receiver were evaluated as functions of the link design. Given the signal and background power at the receiver, it is of interest to evaluate the performance of the link in data communication. In particular, we are interested in a digital link in which the information is first digitally encoded before being modulated onto the optical carrier. In this chapter, the performance of a direct detection optical communication link will be analyzed.

3.1 Digital Encoding Schemes

In a typical M-ary digital communication link the information to be transmitted consists of a sequence of codewords $(d_1,d_2,...d_N)$, where each d_k can take on values between $(0,1,...,M-1)$. A modulator then maps the codewords $\{d_k\}$ into a set of transmitter waveforms $\{s_{d_k}(t)\}$. During each of the transmission intervals, the codeword d_k is transmitted by sending a waveform $s_{d_k}(t)$ over the channel. When $M=2$, the communication system is referred to as a binary channel. For $M>2$, the system is known as an M-ary (block encoded) channel. For a binary channel, each transmitted waveform carries exactly one bit of information. Consequently, the bit data rate R of the binary channel is equal to the rate of transmission of waveforms R_w. For M-ary channels, on the other hand, several bits can be encoded into one transmitted waveform such that, on the average, $\log_2 M$ bits are transmitted per waveform. The bit data rate associated with the M-ary channel is therefore

$$R = R_w \cdot \log_2 M \ . \tag{3.1}$$

The actual symbols transmitted over the channel can be any of the amplitude, frequency, or phase modulated signals. For optical communication systems employing direct detection receivers, only the signal intensity is detected so that the modulation schemes are limited to polarization and intensity modulations. Polarization modulated signals can be detected as an intensity modulated signal by polarization filtering prior to photodetection. Polarization modulation can be accomplished using an electro-optic material, and intensity modulation of the transmitter laser can be achieved by varying the injection current of the laser diode [126], or by using an external electro-optic [127]-[129] or electro-absorption [130]-[132] modulator.

Modulation schemes commonly used in direct detection systems include on-off-keying (OOK) and pulse-position modulation (PPM). In OOK the binary data are transmitted as a sequence of optical pulses such that a binary one is transmitted by the presence of the pulse while the absence of the pulse signals a binary zero. A typical OOK encoded signal is shown in Fig. 3.1. Alternatively, information can be encoded in the PPM scheme. In a M-ary PPM system, each word frame is divided into M time slots. Information is then transmitted by sending an optical pulse in one of the M time slots. Fig. 3.2 shows a typical PPM encoded signal. Because multiple bits of information can be transmitted using a single pulse, PPM is more power efficient than OOK. On the other hand, because each word is divided into many time slots, the bandwidth requirement for a high order PPM system is significantly larger than that of an OOK system. Furthermore, PPM systems are more complicated and more difficult to implement. However, as we shall see, PPM systems offer significant improvements in performance over OOK systems, and are therefore more suitable for spaceborne applications.

3.2 Detector Statistics and MAP Decoding

Because photodetection is essentially a discrete process, output of the photodetector is corrupted by the signal shot noise in addition to the detector thermal noise. The detector photocount rate, which is related to the received optical power by Eq. (2.1), can be written as

$$\lambda(t) = \lambda_S(t) + \lambda_B \tag{3.2}$$

where $\lambda_S(t)$ is the photocount rate associated with the signal $s(t)$, and λ_B is the count rate due to the uniform background radiation. The purpose of the digital demodulator, therefore, is to determine which of the M waveforms $\{s_k(t)\}$ was transmitted given the detector output $r(t)$.

In order to decode the transmitted information, the receiver performs the maximum a posteriori (MAP) decision [133]. Under this criterion, the output of the photodetector is observed over a given symbol period $(0, T_s)$. The decoder then selects the codeword that has the highest probability of having being transmitted. Mathematically, given the output of the photodetector $r(t)$, $0 \le t \le T_s$, the decoder selects the codeword which maximizes

$$P(s_k(t) \mid r(t)) \equiv \text{probability that } s_k(t) \text{ was transmitted given the received signal } r(t)\ . \tag{3.3}$$

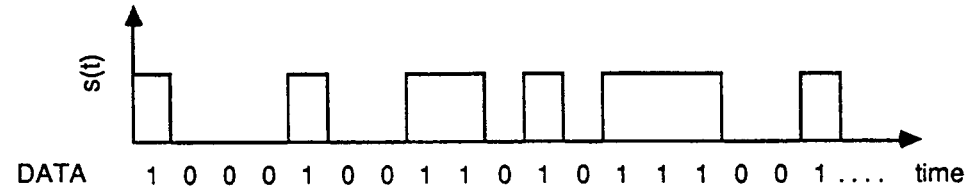


Fig.3.1. Sample bit stream in OOK modulation format.

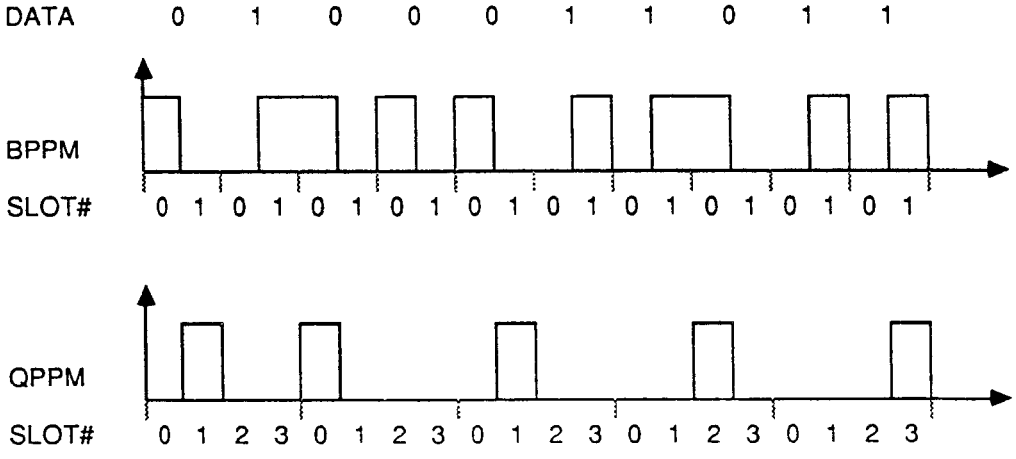


Fig.3.2. Sample bit stream in PPM modulation format.

For a binary system, the decoder outputs a binary one if $P(s_1(t)|r(t))>P(s_0(t)|r(t))$ and decides a binary zero if $P(s_1(t)|r(t))<P(s_0(t)|r(t))$. If $P(s_1(t)|r(t))=P(s_0(t)|r(t))$, then both $s_0(t)$ and $s_1(t)$ are equally likely to be transmitted and a random decision is made to determine the decoder output. Similarly, for an M-ary block encoded system, the decoder outputs the k^{th} codeword if $P(s_k(t)|r(t))$ is the largest among the M posterior probabilities. In the case where several highest posterior probabilities are equal, a random decision is made among these codewords.

For systems where codewords have equal prior probabilities to be transmitted, the MAP decision rule reduces to the maximum likelihood (ML) test [133] where the receiver compares the likelihood function

$$\Lambda_k \equiv P(r(t)|s_k(t)) \ , \tag{3.4}$$

and makes the decision by choosing the codeword corresponding to the maximum likelihood function. The likelihood function in Eq. (3.4) is easier to evaluate than the posterior probability function in Eq. (3.3) and, consequently, for systems with equal priors, the ML test is often used.

In general, given the detector output in Eq. (2.3), the probabilities in Eqs. (3.3) and (3.4) are difficult to calculate. Under two special conditions, however, the calculation of the likelihood function Λ_k can be simplified significantly. The first case is when both thermal noise and detector gain noise can be ignored. In this quantum limited case the detector simply counts the number of photons detected during each interval. Section 3.3 describes the performance of this quantum limited detector. Another special case of interest is when the random avalanche gain is significant such as when APDs are used as the primary photodetectors. In this case, the output of the detector can be approximated by Gaussian statistics and the calculation of prior probabilities Λ_k can be greatly simplified. Section 3.4 discusses the performance of ML decoders when using APDs.

3.3 The MAP Decoder for Photon Counting Receivers

The optimal photodetector measures the arrival time of each photon. The PMTs, with their large internal gains, can best approximate the quantum limited detectors. For systems using a quantum limited detector, the probability of observing k photons during the period $(0, T_s)$ at time $\vec{\tau}=(\tau_1,\tau_2,\ldots,\tau_k)$ is given by [74]

$$P(\vec{\tau}|k) = \left[\prod_{i=1}^{k}\lambda(\tau_i)\right] \exp\left(-\int_0^{T_s}\lambda(\xi)d\xi\right) \ , \tag{3.5}$$

where $\lambda(\xi)$ is the photocount rate at time ξ. For systems where the photocount rate $\lambda(\xi)$ is maintained constant throughout the period $(0, T_s)$, Eq. (3.5) reduces to

$$P(\vec{\tau}|k) = \lambda^k e^{-\lambda T_s} \quad . \tag{3.6}$$

Note that Eq. (3.6) depends only on the number of photons received during the period $(0, T_s)$, and not on the arrival time of individual photons. Consequently, for the MAP decoding of OOK and PPM signals where rectangular pulses are transmitted, the optimal decoder measures only the number of photons received during the period. For this reason, optimal, quantum limited detectors are often known as photon counting detectors.

OOK Systems The maximum a posteriori (MAP) decoder for the OOK system compares the number of received photons with a threshold θ and makes the decision based on the sign of the difference. If the number of received photons is greater than θ, a binary one is decoded. Otherwise, a binary zero is decoded. The decision threshold θ depends on the signal and background photocounts and the probability distribution of the transmitted data. For systems with equal priors, θ is given by [3]

$$\theta = K_S / \log\left[\frac{K_S + K_B}{K_B}\right] \quad , \tag{3.7}$$

where $K_S = \lambda_S T_s$ and $K_B = \lambda_B T_s$ are the expected photocounts due to the signal and background radiations, respectively. In practice, the decoding can be achieved by integrating the photodetector output over a bit period, then comparing the integrator output with a threshold, as depicted in Fig. 3.3. If the detector impulse response time is short compared to the integration period and if thermal noise and detector gain noise can be neglected, the output of the integrator will be proportional to the number of photons received during the pulse period, and the MAP decision can be made accordingly.

For the MAP decision rule given above, a decision error occurs whenever $y<\theta$ and a binary 1 was transmitted, or when $y>\theta$ and a binary 0 was transmitted. Using the fact that receiver photocounts are Poisson distributed, the probability of bit error (PBE) for the OOK system is given by [3]

$$P_E = \frac{1}{2} \sum_{k=\Gamma(\theta)}^{\infty} \frac{K_B^k}{k!} e^{-K_B} + \frac{1}{2} \sum_{k=0}^{L(\theta)} \frac{(K_S + K_B)^k}{k!} e^{-(K_S + K_B)} \quad , \tag{3.8}$$

where $\Gamma(\theta)$ and $L(\theta)$ denote the smallest integer larger than θ and the largest integer smaller than θ, respectively. Figure 3.4 is a plot of probability of bit error for an OOK system using photon counting detectors

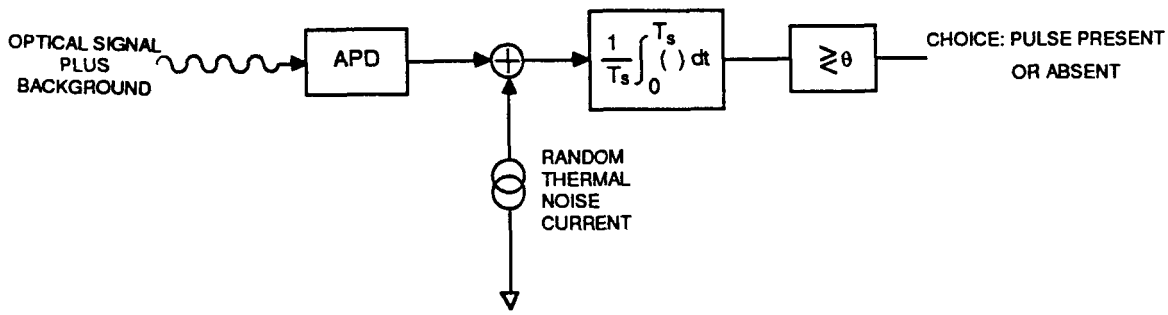


Fig.3.3. Model of an OOK demodulator.

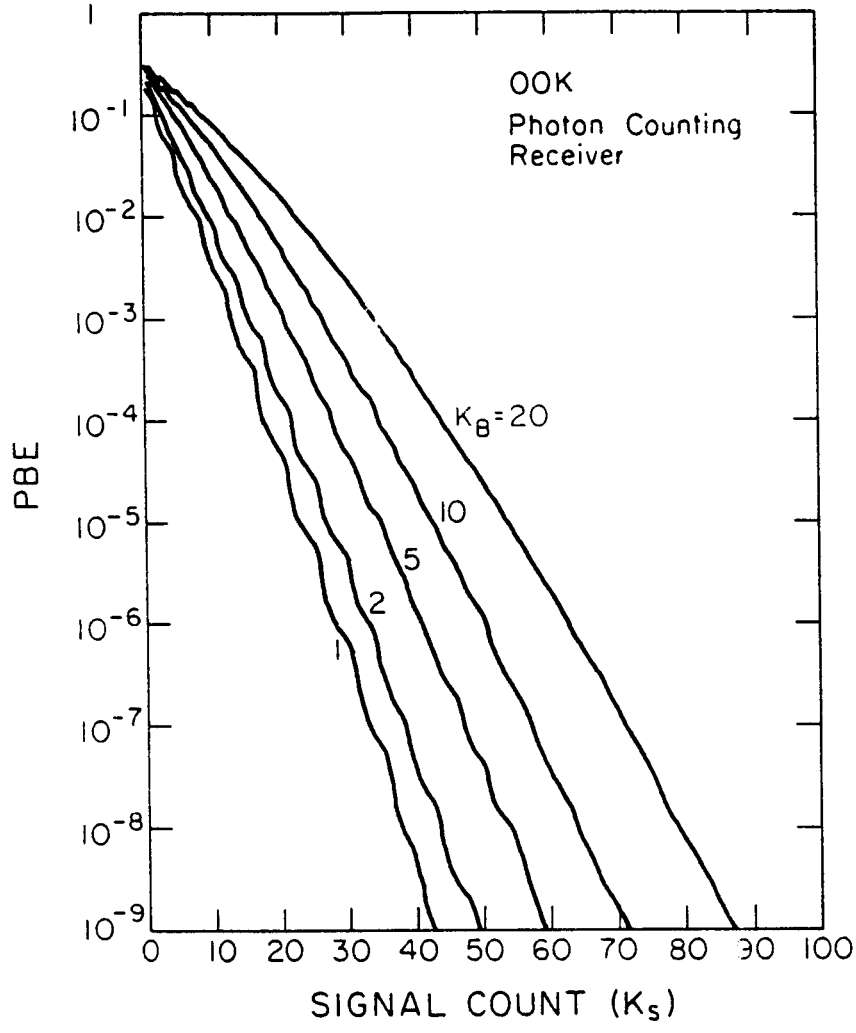


Fig.3.4. Probability of bit error (PBE) of an OOK modulated signal versus the signal photocount K_s.

vs. the expected signal photocount, K_S. The cusps in the curve are due to the fact that K_S and K_B are integer-valued variables.

OOK systems are simple and easy to construct. However, the decision threshold in Eq. (3.7) depends on both the signal and background photocounts. For systems where the signal and background intensities are not precisely known, optimal performance cannot be achieved. Instead, when operating under conditions where the signal and background strength are not known a priori, it is desirable to have a robust decoding algorithm that does not depend on signal and background photocount rates. An obvious alternative is to use the pulse-position modulation scheme.

PPM Systems The optimal PPM decoder counts the number of photons in each of the M time slots and chooses the slot with largest photocount. Figure 3.5 shows a practical implementation of this decoding algorithm. Because photocounts from the M time slots are compared, this decoding algorithm is insensitive to fluctuations in background and signal intensities. The probability of word error (PWE) for an M-ary PPM system using a quantum limited, photon counting detector is given by [3]

$$PWE(K_S,K_B,M) = 1 - \frac{1}{M}e^{-(K_s+MK_B)} - \sum_{k=1}^{\infty}\frac{(K_S+K_B)^k}{k!}e^{-(K_s+K_B)}\left[\sum_{j=0}^{k-1}\frac{K_B^j}{j!}e^{-K_B}\right]^{M-1}\left[\frac{(1+a)^M-1}{Ma}\right], \tag{3.9}$$

where

$$a = \frac{K_B^k}{k!\cdot\sum_{j=0}^{k-1}\frac{K_B^j}{j!}}.$$

The error rate given in Eq. (3.9) is difficult to evaluate. However, when the background count rate is much smaller than the signal count rate, Eq. (3.9) can be approximated by the union bound,

$$PWE(K_S,K_B,M) \le (M-1)\cdot PWE(K_S,K_B,2), \tag{3.10}$$

where $PWE(K_S,K_B,2)$ is the error rate of the binary PPM channel, given by [124]

$$PWE(K_S,K_B,2) = \frac{1}{2}\left[1 + Q(\sqrt{2K_B},\sqrt{2(K_S+K_B)}) - Q(\sqrt{2(K_S+K_B)},\sqrt{2K_B})\right], \tag{3.11}$$

and $Q(a,b)$ is the Marcum's Q function [30], [124],

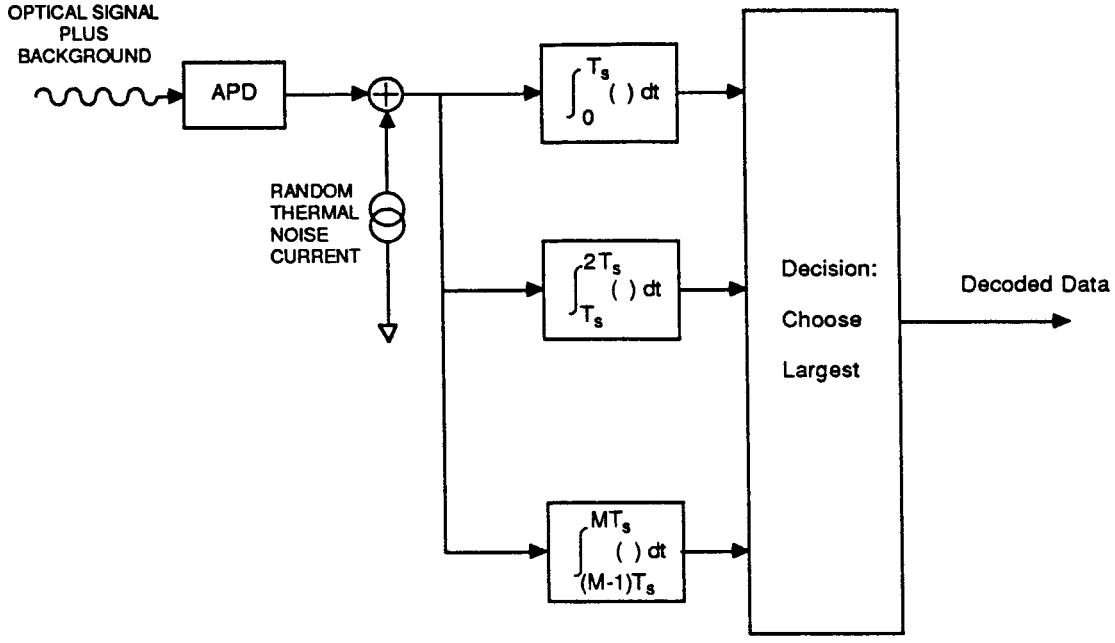


Fig.3.5. Model of a PPM demodulator.

$$Q(a,b) = \int_b^\infty e^{-\frac{a^2+x^2}{2}} I_0(ax)xdx \quad . \tag{3.12}$$

The error rate of a binary PPM channel can be further bounded by using the Chernoff bound [74], [133]

$$P_E(K_S, K_B, 2) \le e^{-(\sqrt{K_s+K_B}-\sqrt{K_B})^2} \quad . \tag{3.13}$$

When both signal and background counts are high, receiver photocounts in both signal and background slots can be approximated by Gaussian random variables and the resulting error probability is given by

$$P_E(K_S, K_B, 2) \approx \frac{1}{2}\text{erfc}\left[\frac{K_S}{\sqrt{2(K_S+2K_B)}}\right] \quad . \tag{3.14}$$

The validity of the union bound can be seen in Fig. 3.6 where the PWE of a 4-ary PPM system calculated using Eqs. (3.9) and (3.10) is plotted against the signal photocount K_S. Note that the union bound gives an excellent approximation of the error rate at high signal-to-noise ratios. A comparison of the binary PPM error rate calculated using the Chernoff bound, the Gaussian approximation, and the exact expression can be seen in Fig. 3.7. It is seen that the Chernoff bound is tight for small background counts, and the Gaussian approximation is good when both signal and background counts are high. The error probabilities given in Eqs. (3.9) and (3.10) are the probabilities of word error (PWE). For communication system design, it is usually desirable to specify the equivalent bit error probability for the system. The probability of bit error (PBE) for an M-ary PPM system can be related to the PWE by determining the probability that a given bit is incorrect after an erroneous decoding. In general, when an error occurs at the decoding, any bit will have a probability of (1/2) to be correctly decoded. However, out of the M possible codewords, one codeword will result in a correct decoding. Consequently, the PBE is related to the PWE by [3]

$$PBE = \frac{M}{2(M-1)} PWE \quad . \tag{3.15}$$

Figure 3.8 is a plot of the PBE vs the PPM order with the equivalent bit data rate and signal power fixed. Notice that the performance of the PPM improves with increasing PPM order, because at a fixed bit data rate R the word rate R_w actually decreases with increasing PPM order. Consequently, at a fixed signal power the signal photocount K_S increases with increasing PPM order. At the same time, since the equivalent slot period decreases with increasing PPM order, the background count per time slot, K_B, decreases. Therefore, the PBE of the M-ary PPM system decreases with increasing M. However, it should be noted that the system

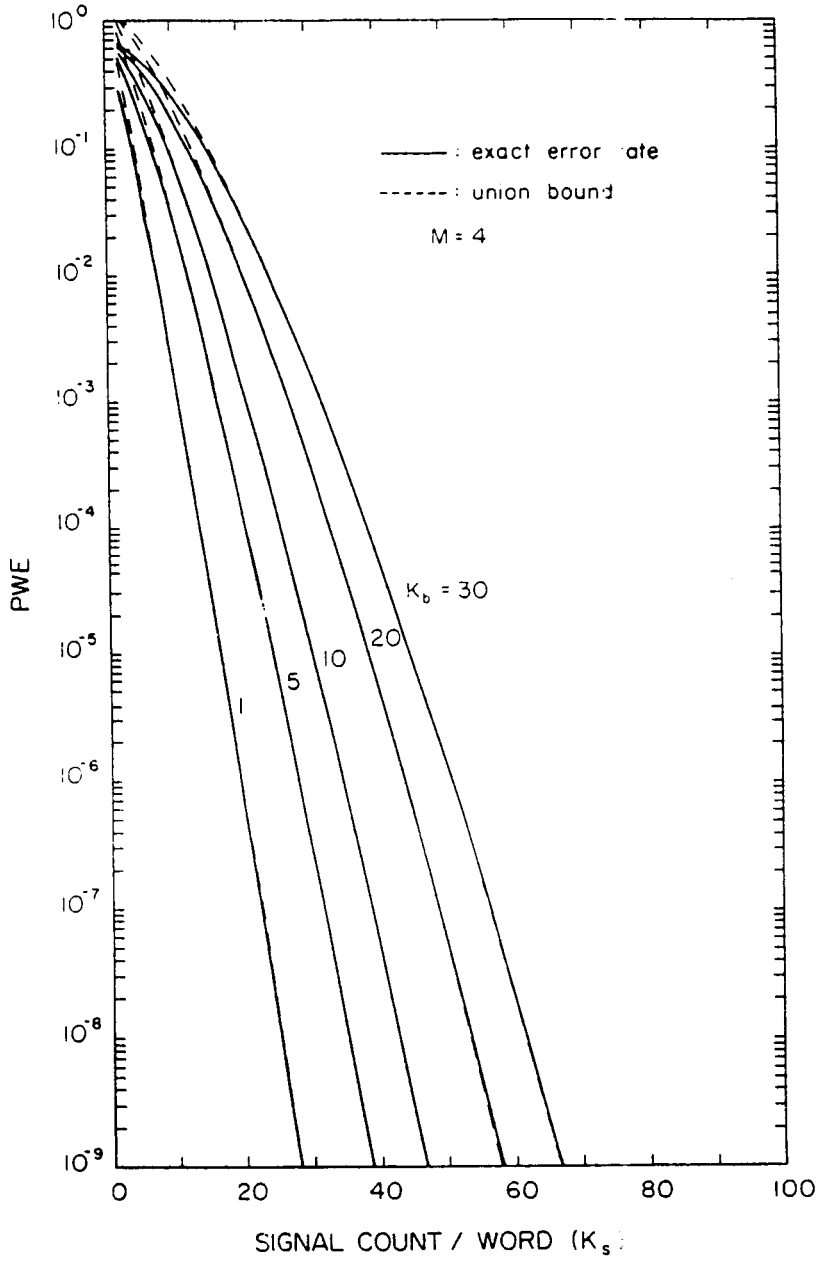


Fig.3.6. Probability of word error (PWE) of a 4-ary PPM system evaluated using the exact expression in Eq.(3.8), and the union bound in Eq.(3.9).

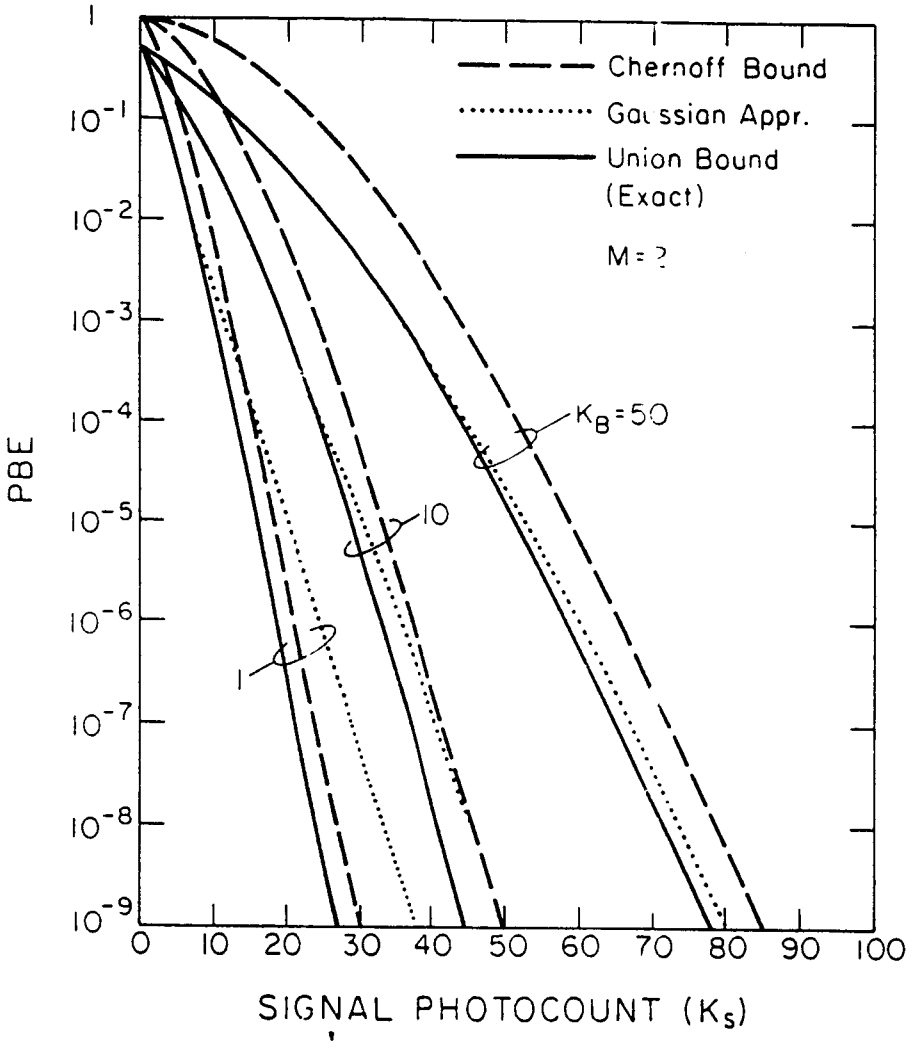


Fig.3.7. Probability of bit error (PBE) of binary PPM systems evaluated using the exact expression in Eq.(3.10), the Chernoff bound in Eq.(3.12), and the Gaussian approximation in Eq.(3.13).

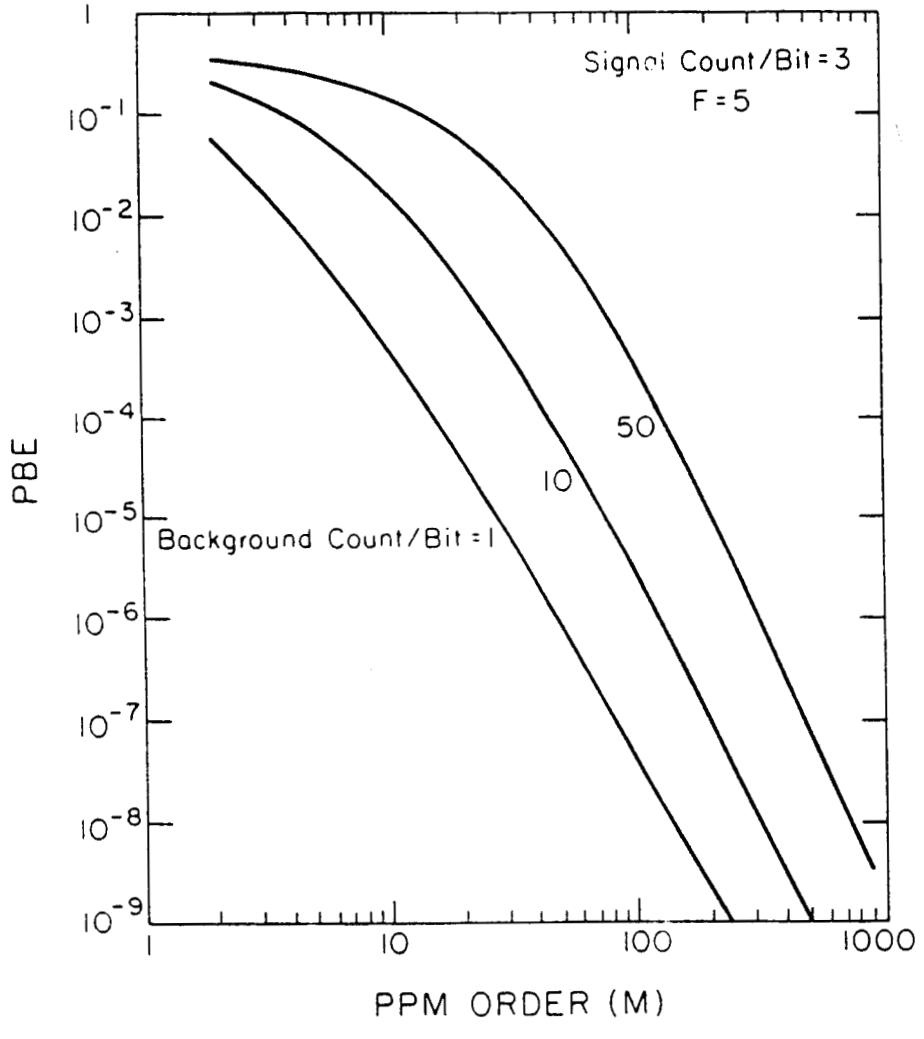


Fig.3.8. Probability of bit error (PBE) of the PPM system versus the PPM order.

bandwidth and decoder complexity required for higher order PPM increase rapidly with increasing M. Furthermore, in order to maintain a constant signal power, the peak transmitter power must increase linearly with increasing PPM order. In practice, both bandwidth and maximum peak power of lasers are limited by physical constraints. Consequently, it is not possible to increase the PPM order indefinitely.

It should also be noted that the improvement in system performance with increasing PPM order is possible only for systems where the transmitter laser has a perfect modulation extinction. In the presence of a finite extinction ratio m, receiver photocounts in the background slots are limited by the residual signal power. Therefore, even though the peak transmitter power increases with increasing PPM order, the background photocount also increases proportionally. Consequently, there is a limit at which further increase in the PPM order will only result in an increasing error rate. Fig. 3.9 is a plot of the PBE versus the PPM order for a system with a constant data rate and fixed signal and background powers. Note that for a nonzero extinction ratio, an optimum PPM order exists which minimizes the PBE.

3.4 Avalanche Detectors

Previous considerations of direct detection receiver performance have assumed that quantum limited, photon-counting detectors are used to detect the incoming optical radiation. In reality, when APDs are used in place of the photodetectors for the receivers shown in Figs. 3.3 and 3.5, the output of the integrator is no longer Poisson distributed because of the random avalanche gain. The detailed statistics of the APD output are difficult to characterize [75]-[77]. However, for communication receivers, the output of the integrator can be effectively approximated using Gaussian statistics [23]-[25]. The mean and variance of the integrator output can be written as [23]

$$E[X] = G(K_S + K_B + \frac{I_b T_s}{e}) + \frac{I_s T_s}{e} \quad , \tag{3.16}$$

$$var(X) = G^2 F(K_S + K_B + \frac{I_b}{e}) + \frac{I_s T_s}{e} + K_{th}^2 \quad , \tag{3.17}$$

where

$$X = \int_0^{T_s} i(t)dt \tag{3.18}$$

is the integral of APD output over the period T_s. I_b, I_s are the gain dependent and gain independent portions of

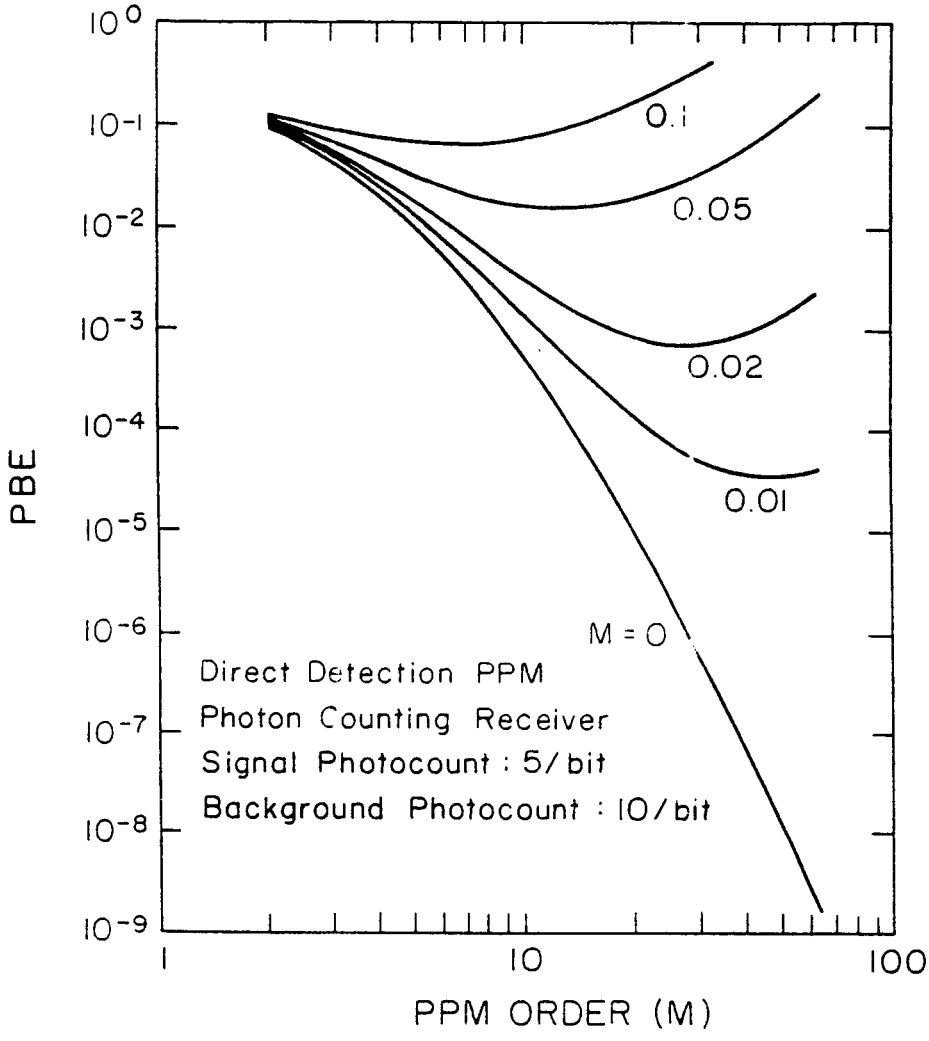


Fig.3.9. Probability of bit error (PBE) versus the PPM order for systems where the transmitter laser has a nonzero extinction ratio.

the detector dark current, and K_{th}^2 is the variance due to thermal noise. K_{th}^2 can be related to the detector parameter by

$$K_{th}^2 = \frac{2k_B T_{eq} T_s}{R_L e^2} \quad . \tag{3.19}$$

By approximating the integrator output as Gaussian random variables, the PBE of an OOK system can be written as

$$P_E = \frac{1/2}{\sqrt{2\pi}\sigma_0}\int_\theta^\infty e^{-(x-\mu_0)^2/2\sigma_0^2}dx + \frac{1/2}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^\theta e^{-(x-\mu_1)^2/2\sigma_1^2}dx \tag{3.20}$$

where

$$\mu_0 = G\left(K_B + \frac{I_b T_s}{e}\right) + \frac{I_s T_s}{e} \ ,$$

$$\mu_1 = G\left(K_S + K_B + \frac{I_b T_s}{e}\right) + \frac{I_s T_s}{e} \ ,$$

$$\sigma_0^2 = G^2 F\left(K_B + \frac{I_b T_s}{e}\right) + \frac{I_s T_s}{e} + K_{th}^2 \ ,$$

$$\sigma_1^2 = G^2 F\left(K_S + K_B + \frac{I_b T_s}{e}\right) + \frac{I_s T_s}{e} + K_{th}^2$$

are the means and variances of the integrator output in the absence and the presence of a signal pulse, respectively, and it has been assumed that both binary zero and binary one are equally likely to be transmitted. The threshold θ can be related to the integrator output by

$$\log\left[\frac{\sigma_1}{\sigma_0}\right] - \frac{(\theta-\mu_0)^2}{2\sigma_0^2} + \frac{(\theta-\mu_1)^2}{2\sigma_1^2} = 0 \ . \tag{3.21}$$

The PBE of an APD-based OOK system is calculated using the APD parameters shown in Table 2.2 and plotted in Fig. 3.10 versus the signal photocount K_S. It is seen that a significantly higher signal power is needed to compensate for the avalanche noise.

For an M-PPM system, each of the M integrator outputs in Fig. 3.5 is Gaussian distributed. Since only one time slot contains the signal pulse, the integrator outputs for the remaining $(M-1)$ time slots are identically distributed. Consequently, the PWE of the M-PPM channel can be written as

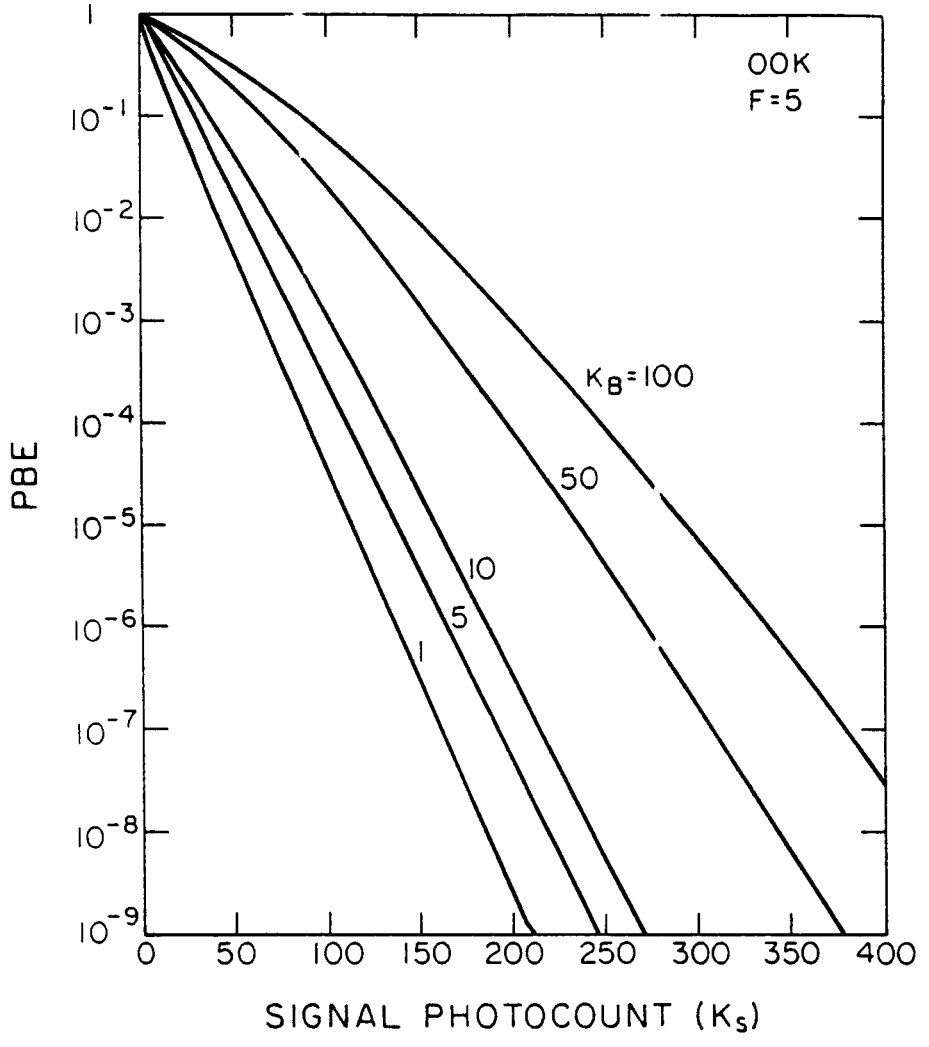


Fig.3.10. Probability of bit error (PBE) of an OOK system using avalanche photodiode.

$$P_E = 1-\left\{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_1}e^{-(x-\mu_1)^2/2\sigma_1^2}\left[\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_0}e^{-(y-\mu_0)^2/2\sigma_0^2}dy\right]^{M-1}dx\right\} . \tag{3.22}$$

The signal-to-noise ratio (SNR) of the APD-based PPM receiver can be defined as

$$\rho = \frac{(\mu_1-\mu_0)^2}{\sigma_1^2+\sigma_0^2} = \frac{G^2K_S^2}{G^2F\,(K_S+2K_B+2\frac{I_bT_s}{e})+2\frac{I_sT_s}{e}+2K_{th}^2} . \tag{3.23}$$

When $M=2$, the PBE of the binary PPM system can be written as

$$P_E = \frac{1}{2}\mathrm{erfc}(\sqrt{\rho/2}) . \tag{3.24}$$

The PBE given by Eq. (3.21) is calculated using the parameters shown in Table 2.2 and and plotted against the signal photocount K_S in Fig. 3.11. Again, it is seen that a much higher signal power is needed to overcome the avalanche noise.

Given the receiver design, the gain of the APD is an important parameter which must be properly specified. The higher the detector gain, the smaller the effect of thermal noise on the receiver SNR. However, higher detector gain also implies larger excess noise factor which in turn can reduce the receiver SNR. Consequently, given the receiver parameter and the signal and background strength, there is an optimum detector gain which maximizes the detector SNR. This optimum detector gain can be calculated by differentiating Eq. (3.23) with respect to the detector gain G, and solving the resulting equation. By using the fact that the excess noise factor F is related to the detector gain G by Eq. (2.6), and the assumption that the detector dark current is small compared to the detector thermal noise, the optimum APD gain can be approximated by

$$G_{\mathrm{opt}} = \left[\frac{8k_BT_{eq}T_s/R_Le^2}{k_{\mathrm{eff}}K_S}\right]^{1/3} , \quad K_B \ll K_S . \tag{3.25}$$

Given the receiver design and the signal strength, an optimum APD gain can be chosen to maximize the system performance.

The effect of a finite modulation extinction ratio on the PPM channel can also be analyzed using Eq. (3.23). In the presence of a finite extinction ratio m, the signal and background photocounts can be related to the average signal and background power by

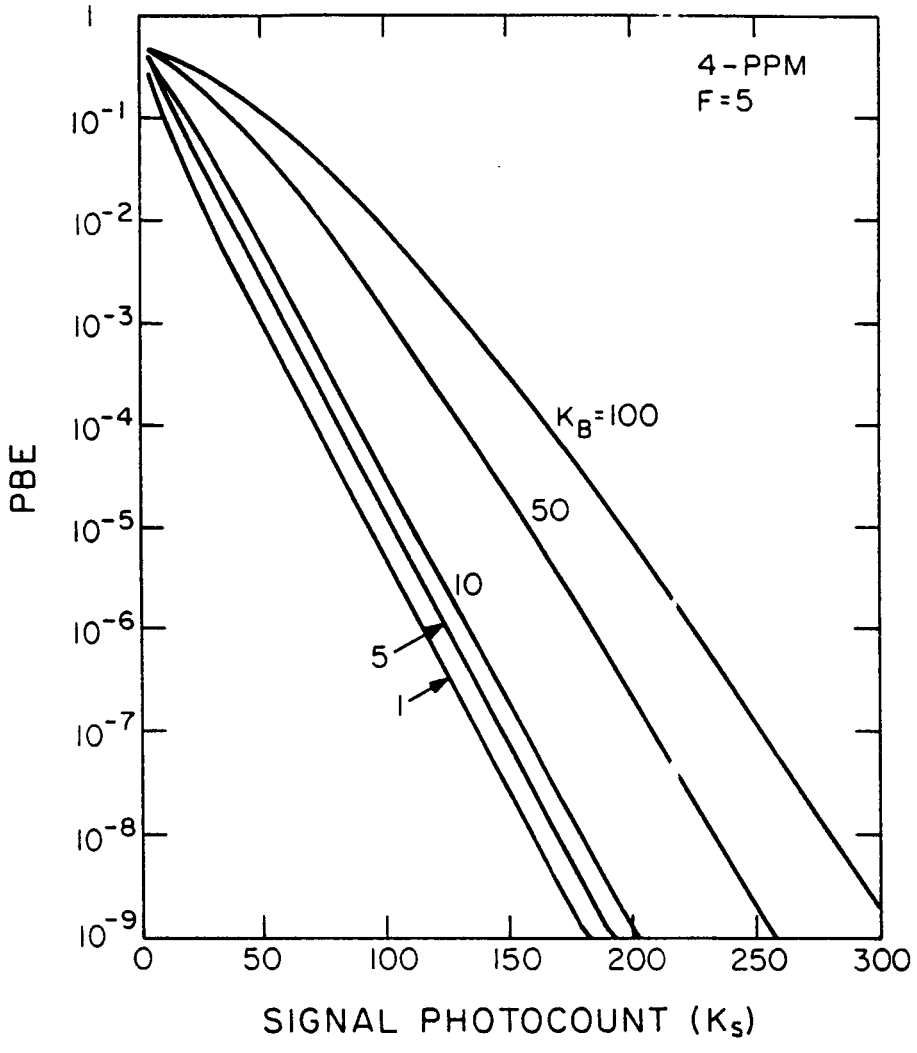


Fig.3.11. Probability of bit error (PBE) of an APD-based 4-ary PPM system versus the signal photocount K_S.

$$K_S = \frac{\alpha P_S T_w}{1+m(M-1)} \quad ,$$

$$K_B = mK_S + \alpha P_B T_s \quad , \tag{3.26}$$

where M is the order of PPM, and $\alpha=\eta/h\nu$ is the detector responsivity. The word period T_w can be written in terms of the data rate R and the order of PPM as

$$T_w = \log_2 M / R \quad . \tag{3.27}$$

By substituting Eqs. (3.26) and (3.27) into Eq. (3.23), the receiver SNR can be written as a function of the signal and background power, as well as the data rate:

$$\rho = \frac{\alpha^2 P_S^2 M \log_2 M / R\,[1+m(M-1)]^2}{F\left[(1+2m)\frac{\alpha P_S M}{1+m(M-1)} + 2\alpha P_B + 2\frac{I_b}{e}\right] + 2\frac{I_b}{e} + 2\frac{2k_B T_{eq}}{R_L e^2}} \quad . \tag{3.28}$$

Note that the receiver SNR increases with increasing signal power and decreasing background power and data rate. When the modulation extinction ratio m is equal to 0, the receiver SNR increases asymptotically as $\log_2 M$. In the presence of a finite extinction ratio, however, the receiver SNR first increases with increasing M until an optimum SNR is reached, and further increase in M will only result in a smaller SNR. In general, this optimum PPM order can be calculated by differentiating Eq. (3.28) with respect to M and solving the resulting equation. In the limit where the detector dark current, thermal noise, and background power are negligible compared to the signal power, the optimum M is seen to satisfy

$$M(\ln M - 1) = \frac{1-m}{m} \quad . \tag{3.29}$$

It should be noted that in deriving Eq. (3.27), it has been assumed that the average signal power is maintained at a constant level regardless of the PPM order. This implies that the peak transmitter power increases linearly with the PPM order M. In practice, because of device limitations, the peak power of lasers is often limited. For such systems, the average transmitter power is a function of PPM order M, and the optimal PPM order is determined not only by the extinction ratio m, but also by the ratio between peak and average laser power.

3.5 Summary

The performance of direct detection systems using OOK and PPM have been investigated. It is seen that PPM systems offer comparable PBE at a much lower signal power requirement. When the transmitter laser has

a perfect modulation extinction, the PBE of the PPM systems is seen to improve with increasing PPM order at constant signal power. Furthermore, the PPM signal can be decoded regardless of the signal and background strength. The combination of these advantages makes PPM highly attractive for spaceborne implementations. For these reason, it will be assumed in the following that PPM is used whenever direct detection systems are employed. It should be noted that the performance analysis for PPM systems carried out in this chapter have taken into account only the detector shot noise and thermal noise. Other noises that are present in a realistic PPM system, such as spatial and temporal tracking errors, have been ignored. In the presence of these errors, the performance of the PPM system can be seriously deteriorated. The performance analysis of PPM systems under spatial and temporal tracking errors will be carried out in Chapters 5 and 6, respectively.

4. HETERODYNE SYSTEMS

The direct detection PPM system described in the last chapter is structurally simple and easy to implement. However, the large power requirement of a practical PPM link (≈150-200 photons for 4-ary PPM at 10^{-9} error rate) presents a serious disadvantage in an ISL application where the available signal power is limited. Alternatively, optical communication systems can be implemented using heterodyne receivers. Compared to direct detection PPM systems, heterodyne systems offer an improved receiver sensitivity and better background noise rejection. In addition, frequency and phase encodings can be used in heterodyne systems to further improve the system performance. However, heterodyne systems are very sensitive to the phase fluctuation in the optical carrier and their performances can be seriously deteriorated in the presence of large carrier phase noise. Consequently, for practical heterodyne systems, the spectral linewidths of lasers must be carefully controlled. The purpose of this chapter is to analyze the performance of the heterodyne system both in the absence and the presence of carrier phase noise.

4.1 Heterodyne Modulation Schemes

Information encoding in optical heterodyne systems is similar to its counterpart in RF carrier modulation systems. Modulation schemes such as optical amplitude-shift keying (ASK), frequency-shift keying (FSK), and phase-shift keying (PSK) can be used to transmit messages over the heterodyne channel.

Amplitude shift keying (ASK) is the simplest form of digital carrier modulation. The information is encoded onto the optical carrier by varying the amplitude of the transmitted signal. The ASK schemes commonly used in digital optical transmissions include on-off keying (OOK) and pulse-position modulation (PPM). The transmitted signal corresponding to these signaling schemes can be modeled as:

$$s(t)=\begin{cases} A\sin(2\pi f_c t+\phi_c) & 0\le t\le T_s \quad \text{for a binary one, and} \\ 0 & 0\le t\le T_s \quad \text{for a binary zero,} \end{cases} \qquad \text{(OOK)} \quad (4.1)$$

$$s_k(t)=\begin{cases} A\sin(2\pi f_c t+\phi_c) & kT_s\le t\le (k+1)T_s \\ 0 & \text{elsewhere} \end{cases} \qquad 0\le t\le MT_s \ , \qquad \text{(PPM)} \quad (4.2)$$

where $s_k(t)$ is transmitted when the message codeword is equal to k, and A, f_c and ϕ_c are the amplitude, frequency and phase of the transmitted carrier, respectively.

In practice, ASK signals can be encoded by either externally modulating a frequency stabilized laser, or by directly modulating the injection current density of a semiconductor laser. External intensity modulation can be implemented using the electrooptic [127]-[129] or electro-absorption (Franz-Keldysh) effect [130]-[132]. External modulation of lasers provides a better frequency stability but, because the modulator severely attenuates the signal, a smaller signal power is available at the output. In contrast, direct modulation of semiconductor lasers provides a higher output power, a better modulation efficiency, and a larger modulation bandwidth. However, because of the spurious frequency modulation due to changes in the injection current density [134], direct modulation of semiconductor lasers suffers from the poor temporal coherence of the transmitted signal. A frequency stabilized external post amplifier can be used to reduce the phase fluctuation [135], but at the cost of increased system complexity. Another disadvantage of direct modulation is that, in order to facilitate rapid switching, the device is usually biased just below the lasing threshold. This leads to a small leakage of the optical signal and a finite extinction ratio. The presence of this finite extinction ratio imposes a ceiling on the maximum SNR that can be achieved at the decoder.

Frequency shift keying (FSK) is by far the preferred method for implementing the optical heterodyne system. Unlike the ASK system, the information is encoded by varying the signal frequency. The transmitted signal for an FSK system can be modeled as

$$s_k(t) = A\sin(2\pi f_k t + \phi_c)\ ,\ 0 \le t \le T_s\ , \qquad \text{(FSK)} \quad (4.3)$$

where the signal frequency is chosen from a set $\{f_0 f_1, \ldots, f_{M-1}\}$ of frequencies such that f_k is transmitted when the message codeword is equal to k. The correlation coefficients ρ_{ij} between the i^{th} and j^{th} signal waveforms can be defined as

$$\rho_{ij} = \int_0^{T_s} s_i(t)s_j(t)dt / \sqrt{E_i E_j}\ , \qquad (4.4)$$

where

$$E_i = \int_0^{T_s} s_i^2(t)dt \qquad (4.5)$$

is the energy of the i^{th} signal waveform. For an M-ary system, it is desirable that signal waveforms corresponding to different codewords are orthogonal. The correlation coefficients between signals for an orthogonal system are equal to zero if $i \neq j$, and equal to 1 if $i = j$. Orthogonality can be realized for an MFSK

system if the tone spacing between signaling frequencies is a multiple of the symbol frequency $f_s=1/T_s$ [28]-[31].

Frequency modulation of the optical carrier can be accomplished in practice by either internal or external modulation techniques. External frequency modulation of laser output can be realized using electrooptic [136] or acoustooptic [137] effects. As was the case with external amplitude modulation, these techniques suffer from large insertion losses and limited modulation bandwidths. Direct frequency modulation of semiconductor lasers can be achieved using electrooptic [138] or photoelastic [139] modulations or by directly modulating the injection current density [140], [141]. These different frequency modulation techniques are compared in Table 4.1 [32]. Among them, the injection current modulated system is the easiest to implement.

In addition to ASK and FSK, the optical carrier can also be phase modulated. By changing the index of refraction of the medium with an external voltage, the phase of the transmitted signal can be controlled using the electro-optic effect [142]-[144]. The transmitted signal for a PSK system can be modeled as

$$s_k(t) = A\sin(2\pi f_c t+\phi_k)\ , \qquad \text{(PSK)} \quad (4.6)$$

where ϕ_k is chosen from a set of phases $\{\phi_0,\phi_1,\ldots,\phi_{M-1}\}$. When $M=2$ and $\phi_1-\phi_0=\pi$, the amplitude of the signal when a binary one is transmitted is the negative of that when a binary zero is sent. This particular BPSK scheme is also known as the antipodal ASK, which has the property that the correlation coefficient between the two signals is equal to -1.

4.2 Demodulation in the Absence of Phase Noise

At the receiver, the optical signal is mixed with a strong LO output and then photodetected. This heterodyne process shifts the spectrum of the input signal from the optical frequency down to an IF frequency, f_{IF}, where it can be easily processed using conventional RF electronics. When both temporal and spatial trackings can be perfectly realized, the output of the photodetector is often modeled as the sum of an IF signal, $s(t)$, and an additive shot noise, $n_s(t)$,

$$r(t) = s(t)+n_s(t),\ 0\le t\le T_s\ . \qquad (4.7)$$

At sufficiently high LO power, the additive noise, $n_s(t)$, which is is due primarily to the LO shot noise, can be

Table 4.1 Comparison of FSK schemes.

Method	External Modulators		Internal Modulation		
	Electrooptic	Acoustooptic	Injection Current	Electrooptic	Photoelastic
Modulation frequency	1-2 GHz (normal type) >2 GHz (travelling wave)	50-100 MHz (two dimensional) 200-300 MHz (three dimensional)	<2-3 GHz	<2 GHz	<3 GHz
Frequency shift	500 MHz - 1 GHz	1 GHz	100-400 MHz/mA	0.7-3 GHz/V	200 MHz/V
Problems	Complex electrode configuration Largee insertionn loss	Large insertion loss Limited modulation bandwidth Deflection angle modulation	Strong frequency characteristics Relaxation oscillation Spurious AM	Complicated structure Spurious AM and FM	Baseband modulation not feasible

modeled as an additive white Gaussian noise (AWGN) with power spectral density

$$N_s \simeq \eta P_{LO}/h\nu \ . \tag{4.8}$$

Consequently, the heterodyne channel can be effectively modeled as an AWGN channel, and the demodulation of the IF signal is similar to that of an RF carrier modulation system. In the case where an M-ary signaling scheme is used to modulate the optical signal, $s(t)$ can be any one of the M signal waveforms $\{s_k(t)\}$. The purpose of the IF demodulator, therefore, is to determine which of the signals $\{s_k(t)\}$ was transmitted given the receiver output $r(t)$, $0 \le t \le T_s$.

In general, the carrier modulated signal can be demodulated using either phase coherent or incoherent techniques. Coherent demodulation of the IF signal can be achieved by using a matched filter or, equivalently, the correlation detector [28]-[31]. Both methods can be implemented directly at the carrier frequency, or at the baseband frequency by first demodulating the IF signal using a coherent reference carrier. Figure 4.1 shows the alternative implementations of phase coherent correlation detectors. For M-ary phase coherent systems, M correlators are used to demodulate the IF signal, and the decoder simply picks the channel corresponding to the largest correlator output. Figure 4.2 shows the block diagram of a M-ary phase coherent receiver implemented using correlation detectors. The probability of error for the coherent demodulator can be written as

$$P_E = \sum_{k=0}^{M-1} \pi_k P_E(k) \ , \tag{4.9}$$

where $P_E(k)$ is the probability of decoding error given that $s_k(t)$ was transmitted, and π_k is the prior probability of transmitting $s_k(t)$. The conditional probability $P_E(k)$ can be evaluated by realizing that the outputs of the correlators are Gaussian distributed with statistics

$$\begin{aligned} E[x_i \mid s_k(t)] &= \int_0^{T_s} s_i(t) s_k(t) dt = \sqrt{E_i E_k}\,\rho_{ik} \ , \\ Var(x_i \mid s_k(t)) &= E_i N_s \ , \end{aligned} \tag{4.10}$$

where x_i is the output of the i^{th} correlator. A decoding error occurs when the output of the j^{th} correlator, x_j, $j \neq i$, is greater than x_k. Using the fact that $\mathbf{x}=(x_0, x_1, \cdots, x_{M-1})$ are jointly Gaussian, $P_E(k)$ can be written as [29]

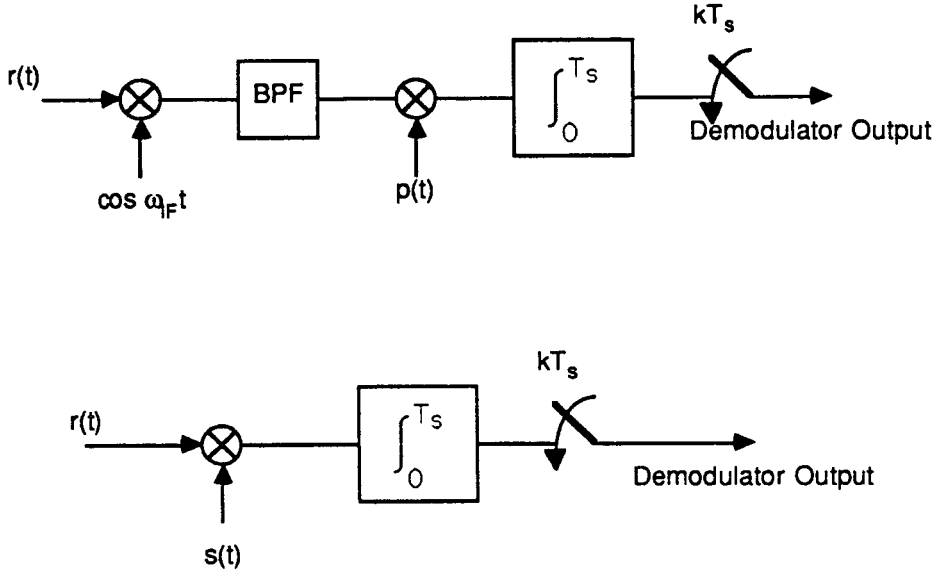


Fig.4.1. Alternative implementations of the correlation detector.

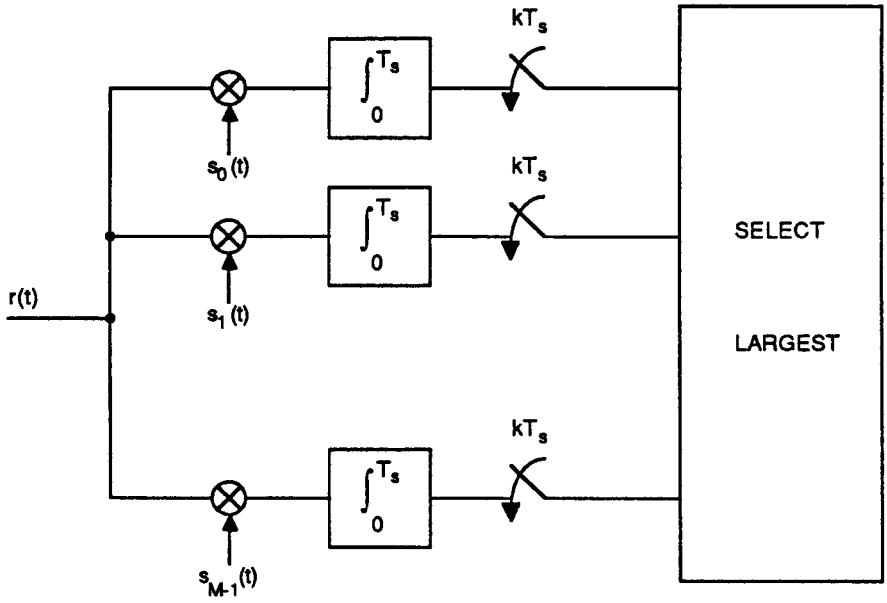


Fig.4.2. Block diagram of an M-ary coherent receiver.

$$P_E(k) = 1 - \int_{-\infty}^{\infty} dx_i \cdot \prod_{i \neq k}^{M} \int_{-\infty}^{x_i + \sqrt{E_i/N_s}(1-\rho_{ik})} dx_k \frac{1}{(2\pi)^{M/2} |\rho_{ij}|^{1/2}} e^{-\frac{1}{2}\mathbf{x}\cdot\boldsymbol{\rho}^{-1}\cdot\mathbf{x}^T} dx_1\, dx_2, \ldots, dx_M \quad , \tag{4.11}$$

where $|\rho_{ik}|$ is the determinant of the correlation matrix

$$\rho = \begin{bmatrix} \rho_{0,0} & \rho_{1,0} & \cdots & \rho_{M-1,0} \\ \rho_{0,1} & \rho_{1,1} & \cdots & \rho_{M-1,1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \rho_{0,M-1} & \rho_{1,M-1} & \cdots & \rho_{M-1,M-1} \end{bmatrix} \quad . \tag{4.12}$$

A special case of interest occurs when the M waveforms, $\{s_k(t)\}$, are orthogonal and have equal energy and equal priors. In this case $\rho_{ik}=\delta_{ik}$, $E_k=E$, and the probability of error reduces to [29]

$$P_E = 1 - \int_{-\infty}^{\infty} \frac{1}{(\sqrt{2\pi})^M} e^{-x^2/2} \cdot \left[\int_{-\infty}^{x+\sqrt{2\beta_c}} e^{-y^2/2} \right]^{M-1} dx \quad , \qquad \text{(Orthogonal Systems)} \tag{4.13}$$

where

$$\beta_c = E/2N_s = A^2 T_s / 4N_s \tag{4.14}$$

is the signal-to-noise ratio of the coherent demodulator. For binary orthogonal systems, (4.14) can be further reduced to

$$P_E = \frac{1}{2}\text{erfc}(-\sqrt{\beta_c/2}) \quad . \tag{4.15}$$

Another important system employing the coherent demodulator is the binary antipodal system (BPSK). For this system the correlation matrix has zero determinant and, consequently, the two correlator outputs are degenerated. Decoding of the binary antipodal signaling scheme can therefore be accomplished using only one correlator. Figure 4.3 depicts the block diagram of the BPSK decoder. Note that a decoding error occurs if $x<0$ when $s(t)$ is transmitted, or if $x>0$ when $-s(t)$ is transmitted. Using the fact that x has mean $E[x \mid \pm s(t)] = \pm E$ and variance $\sigma_x^2 = N_s E$, the probability of error for the BPSK decoder can be written as

$$P_E = \frac{1}{2}\text{erfc}(-\sqrt{\beta_c}) \quad . \tag{4.16}$$

Comparison of Eqs. (4.15) and (4.16) shows that the antipodal system (BPSK) offers a 3 dB improvement in SNR over the orthogonal system (BFSK).

Table 4.2 is a list of probability of errors for various coherently demodulated heterodyne systems in terms of the signal to noise ratio β_c. The detection threshold for the OOK system has been chosen to minimize

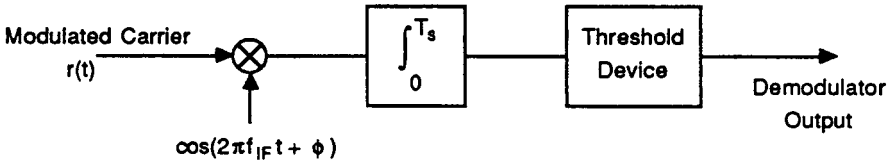


Fig.4.3. Block diagram of a BPSK receiver.

Table 4.2 Error probabilities of the coherent heterodyne channel.

Modulation Format	Error Probability
OOK	$\frac{1}{2}\mathrm{erfc}(\sqrt{\beta_c/4})$
MPPM	$1-\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}.[\int_{-\infty}^{x+\sqrt{2\beta_c}}e^{-y^2/2}dy\,]^{M-1}dx$
MFSK	$1-\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}.[\int_{-\infty}^{x+\sqrt{2\beta_c}}e^{-y^2/2}dy\,]^{M-1}dx$
BPSK	$\frac{1}{2}\mathrm{erfc}(\sqrt{\beta_c})$

the error probability. For a heterodyne MPPM system, it is assumed that the correlation detection is carried out for each of the M time slots. Since only one slot contains the signal pulse, the performance of the MPPM system is similar to that of an MFSK system. Note that by substituting the expressions for signal amplitude and noise density from (2.30) and (4.8) into (4.14), the IF SNR can be written as

$$\beta_c = \eta_{het}\left[\frac{\eta}{h\nu}\right] P_s T_s = \eta_{het} K_s \ , \tag{4.17}$$

which is just the product of the heterodyne efficiency and the number of (expected) signal photons detected over the period T_s.

Phase coherent demodulation is optimal in the sense that both signal amplitude and phase are used to decode the transmitted data. However, in order to perform the correlation detection, the receiver must have a prior knowledge of the carrier's frequency and phase so that the local reference signals $\{s_k(t), k=0,1,...,M-1\}$ can be generated. This carrier reference must be acquired through a separate synchronization circuit. Alternatively, carrier modulated signals can be demodulated using noncoherent techniques such as envelope or differentially coherent detections. These techniques do not require the presence of a reference carrier. Therefore, the receiver design can be considerably simplified. In addition, for systems where the carrier reference is difficult to generate, such as in the presence of a large, randomly fluctuating phase noise, noncoherent systems have the distinct advantage since they require no phase information to perform the decoding.

Figure 4.4(a) shows the structure of an envelope detector centered at frequency f_k. The output of the detector is given by

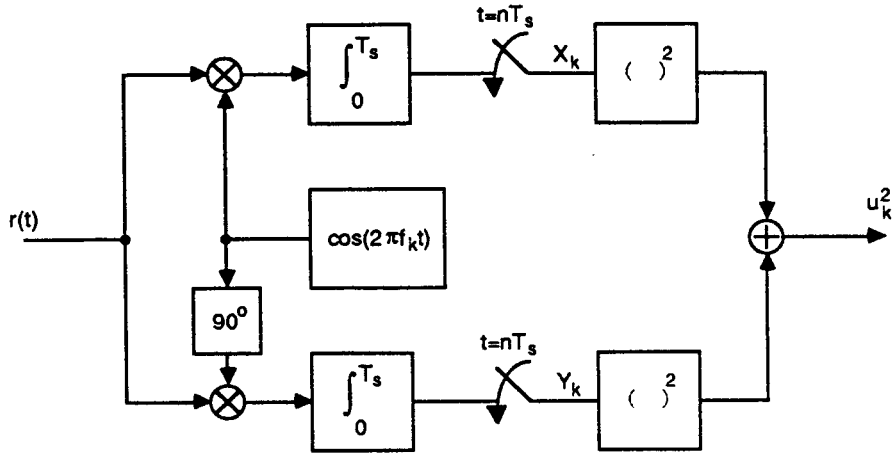
$$u_k^2 = X_k^2 + Y_k^2 \ , \tag{4.18}$$

where

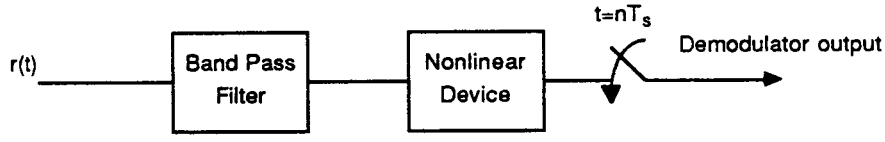
$$X_k = \frac{2}{T_s}\int_0^{T_s} r(t)\cos 2\pi f_k t \ dt \ ,$$

$$Y_k = \frac{2}{T_s}\int_0^{T_s} r(t)\sin 2\pi f_k t \ dt \ . \tag{4.19}$$

Note that u_k^2 can be written as



(a)



(b)

Fig.4.4. Alternative implementations of the envelope detector.

$$u_k^2 = \left| \frac{2}{T_s} \int_0^{T_s} r(t) e^{-i 2\pi f_k t} dt \right|^2 , \tag{4.20}$$

which is simply the Fourier transform magnitude square of $r(t)$ at $f = f_k$. This spectral analyzer decoder has been shown to be optimal for decoding MFSK signals when the carrier phase is unavailable [28], [145]. In practice, envelope detectors are usually implemented using the structure shown in Figure 4.4(b). In the absence of a rapidly fluctuating carrier phase, these alternative configurations are equivalent.

The probability distribution of u_k, given that $s_j(t) = A \sin(2\pi f_j t + \phi)$ was transmitted, can be written as:

$$P_k(u_k) = \frac{u_k}{N_0} e^{-\frac{A_{kj}^2 + u_k^2}{2N_0}} I_0(u_k A_{kj} / N_0) , \tag{4.21}$$

where

$$N_0 = 2N_s / T_s . \tag{4.22}$$

The parameter A_{kj}^2 in (4.21) is given by

$$\begin{aligned} A_{kj}^2 &= \left[\frac{2}{T_s} \int_0^{T_s} A \cos(2\pi f_j t + \phi) \cos 2\pi f_k t dt \right]^2 + \left[\frac{2}{T_s} \int_0^{T_s} A \cos(2\pi f_j t + \phi) \sin 2\pi f_k t dt \right]^2 \\ &= \left[\frac{1}{T_s} \int_0^{T_s} A \cos[2\pi \nu_{kj} t + \phi] dt \right]^2 + \left[\frac{1}{T_s} \int_0^{T_s} A \sin[2\pi \nu_{kj} t + \phi] dt \right]^2 , \end{aligned} \tag{4.23}$$

where $\nu_{kj} = f_k - f_j$ is the tone spacing between j^{th} and k^{th} frequencies, and we have assumed that the sum frequency terms have negligible contributions to the integral. When $A_{kj}^2 \neq 0$, the above probability density function is known as the Rice-Nakagami distribution, or simply the Rician distribution. When either the signal amplitude is zero, or when the transmitted signal frequency is orthogonal to the detector center frequency, the parameter A_{kj}^2 is equal to zero. In this case $P(u_k)$ reduces to a Rayleigh distribution.

Given the probability density function of the envelope detector output, the probability of error for an OOK system can be written as

$$\begin{aligned} P_E &= \pi_0 \int_\theta^\infty \frac{u_k}{N_0} e^{-u_k^2 / 2N_0} du_k + \pi_1 \int_0^\theta \frac{u_k}{N_0} e^{-\frac{u_k^2 + A_{kk}^2}{2N_0}} I_0(u_k A_{kk} / N_0) du_k \\ &= \pi_0 e^{-\theta^2 / 2N_0} + \pi_1 \left[1 - Q\left(\frac{A_{kk}}{\sqrt{N_0}}, \frac{\theta}{\sqrt{N_0}} \right) \right] , \end{aligned} \tag{4.24}$$

where θ is the decision threshold, and $Q(\alpha,\beta)$ is the Marcum's Q function [30]. The first term on the right-handed side of Eq. (4.24) is the probability that $u_k>\theta$ when no signal is transmitted, and the second term is the probability that $u_k<\theta$ when a binary one is sent. The PBE in Eq. (4.24) can be minimized by choosing the threshold θ to satisfy

$$\theta = \frac{N_0}{A_{kk}} I_0^{-1}\left[\frac{\pi_0}{\pi_1} e^{A_{kk}^2/2N_0}\right] . \tag{4.25}$$

For M-ary FSK systems, M envelope detectors, each tuned to a particular signal frequency, are used to demodulate the signal. The decoder then selects the channel with the largest detected envelope. The probability of error for this system, given that f_k was transmitted, can be written as

$$P_E(k) = 1 - \int_0^\infty P_k(u_k)\left[\prod_{k\neq i}\int_0^{u_k} P_i(u_i)du_i\right] du_k . \tag{4.26}$$

When the frequencies transmitted are orthogonal, the outputs of the envelope detectors not matched to the signal frequency are identically distributed with Rayleigh densities. The probability of decoding error can therefore be evaluated as

$$P_E(k) = 1 - \int_0^\infty \frac{u_k}{N_0} e^{-\frac{A_{kk}^2+u_k^2}{2N_0}} I_0\left(\frac{u_k A_{kk}}{N_0}\right)\cdot\left[\int_0^{u_k}\frac{u_i}{N_0} e^{-\frac{u_i^2}{2N_0}} du_i\right]^{M-1} du_k$$

$$= 1 - \int_0^\infty e^{-(x+\beta_{kk})} I_0\left(\sqrt{4\beta_{kk}x}\right)\cdot[1-e^{-x}]^{M-1} dx \tag{4.27}$$

$$= \sum_{j=1}^{M-1} \frac{(-1)^{j+1}}{j+1}\binom{M-1}{j} e^{-j\beta_{kk}/(j+1)} ,$$

where

$$\beta_{kk} = A_{kk}^2 T_s/4N_s . \tag{4.28}$$

For systems with equal energy pulses, $\beta_{kk}=\beta\equiv A^2T_s/4N_s$, $k=0,1,\ldots,M-1$, and the unconditioned probability of error is equal to the $P_E(k)$ given by Eq. (4.27). When $M=2$, the probability of error for the binary system can be further reduced to

$$P_E = \frac{1}{2} e^{-\beta/2} \tag{4.29}$$

For nonorthogonal FSK systems, the probability of error given by Eq. (4.26) is in general difficult to calculate. However, when $M=2$, the error probability can be written in a closed-form [30]:

$$P_E(k) = \frac{1}{2}[1-Q(\sqrt{\beta_{kk}},\sqrt{\beta_{jk}})+Q(\sqrt{\beta_{jk}},\sqrt{\beta_{kk}})] \quad j,k=0,1\ ;\ \ j\neq k\ , \tag{4.30}$$

where $\beta_{ij}=A_{jk}^2 T_s/4N_s$. Note that for an orthogonal system, $\beta_{jk}=0$ when $j\neq k$, and Eq. (4.30) reduces to the PBE of an orthogonal BFSK given by Eq. (4.29).

In addition to the envelope detection, noncoherent demodulation can also be realized using a differentially coherent technique. Figure 4.5 shows the block diagram of a differentially coherent demodulator. The received signal is delayed by one bit period and used as the carrier reference to demodulate the next waveform received. Because no absolute carrier reference is available, the information must be encoded in the differential changes of the carrier phase. This system is commonly known as the differential phase-shift keying (DPSK) system. The analysis of binary DPSK receivers in AWGN can be found in standard communication theory texts [28]-[31], [146], and the probability of decoding error can be given by

$$P_E = \frac{1}{2}e^{-\beta}\ . \tag{4.31}$$

By comparing Eq. (4.31) and (4.29), it is seen that the binary DPSK system offers a 3dB improvement in SNR over the envelope detected BFSK system.

Table 4.3 is a list of the probability of errors for various systems using noncoherent detectors. Note that by substituting expressions for the signal amplitude and the noise spectral density into Eq. (4.21), the noncoherent SNR β can also be reduced to $\eta_{het}K_S$.

Figure 4.6 is a plot of error probabilities for various heterodyne systems versus the signal-to-noise ratios β. Note that the PSK system requires 3dB less signal power to achieve comparable error rate than the FSK and PPM systems. Similarly, FSK and PPM systems have a 3dB power advantage over the OOK system. Also shown in Figure 4.6 is the error probability of an ideal, quantum limited, photon counting direct detection binary PPM system with no background noise interference. It is seen from the figure that, when operating under ideal conditions, the performance of the direct detection BPPM system is comparable to that of the heterodyne DPSK system, and is superior to both heterodyne FSK and PPM systems. It should be noted, however, that although quantum limited detection can be easily achieved for heterodyne systems, direct

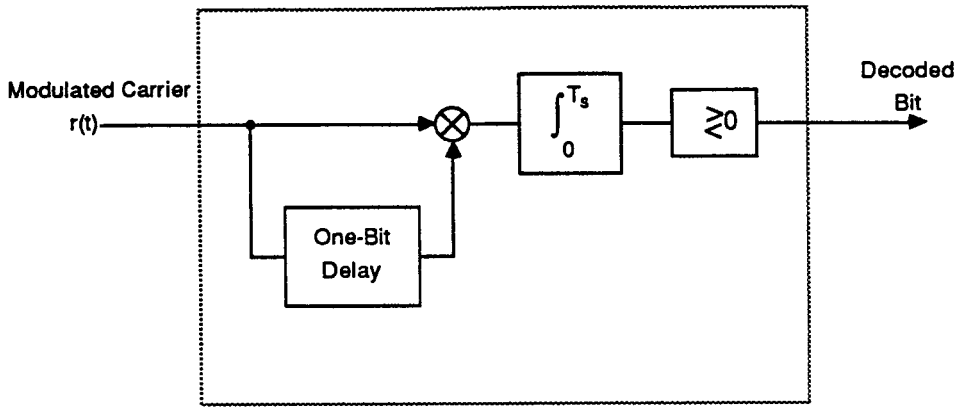


Fig.4.5. Block diagram of a DPSK receiver.

Table 4.3 Error probabilities for the noncoherent heterodyne channel.

Modulation Format	Error Probability
OOK	$\frac{1}{2}\left[1-Q(\sqrt{2\beta},\sqrt{2\theta})+e^{-\theta}\right]$
MPPM	$\sum_{k=1}^{M-1}\frac{(-1)^{k+1}}{k+1}\binom{M-1}{k}e^{-k\beta/(k+1)}$
MFSK	$\sum_{k=1}^{M-1}\frac{(-1)^{k+1}}{k+1}\binom{M-1}{k}e^{-k\beta/(k+1)}$
DPSK	$\frac{1}{2}e^{-\beta}$

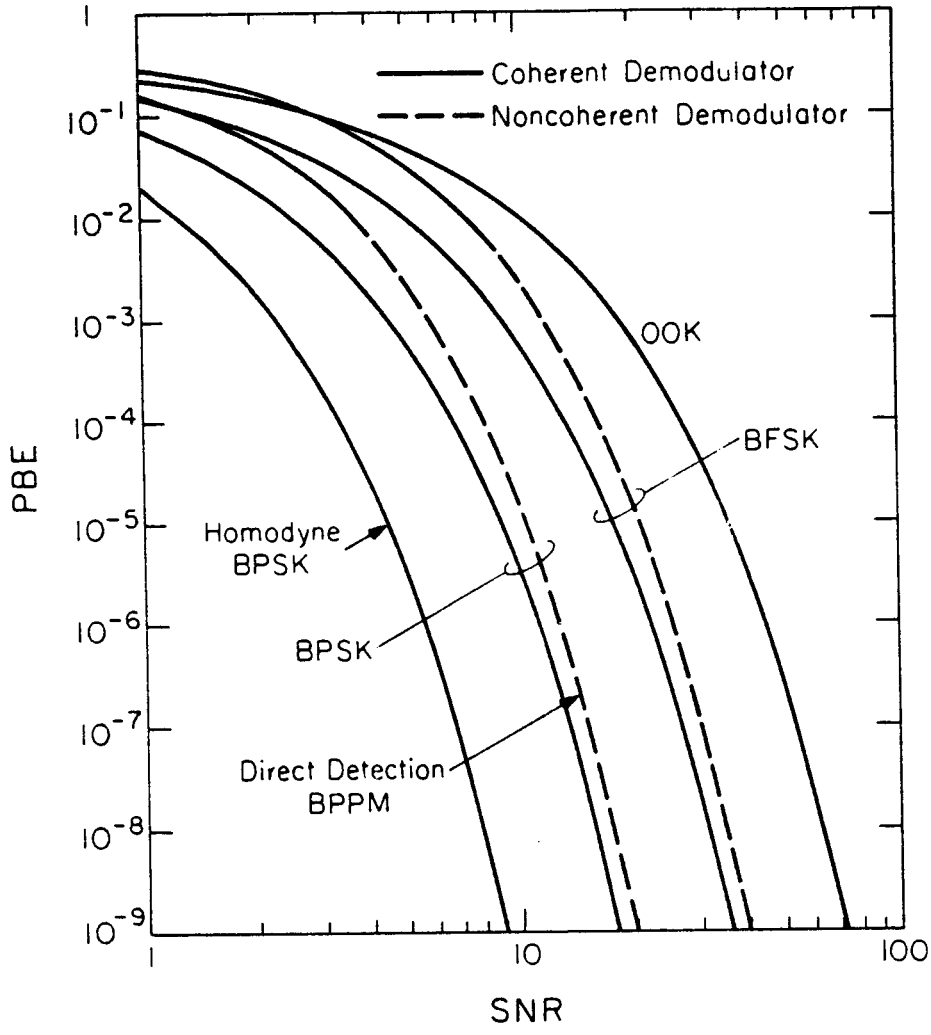


Fig.4.6. Probability of bit error (PBE) versus the SNR for various binary heterodyne systems. Also plotted in the figure is the PBE of an ideal direct detection BPPM system versus the signal photocount K_S.

detection systems are usually not shot noise limited. Incoherent background interference, thermal noise, and avalanche noise that can be ignored in heterodyne systems will have a significant impact on the performance of direct detection systems. It is shown in Figure 4.6 that noncoherent systems require a slightly higher signal power to achieve a given error rate than the comparable coherent systems. However, it should be realized that the analysis leading to the error rate expressions has assumed that the heterodyne channel can be modeled as an AWGN channel with ideal carrier synchronization and local reference recovery. In practice, the output of the heterodyne receiver is contaminated with LO intensity noise and laser phase noise in addition to the LO shot noise. The spurious amplitude modulation at the output of the transmitter laser can also have a measurable impact on the error performance of heterodyne receivers. Intensity noise at the output of a semiconductor laser has been characterized as a Gaussian process [147], and may be considered as part of the additive noise. The impact of intensity noise on the performance of heterodyne FSK systems has been studied. It is shown that, at a data rate of 1 Gb/s, the power penalty due to the intensity noise is approximately 1.5 dB [148]. The phase noise, on the other hand, modulates the phase of the received IF signal and cannot be modeled as an additive noise. In the presence of the carrier phase noise, carrier recovery that is required for the coherent demodulator cannot be perfectly realized. Consequently, higher signal power is needed to maintain the system performance. Noncoherent demodulators, on the other hand, do not require a carrier reference to perform the decoding, and are therefore less sensitive to the carrier phase noise.

4.3 Statistics of the Laser Phase Noise

The presence of carrier phase noise at the laser output results in the broadening of the observed laser linewidth. For systems using gas or solid state lasers, the observed linewidth broadening is very small compared to the data rate. Consequently, the effect of carrier phase noise can be practically ignored for these lasers. Semiconductor lasers, on the other hand, suffer from large frequency and phase fluctuations that can seriously affect the performance of heterodyne systems. Even when mode stabilized lasers are used, the observed linewidths for semiconductor lasers are typically on the order of 1-50 MHz. At these linewidths, the effect of carrier phase noise on the performance of the communication link cannot be neglected. The linewidth broadening effect of the semiconductor laser has been attributed to the changes in the index of refraction of the cavity due to the fluctuation in carrier density [42].

Experimentally, the spectral lineshape of semiconductor lasers has been characterized as Lorentzian [149]. The frequency spectrum of the laser output has also been carefully measured [43]. It is shown that the spectrum of the laser phase noise consists of three major components: a $1/f$ component at low frequencies, one or several resonance peaks at high (1-10GHz) frequencies and, in between, the laser exhibits a white frequency noise spectrum which is dominant throughout the spectral range of interest. Figure 4.7 shows a typical frequency spectrum of a semiconductor laser [43]. For a well designed system, the $1/f$ noise can be tracked out by the receiver and the resonant peaks are usually far above the system bandwidth. Consequently, only the white frequency component remains in the detected IF phase noise.

Motivated by these experimental results, the phase noise of the semiconductor laser can be modeled as a Wiener process $\phi(t)$, and instantaneous frequency $\dot{\phi}(t)$ satisfies

$$<\dot{\phi}(t)\dot{\phi}(\tau)> = \Delta_f \delta(t-\tau) \ , \tag{4.32}$$

where Δ_f is the power spectral density of the instantaneous frequency. The phase of the laser output is randomly modulated by the carrier phase noise $\phi(t)$ such that the carrier signal $s(t)$ can be modeled as

$$s(t) = A \sin(2\pi f_S t + \phi(t) + \phi_0) \ , \tag{4.33}$$

where ϕ_0 is the initial phase angle of the laser output. The power spectra of $\phi(t)$ and $s(t)$ can be deduced from (4.32) and (4.33) to be

$$S_\phi(f) = \Delta_f / \pi f^2 \ , \tag{4.34}$$

$$S_{IF}(f) = \frac{2\pi\Delta_f}{(f-f_{IF})^2 + (\Delta_f/2)^2} + N_s \ . \tag{4.35}$$

Note that the spectral lineshape of the IF signal derived from the white frequency noise assumption is indeed Lorentzian with linewidth Δ_f. For a symmetric heterodyne system where both signal and LO lasers have nonzero linewidths, the observed IF phase fluctuation will be the difference of the two laser phases and, consequently, the linewidth of the IF signal will be the sum of individual laser linewidths.

Because the carrier phase noise is nonstationary, the ideal maximum likelihood demodulator for the heterodyne signal must consider the past history of the phase [150], [151]. However, such receivers are very difficult to implement in practice. Instead, phase coherent and noncoherent receivers introduced in Section 4.2

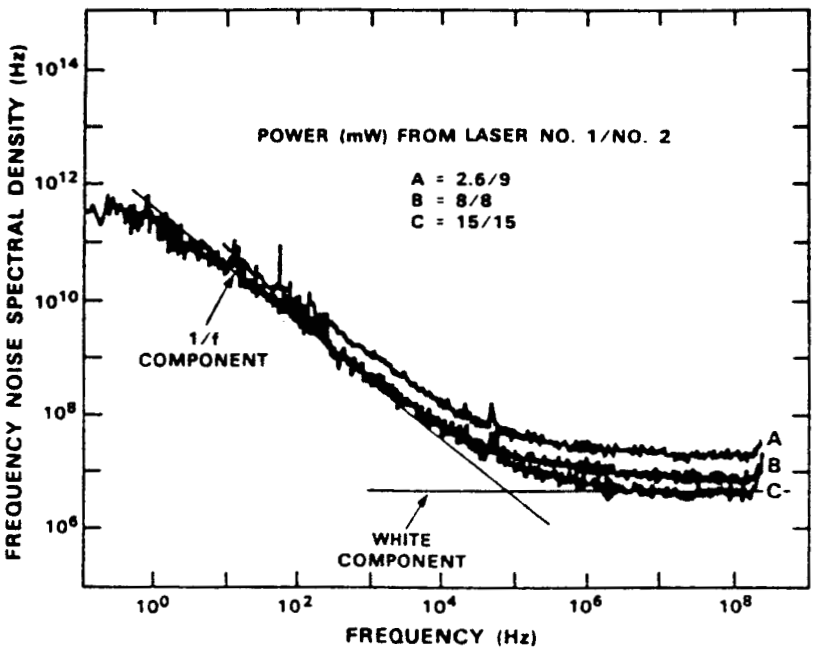


Fig.4.7. Power spectrum of the laser frequency noise for typical semiconductor laser diodes. (From Jeromin and Chan, GLOBECOM'84)

are often used to demodulate the IF signal.

4.4 Carrier Synchronization and Coherent Demodulation of IF Signals

Successful demodulation using a phase coherent receiver demands a receiver capable of estimating the frequency and phase of the transmitted signals. In practice, phase coherent demodulation requires the presence of a local reference carrier that is synchronized to the transmitter carrier. Before the coherent demodulation can take place, this carrier reference must be properly established.

At sufficiently high signal levels, carrier synchronization can be accomplished by injection locking [152] a local oscillator laser with the incoming signal. However, injection locking typically requires a high signal power (-20-30 dBm) [153] that is not achievable in ISL links. Alternatively, phase-locked loops (PLL) [28], [29] can be used to recover the transmitted carrier. Shown in Figure 4.8, a PLL consists of a local oscillator (LO) that can be frequency tuned by an externally supplied signal, a phase comparator with an output that is proportional to the phase difference between the signal and LO, and a loop filter. The applications of PLLs to RF carrier synchronizations have been well documented [28]-[31]. When the input signal to the PLL contains a spectral component at the desired lock-in frequency, the PLL can be used to recover the frequency and phase of the signal. For signals that do no contain a spectral component at the desired lock-in frequency, such as in a suppressed carrier transmission system, the simple PLL shown in Figure 4.8 cannot be used. Instead, one of several suppressed carrier tracking loops must be employed to establish the carrier synchronization. Figure 4.9 shows the different configurations of suppressed carrier tracking loops. These loops in general employ nonlinearities to generate the desired spectral component at the lock-in frequency.

Carrier synchronization in optical heterodyne systems is similar to that of RF systems [154]-[156]. The VCO is implemented using a local oscillator laser which can be frequency tuned by changing the supplied current or voltage. The phase comparator can be implemented using a heterodyne receiver. For suppressed carrier systems, an optical hybrid can be used to replace both in-phase and quadrature arms of an I-Q loop [156]. Optical PLLs have been successfully used to track the incoming signal frequencies at a signal level of -50-65 dBm [153].

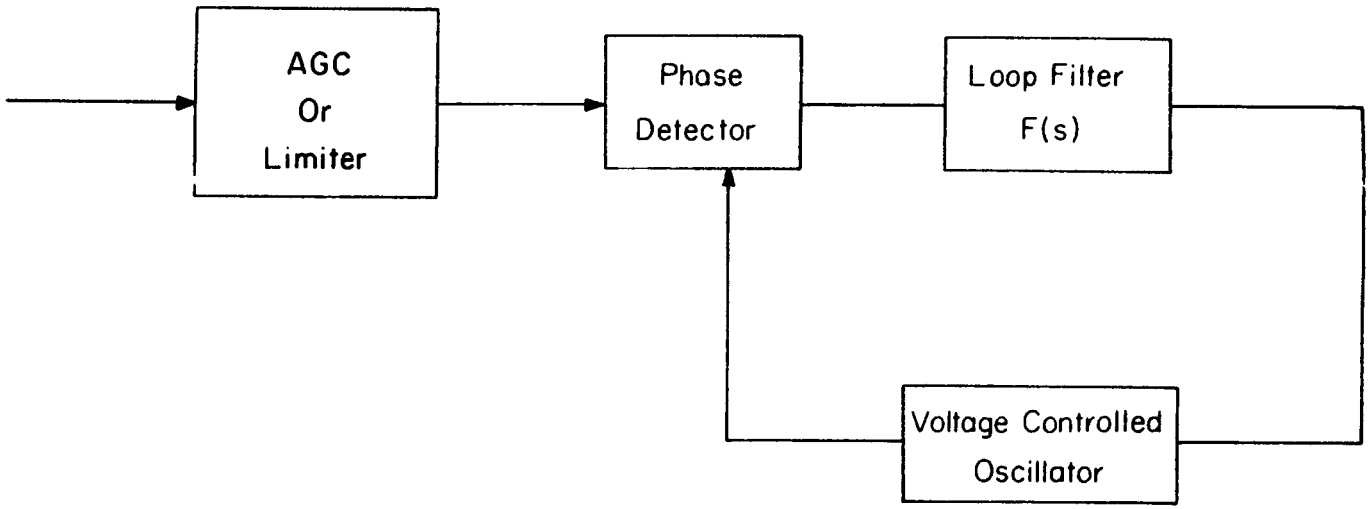
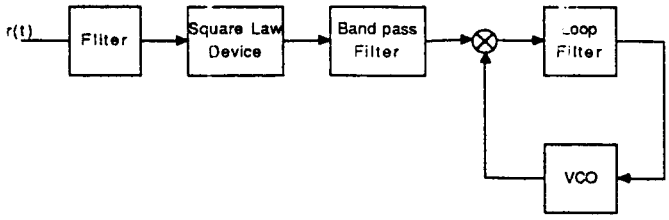
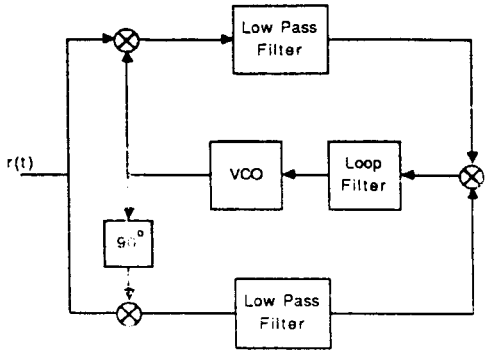


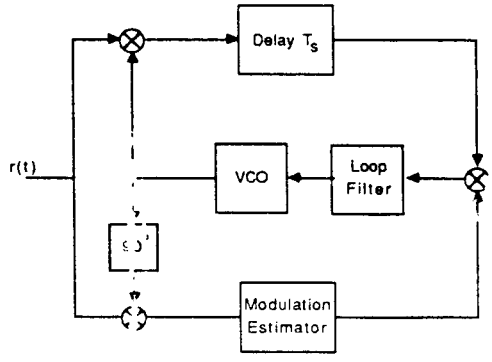
Fig.4.8. Block diagram of a simple phase-locked loop.



(a)



(b)



(c)

Fig.4.9. Suppressed carrier tracking loops: (a) the squaring loop, (b) the Costas I-Q loop, and (c) the decision feed-back loop.

The performance of the carrier tracking PLL depends on the properties of the input signal and that of the loop design. In general, when tracking is established, the phase difference ψ between the signal and LO output can be modeled as a Gaussian random process. The variance of ψ can be written as the sum of contributions from the shot noise and phase noise:

$$\sigma_{\psi}^2 = \sigma_{SN}^2 + \sigma_{PN}^2 \ . \tag{4.36}$$

The shot noise contribution to the phase noise variance can be deduced from the linearized PLL theory [156] as

$$\sigma_{SN}^2 = 2N_s B_c / A^2 \ , \tag{4.37}$$

where A is the amplitude of the IF signal, N_s is the noise spectral density, and B_c is the (two-sided) closed-loop bandwidth of the PLL. B_c can be related to the closed-loop transfer function $H_c(f)$ by

$$B_c = \int_{-\infty}^{\infty} |H_c(f)|^2 df \ . \tag{4.38}$$

The contribution of carrier phase noise to the synchronization error ψ can be written as [31]

$$\sigma_{PN}^2 = \int_{-\infty}^{\infty} S_{\phi}(f) |1 - H_c(f)|^2 df \ , \tag{4.39}$$

where $S_{\phi}(f)$ is the power spectrum of the phase noise. Note that the phase noise contribution to the synchronization error in (4.39) can be interpreted as the amount of the phase noise power not filtered by the tracking loop of transfer function $H_c(f)$. By approximating $H_c(f)$ as an ideal low-pass filter with bandwidth B_c, the phase noise contribution to the synchronization error can be simplified to

$$\sigma_{PN}^2 = 2\Delta_f / \pi B_c \ . \tag{4.40}$$

For a more realistic loop design, (4.40) should be multiplied by a constant ζ which, for a second-order PLL with a damping factor 0.7, is equal to 3.7 [31], [156].

By combining Eqs. (4.37) and (4.40), it is seen that the variance of phase noise contains a term that is proportional to the loop bandwidth and a term that is inversely proportional to B_c. Consequently, for a given linewidth Δ_f and the signal-to-noise ratio, an optimal PLL bandwidth exists which minimizes the phase error variance. This optimal PLL bandwidth can be evaluated to be [156]

$$B_{opt} = \sqrt{4\zeta K_S \Delta_f T_s / \pi} f_s \ , \tag{4.41}$$

where $f_s=1/T_s$ is the symbol rate of the channel.

When the bandwidth of the PLL is small compared to f_s, the synchronization error ψ is slowly varying with respect to the symbol period. In this case the synchronization error ψ may be approximated by a constant during each symbol period. The conditional error probability for a coherent heterodyne receiver, given the signal to noise ratio β_c and the synchronization phase error ψ, can be written as

$$P_E(\beta_c,\psi) = P_E(\beta_c \cos^2\psi) \ , \tag{4.42}$$

where β_c is the SNR defined in Eq. (4.10). The unconditioned PBE for the system is therefore the expectation of Eq. (4.42) with respect to the distribution of ψ:

$$P_E(\beta_c,\sigma_\psi^2) = E\left[P_E(\beta_c,\psi)\right] = \frac{1}{\sqrt{2\pi}\sigma_\psi}\int_{-\infty}^{\infty} P_E(\beta_c \cos^2\psi) e^{-\psi^2/2\sigma_\psi^2} d\psi \ . \tag{4.43}$$

In writing (4.43), it is assumed that the carrier synchronization has been properly established and that the instantaneous carrier tracking error ψ is Gaussian distributed.

The effects of phase noise and synchronization error on a coherent demodulator can be seen in Figures 4.10-4.12. In Figure 4.10, the probability of error for a BPSK system is plotted against the SNR for several values of the IF linewidth. The loop bandwidth B_c is equal to $10^{-1}f_s$. Note that for $\Delta_f T_s \geq 0.005$ an error rate "floor" is visible which is independent of the SNR. This is because at a high signal power and a constant B_c, the probability of synchronization error is dominated by the phase noise contribution which cannot be reduced by increasing the signal power. However, if the loop bandwidth is adjusted to minimize the synchronization error, the performance of the demodulator becomes much less sensitive to the laser phase noise. Figure 4.11 is a plot of the PBE versus the SNR for a BPSK system where the loop bandwidth is chosen to minimize σ_ψ^2. It is seen that, at $PBE=10^{-9}$, an IF linewidth as large as 2% of the symbol rate can be tolerated with a 3dB increase in the required SNR. The power penalties associated with the synchronization error are shown in Figures 4.12(a) and 4.12(b) for a binary PSK system at $PBE=10^{-9}$. Figure 4.12(a) is plotted assuming that the synchronization bandwidth is constant, and Figure 4.12(b) is plotted assuming that B_c has been optimized using Eq. (4.41). It is seen from these figures that the power penalty associated with the laser phase noise increases rapidly when the laser linewidth becomes a significant fraction of the loop bandwidth. For systems where the loop bandwidth is optimally adjusted, the power penalty is less sensitive to the IF phase noise. However, it should be noted that the assumption leading to the expression of error rate is valid only when the

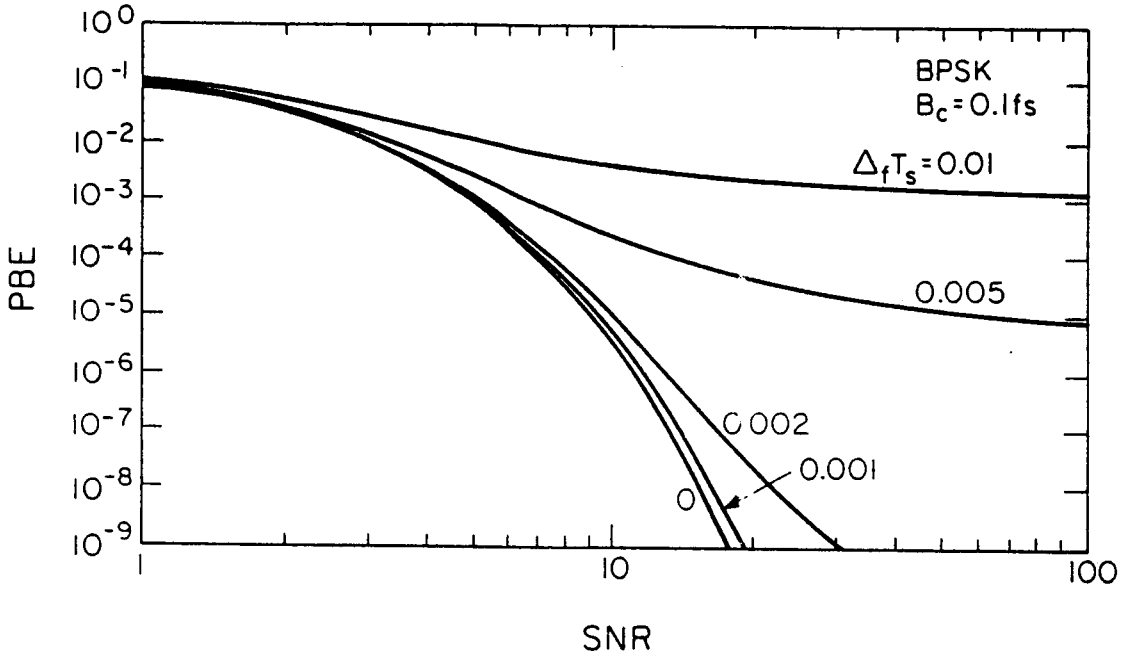


Fig.4.10. Probability of bit error versus the receiver SNR for a BPSK system at several values of the IF linewidth. The bandwidth of the carrier tracking loop, B_c, is equal to $0.1 f_s$.

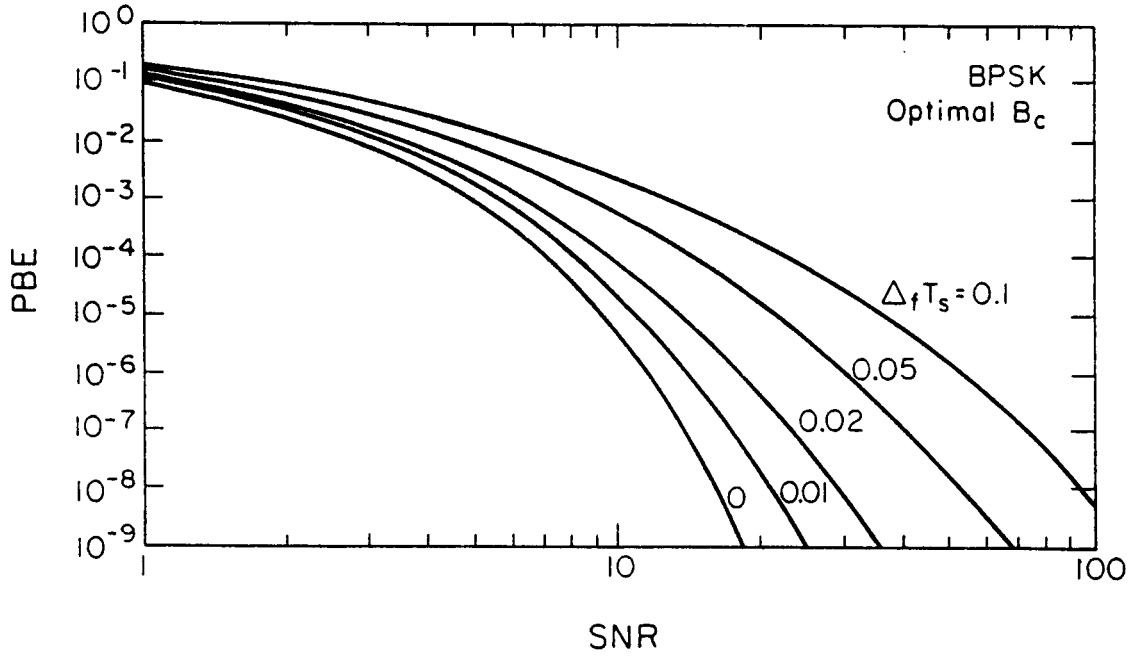
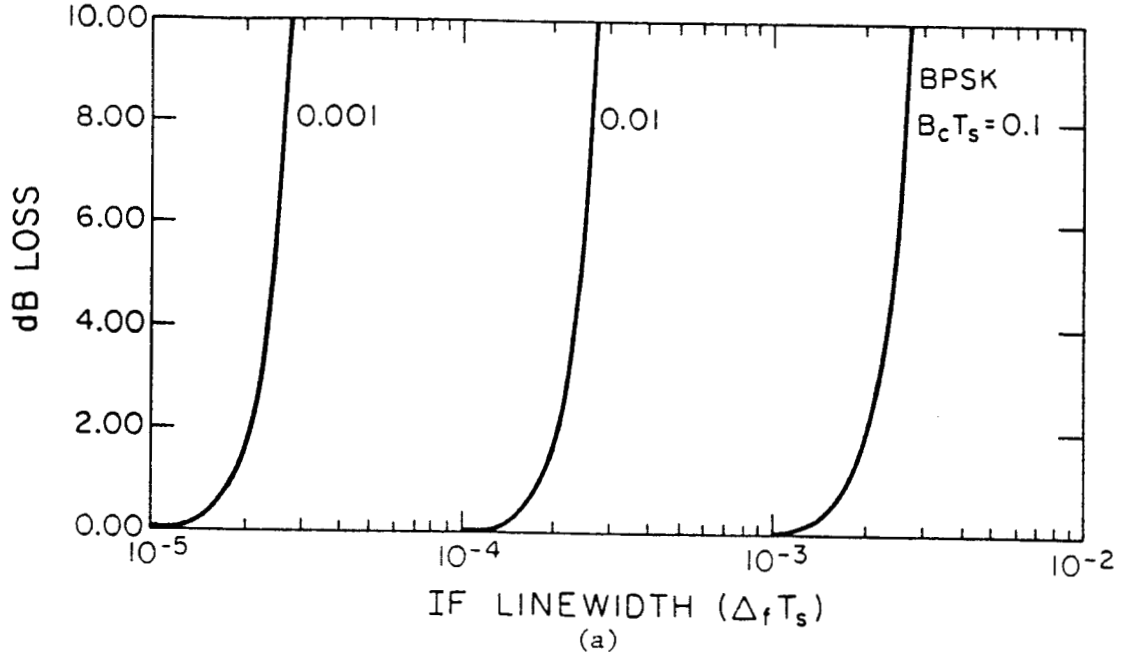


Fig.4.11. Probability of bit error versus the receiver SNR for a BPSK receiver in which the bandwidth of the carrier tracking loop is chosen to minimize the variance of phase error.



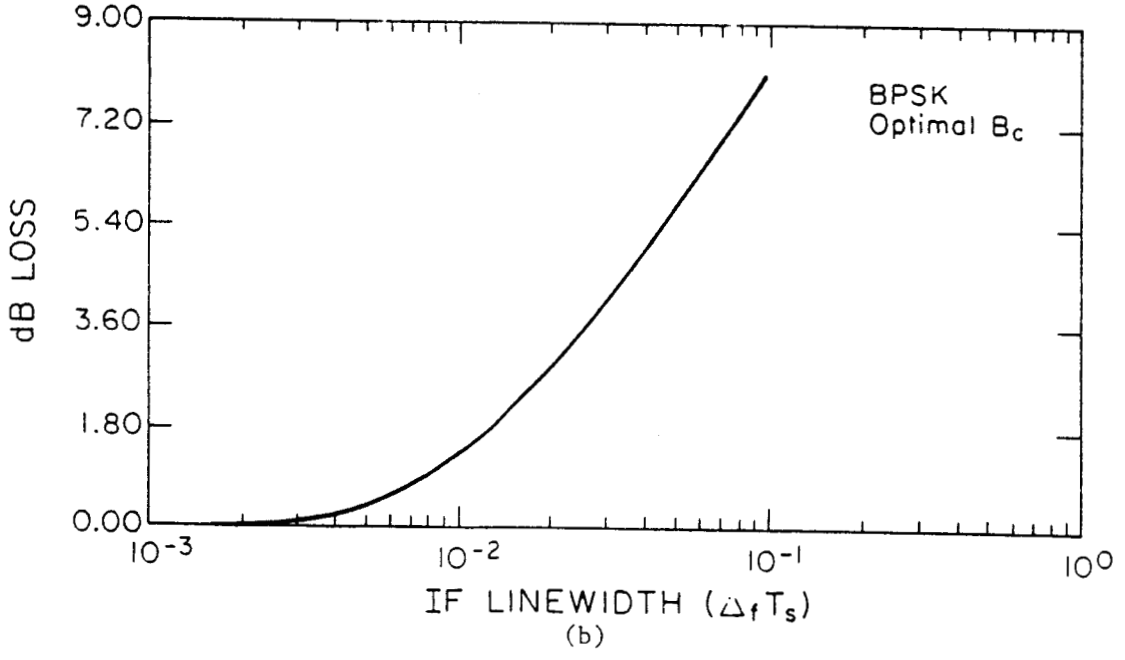


Fig.4.12. Power penalty due to carrier phase noise versus the IF linewidth for a BPSK receiver at 10^{-9} error rate. The bandwidth of the tracking loop, B_c is equal to $0.1f_s$ in Fig.4.12(a), and is chosen to minimize the variance of phase error in Fig.4.12(b).

loop bandwidth is small compared to the data rate. At large IF linewidths, the loop bandwidth required to achieve the optimal performance can be a significant fraction of the data rate. Such a large synchronization bandwidth can lead to a significant phase fluctuation within a given codeword period. For such systems, a much higher signal power is needed to maintain the system performance [157].

4.5 Noncoherent Demodulation in the Presence of Phase Noise

Because optical phase synchronization requires a much higher signal power than can be achieved through an ISL, coherent heterodyne systems are difficult to implement in practice. Instead, noncoherent systems that do not require the carrier phase synchronization are more suited for the application.

M-ary orthogonal FSK is one of the most efficient noncoherent schemes. Unfortunately, because of the nonstationary nature of the carrier phase noise, performance analysis of the noncoherent MFSK system is very difficult to accomplish, and some assumptions are needed to simplify the analysis. A simple, yet effective approximation of the phase error is to assume that the phase diffusion is introduced by a random frequency noise ν [52] such that the phase diffusion occurred during the observation interval $(0,T_s)$ are approximated by

$$\phi(t)-\phi(0) \approx 2\pi\nu t \quad , \; 0\le t\le T_s \quad . \tag{4.44}$$

Because the carrier phase noise can be modeled as a Wiener process, the frequency ν can be modeled as a Gaussian random variable. The variance of ν can be calculated from Eqs. (4.33) and (4.44) to be:

$$Var(\nu)=\frac{\Delta_f}{2\pi T_s} \quad . \tag{4.45}$$

Recall from Section 4.2 that the performance of noncoherent FSK systems is determined by the parameters β_{ij} where $\beta_{ij}=A_i^2T_s/4N_s$. By substituting Eq. (4.44) into Eq. (4.23), it is seen that

$$\beta_{kj} = \beta\left[\frac{\sin[2\pi(\nu+\nu_{kj})T_s/2]}{[2\pi(\nu+\nu_{kj})T_s/2]}\right]^2 , \tag{4.46}$$

where $\nu_{kj}=f_k-f_j$ is the tone spacing between the j^{th} and the k^{th} signals, and ν is the random frequency noise. Note that because of the frequency noise ν, β_{kj} is in general not equal to zero even when ν_{kj} is an integer multiple of the symbol frequency. Consequently, in the presence of the carrier phase noise the M signaling frequencies of the MFSK system are no longer orthogonal. This channel cross talk is the major source of error for the MFSK systems.

Given the random frequency noise ν, the PBE of the binary FSK system can be calculated using Eq. (4.30). Consequently, by averaging over the distribution of ν, an approximation of the error rate can be obtained for a binary FSK system operating with finite laser phase noise. Figure 4.13 is a plot of the PBE for a binary FSK system versus the receiver SNR. The tone spacing between the frequencies ν_{01} is equal to the symbol rate f_s. Note that the presence of carrier phase noise introduces an error rate "floor" at the decoder which is clearly visible for $\Delta_f T_s > 0.05$. This error rate floor has been estimated [36] to be

$$P_E \geq \frac{1}{2}\text{erfc}\left[\frac{\nu_d T_s}{\sqrt{4\Delta_f T_s/\pi}}\right] . \tag{4.47}$$

By comparing the performance shown in Figure 4.13 with the performance of the coherent BPSK system shown in Figure 4.10, it is seen that the noncoherent system is far less sensitive to carrier phase noise than the phase coherent system. It is intuitive that the channel cross talk due to phase noise can be minimized by increasing the tone spacing. Figure 4.14 is a plot of the PBE versus the tone spacing for several values of the IF linewidth. Note that when the phase noise can be ignored, the performance of the FSK system is optimized when the tone spacing is equal to an integer multiple of f_s. However, at nonzero linewidth, the orthogonality is destroyed and the performance becomes a function of the tone spacing in addition to the SNR and the IF linewidth. The power penalties associated with the phase noise are plotted in Figure 4.15 versus the IF linewidth for a binary FSK system at different values of the tone spacing and PBE. Note that the power penalty decreases with increasing tone spacing and decreasing IF linewidth. Furthermore, because of the error rate "floor," the power penalty is seen to increase sharply when the error rate floor due to the carrier phase noise becomes greater than the desired PBE. It should be noted that the power penalty curves shown in Figure 4.15 are derived using the assumption that the carrier phase noise can be approximated by a random frequency noise. The actual power penalty incurred by the phase noise is higher than those shown in Figure 4.15. In practice, care must be taken to allocate sufficient power margins to account for this error in approximation. Table 4.4 shows the power penalty due to carrier phase noise for a binary FSK system at different values of the PBE and IF linewidths. Note that the power penalty increases rapidly with increasing IF linewidth.

Another noncoherent system commonly used in RF systems is the differentially coherent PSK system (DPSK). Unlike the NCFSK system, which uses the envelope detector, the DPSK receiver performs the decoding by correlating the signal with the waveform received over the previous bit period. Unfortunately, as

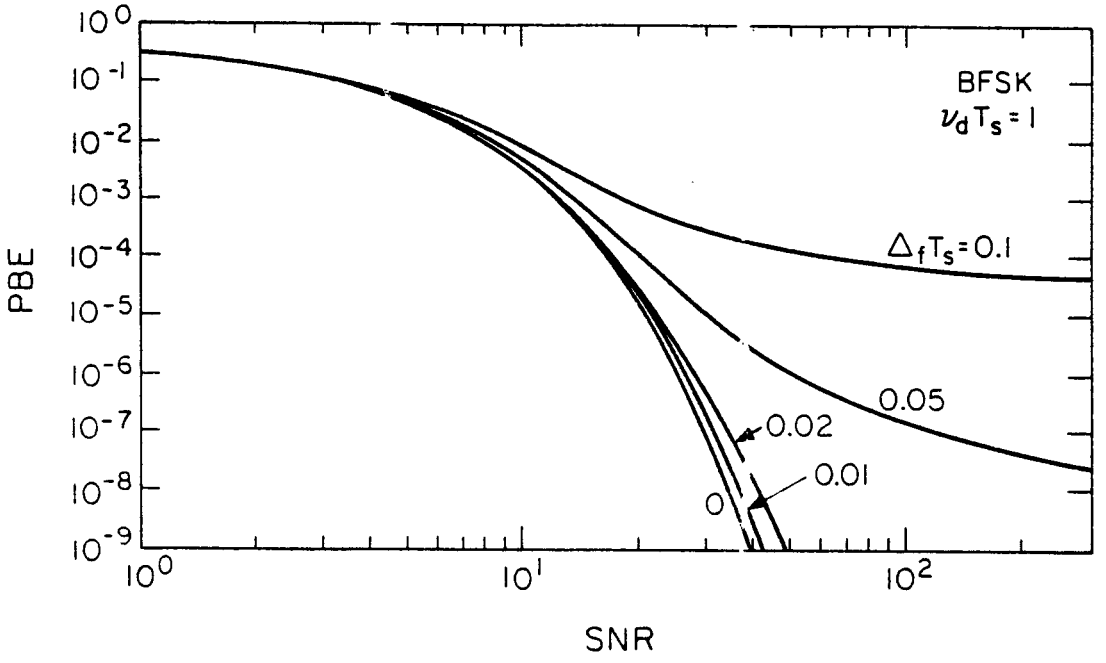


Fig.4.13. Probability of bit error for a binary NCFSK receiver versus the receiver SNR at several values of the IF linewidth Δ_f.

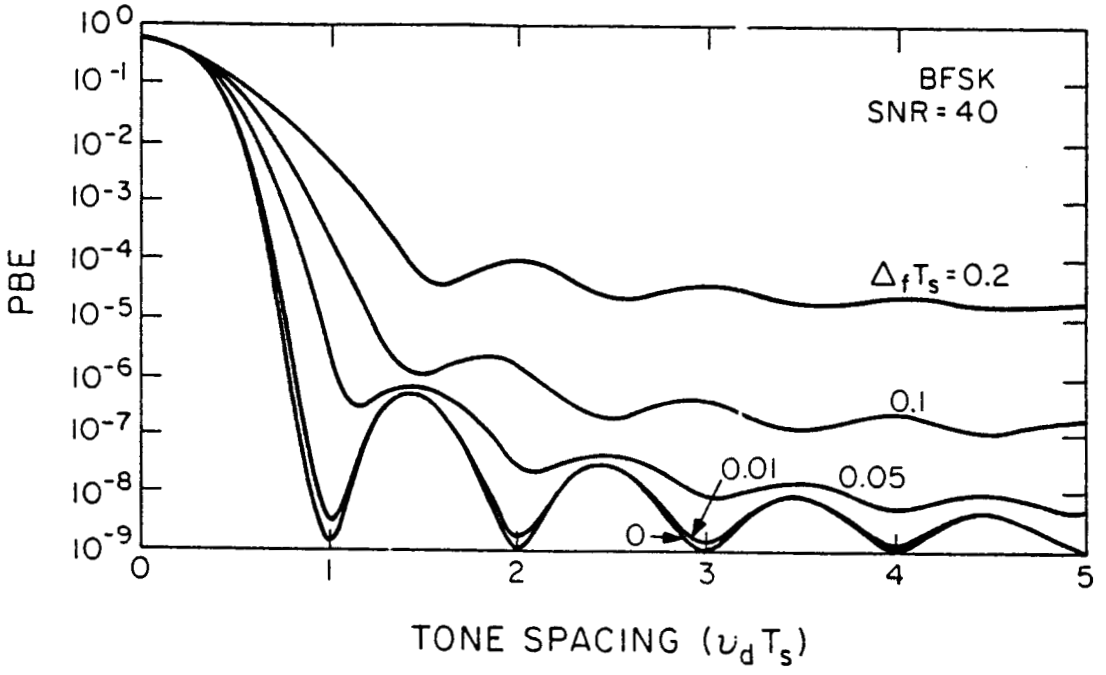
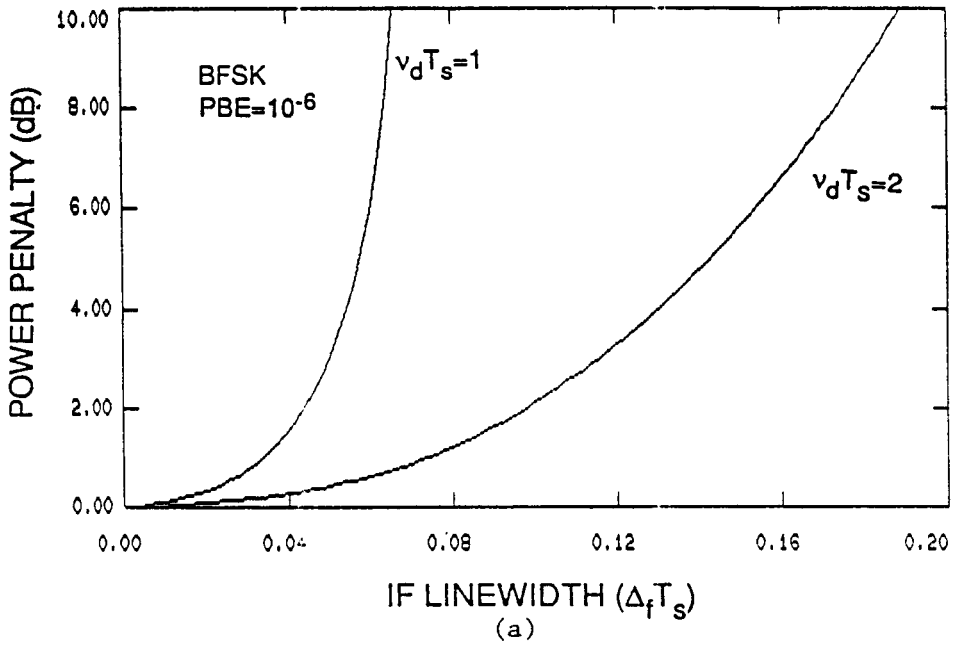


Fig.4.14. Probability of bit error for a binary NCFSK system versus the tone spacing at different values of the IF linewidth.



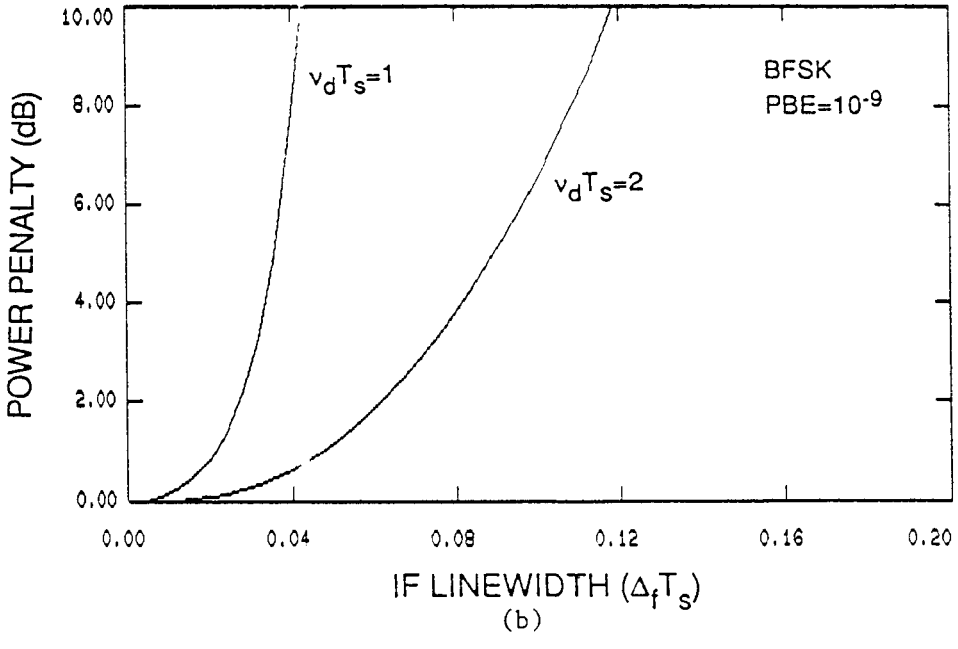


Fig.4.15. Power penalty due to phase error versus the IF linewidth for a BFSK receiver at (a) PBE=10^{-6} and (b) PBE=10^{-9}.

Table 4.4 Power penalty due to laser phase noise for a noncoherent BFSK system.

PBE	IF Linewidth ($\Delta_f T_s$)	Tone Spacing	
		$\nu_d T_s$=1	$\nu_d T_s$=2
10^{-6}	0.03	0.8 dB	0.2 dB
	0.05	3.0 dB	0.5 dB
	0.1	>10.0 dB	2.1 dB
10^{-9}	0.03	3.0 dB	0.5 dB
	0.05	>10.0 dB	1.3 dB
	0.1	> 10.0 dB	6.6 dB

was the case with NCFSK receivers, the performance analysis becomes very difficult when the phase noise is considered. When the laser linewidth is sufficiently small, however, the probability of error for a DPSK system with carrier phase noise can be approximated by (Appendix A):

$$P_E \approx e^{-\beta} + \frac{1}{2}\mathrm{erfc}\left(\frac{\pi/2}{\sqrt{2/\beta + 8\pi\Delta_f T_s/3}}\right) . \tag{4.48}$$

Figure 4.16 is a plot of the PBE evaluated using Eq. (4.48) versus the SNR. Also plotted in the figure is the PBE of an ideal DPSK receiver evaluated using Eq. (4.31). Note that at small laser linewidths, the approximation in Eq. (4.48) gives a higher error probability than the actual PBE. However, at PBE = 10^{-9}, the difference in the required signal power calculated using Eq. (4.48) instead of Eq. (4.31) is small. Note that as $\Delta_f T_s \approx 0.01$ the required signal power needed to achieve 10^{-9} error rate begins to increase rapidly. The power penalty associated with the carrier phase noise is plotted in Figure 4.17 versus the IF linewidth. It is seen that a sharp increase in the power penalty in Figure 4.17 occurs when the error rate "floor" exceeds the desired PBE. The error rate "floor" can be calculated from Eq. (4.48) by noting that at $\beta \to \infty$, the probability of error is bounded by [158]

$$P_E \geq \frac{1}{2}\mathrm{erfc}\left(\sqrt{3\pi/32\Delta_f T_s}\right) . \tag{4.49}$$

Note that the error rate "floor" is greater than 10^{-9} when the IF linewidth is greater than 2% of the data rate.

4.6 Summary

In the absence of carrier phase noise and spatial tracking error, the heterodyne channel can be modeled as an AWGN channel. Under this ideal condition, the coherently demodulated heterodyne channel offers a superior performance compared to the noncoherent channel. However, when carrier phase noise is present, such as when semiconductor lasers are used, the coherent heterodyne channel becomes very sensitive to the carrier synchronization error. When operating under extremely low signal power such as in an ISL, the signal power needed to achieve coherent carrier synchronization may be difficult to achieve. Noncoherent heterodyne systems, on the other hand, do not require the presence of a synchronized local reference and, therefore, can be more easily implemented. In particular, the noncoherent FSK (NCFSK) system is less sensitive to the phase noise compared to the DPSK system. For these reasons, the analysis of the heterodyne ISL will be concentrated on the NCFSK system.

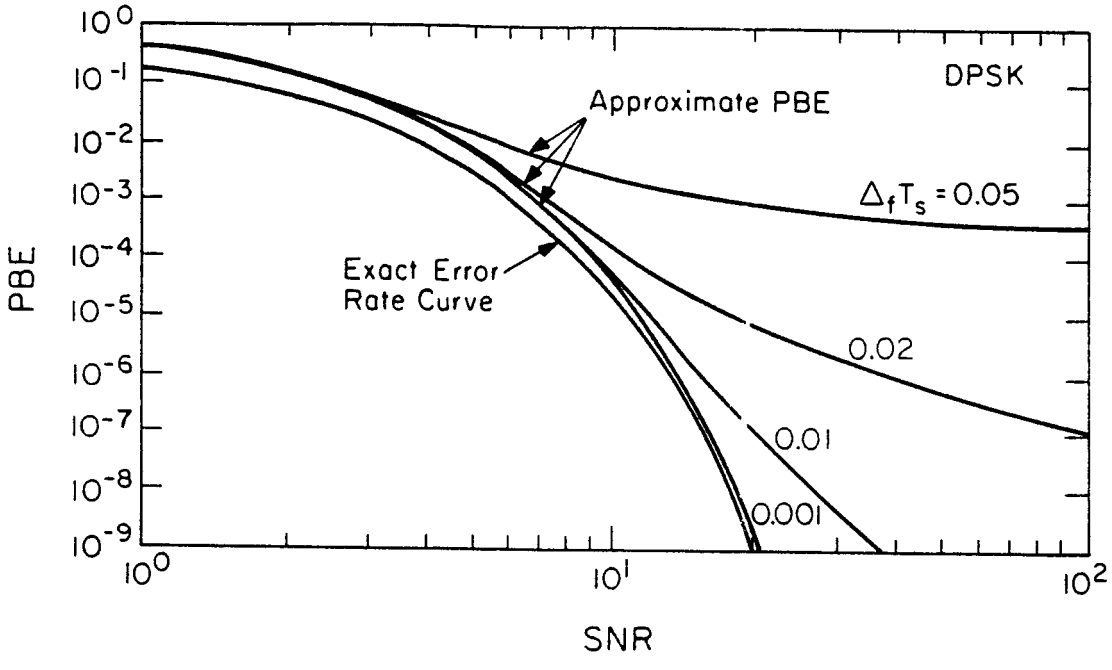


Fig.4.16. Probability of error for a DPSK receiver versus the receiver SNR at several values of the IF linewidth Δ_f.

C-2

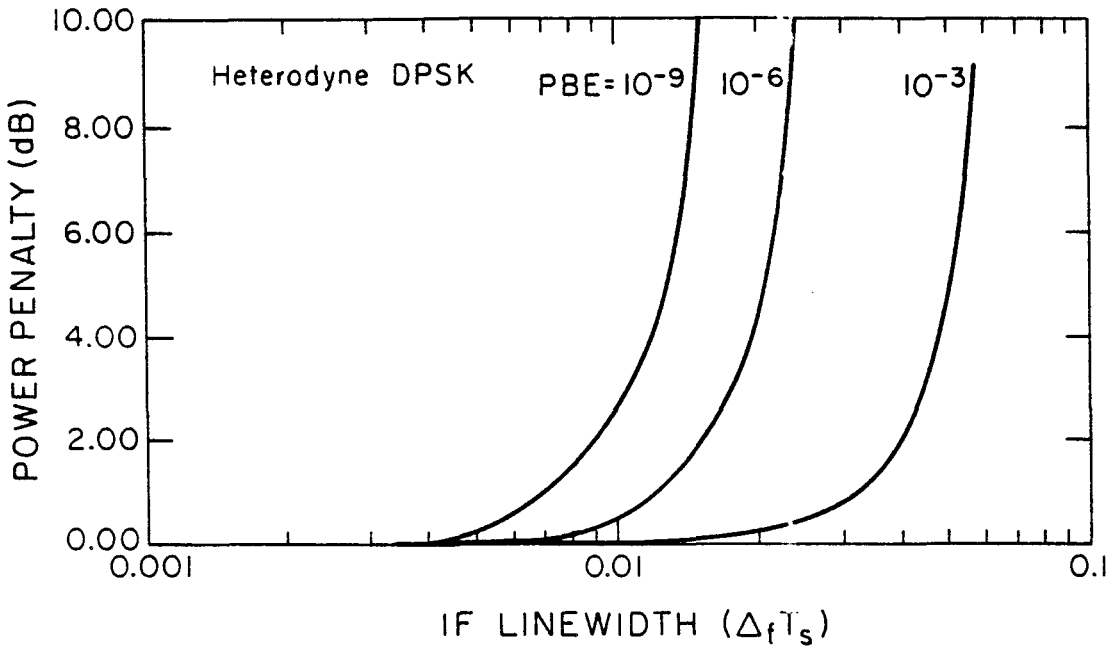


Fig.4.17. Power penalty due to the carrier phase noise versus the IF linewidth for a DPSK receiver.

5. SPATIAL TRACKING

Because of the small beam divergence and the low available signal power, the spatial tracking requirement for an optical ISL is more stringent than that of a conventional microwave link. An optical ISL operating at 850 nm wavelength using a 10 cm diameter transmitter telescope will have an effective angular beamwidth of only 10 μrad compared to a typical microwave beamwidth of several milliradians. When operating at such a narrow beamwidth, line-of-sight (LOS) tracking between the receiver and transmitter must be properly established so that the main lobe of the transmitted signal will fall on the receiver aperture. The purpose of the tracking subsystem, therefore, is to acquire the mutual positions of the transmitter and receiver and maintain the LOS tracking for the duration of the communication.

5.1 Spatial Acquisition

Before high speed data transfer can take place over the ISL, the transmitter and receiver optics must achieve mutual LOS tracking of each other. Initially, by using the known satellite trajectory and the telemetry data, relative angular positions of the satellites can be stabilized to within 1 mrad. The angular width of a typical transmitted signal, however, is only a few microradians wide. Consequently, spatial acquisition must be carried out to reduce the angular uncertainty.

In order for the receiver to acquire the angular position of the transmitter, a beacon signal from the remote transmitter must illuminate the receiver. In general, illumination of the initial angular uncertainty zone can be accomplished using either a parallel or a serial illumination scheme. In a parallel illumination scheme, the entire uncertainty zone is illuminated simultaneously. Because of the larger beam divergence needed to cover the search zone, smaller signal power is available at the receiver, and longer observation time is needed to collect sufficient signal photons for an acceptable acquisition accuracy. In contrast, in a serial illuminated system, the angular uncertainty zone is divided into smaller cells, and the transmitter illuminates each cell sequentially. Acquisition is possible only when the spatial cell containing the receiver is illuminated. Since each cell is now illuminated with a higher signal intensity, however, a shorter observation period is needed to achieve the desired acquisition accuracy.

At the receiver, the angular position of the incident optical signal is extracted and the resulted angular position estimate is used to control the servo system which aligns the optical LOS. Estimation of the incident

angle for an optical signal is equivalent to determining the position of the signal image at the receiver focal plane. Fig. 5.1 illustrates the relationship between the incident angle of the signal and the image position. Position acquisition at the receiver focal plane is usually accomplished by dividing the focal plane area into smaller acquisition cells where each cell can be individually photodetected [55], [159]. For a typical uncertainty zone of 1 mrad, and a desired angular resolution of several microradians, the required acquisition array is very large (several hundred squared). Charge-coupled devices (CCD) with high pixel density can be used to detect this focal plane intensity pattern and estimate the centroid of the diffraction pattern [160]. Smaller detector arrays can also be used to achieve comparable acquisition accuracy by either mechanically scanning the detector over the entire focal plane area or using a zooming algorithm. In a zooming algorithm, the acquisition is divided into several stages. At each stage, the remaining uncertainty zone is divided into a number of acquisition cells, and the acquisition is carried out among these cells. This process is then repeated until the desired acquisition accuracy is achieved.

Both direct and heterodyne detection schemes can be used to detect the image position at the receiver focal plane. The direct detection acquisition sensor is simpler to operate, but is more sensitive to the background and thermal noise. Consequently, systems employing direct detection acquisition sensors generally require a longer acquisition time to achieve a comparable acquisition accuracy [54], [55]. On the other hand, heterodyne acquisition systems are relative immune to the background and thermal noise, but require a much more complicated optical structure.

The performances of spatial acquisition systems using different illumination strategies and detection schemes have been studied for the ISL application [55]. It is seen in the study that acceptable acquisition accuracy can be achieved within a reasonable amount of time (1-10 sec) using a relatively small laser power ($\approx$10 mW). In the following, it will be assumed that spatial acquisition procedure is completed and that both the transmitter and receiver have achieved mutual LOS tracking of each other.

5.2 The LOS Tracking

After the angular position of the remote satellite has been properly acquired, the optical assembly can then be rotated to point along the LOS. Because of the relative motion and the on-board mechanical noise of the satellites, pointing of the transmitter optics is subjected to some continuous, time-varying fluctuations even after

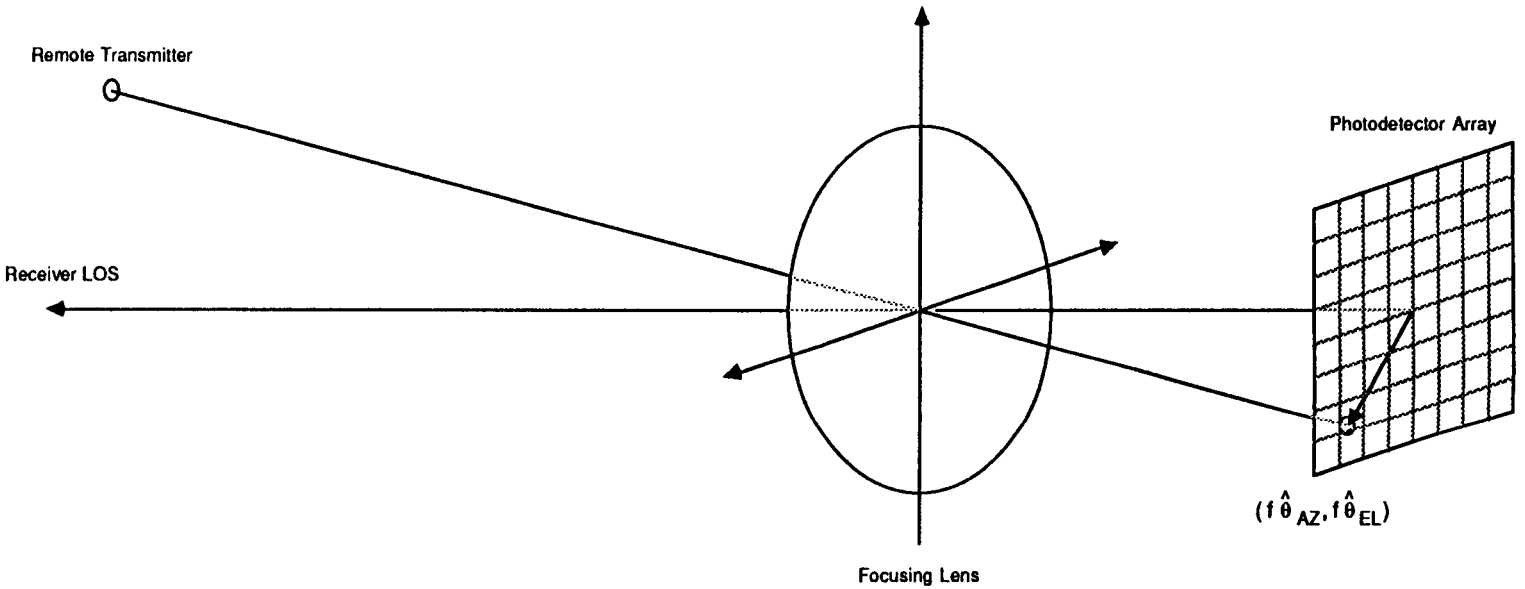


Fig.5.1. Geometrical relationship between the angular position of the remote transmitter and its image on the focal plane detector array.

the initial spatial acquisition has been achieved. For a heterodyne system, these mechanical perturbations can also affect the optical alignment between the incoming signal and the LO and result in a degraded receiver performance. In a practical ISL, the initial alignment of the optical LOS and the subsequent fine tracking of the LOS perturbations are usually achieved using a servo-driven opto-mechanical spatial tracking loop.

Figure 5.2 shows the block diagram of a spatial tracking loop. The beacon signal from the remote transmitter is detected using an appropriate tracking sensor. The beacon signal can be a separate, low data rate signal or, for the case of an optical ISL, the information-carrying signal itself. The tracking sensor generates a signal output that is proportional to the angular deviation between the current receiver LOS and the angular position of the remote transmitter. This angular estimate is then used to control the tracking servo and align the receiver LOS. For an optical ISL between LEO and GEO satellites, transmitter point-ahead must be added to the receiver LOS to compensate for the long path delay and the large relative velocity. In order to achieve a high degree of accuracy in spatial tracking, the servo system usually employs a nested loop design [56]-[59], which consists of a low bandwidth coarse-gimbal loop that tracks the lower frequency perturbations of the LOS, and a high bandwidth inner loop, which compensates for the higher frequency motion disturbances. The use of nested servo loops essentially decouples the tracking of high frequency, mechanical perturbations from the tracking of low frequency LOS motions. For the spatial tracking system just described, tracking error arises primarily from the noise associated with the tracking sensor, and the disturbances associated with base motion and mechanical vibration of the satellite. The point-ahead calculation can be achieved to a high degree of accuracy using the known ephemerides data, and other systematic errors can be minimized by a proper design.

The effect of the sensor noise on the tracking system performance can be characterized by its noise equivalent angle (NEA), which is defined as the standard deviation in estimating the current angular position. The problem of estimating the angular deviation of the received signal is equivalent to acquiring the incident angle of the received signal, and can be accomplished using a focal plane detector array. Because of the small angular deviation expected after the spatial acquisition procedure is completed, only a small number of spatial cells are needed to detect the focal plane intensity variation. A quadrant photodetector is commonly used to derive an angular position estimate of a closely tracked signal [56], [57], [158], [161]. Shown in Fig. 5.3, the quadrant detector contains four separate photodetector cells that are carefully matched in detection sensitivities. The use of quadrant detectors allows the azimuthal and elevation angular position estimates to be separated, and

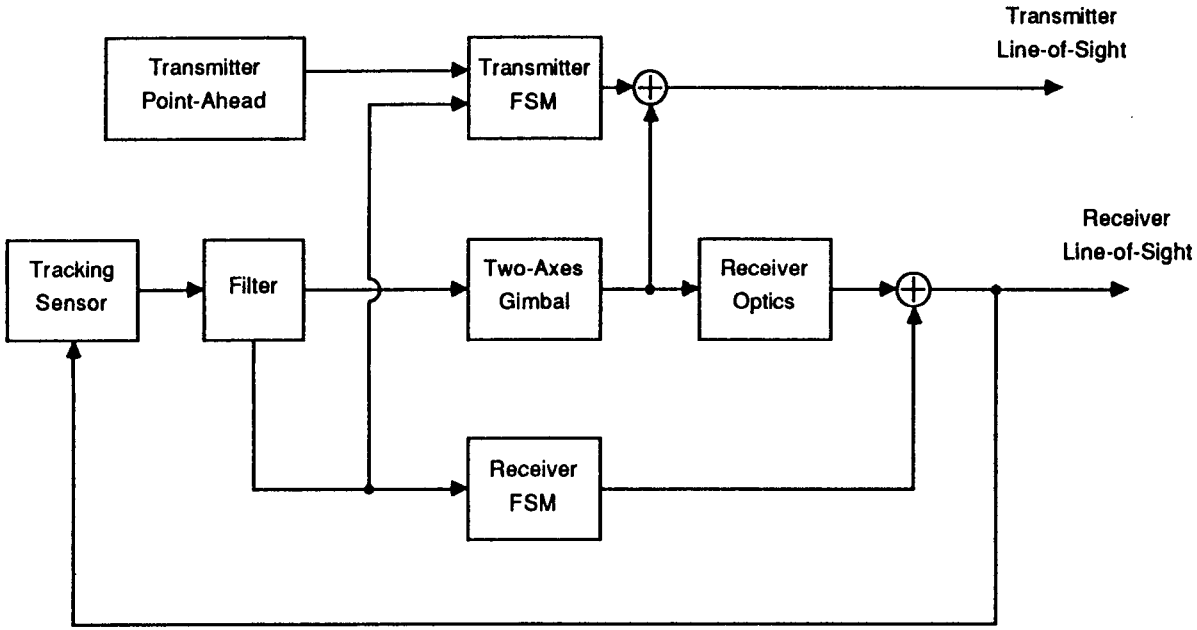


Fig.5.2. Block diagram of a spatial tracking circuit.

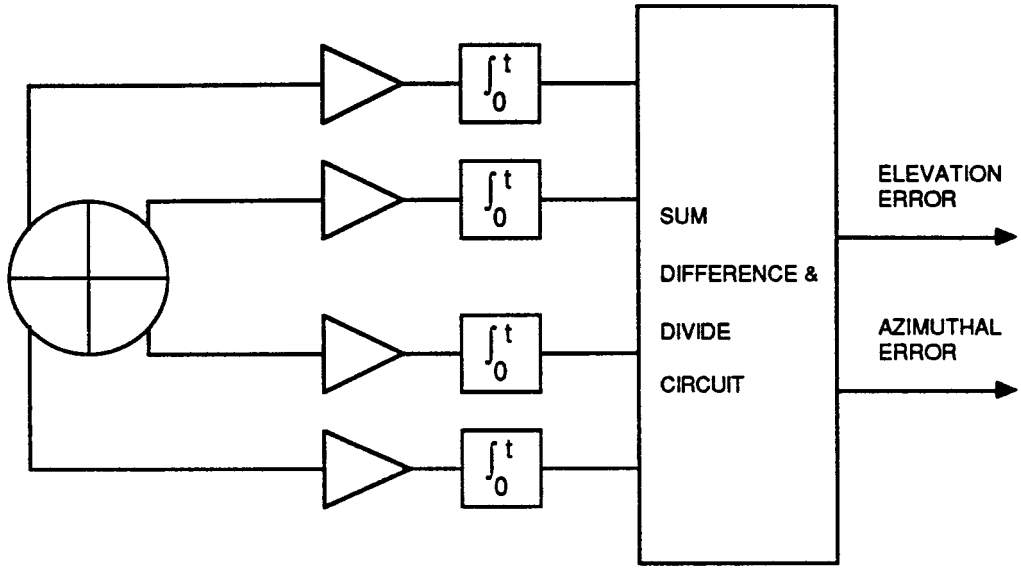


Fig. 5.3. Quadrant photodetector array and the angular position estimator.

therefore considerably simplifies the tracking loop design. For spatial tracking systems using direct detection detectors, the angular position deviations can be estimated using [56], [57]

$$\hat{\theta}_{AZ} = \left[1.22\frac{\pi\lambda}{4d_R}\right]\frac{K_1-K_2-K_3+K_4}{K_T}$$
$$\hat{\theta}_{EL} = \left[1.22\frac{\pi\lambda}{4d_R}\right]\frac{K_1+K_2-K_3-K_4}{K_T} \tag{5.1}$$

where $\hat{\theta}_{AZ}$ and $\hat{\theta}_{EL}$ are the azimuthal and elevation angular deviation estimate, d_R is the diameter of the receiver aperture, K_1,K_2,K_3,K_4 are photocounts from the four quadrants, respectively, and $K_T=K_1+K_2+K_3+K_4$ is the total photocount from the tracking detector. When the angular deviation is small and when the background noise can be neglected, the estimator shown in Eq. (5.1) can be shown to be unbiased. The variance of the estimator error (θ^2_{NEA}) can be calculated directly from Eq. (5.1) to be [56], [57]:

$$\theta^2_{NEA} = Var(\hat{\theta}) = \left[1.22\frac{\pi\lambda}{4d_R}\right]^2\frac{F}{K_S} = \left[1.22\frac{\pi\lambda}{4d_R}\right]^2 F\frac{B_T}{\lambda_S} \tag{5.2}$$

where B_T is the bandwidth of the tracking loop, K_S is the number of signal photons collected, λ_S is the signal photocount rate, and F is the excess noise factor of the APD. Equation (5.2) was derived using the assumption that the intensity pattern on the detector surface can be approximated by a uniformly illuminated circular disk with radius $1.22\lambda f/d_R$. Similar results for θ^2_{NEA} can be derived for an Airy intensity pattern with only a small change in the coefficient. Given the bandwidth of the tracking loop, the NEA due to sensor noise can be evaluated from Eq. (5.2). Because of the small tracking bandwidth, the amount of signal photons received during the observation period is generally very large. Consequently, for a practical spatial tracking system, sensor noise contribution to the tracking error is practically negligible.

In addition to the direct detection tracking sensor, angular position discrimination can also be accomplished using heterodyne sensors [54]. Ideally the heterodyne detector offers a quantum limited detection sensitivity and a higher SNR compared to the direct detection tracking sensor. In practice, the performance of the heterodyne spatial tracker is very sensitive to the LO alignment and the carrier phase noise. The sensitivity to the LO alignment error can be minimized by using a smaller receiver aperture, and the susceptibility to carrier phase noise can be reduced by using a noncoherent envelope detector. However, these measures can reduce the SNR and affect the detector NEA. In addition, heterodyne detectors are much more complicated and offer only

a marginal improvement in performance. Consequently, for most practical purposes, only direct detection tracking sensors are used.

Base motion and mechanical noise are the dominant sources of spatial tracking error. Because of the noises due to satellite motion and gimbal friction, some small motions in the optical LOS can be expected even when the sensor noise can be neglected. The effect of these noises on the tracking servo system can be more effectively analyzed using the simplified block diagram shown in Fig. 5.4. Note that the mechanical perturbation appears as an additive noise in this block diagram. Using the linearized PLL theory, the residual receiver LOS tracking error can be written as [56], [162]

$$\sigma_{\theta_R}^2 = \frac{1}{2\pi}\int S_\theta(\omega)\,|1-H_T(\omega)|^2 d\omega \tag{5.3}$$

where $S_\theta(\omega)$ is the power spectral density of the residual mechanical noise, and $H_T(\omega)$ is the closed-loop transfer function of the spatial tracking loop. The residual mechanical noise given by Eq. (5.3) can be interpreted as the amount of motion disturbance that falls outside the tracking loop bandwidth. The effect of motion and mechanical noise on the spatial tracking accuracy has been analyzed using the base motion disturbance record of the LANDSAT [56], [57], [59]. It is seen that, for a well-designed servo loop, an RMS tracking error of 0.5-1 μrad can be expected.

In addition to the sensor noise and the base motion disturbances, the accuracy of the LOS tracking is also affected by the friction of the gimbal, the error in point-ahead calculation, and other systematic errors. However, detailed studies of these noises [56], [57], [59] have shown that, with a careful implementation, the effects of these noises are small compared to that of the base motion disturbances. Consequently, for the analysis of the spatial tracking system, it is assumed that an RMS tracking accuracy of ≈0.5-1 μrad can be achieved. In particular, for the analysis of the system performance and link budget allocation, an RMS pointing error of 1μrad will be assumed. Note that since the rms tracking error is dominated by the residual mechanical vibration of the satellite, it cannot be minimized by simply increasing the signal power. Furthermore, since the azimuth and elevation trackings are decoupled by the tracking sensor, it will be assumed that the tracking errors in the azimuth and elevation directions are independent and identically Gaussian distributed. The resulting radial tracking error is therefore Rayleigh distributed with probability distribution:

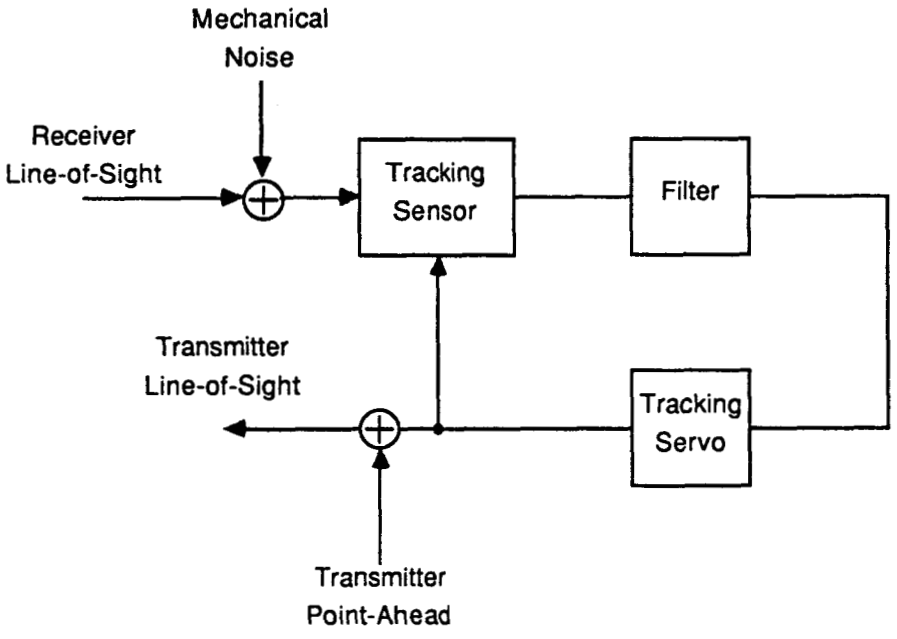


Fig.5.4. Noise equivalent block diagram of the spatial tracking circuit.

$$p(\theta) = \frac{\theta}{\sigma_\theta^2} e^{-\theta^2/2\sigma_\theta^2} \quad . \tag{5.4}$$

Finally, because the point-ahead calculation can be achieved with a high degree of accuracy, it will be assumed that, for a heterodyne system, the LO alignment error and the transmitter pointing error are identically distributed.

5.3 The Design of ISL in the Presence of Pointing and Tracking Errors

Given the performance of the spatial tracking system, the design of the communication link must be optimized to account for the tracking error. For direct detection systems, this implies tailoring the angular width of the transmitted signal in order to minimize the power requirement for a given error rate. For heterodyne systems, the finite tracking error for the LO also implies that the receiver aperture diameter must be carefully selected in order to optimize the signal-to-noise ratio.

5.3.1 Direct Detection PPM Systems

Given the receiver design, the number of signal photons collected at the receiver is a function of the transmitter power P_L, the transmitter antenna gain G_T, and the instantaneous pointing error θ_T [60]. For an M-ary PPM system operating at wavelength λ over a link distance z, the signal photocount K_S can be written as

$$K_S(P_L, G_T, \theta_T) = \eta_T G_T \left(\frac{\lambda}{4\pi z}\right)^2 \eta_R G_R L_T(G_T, \theta_T) M \left(\frac{\eta}{h\nu}\right) P_L T_s \tag{5.5}$$

where η_T, η_R are efficiencies of the transmitter and receiver optics, G_T, G_R are the transmitter and receiver antenna gains, $L_T(G_T, \theta_T)$ is the pointing loss factor associated with the pointing error θ_T, η is the quantum efficiency of the detector, $h\nu$ is the photon energy, P_L is the average transmitter power, T_s is the time slot period, and it has been assumed that the peak transmitter power is M times the average power. The receiver antenna gain G_R can be related to the receiver aperture diameter d_R and the wavelength λ by

$$G_R = \left(\frac{\pi d_R}{\lambda}\right)^2 \quad . \tag{5.6}$$

For a transmitter with a uniformly illuminated, unobscured circular aperture of diameter d_T, the transmitter antenna gain and the pointing loss factor can be written as

$$G_T = (\frac{\pi d_T}{\lambda})^2 , \tag{5.7}$$

$$L_T(G_T,\theta_T) = \left[\frac{2J_1(\pi\theta_T d_T/\lambda)}{\pi\theta_T d_T/\lambda}\right]^2 = \left[\frac{2J_1(\sqrt{G_T}\theta_T)}{\sqrt{G_T}\theta_T}\right]^2 . \tag{5.8}$$

In practice, the intensity cross-section of a single mode laser output can be better approximated by a Gaussian beam. When the intensity cross section of the signal at the transmitter aperture is Gaussian with RMS width W, G_T and L_T are given by

$$G_T = (\frac{2\pi W}{\lambda})^2 , \tag{5.9}$$

$$L_T(G_T,\theta_T) = \exp\left[-G_T\theta_T^2\right] . \tag{5.10}$$

Because the instantaneous pointing error fluctuates in time, the signal photocount K_S and the probability of bit error (PBE) through the ISL will be functions of time. For a given RMS pointing error, σ_{θ_T}, the average PBE at the receiver can be written as:

$$PBE = \int_0^{\infty} P_E[K_S(P_L,G_T,\theta_T),K_B]\frac{\theta_T}{\sigma_{\theta_T}^2} e^{-\theta_T^2/2\sigma_{\theta_T}^2} d\theta_T \tag{5.11}$$

where $P_E(K_S,K_B)$ is the PBE of the direct detection PPM system given the signal photocount K_S and the background photocount K_B, and it has been assumed that the pointing error θ_T is Rayleigh distributed. Equation (5.11) shows that the PBE is a function of P_L, G_T, and σ_{θ_T}. By normalizing the instantaneous pointing error θ_T with respect to its RMS value, however, the signal photocount K_S in Eq. (5.5) can be written as

$$K_S(P_L,G_T,\theta_T) = \left[\frac{\alpha_0 P_L}{\sigma_{\theta_T}^2}\right]\left[G_T\sigma_{\theta_T}^2\right] L_T(G_T\sigma_{\theta_T}^2, \frac{\theta_T}{\sigma_{\theta_T}}) \tag{5.12}$$

where $\alpha_0 = \eta_T\eta_R(\frac{\eta}{h\nu})(\frac{\lambda}{4\pi z})^2(\frac{\pi d_R}{\lambda})^2 MT_s$ is a constant that is independent of the transmitter antenna gain and the signal power. Note that the parameter $\alpha_0 P_L$ represents the signal photocount per unit transmitter antenna gain at the detector in the absence of spatial tracking error. By substituting Eq. (5.12) into Eq. (5.11) and carrying out the substitution of variable $u=\theta_T/\sigma_{\theta_T}$, it is seen that, instead of three independent parameters, the PBE for a direct detection PPM systems depends only on the parameters $\alpha_0 P_L/\sigma_{\theta_T}^2$ and $G_T\sigma_{\theta_T}^2$. Consequently, a constant PBE can be maintained at different values of the RMS pointing error provided that the transmitter

power P_L and the transmitter antenna gain G_T are scaled such that the parameters $\alpha_0 P_L/\sigma_{\theta_T}^2$ and $G_T\sigma_{\theta_T}^2$ remain constant.

For a given σ_{θ_T}, there is a finite probability that a large instantaneous pointing error can occur which results in a small received signal photocount and a high probability of error regardless of the transmitter power. Thus, for a given transmitter power, the PBE increases rapidly when the RMS pointing jitter becomes a significant fraction of the transmitted beamwidth. This fact is demonstrated in Fig. 5.5 where the PBE of a 4-ary PPM system is plotted against the signal power for several values of the RMS pointing error σ_{θ_T}. Note that when the RMS pointing jitter exceeds ≈15% of λ/d_T, a much higher signal power is needed to maintain the PBE at 10^{-9}. This is because at large σ_{θ_T}, the probability that the main lobe of the transmitted signal does not cover the receiver aperture can exceed the desired PBE and, therefore, result in a large power penalty. The power loss due to pointing error can be minimized by increasing the angular width of the transmitted signal or, equivalently, reducing the size of the transmitter aperture. However, even though a wider transmitted signal reduces the loss due to pointing error, a larger angular width of the transmitted signal implies a smaller signal intensity at the receiver and, consequently, a smaller detector SNR and a higher error rate. Given the receiver design, therefore, there is a trade-off between the size of the transmitter aperture and the power of the transmitter laser. Figure 5.6 is a plot of the transmitter power needed to achieve 10^{-9} error rate versus the transmitter antenna G_T for a 4-ary PPM system operating at 200 Mbps. Note that, when the spatial tracking can be perfectly realized ($\sigma_{\theta_T} = 0$), the required transmitter power is inversely proportional to G_T. When $\sigma_{\theta_T} \neq 0$, however, the required signal power first decreases with increasing G_T until an optimum antenna gain is reached, and any further increase in the transmitter antenna gain will only result in a higher power requirement. This is because at large transmitter antenna gains the probability of error due to the transmitter pointing error becomes more significant as the beamwidth decreases with an increasing aperture size. Consequently, for a given RMS pointing jitter and PBE, an optimal antenna gain exists which minimizes the laser power requirement.

It is of interest to investigate the optimum transmitter antenna gain and the required signal power as functions of the RMS pointing jitter. By differentiating both sides of Eq. (5.11) with respect to the antenna gain G_T while maintaining a constant PBE, it is seen that

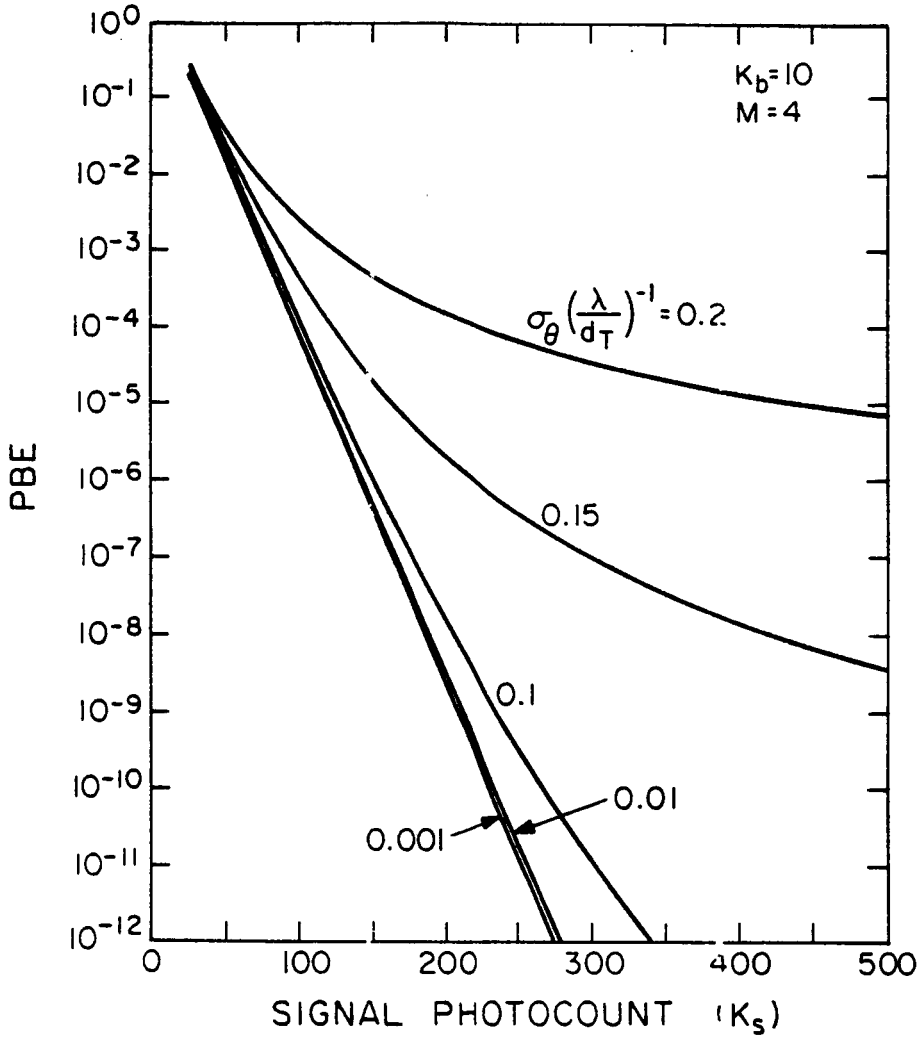


Fig.5.5. PBE versus the signal power K_S for a 4-ary PPM system at several values of the rms pointing error σ_θ. The photodetector is assumed to be an APD with excess noise factor F=5.

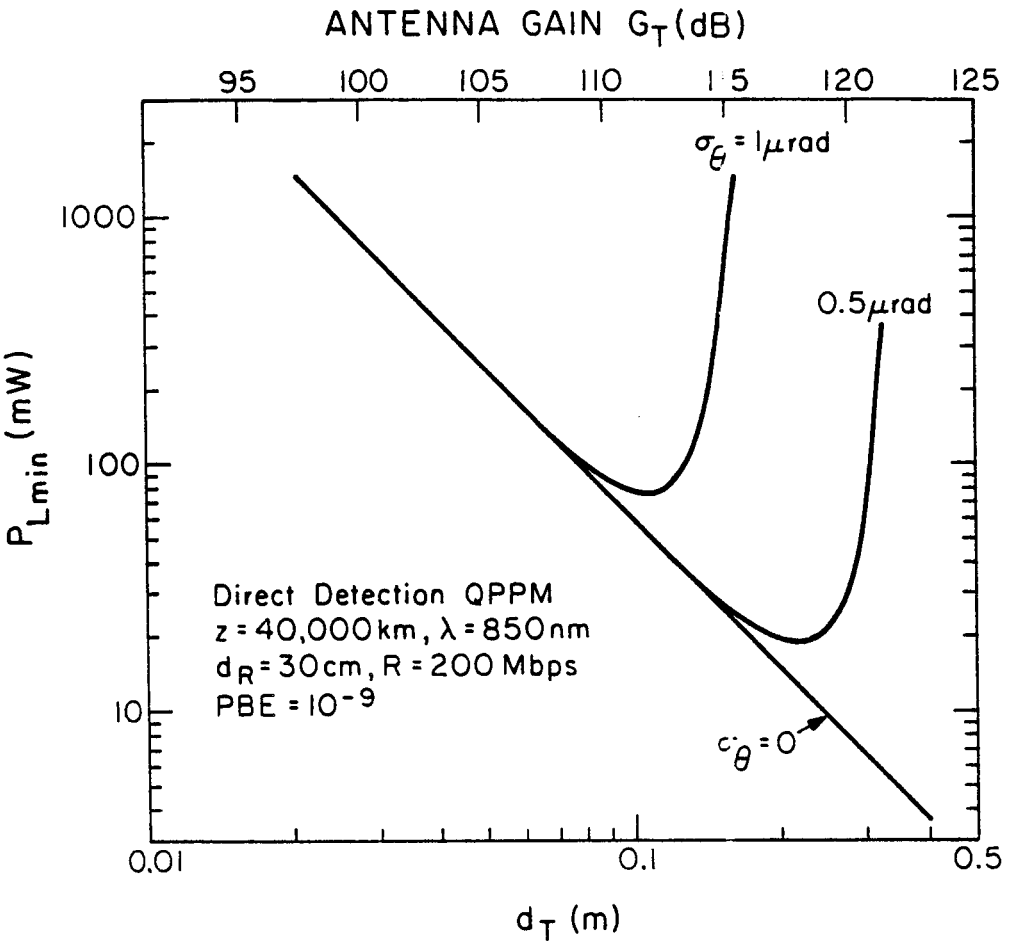


Fig.5.6. Required transmitter power at $PBE=10^{-9}$ versus the diameter of the transmitter aperture d_T and the transmitter antenna gain G_T for a 4-ary PPM system with rms pointing errors of 0, 0.5 and 1μrad.

$$\frac{d\,PBE}{dG_T} = 0 = \int_0^{\infty} \frac{\partial P_E(K_S,K_B)}{\partial K_S}\left[\frac{\partial K_S}{\partial G_T} + \frac{\partial K_S}{\partial P_L}\frac{dP_L}{dG_T}\right] u e^{-u^2/2} du \tag{5.13}$$

where K_S is an implicit function of $P_L/\sigma_{\theta_T}^2$, $G_T\sigma_{\theta_T}^2$, and $u=\theta_T/\sigma_{\theta_T}$. At the optimal antenna gain $G_{T_{opt}}$ which minimizes the required signal power, $dP_L/dG_T=0$, and it follows from Eq. (5.13) that

$$\int_0^{\infty} \frac{\partial P_E(K_S,K_B)}{\partial K_S}\frac{\partial K_S(\alpha_0 P_L/\sigma_{\theta_T}^2, G_T\sigma_{\theta_T}^2, u)}{\partial G_T} u e^{-u^2/2} du = 0 \ , \ G_T=G_{T_{opt}}, \ P_L=P_{L_{min}} \ . \tag{5.14}$$

Given PBE and modulation format, Eqs. (5.11) and (5.14) are a pair of equations that can be solved simultaneously for the parameters $\alpha_0 P_L/\sigma_{\theta_T}^2$ and $G_T\sigma_{\theta_T}^2$. Since Eqs. (5.11) and (5.14) depend on the RMS pointing error only through $\alpha_0 P_L/\sigma_{\theta_T}^2$ and $G_T\sigma_{\theta_T}^2$, it follows that, if $\gamma_T=G_{T_{opt}}\sigma_{\theta_T}^2$ and $\gamma_P=\alpha_0 P_{L_{min}}/\sigma_{\theta_T}^2$ can be found to satisfy both Eqs. (5.11) and (5.14), then at all values of σ_{θ_T} the optimal aperture and the minimal power requirement can be related to the RMS pointing error by

$$G_{T_{opt}} = \gamma_T/\sigma_{\theta_T}^2 \tag{5.15}$$

and

$$P_{L_{min}} = \left[\frac{\gamma_P}{\alpha_0}\right]\sigma_{\theta_T}^2 \ . \tag{5.16}$$

That is, given the PBE and the modulation format, the optimal transmitter antenna gain is inversely proportional to the square of the RMS pointing error and the required transmitter power $P_{L_{min}}$ is proportional to the square of the RMS pointing error. Note that the values of γ_T and γ_P are independent of the RMS pointing jitter. In practice, given the transmitter pointing error, γ_T and γ_P can be calculated numerically. Plotted in Fig. 5.7 is the value of γ_T versus the PBE for a 4-ary PPM system calculated using the ISL parameters shown in Table 2.1. It is seen that, for systems with a uniformly illuminated circular aperture, the value of γ_T ranges from approximately 0.15 rad^2 at $PBE=10^{-9}$ to about 0.5 rad^2 at $PBE=10^{-3}$. For systems operating at λ=850 nm and RMS pointing error of 1 μrad, this is equivalent to an optimum transmitter aperture diameter of ≈10 cm at PBE=10^{-9} and ≈17 cm at PBE=10^{-3}. The value of γ_T is slightly higher for systems where the transmitted signal can be approximated by Gaussian beams.

Figure 5.8 is a plot of γ_P versus the PBE for a 4-ary PPM ISL. The value of the parameter γ_P ranges from ≈200/rad^2 at PBE = 10^{-3} to approximately 1500/rad^2 at PBE = 10^{-9}. Note that the transmitter power needed to

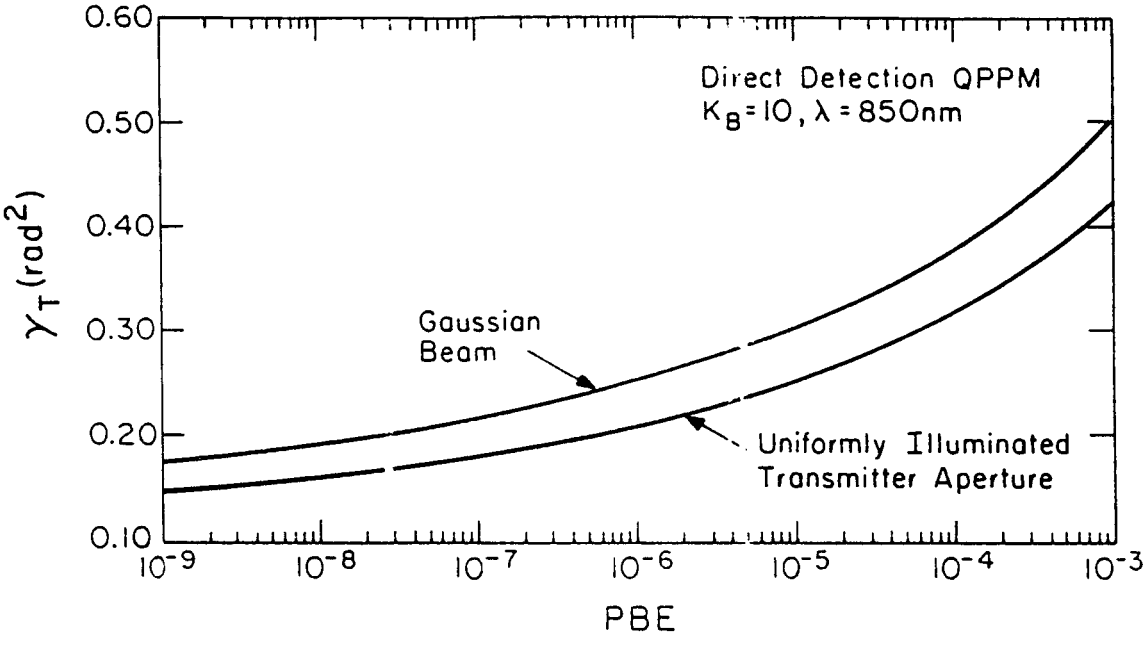


Fig.5.7. The parameter $\gamma_T = G_{T_{opt}} \sigma_\theta^2$ versus the PBE for a 4-ary PPM system.

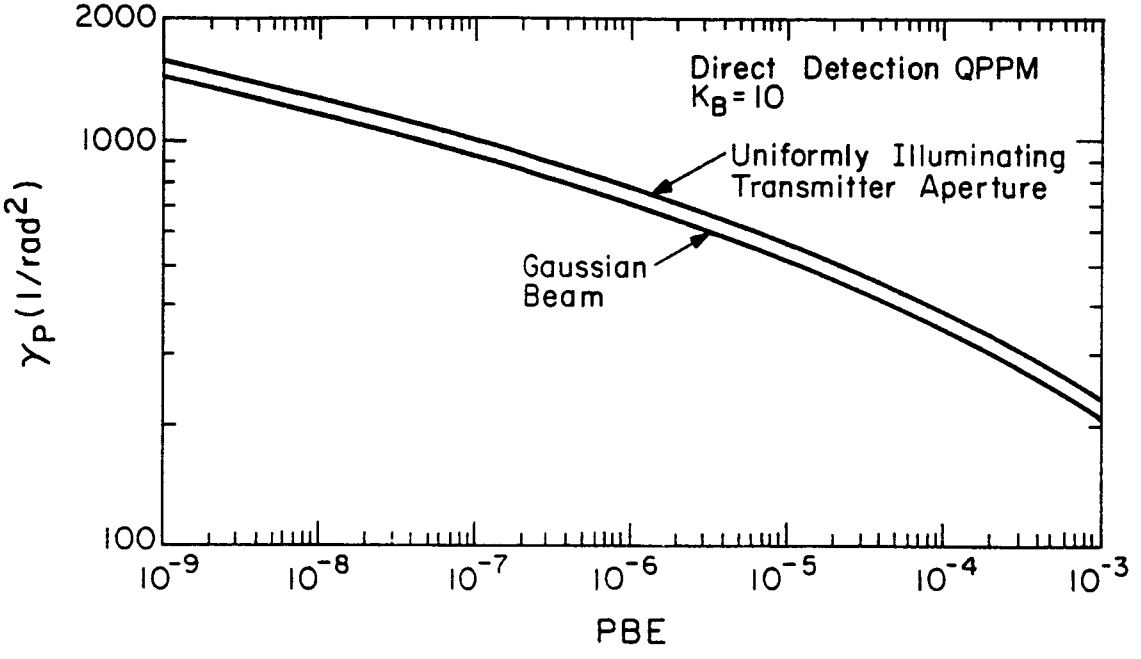


Fig.5.8. The parameter γ_P versus PBE for a 4-ary PPM system.

achieve a given PBE is smaller for systems where the intensity cross section of the signal is Gaussian distributed than for systems with uniformly illuminated apertures.

The power penalty associated with the spatial tracking error can be calculated by comparing the signal power needed to achieve a fixed error rate in the presence of spatial tracking errors with the required signal level when the spatial tracking can be perfectly realized. The presence of this power penalty can be seen from Fig. 5.6. Note that, even when the optimal transmitter aperture is used, the signal power required to achieve a given error rate is about 1.6 dB higher than when the RMS pointing error can be ignored. By combining Eqs. (5.15) and (5.16), the required transmitter power can be related to the optimal transmitter antenna gain by

$$P_{L_{\min}} = \left[\frac{\gamma_P \gamma_T}{\alpha_0}\right] \frac{1}{G_{T_{opt}}} \quad . \tag{5.17}$$

For systems with perfect spatial tracking, the signal power needed to maintain a constant PBE can also be inversely related to the transmitter antenna gain G_T by inverting Eq. (5.5). Since the required transmitter power is inversely proportional to the transmitter antenna gain in both cases, it follows that for a given antenna gain, the ratio between the required signal power in the presence of random pointing error and that when the spatial tracking can be perfectly realized is a constant which is independent of G_T. Furthermore, because both γ_T and γ_P are independent of σ_{θ_T}, the power penalty due to spatial tracking error is independent of the RMS pointing error. Figure 5.9 is the plot of this power penalty as a function of the PBE for a 4-ary PPM system. Note that the power penalty ranges from 1.6 dB at PBE=10^{-9} to about 2.3 dB at PBE=10^{-3}. For systems with Gaussian transmitted beams, the power penalty is slightly higher. This is because the optimal transmitter antenna gain is smaller for an uniformly illuminated aperture than for a Gaussian beam. Therefore, a larger transmitter power is needed to achieve a given PBE when $\sigma_{\theta_T} = 0$, and hence a smaller power penalty.

Table 5.1 shows a typical link budget for a 4-ary direct detection PPM system calculated using the ISL parameters shown in Table 2.1. Note that for direct detection systems, there is no constraint on the size of the receiver aperture. The receiver aperture diameter of 30 cm was chosen because of the size and weight constraints. In contrast, the presence of a random transmitter pointing error results in an optimal transmitter aperture diameter which minimizes the required signal power. Compared to the system using a 30 cm diameter, uniformly illuminated, transmitter telescope, the ISL design that is optimized for an RMS pointing error of 1 μrad has a much smaller transmitter antenna gain (≈7.5 dB at σ_{θ_T}=1 μrad and PBE=10^{-6}). In addition, a power

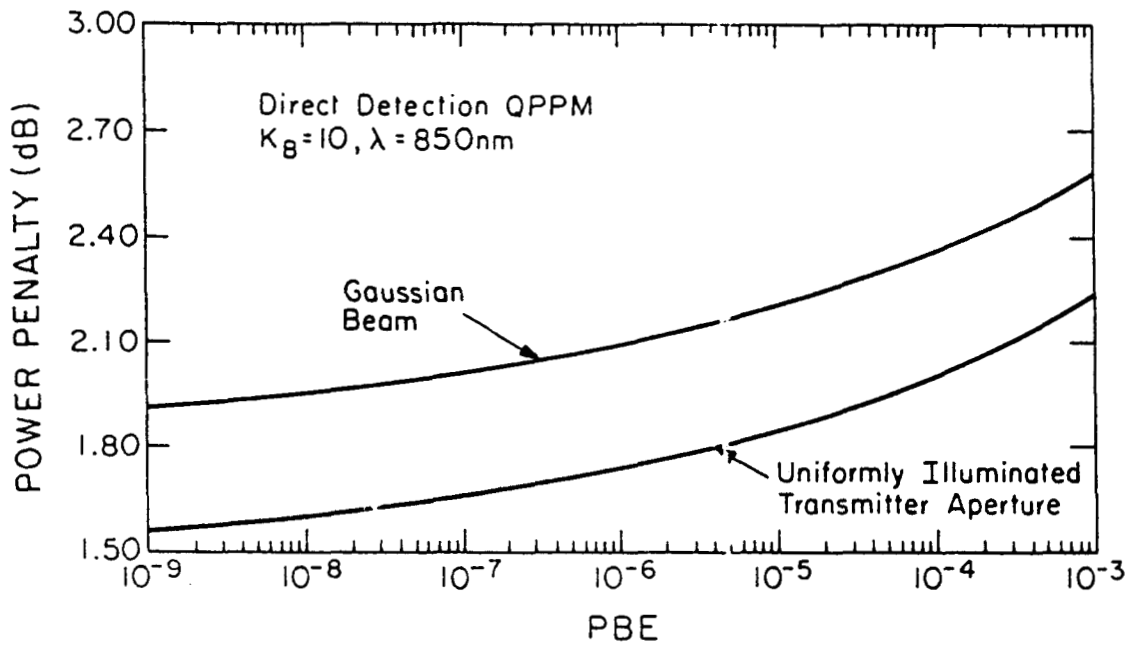


Fig.5.9. Power penalty due to the transmitter pointing error versus the PBE for a 4-ary PPM system.

Table 5.1 Typical link budget of a direct detection QPPM system with 1 μrad RMS pointing error.

Modulation Format	4-PPM	
Data Rate	200 Mbps	
Laser Power	100 mW	20.0 dBm
Path Loss	40,000km	-295.4 dB
Antenna Gain		
Transmitter†	12.6 cm	113.4 dB
Receiver	30 cm	120.9 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-6}	73 Photons/bit	-53.1 dBm
Pointing Loss		-1.8 dB
Synchronization Loss		-0.1 dB
Link Margin		2.5 dB

† Optimum for $\sigma_{\theta_T} = 1$ μrad.

penalty (≈1.8 dB at PBE=10^{-6}) must be allocated to account for the random pointing error. Compared to the small power penalty due to synchronization error (0.1 dB) [66], it is seen that the spatial tracking error dominates the power allocation in the direct detection ISL.

5.3.2 Heterodyne NCFSK Systems

For heterodyne systems, spatial tracking errors affect not only the pointing accuracy of the transmitter, but also the alignment between the received signal and the LO [3]. Given an instantaneous LO tracking error of θ_R and a transmitter pointing error of θ_T, the signal-to-noise ratio (SNR) at the heterodyne receiver can be written as

$$\beta(P_L,G_T,G_R,\theta_T,\theta_R) = \eta_T(\frac{\pi d_T}{\lambda})^2(\frac{\lambda}{4\pi z})^2\eta_R(\frac{\pi d_R}{\lambda})^2\left[\frac{2J_1(\pi\theta_T d_T/\lambda)}{\pi\theta_T d_T/\lambda}\right]^2\left[\frac{2J_1(\pi\theta_R d_R/\lambda)}{\pi\theta_R d_R/\lambda}\right]^2\left[\frac{\eta P_L}{h\nu}\right]T_s$$
$$= \frac{\alpha_1 P_L}{\sigma_{\theta_T}^2\sigma_{\theta_R}^2}(G_T\sigma_{\theta_T}^2)(G_R\sigma_{\theta_R}^2)L_T(G_T\sigma_{\theta_T}^2,\theta_T/\sigma_{\theta_T})L_R(G_R\sigma_{\theta_R}^2,\theta_R/\sigma_{\theta_R}) \tag{5.18}$$

where $\alpha_1=\eta_T\eta_R(\frac{\eta}{h\nu})(\frac{\lambda}{4\pi z})^2T_s$ is a constant which is independent of the transmitter power and the pointing and tracking errors. Note that $\alpha_1 P_L$ is simply the receiver SNR per unit transmitter and receiver antenna gains. In deriving Eq. (5.18), it has been assumed that the transmitter aperture is uniformly illuminated and that both signal and LO can be approximated by plane waves that cover the receiver aperture of diameter d_R. In addition to the pointing loss factor, $L_T(G_T\sigma_{\theta_T}^2,\theta_T/\sigma_{\theta_T})$, the presence of the LO tracking error results in a tracking loss factor $L_R(G_R\sigma_{\theta_R}^2,\theta_R/\sigma_{\theta_R})$ which, when both signal and LO are plane waves, is given by

$$L_R(G_R\sigma_{\theta_T}^2,\theta_R/\sigma_{\theta_T}) = \left[\frac{2J_1(\pi\theta_R d_R/\lambda)}{\pi\theta_R d_R/\lambda}\right]^2 = \left[\frac{2J_1(\sqrt{G_R\sigma_{\theta_T}^2}\theta_R/\sigma_{\theta_T})}{\sqrt{G_R\sigma_{\theta_T}^2}\theta_R/\sigma_{\theta_T}}\right]^2. \tag{5.19}$$

By normalizing the pointing and tracking errors by their respective RMS values, the receiver SNR depends on pointing and tracking errors only through $(\alpha_1 P_L/\sigma_{\theta_T}^2\sigma_{\theta_R}^2)$, $G_T\sigma_{\theta_T}^2$, $G_R\sigma_{\theta_R}^2$, $\theta_T/\sigma_{\theta_T}$, and $\theta_R/\sigma_{\theta_R}$.

The probability of error for a binary heterodyne NCFSK system, conditioned on the IF SNR, can be written as [28]

$$P_E = \frac{1}{2}e^{-\beta/2}. \tag{5.20}$$

Equation (5.20) was derived assuming that the carrier phase noise has a negligible impact on the performance of

the NCFSK system. In the presence of carrier phase noise, the PBE depends not only on the receiver SNR, but also on the linewidths of the transmitter and LO lasers. The unconditional probability of error for the heterodyne NCFSK channel can be written as

$$PBE = \iint P_E\left[\beta\left(\frac{\alpha_1 P_L}{\sigma_{\theta_T}^2\sigma_{\theta_T}^2}, G_T\sigma_{\theta_T}^2, G_R\sigma_{\theta_R}^2, u, v\right)\right] ue^{-u^2/2}\, ve^{-v^2/2}\, du\, dv \tag{5.21}$$

where $u=\theta_T/\sigma_{\theta_T}$ and $v=\theta_R/\sigma_{\theta_R}$ are the normalized pointing and tracking errors, respectively. Note that the unconditional PBE depends only on three parameters $(\alpha_1 P_L/\sigma_{\theta_T}^2\sigma_{\theta_R}^2)$, $G_T\sigma_{\theta_T}^2$, and $G_R\sigma_{\theta_R}^2$. Because of the trade-off between the pointing loss factor and the beam divergence, it can be argued that, for a given RMS pointing error and the desired PBE, an optimal transmitter antenna gain exists which minimizes the required signal power. For a heterodyne system, the presence of an LO tracking loss factor suggests a similar trade-off between the receiver aperture diameter and the signal-to-noise ratio. By inspecting Eq. (5.18) it is seen that, when the LO tracking can be perfectly realized, the receiver SNR is proportional to the receiver antenna gain G_R. In the presence of the LO tracking error, however, the SNR decreases rapidly when the argument of the Bessel function, $\sqrt{G_R}\theta_R$, approaches 3.84. Consequently, the receiver aperture diameter cannot be increased indefinitely, and an optimal receiver antenna gain exists which maximizes the system performance.

By differentiating both sides of Eq. (5.21) with respect to G_T and G_R, and realizing that $dP_L/dG_T=dP_L/dG_R=0$ at optimal antenna gains, it is seen that the optimal antenna gains must satisfy

$$\int \frac{dP_E(\beta)}{d\beta}\frac{\partial\beta\left(\frac{\alpha_1 P_L}{\sigma_{\theta_T}^2\sigma_{\theta_R}^2}, G_T\sigma_{\theta_T}^2, G_R\sigma_{\theta_R}^2, u, v\right)}{\partial G_T} ue^{-u^2/2}\, ve^{-v^2/2}\, du\, dv = 0, \tag{5.22}$$

$$\int \frac{dP_E(\beta)}{d\beta}\frac{\partial\beta\left(\frac{\alpha_1 P_L}{\sigma_{\theta_T}^2\sigma_{\theta_R}^2}, G_T\sigma_{\theta_T}^2, G_R\sigma_{\theta_R}^2, u, v\right)}{\partial G_R} ue^{-u^2/2}\, ve^{-v^2/2}\, du\, dv = 0, \quad G_T=G_{T_{opt}},\ G_R=G_{R_{opt}}. \tag{5.23}$$

Given the PBE, the optimal antenna gains and the required signal power can be calculated by solving Eqs. (5.21)-(5.23). Note that these equations depend only on $\alpha_1 P_L/(\sigma_{\theta_T}^2\sigma_{\theta_R}^2)$, $G_T\sigma_{\theta_T}^2$, and $G_R\sigma_{\theta_R}^2$. Therefore, if $\gamma_T=G_{T_{opt}}\sigma_{\theta_T}^2$, $\gamma_R=G_{R_{opt}}\sigma_{\theta_R}^2$, and $\gamma_P=\alpha_1 P_{L_{min}}/(\sigma_{\theta_T}^2\sigma_{\theta_R}^2)$ can be found to satisfy Eqs. (5.21)-(5.23) for a given PBE, then for all values of the RMS pointing and tracking errors the optimal antenna gains can be related to the RMS pointing and tracking jitters by

$$G_{T_{opt}} = \gamma_T / \sigma_{\theta_T}^2 \ ,$$
$$G_{R_{opt}} = \gamma_R / \sigma_{\theta_R}^2 \ . \tag{5.24}$$

Furthermore, the required transmitter power can be related to the RMS pointing and tracking errors by

$$P_{L_{\min}} = \left[\frac{\gamma_P}{\alpha_1} \right] \sigma_{\theta_R}^2 \sigma_{\theta_T}^2 \ . \tag{5.25}$$

When $\sigma_{\theta_T} = \sigma_{\theta_R}$, the symmetry of Eq. (5.18) between G_T and G_R suggests that the optimal transmitter and receiver antenna gains are equal to each other. As a result, $\gamma_T = \gamma_R$ for all values of σ_{θ_T} and σ_{θ_R}.

The parameter γ_T can be calculated by solving Eqs. (5.21)-(5.23). The resulting value of γ_T is plotted in Fig. 5.10 versus the PBE for a binary NCFSK system. By comparing Fig. 5.10 and Fig. 5.7, it is seen that, for a similar RMS transmitter pointing error, the performance of the heterodyne system is optimized at a slightly higher transmitter antenna gain than the direct detection PPM system. However, it should be noted that for heterodyne systems, the receiver antenna gain is constrained by the RMS LO tracking error while no such constraint exists for the direct detection PPM system. As a result, the product of the transmitter and receiver antenna gains is usually smaller for the heterodyne system than for a comparable direct detection system.

Figure 5.11 is a plot of γ_P versus PBE for a heterodyne 4-ary NCFSK system. The value of γ_P ranges from $\approx$160/rad^4 at PBE = 10^{-3} to approximately 2,600/rad^4 at PBE = 10^{-9}. Note that the required transmitter power for a heterodyne NCFSK system is proportional to both $\sigma_{\theta_T}^2$ and $\sigma_{\theta_R}^2$. In contrast, the required signal power for a direct detection PPM system is proportional to $\sigma_{\theta_T}^2$. Consequently, the heterodyne channel is more sensitive to spatial tracking errors. For systems with large RMS pointing and tracking jitters, the amount of transmitter power needed to achieve a given PBE can be significantly larger for the heterodyne system despite the higher receiver sensitivity. For example, given the link parameters shown in Table 2.1 and the parameter γ_P plotted in Fig. 5.11, the required transmitter power needed to achieve 10^{-6} error rate is approximately 0.1 mW at $\sigma_{\theta_T} = \sigma_{\theta_R}$ = 0.2 μrad). However, at large pointing and tracking jitters, the required transmitter power is unacceptably high ($\approx$1 W for $\sigma_{\theta_T} = \sigma_{\theta_R}$ = 2 μrad).

In addition to imposing constraints on the antenna gains, random pointing and tracking errors also incur a power penalty in the link budget. By combining Eqs. (5.24) and (5.25), the required transmitter power in the presence of pointing and tracking errors can be related to the optimal transmitter and receiver antenna gains by

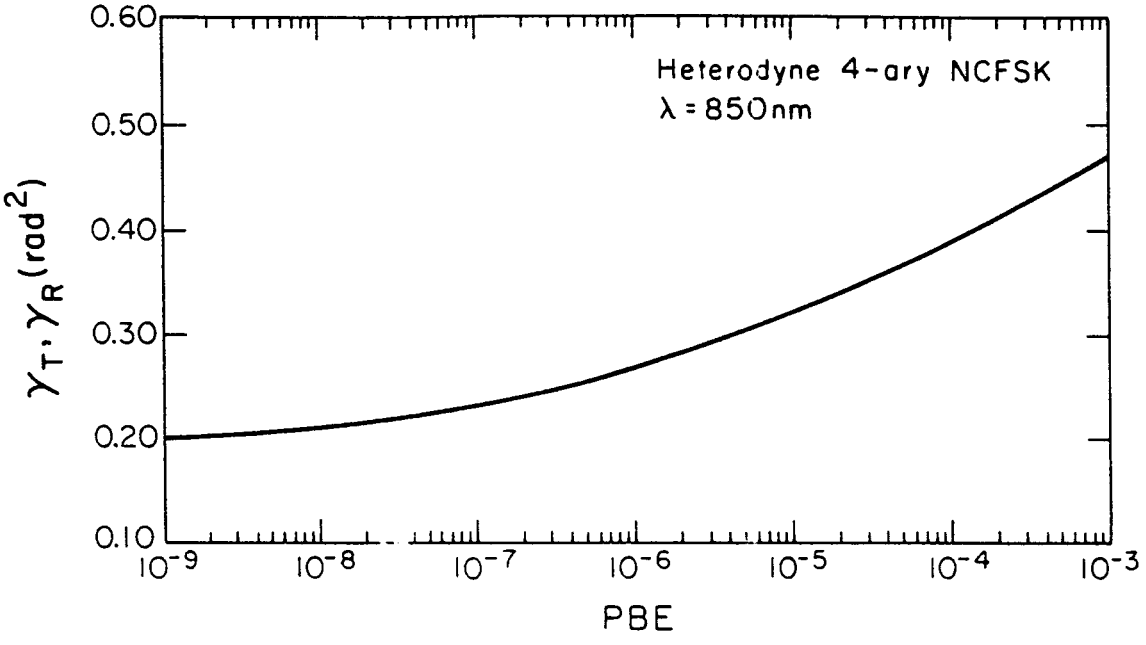


Fig.5.10. The parameter $\gamma_T = G_{T_{opt}} \sigma^2_{\theta_T}$ and $\gamma_R = G_{R_{opt}} \sigma^2_{\theta_R}$ versus the PBE for a 4-ary NCFSK system. The transmitter aperture is assumed to be uniformly illuminated by the signal.

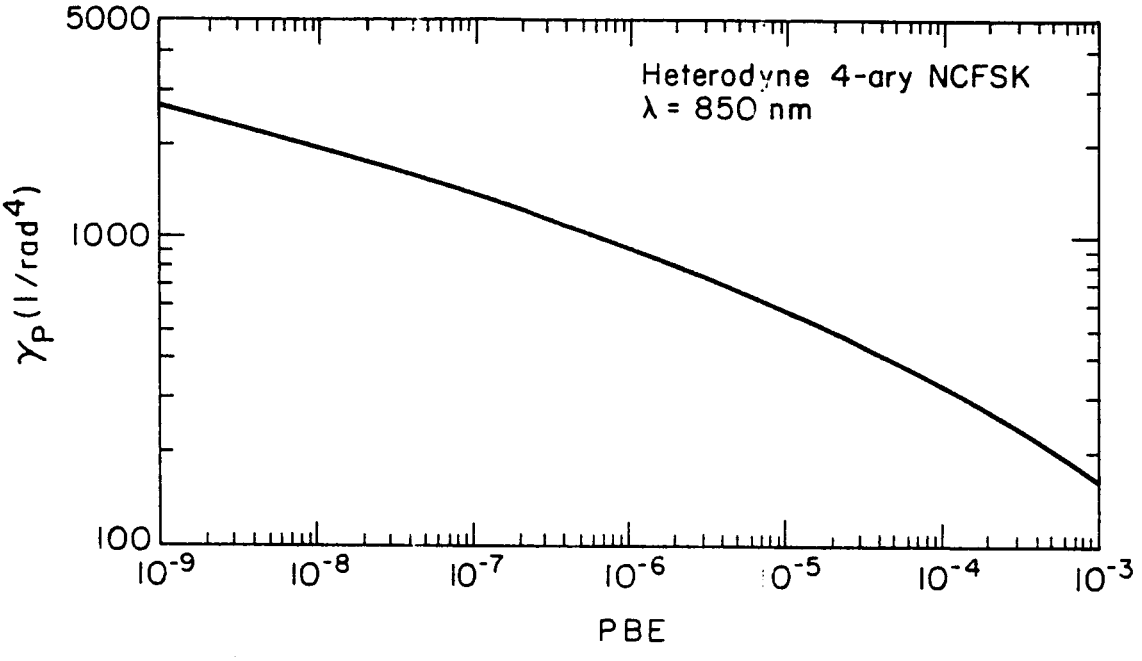


Fig.5.11. The parameter γ_p versus PBE for a 4-ary NCFSK system.

$$P_{L_{\min}} = \left[\frac{\gamma_P \gamma_T \gamma_R}{\alpha_1} \right] \frac{1}{G_{T_{opt}} G_{R_{opt}}} \quad . \tag{5.26}$$

In the absence of pointing and tracking errors, it is seen from Eq. (5.18) that, in order to maintain a constant PBE, the required transmitter power is also inversely related to the product of G_T and G_R. Consequently when both transmitter and receiver antenna gains are optimized, the power penalty, which is defined as the ratio between the required signal power in the presence of pointing and tracking errors and that when spatial tracking can be perfectly realized, is independent of the antenna gains. It follows from Eq. (5.23) that the power penalty is independent of the RMS pointing and tracking errors. The power penalty for a binary FSK system is plotted in Fig. 5.12 versus the PBE. Note that the power penalty associated with the spatial tracking error is approximately 4 dB for $PBE=10^{-9}$, and increases to about 5 dB for $PBE=10^{-3}$. The power penalty shown in Fig. 5.12 is calculated assuming that both transmitter and receiver antenna gains are optimized. Therefore, despite the relatively small increase in the power penalty (2-3 dB) over the direct detection system, the actual power margin in a heterodyne system is much smaller since a smaller receiver antenna must be used.

Table 5.2 shows the typical link budget for a 4-ary NCFSK system designed with 1 μrad RMS transmitter pointing and LO tracking errors. The link budget is calculated using the ISL parameters shown in Table 2.1. Compared to an ideal system using 30 cm transmitter and receiver telescopes, the design optimized for RMS pointing and tracking errors of 1 μrad has much smaller transmitter and receiver antenna gains. In addition, a power penalty (≈4 dB) must be allocated in the link budget because of the random pointing and tracking errors. By comparing Tables 5.1 and 5.2, it is seen that the spatial tracking error affects the heterodyne channel much more severely than the direct detection channel. In fact, the advantage of using a heterodyne system because of the higher detection sensitivity (≈5-10 dB) is quickly offset by the smaller receiver antenna gain and the larger power penalty due to spatial tracking error. The link budget shown in Table 5.2 was calculated assuming that a signal power of 100 mW is available at the transmitter. In practice, because heterodyne systems are very sensitive to carrier phase noise, mode-stabilized semiconductor lasers with dynamic linewidths on the order of several megahertz are required to maintain a reasonable detection sensitivity. Such lasers are not readily available at power output exceeding 50 mW. It should be emphasized, however, that even though the power budget given in Table 5.2 shows a smaller power margin than the direct detection PPM system, heterodyne systems offer better background noise immunity and narrower field-of-view. In the presence of strong

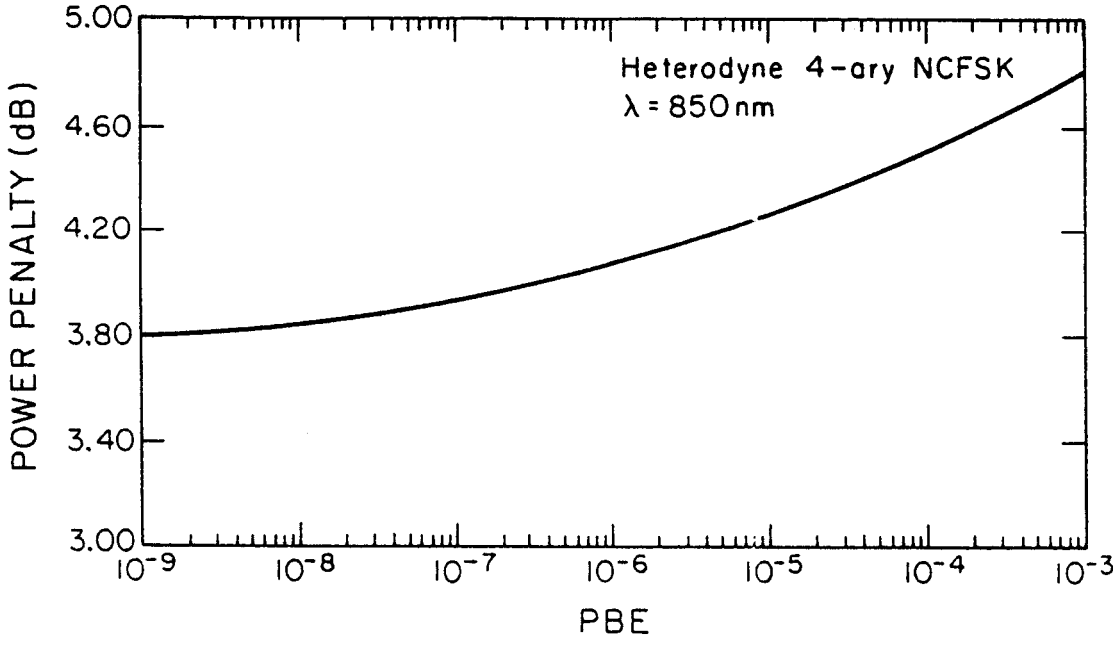


Fig.5.12. Power penalty associated with the spatial tracking error versus the PBE for a binary NCFSK system.

Table 5.2 Link Budget of a heterodyne QFSK system with $\sigma_{\theta_T} = \sigma_{\theta_R} = 1$ μrad.

Modulation Format	4-FSK	
Data Rate	200 Mbps	
PBE	10^{-6}	
Laser Power	100 mW	20.0 dBm
Path Loss	40,000km	-295.4 dB
Antenna Gain		
Transmitter†	14.1 cm	114.3 dB
Receiver†	14.1 cm	114.3 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-6}	14 Photons/bit	-60.4 dBm
Pointing and Tracking Loss		-4.0 dB
Synchronization Loss		-0.1 dB
Phase Noise @ PBE=10^{-6} $\Delta_f T_s$=0.03, and $\nu_d T_s$=1		-0.8 dB
Link Margin		1.1 dB

Δ_f = IF linewidth, ν_d = FSK tone spacing.

† Optimum for $\sigma_{\theta_T} = \sigma_{\theta_R} = 1$ μrad.

background noise and hostile jamming, heterodyne systems can have a superior performance. Furthermore, the power budget shown in Table 5.2 was calculated for RMS pointing and tracking errors of 1 μrad. By reducing the RMS errors from 1 μrad to 0.5 μrad, the power margin of the heterodyne system can be increased by 12 dB compared to a 6 dB increase in power margin for the direct detection PPM system.

5.3.3 Suboptimal Design

Because the optimal antenna gains are inversely related to the RMS pointing and tracking errors, antenna diameters at which the system performance is optimized can be very large for systems with small RMS pointing errors. For instance, the optimal antenna diameter is greater than 50 cm for an optical ISL operating with 0.2 μrad RMS pointing errors. For such systems, the sizes of the transmitter and receiver antennas are limited by the weight and budget constraints rather than performance considerations. In practice, these systems usually employ a suboptimal design in which the antenna gains are chosen to be smaller than the optimal values. Because the antenna gains are not optimized, higher signal power is needed to maintain the system performance. However, since the transmitted beamwidth increases with decreasing transmitter antenna gain, such systems are less sensitive to the random pointing errors. Similarly, a smaller receiver antenna gain implies that the receiver SNR is much less sensitive to the LO tracking error. In fact, an inspection of Fig. 5.6 shows that, when the antenna gains are chosen to be smaller than their optimal value, a smaller power penalty due to spatial tracking error is expected. Figures 5.13 and 5.14 are plots of power penalty versus antenna gain for a 4-ary PPM system and a binary NCFSK system, respectively. Note that the power penalty increases slowly until the optimal antenna gain is reached, and then increases rapidly for $G_T > G_{T_{opt}}$.

Table 5.3 summarizes the link budgets for a direct detection QPPM and a heterodyne QFSK systems with RMS pointing and tracking errors of 0.2 μrad and transmitter and receiver antenna diameters of 20 cm. Note that since the antenna gains are much smaller than the optimal value ($G_T < 0.1 G_{T_{opt}}$), only a small power penalty needs to be allocated for pointing and tracking errors (0.2 dB for direct detection PPM systems, and 0.4 dB for heterodyne FSK systems). Consequently, in the limit of small pointing and tracking jitters, the heterodyne channel is superior to the direct detection PPM channel because of the higher receiver sensitivity.

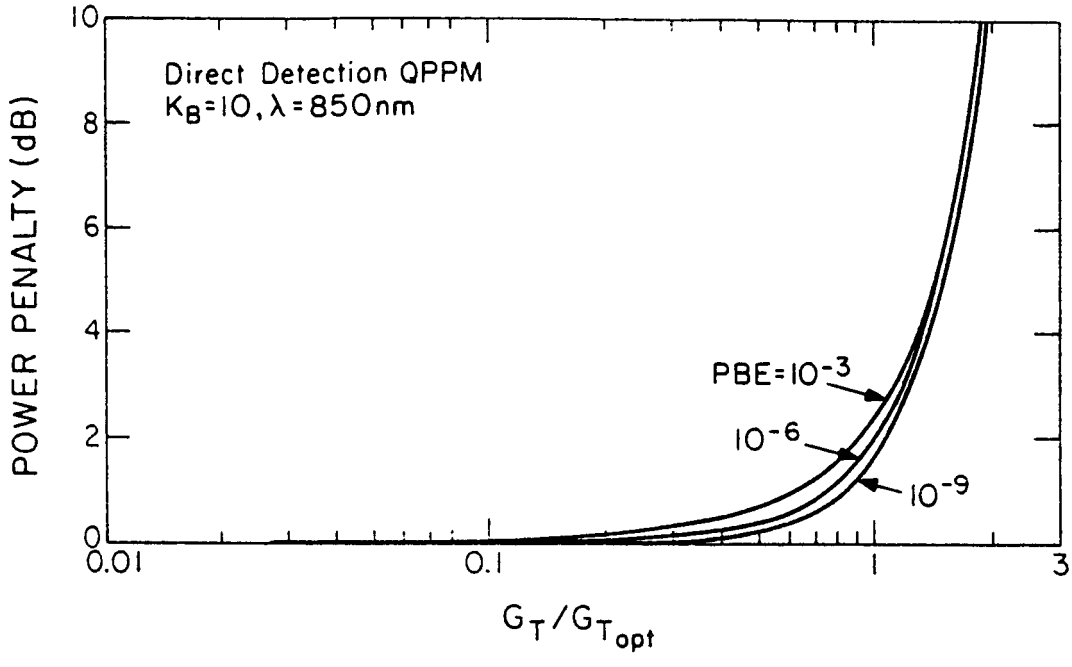


Fig.5.13. Power penalty versus antenna gain for a 4-ary PPM system.

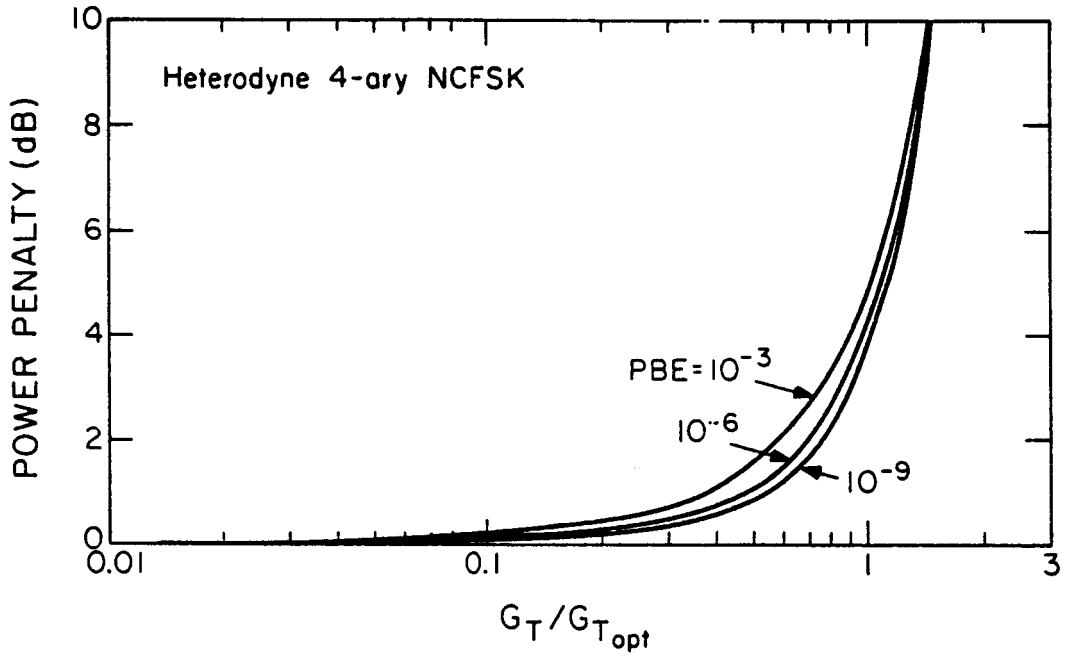


Fig.5.14. Power penalty versus antenna gain for a 4-ary NCFSK system.

Table 5.3 Link budgets for QPPM and QFSK systems with 0.2 μrad RMS tracking jitters and 20 cm diameter transmitter and receiver apertures.

	Direct Detection 4-PPM	Heterodyne 4-FSK
Data Rate	200 Mbps	200 Mbps
PBE	10^{-6}	10^{-6}
Laser Power	20.0 dBm	20.0 dBm
Path Loss	-295.4 dB	-295.4 dB
Antenna Gain		
Transmitter†	117.4 dB	117.4 dB
Receiver†	117.4 dB	117.4 dB
Optics Efficiencies		
Transmitter	-3.0 dB	-3.0 dB
Receiver	-4.6 dB	-4.6 dB
Receiver Sensitivity	-53.1 dBm	-60.4 dBm
Pointing and Tracking Losses	-0.2 dB	-0.4 dB
Temporal Tracking Loss	-0.1 dB	-0.1 dB
Phase Noise‡	-	-0.8 dB
Link Margin	4.6 dB	10.9 dB

† Assumes 20 cm diameter apertures.
‡ Calculated at PBE=10^{-6}, $\nu_d T_s$=1, and $\Delta_f T_s$=0.03.

5.4 Impact of Static Pointing Error

Thus far the analysis on the impact of spatial tracking errors has assumed that tracking errors in each of the two axes can be modeled as zero-mean Gaussian distributed random variables such that the resulting radial tracking error is Rayleigh distributed. The impact of Rayleigh distributed pointing and tracking errors on the design of ISLs is seen by the existence of optimal antenna gains and the additional power penalty. In some systems, however, because of systematic noises, there may be a static bias at the output of the spatial tracking circuit. The presence of this bias results in a static pointing error at the transmitter and receiver in addition to the random, zero-mean Gaussian pointing jitters. When both static and random pointing errors are present at the transmitter, the probability distribution of the instantaneous pointing error θ_T is no longer Rayleigh distributed. Instead, for ISLs with a static pointing error of θ_0, the instantaneous pointing error θ_T is Rician distributed with density function

$$p(\theta_T) = \frac{\theta_T}{\sigma_{\theta_T}^2} e^{-(\theta_T^2+\theta_0^2)/2\sigma_{\theta_T}^2} I_0\left(\frac{\theta_T \theta_0}{\sigma_{\theta_T}^2}\right) \tag{5.27}$$

where σ_{θ_T} is the variance of the random pointing jitter, and $I_0(x)$ is the modified Bessel function of the first kind. In the presence of static and random pointing errors, the PBE can be calculated by substituting Eq. (5.27) for the density of θ_T in Eq. (5.11) and carrying out the integration. Note that by normalizing the random pointing jitter by its RMS value, the PBE can be written as a function of $\theta_0/\sigma_{\theta_T}$, $P_L/\sigma_{\theta_T}^2$ and $G_T\sigma_{\theta_T}^2$. Consequently, in the presence of static pointing errors, both the optimal antenna gain and the minimal signal power are functions of the static pointing error. Both $G_{T_{opt}}$ and $P_{L_{min}}$ can be calculated numerically. Figure 5.15 is a plot of the optimal transmitter antenna gain versus the RMS pointing jitter for a 4-ary PPM system operating at PBE=10^{-6}. Note that in the presence of a static pointing error, the optimal transmitter antenna gain is smaller than that when θ_0=0. At small RMS pointing errors, the optimal antenna gain is much smaller for systems with large static pointing errors. However, at large RMS pointing jitters, the effect of static pointing error is practically negligible. Similarly, given θ_0, and σ_{θ_T}, the required transmitter power, $P_{L_{min}}$, can be calculated. Figure 5.16 is a plot of $P_{L_{min}}$ versus the RMS pointing jitter for a 4-ary PPM system. Note that at small RMS pointing jitters, static pointing jitter can result in a large increase in the required signal power. As the RMS pointing jitter increases, however, the impact of static pointing error becomes less significant and the required transmitter power converges to that of a system with θ_0=0.

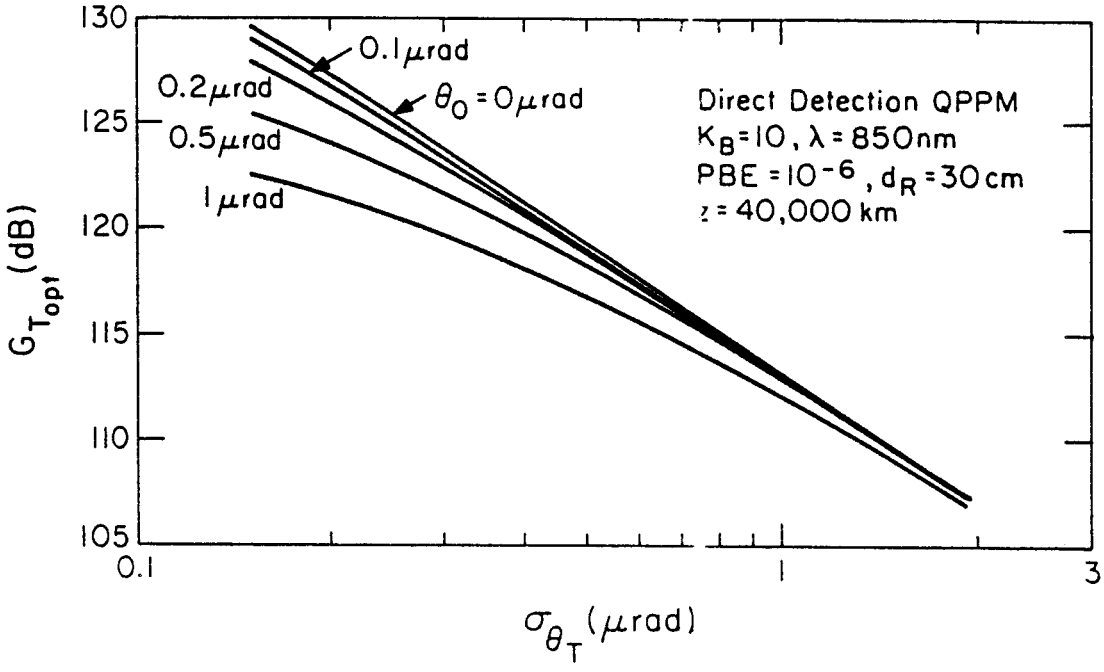


Fig.5.15. Optimal transmitter antenna gain versus the RMS pointing jitter for a 4-ary PPM system with both static and random pointing errors.

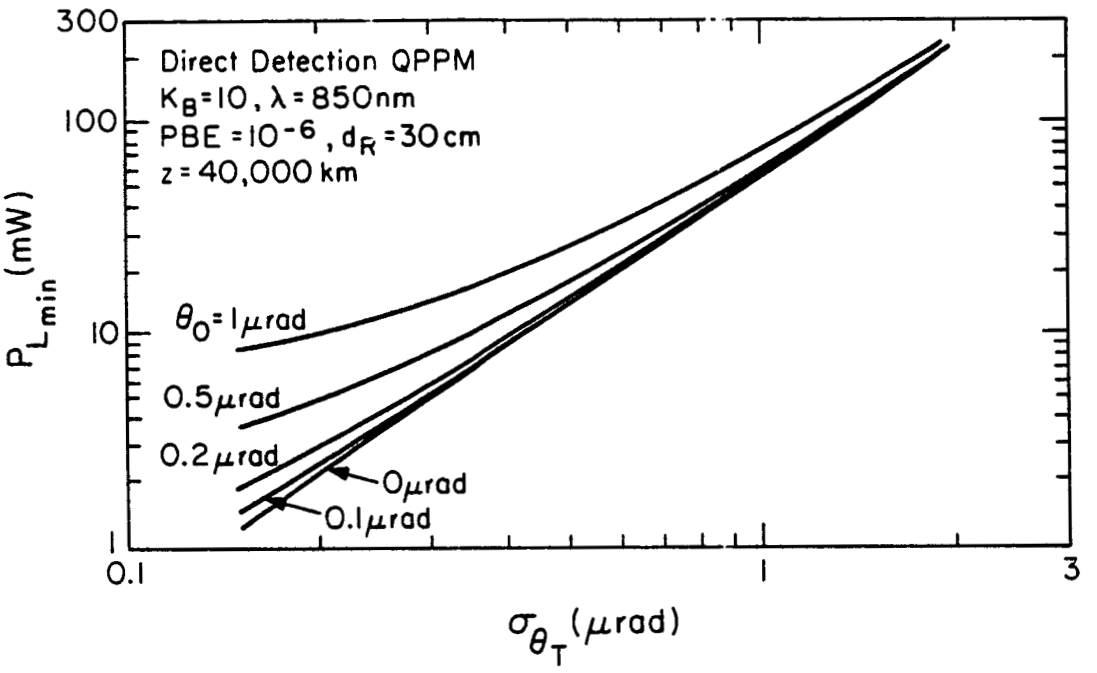


Fig.5.16. Required transmitter power versus the RMS pointing jitter for a 4-ary PPM system with both static and random pointing errors.

Detailed analysis of the system performance in the presence of static and random pointing errors is in general very difficult. Under two special conditions, however, the performance analysis can be simplified considerably. The first case occurs when the RMS pointing error is much greater than the static pointing error, $\sigma_{\theta_T} \gg \theta_0$. In this case the probability distribution of θ_T given in Eq. (5.27) can be approximated by the Rayleigh density shown in Eq. (5.4), and the analysis for the spatial tracking error can be carried out by considering only the random pointing error. The second case occurs when the RMS pointing error σ_{θ_T} is much smaller than the static pointing error θ_0. In this case the probability distribution of θ_T is concentrated around $\theta_T=\theta_0$, and the analysis can be carried out assuming that only the static pointing error is present.

In the presence of static pointing error, the transmitter power is related to the received photocount and the transmitter parameters by

$$P_L=K_S/\eta_T\eta_T G_T G_R(\frac{\lambda}{4\pi z})^2 M(\frac{\eta}{h\nu})L_p(G_T,\theta_0) \tag{5.28}$$

where $L_p(G_T,\theta_0)$ is the pointing loss factor given by Eq. (5.11). For given K_S and θ_0, the required transmitter power can be minimized by maximizing the product

$$G_T L_p(G_T,\theta_0)=\left[\frac{2J_1(\sqrt{G_T}\theta_0)}{\theta_0}\right]^2 . \tag{5.29}$$

Since the function $J_1(x)$ is maximized at $x\approx 1.84$, the required transmitter power can be minimized if the transmitter antenna gain is chosen such that

$$G_{T_{opt}}\theta_0^2 = (1.84)^2 = 3.38 . \tag{5.30}$$

Notice that, in the presence of a static pointing error, the optimal antenna gain depends only on the static pointing error. The pointing loss factor $L_p(G_T,\theta_0)$ at the optimal antenna gain is given by

$$L_p(G_{T_{opt}},\theta_0) \approx 0.4 \equiv -4\text{dB} , \; G_{T_{opt}} = 3.38/\theta_0^2 . \tag{5.31}$$

In other words, in the presence of a static pointing error, a design which optimizes the transmitter antenna gain will require 4 dB higher transmitter power to maintain the system performance. Compared to the value of $G_{T_{opt}}$ calculated from Fig. 5.7, it is seen that the optimal antenna gain in the presence of a fixed pointing error is much higher than that when the pointing error is randomly distributed.

Similarly, for heterodyne systems, the presence of static pointing and tracking errors implies that optimal transmitter and receiver antenna gains can be found to optimize the system performance. These optimal antenna gains can be found by maximizing the product of the tracking loss factor and the antenna gain. For systems where the pointing and tracking loss factors can be approximated by the Airy patterns given in Eqs. (5.8) and (5.19), the signal power can be minimized by choosing the antenna gains to satisfy

$$G_{T_{opt}} = 3.38/\theta_{0_T}^2 \ , G_{R_{opt}} = 3.38/\theta_{0_R}^2 \ , \tag{5.32}$$

where θ_{0_T} and θ_{0_R} are static pointing and tracking errors at the transmitter and receiver, respectively. When both transmitter and receiver antenna gains are optimized for static pointing errors, the power penalty, which is the product of pointing and tracking loss factors at optimal antenna gains, is approximately 0.16 or, equivalently, -8 dB.

5.5 Summary and Discussion

The performance of the spatial tracking system is dominated by the residual mechanical noise which cannot be reduced by increasing the SNR. The presence of these residual spatial pointing and tracking errors imposes additional constraints in the link design. In contrast to the perfectly tracked system where the link performance improves with increasing transmitter and receiver aperture diameters, the presence of random pointing and tracking errors affects the ISL design in two respects. First, in the presence of random pointing and tracking errors, system performance is optimized at particular values of the transmitter and receiver antenna gains. When tracking errors can be modeled as Rayleigh distributed random variables, these optimal antenna gains are inversely proportional to the mean square pointing and tracking errors. Second, the presence of pointing and tracking errors imposes an additional power penalty in the ISL link budget. The power penalty was shown to be approximately 1.5 to 3 dB for a direct detection PPM system, and between 3 and 5 dB for a heterodyne NCFSK system.

The impact of random pointing and tracking errors on the design of ISLs can be summarized by the ISL link budgets shown in Tables 5.1-5.3. In the absence of spatial tracking error, the heterodyne channel offers a 5-10 dB advantage in detection sensitivity over the comparable direct detection PPM system. However, the heterodyne system is much more sensitive to spatial tracking errors at the transmitter and receiver and the advantage gained in detection sensitivity can be quickly offset by the smaller receiver antenna gain (≈6 dB at

σ_{θ_R}=1μrad) and a higher power penalty (≈2 dB) due to randomly distributed tracking and pointing errors. As a result, for systems with large pointing and tracking errors, the direct detection channel is preferred over the heterodyne channel despite the smaller receiver sensitivity. For systems with small pointing and tracking jitters, on the other hand, the antenna gains are limited by the size and weight of the optical system rather than performance considerations. Because the antenna gains are smaller than the optimal values, these systems are less sensitive to the pointing and tracking jitters. Consequently, in the limit of small pointing and tracking jitters, the heterodyne channel demonstrates a superior performance because of the higher receiver sensitivity.

Thus far the analysis of the system performance was carried out for an un-encoded system. In practice, the performance of the link can be improved considerably by using error control techniques such as source encoding and forced retransmission. Because the tracking and pointing errors are due primarily to the residual mechanical noise that has a typical frequency of several kilohertz, it is evident that errors introduced by the tracking errors generally occur in bursts with a typical burst period of several hundred microseconds. The error control techniques, therefore, must effectively reduce this burst error probability [60]-[63]. Several methods are effective for controlling the burst error. For systems demonstrating short error bursts, source encoding can be used to facilitate error detection and error recovery [163], [164]. However, because of the high data rate of the channel, error bursts due to spatial tracking errors typically span a period of several hundreds or even thousands of bits. Burst error correction for such a long bit stream is impractical because of the complexity of the decoder. Alternatively, burst errors can be effectively controlled by requiring the receiver to acknowledge (ACK) the transmitter upon receiving an error free transmission. In the case where errors occur during the transmission, a no-acknowledgement (NAK) is sent and the transmitter will then re-transmit the previous data packet. When coupled with an effective error detection code, this acknowledge-no acknowledge (ACK-NAK) scheme can be used to ensured that only valid data are received at the receiver. However, because of the long round trip delay for an ISL, a large buffer memory is needed to store the data before an acknowledgement is received. Consequently, for channels demonstrating large burst error probabilities, such a scheme can severely reduce the effective rate of transmission through the channel.

6. TIMING SYNCHRONIZATION

In order to implement an efficient decoding scheme, the receiver requires the presence of a clock signal that is synchronized to the transmitted data stream. Receiver synchronization is also required for spaceborne systems where the actual data rate is affected by the Doppler shift due to the relative motion. Furthermore, receiver synchronization is needed to control the timing jitter and frequency drift corresponding to the oscillator noise.

In some systems, the receiver synchronization is maintained by transmitting a separate timing signal on top of the message waveform. The receiver then extracts the necessary sync information to perform the decoding. For such systems, the receiver synchronization is relatively easy to implement. However, additional signal energy and bandwidth are required by the sync channel. Alternatively, the receiver timing can be recovered directly from the received data bearing signal such that no additional timing signal is necessary. This method has the advantage of power and bandwidth conservation and, consequently, is preferred for applications where the signal power and available bandwidth are limited.

6.1 Synchronization of Direct Detection PPM Systems

The problem of symbol synchronization from a data bearing signal deals with the estimation of the time at which the modulation changes its state. In the case of a direct detection PPM system, this implies estimating and tracking the slot timing from the output of the photodetector $r(t)$, where [3], [74]

$$r(t) = \sum_{\{\tau_j\}} eG_j h(t-\tau_j) \ . \tag{6.1}$$

In Eq. (6.1), τ_j and G_j are the arrival time and gain associated with the j^{th} photon event, $h(t)$ is the impulse response of the detector, and the thermal noise has been assumed to be negligible compared to the signal shot noise. The photocount rate of the detector is proportional to the total power received at the photodetector and, for an M-ary PPM system, can be modeled as

$$\lambda(t) = \lambda_B + \lambda_S \sum_{k=0}^{N_T} p(t-kT_w-d_kT_s) \tag{6.2}$$

where λ_B, λ_S are the photocount rates due to the background and signal, $T_w=MT_s$ is the word period, $N_T=T/T_w$ is the number of codewords in the observation period, d_k is the k^{th} data word, and $p(t)$ is the

transmitted pulse shape. The purpose of the synchronization system, therefore, is to recover the slot timing given the photodetector output in Eq. (6.1). After the slot timing has been recovered, frame and word synchronizations can then be achieved using a coding technique [165], [166].

6.1.1 The MAP Estimator

Given the received signal $r(t)$, the optimal symbol timing estimator is the one which maximizes the *a posteriori* probability $p(\tau | r(t))$, where τ is the timing offset between the receiver and the transmitter. Mathematically, the output of the optimal estimator can be written as [3], [74]

$$\hat{\tau}_{MAP} = arg \max_{\tau} p(\tau | r(t), 0 \le t \le T) = arg \max_{\tau} [p(r(t), 0 \le t \le T | \tau) p(\tau)] \ , \ 0 \le \tau \le T \ . \tag{6.3}$$

The estimator which implements Eq. (6.3) is known as the MAP estimator. When the *a priori* probability distribution of the pulse delay is unknown, the probability can be assumed to be equally distributed over the observation interval. In this case the MAP decision reduces to an ML estimator which maximizes $p(r(t), 0 \le t \le T | \tau)$.

For a general receiver output given in Eq. (6.1), the probability $p(r(t), 0 \le t \le T | \tau)$ is very difficult to calculate. The calculation can be simplified considerably, however, if it is assumed that the detector impulse response can be modeled as a delta function, and that the detector gain G_j is constant. Under these assumptions, the ML estimator is given by [3]

$$\hat{\tau}_{ML} = arg \max_{\tau} E_{\lambda} \left\{ \exp\left[\int_0^T r(t) \log \lambda(t+\tau) dt - \int_0^T \lambda(t+\tau) dt \right] \right\} \ , 0 \le \tau \le T \ , \tag{6.4}$$

where the expectation is taken with respect to the random code sequence $\lambda(t)$. The term involving the integral of $\lambda(t+\tau)$ is the total expected photocount during the period $(0, T)$. For a long observation interval, the fluctuation in this term will be small compared to the total received photocount. Therefore, the ML timing estimator is effectively

$$\hat{\tau}_{ML} = arg \max_{\tau} E_{\lambda} \left\{ \exp\left[\int_0^T r(t) \log \lambda(t+\tau) dt \right] \right\} \ . \tag{6.5}$$

By substituting the expression of the PPM photocount rate from Eq. (6.2) into (6.5) and taking the expectation with respect to d_k, the ML pulse delay estimator can be reduced to the one which maximizes

$$f(r(t),\tau)=\sum_{k=0}^{N_r}\log\left\{\sum_{j=0}^{M-1}\exp\left[\int_{kT_w+jT_s+\tau}^{kT_w+(j+1)T_s+\tau} r(t)\log[\lambda_B+\lambda_S p(t-kT_w-jT_s+\epsilon)]dt\right]\right\} . \tag{6.6}$$

The ML timing estimator that implements the decision rule is shown in Figure 6.1. Judging from the complexity of the estimator, it is apparent that the ML timing estimator, although it maximizes the probability of choosing the correct estimate, is far from practical. A large number of correlators are needed to provide an acceptable timing estimate. Consequently, in order to implement pulse tracking, some simplification to the ML decision rule must be devised.

6.1.2 Early-Late Gate Tracking Loop

A direct and straightforward simplification of the MAP estimator is to reduce the number of correlators required by using a tracking loop. Recall that the ML estimator maximizes the likelihood function $f(r(t),\tau)$ given by (6.6). If it is assumed that $f(r(t),\tau)$ is differentiable with respect to the delay time τ, then the derivative of $f(r(t),\tau)$ can be used as an indication for the timing offset. Note that $\partial f(r(t),\tau)/\partial\tau$ is given by

$$\begin{aligned}\frac{\partial f(r(t),\tau)}{\partial\tau}=\sum_{k=1}^{N_r}\Bigg\{&\sum_{j=0}^{M-1}\exp\left[\int_{kT_w+jT_s+\tau}^{kT_w+(j+1)T_s+\tau} r(t)\log[\lambda_B+\lambda_S p(t-kT_w-jT_s+\tau)]dt\right]\\ &\times\int_{kT_w+jT_s+\tau}^{kT_w+(j+1)T_s+\tau} r(t)\frac{\partial}{\partial\tau}\log[\lambda_B+\lambda_S p(t-kT_w-jT_s+\tau)]dt\\ &\times\left[\sum_{j=0}^{M-1}\exp\left[\int_{kT_w+jT_s+\tau}^{kT_w+(j+1)T_s+\tau} r(t)\log[\lambda_B+\lambda_S p(t-kT_w-jT_s+\tau)]dt\right]^{-1}\right\} .\end{aligned} \tag{6.7}$$

For a given initial estimate of the timing delay $\hat{\tau}$, Eq. (6.7) can be used to provide an indication of the search direction. At $\hat{\tau}>\tau$, $\partial f(r(t),\tau)/\partial\tau|_{\hat{\tau}}$ is negative, while at $\hat{\tau}<\tau$, the value of Eq. (6.7) is positive. Consequently, it is possible to design a closed-loop tracking circuit which makes use of this property. Figure 6.2 shows an example of such an implementation. The phase of the timing estimator is updated every word period by the amount that is proportional to the derivative of the likelihood function.

The timing loop implemented using Eq. (6.7), however, is still very complicated and further simplifications are needed to simplify the design. One such simplification is to approximate the second term

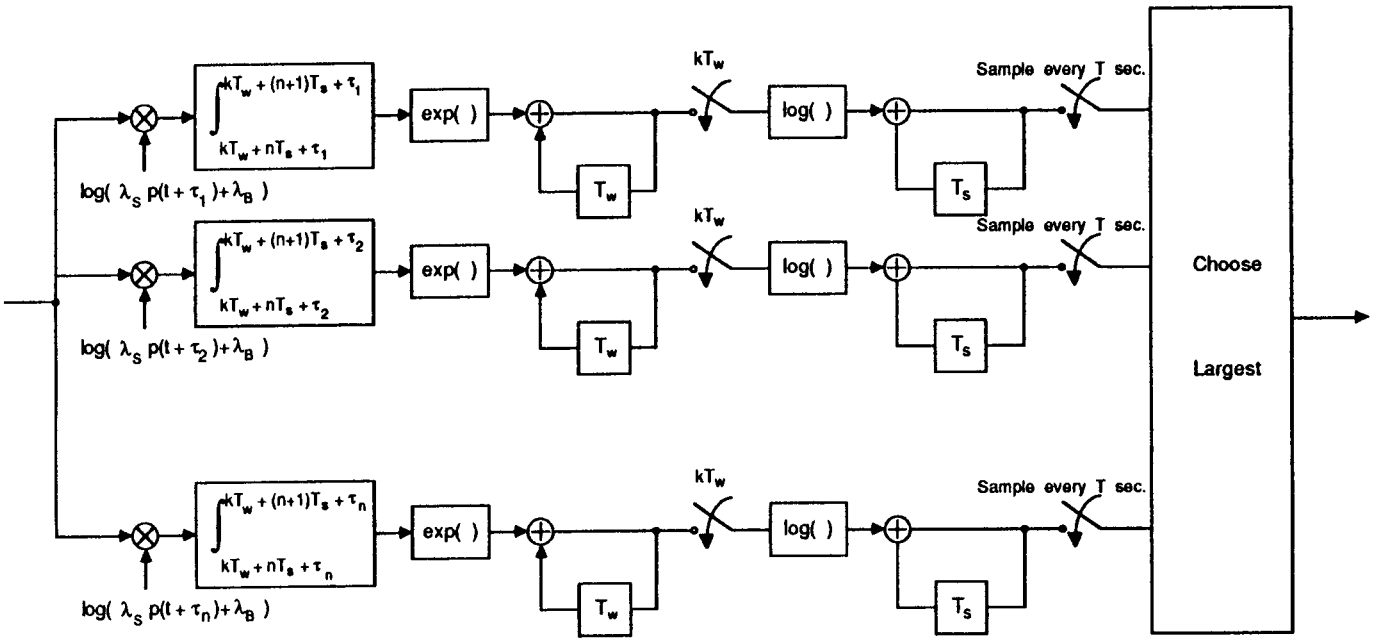


Fig.6.1. Maximum likelihood (ML) timing error estimator for the M-ary PPM channel.

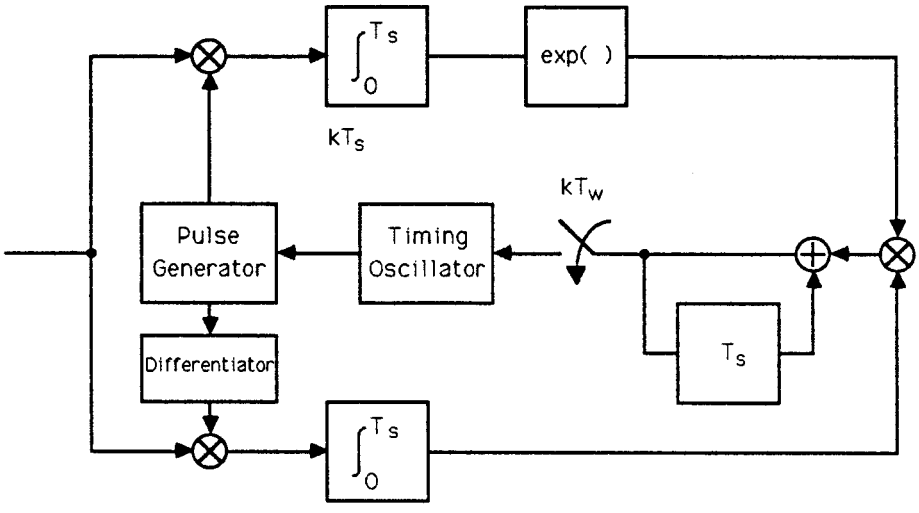


Fig.6.2. Tracking loop which approximates the ML synchronizer.

on the right-handed side of Eq. (6.7) by a convolution of the received signal pulse $r(t)$ with an early-late gate pulse $g(t)$ where

$$g(t) = \begin{cases} +1 & , 0 \le t \le T_s/2 \\ -1 & , T_s/2 \le t \le T_s \end{cases} . \tag{6.8}$$

Another simplification can be made by noting the value computed by the upper branch in Figure 6.2 strongly emphasizes the early-late gate output during the slot which contains the actual signal pulse. Consequently, it is possible to replace the upper branch in Figure 6.2 by the decision value of the decoder. That is, only the early-late gate output during the signal slot will be used to update the timing oscillator. Figure 6.3 shows a block diagram of this simplified decision-driven early-late gate tracking loop.

The application of the decision-driven early-late-gate loop in symbol synchronization has been studied previously [3]. It has been shown that, at high SNR, the variance of the timing error for a binary PPM system can be given by

$$\sigma_\varepsilon^2 = \frac{1}{(2\pi)^2} \frac{B_L T_s}{K_S} , \tag{6.9}$$

where $\varepsilon = \tau/T_s$ is the normalized timing error, B_L is the bandwidth of the tracking loop, T_s is the slot width, and K_S is the signal photocount received over the slot period.

6.1.3 Phase-locked loops

The early-late gate described in the previous section is an approximation to the optimal MAP timing estimator. Consequently, it represents a lower bound on the performance of a synchronization system. However, when used in a direct detection PPM system, the early-late gate synchronizer requires a complicated decision feedback structure which may be difficult to implement. Alternatively, when the input signal to the synchronization subsystem contains a spectral component at the desired lock-in frequency, a simple phase-locked loop (PLL) shown in Figure 4.10 can be used to recover the pulse timing.

The applications of PLLs in radio and optical communication systems have been studied extensively [28]-[31], [63]-[68], [167]-[169]. Gagliardi and Haney [167] and Snyder and Forrester [168] analyzed the probability density of tracking error for a PLL under shot noise input. Mengali and Pezzani [67] studied the phase error variance of a PLL driven by photodetector current in an optical pulse amplitude modulation

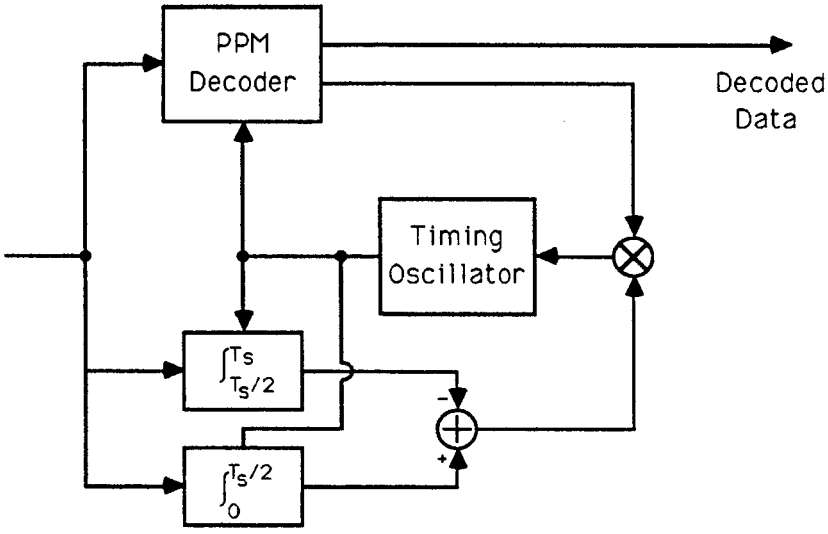


Fig. 6.3 Decision driven early-late gate tracking loop for the PPM system

system. Marshall [65] studied the performance of PLL tracked PPM systems with pulses occupying a fraction of the time slot. These studies have shown that PLLs can be used to track the transmitter timing provided that the transmitted signal contains a frequency component at the desired lock-in frequency.

Given the photodetector output $r(t)$, the time-averaged power spectrum of $r(t)$ can be defined as [3]

$$\begin{aligned} S_r(\omega) &= \lim_{T\to\infty}\frac{1}{2T}E\left[\,\left|\int_T^T r(t)e^{-i\omega t}dt\,\right|^2\right] \\ &= \lim_{T\to\infty}\frac{1}{2T}E_\lambda\left[\int_T^T\int_T^T E\left[r(t)r(\tau)\,|\,\lambda(t)\right]e^{-i\omega(t-\tau)}dtd\tau\right] \quad . \end{aligned} \tag{6.10}$$

The expectation in Eq. (6.10) can be evaluated by first taking the expectation with respect to the shot noise process, conditioned on $\lambda(t)$, then taking the expectation of the resulting expression with respect to $\lambda(t)$. The first expectation can be evaluated by differentiating the joint characteristics function of the filtered Poisson process, given by (Appendix B)

$$\Phi(\omega_1,\omega_2) = E\left[e^{-i\left[\omega_1 r(t)+\omega_2 r(\tau)\right]}\right] = \exp\left\{\int_0^T\lambda(\xi)\left[\Phi_G(\omega_1 h(t-\xi)+\omega_2 h(\tau-\xi))-1\right]d\xi\right\} \quad , \tag{6.11}$$

where $\Phi_G(\omega)=E\left[e^{i\omega G}\right]$ is the characteristics function of the detector gain G. The resulting joint expectation of the photodetector output can be written as

$$E\left[r(t)r(\tau)\,|\,\lambda(t)\right] = \langle G^2\rangle\int\lambda(\xi)h(t-\xi)h(\tau-\xi)d\xi + \langle G\rangle^2\left[\int\lambda(\xi)h(t-\xi)d\xi\right]\left[\int\lambda(\eta)h(\tau-\eta)d\eta\right] \quad , \tag{6.12}$$

where the brackets $\langle G^n\rangle$ denote the ensemble average of G^n. By substituting (6.12) into (6.10) and taking the expectation with respect to the codewords $\{d_k\}$, assuming that $\{d_k\}$ are independent and uniformly distributed over $(0,1,\ldots,M-1)$, the resulting power spectrum of the photodetector output can be written as (Appendix C)

$$\begin{aligned} S_r(\omega) = G^2|H(\omega)|^2\Bigg[& F\left(\lambda_B+\frac{\lambda_S}{M}\right)+2\pi\delta(\omega)\left(\lambda_B^2+\frac{2\lambda_B\lambda_S}{M}\right) \\ & +\frac{\lambda_S^2}{T_w}|P(\omega)|^2(1-|R(\omega)|^2) \\ & +\frac{2\pi}{T_s}\lambda_S^2|P(\omega)|^2\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{2\pi k}{T_s}\right)\Bigg] \quad , \end{aligned} \tag{6.13}$$

where F denotes the excess noise factor of the detector,

$$R(\omega)=E[\exp(-jC_k\omega T_s)]=\frac{1}{M}\sum_{k=0}^{M-1}e^{-j\omega kT_s} \quad , \tag{6.14}$$

and $P(\omega)$, $H(\omega)$ are the Fourier transforms of the pulse shape and the impulse response of the detector-filter, respectively. If $p(t)$ is a rectangular pulse with width T_s, its Fourier transform $P(\omega)$ will be zero at all integer multiples of the slot frequency, $2\pi/T_s$. In this case, the detector output spectrum in Eq. (6.13) contains no discrete frequency components at the slot frequency or its harmonics. Consequently, the PLL cannot track the output of the photodetector directly. Preprocessing of the detector output is necessary to generate a frequency component at the slot frequency.

Only nonlinear processing needs to be considered, because any linear filtering will simply amount to multiplying the power spectrum in Eq. (6.13) by the magnitude square of the filter transfer function, and will not affect the absence of the spectral component at the slot frequency. One approach is to filter the signal and then square the filter output, as depicted in Figure 6.4. The output of this preprocessing circuit can be written as

$$i_2(t)=r^2(t)=\left[\sum_{\tau_j}G_j h(t-\tau_j)\right]^2 \quad , \tag{6.15}$$

where $h(t)$ now denotes the combined impulse response of the detector and the preprocessing filter. The expectation value of the preprocessor output can be evaluated by first differentiating its conditional characteristics function in Eq. (6.11), then taking the expectation with respect to the codewords $\{d_k\}$. The resulting expression for the expected preprocessor output is given by

$$\begin{aligned} E[i^2(t)] &= G^2F\left[\lambda_B+\frac{\lambda_S}{M}\sum_{k=-\infty}^{\infty}p(t-kT_s)\right]*h^2(t) \\ &+G^2\left\{\left[\lambda_B+\frac{\lambda_S}{M}\sum_{k=-\infty}^{\infty}p(t-kT_s)\right]*h(t)\right\}^2 \\ &-\frac{G^2\lambda_S^2}{M^2}\sum_{k=-\infty}^{\infty}\sum_{m,n=0}^{M-1}[h(t)*p(t-kT_w-nT_s)]\cdot[h(t)*p(t-kT_w-mT_s)] \\ &+\frac{G^2}{M}\lambda_S^2\sum_{k=-\infty}^{\infty}[h(t)*p(t-kT_s)]^2 \quad , \end{aligned} \tag{6.16}$$

where we have used the fact that

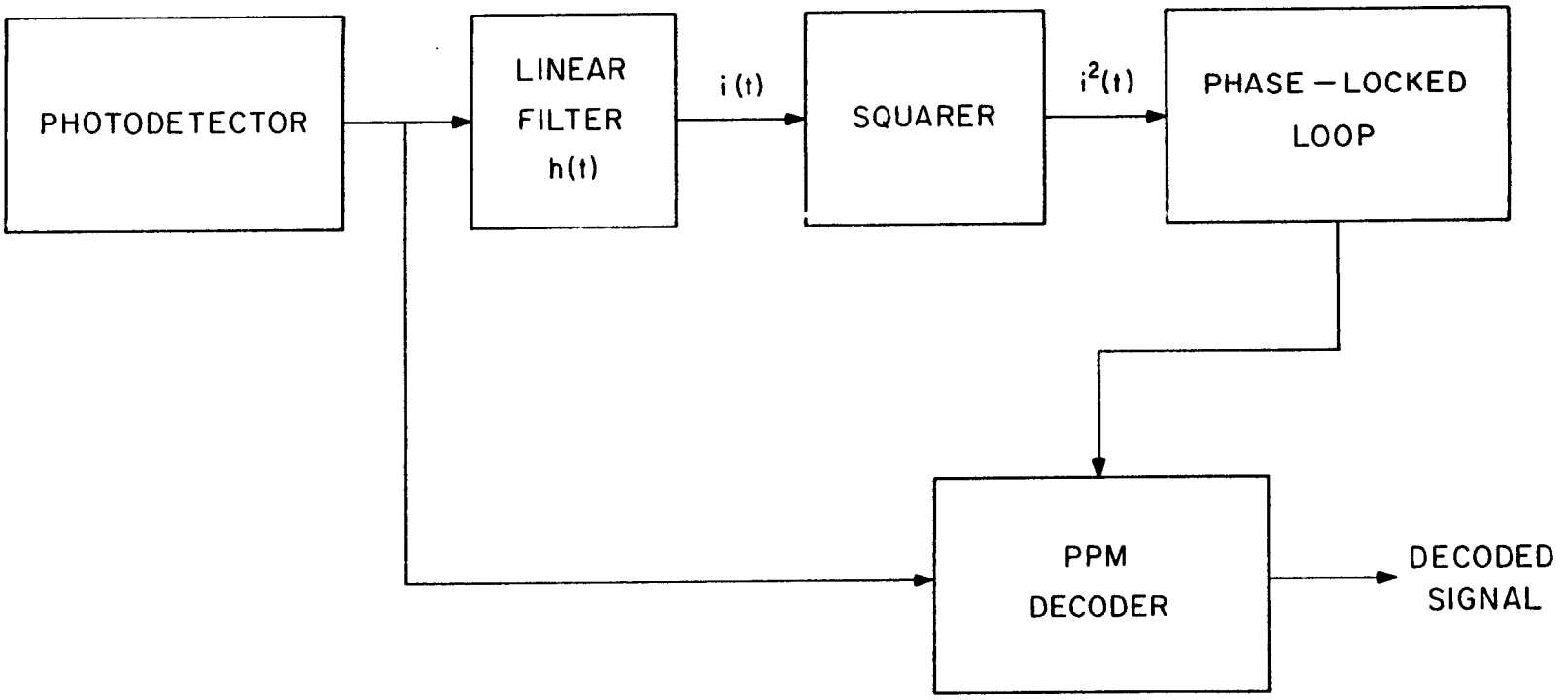


Fig.6.4. The non-linear PLL pulse synchronizer.

$$\sum_{k=-\infty}^{\infty}\sum_{j=0}^{M-1} p(T-kT_w-jT_s) = \sum_{k=-\infty}^{\infty} p(t-kT_s) \ . \tag{6.17}$$

By examining Eq. (6.16), it is seen that the last term at the right-handed side is periodic with period T_s. Therefore, in contrast to the photodetector output, the squared current does have a strong spectral component at the slot frequency. The output of the preprocessing circuit can therefore be regarded as the sum of a periodic signal $s(t)$ and an additive noise term $n(t)$

$$i_2(t) = s(t)+n(t) \ , \tag{6.18}$$

where from (6.16)

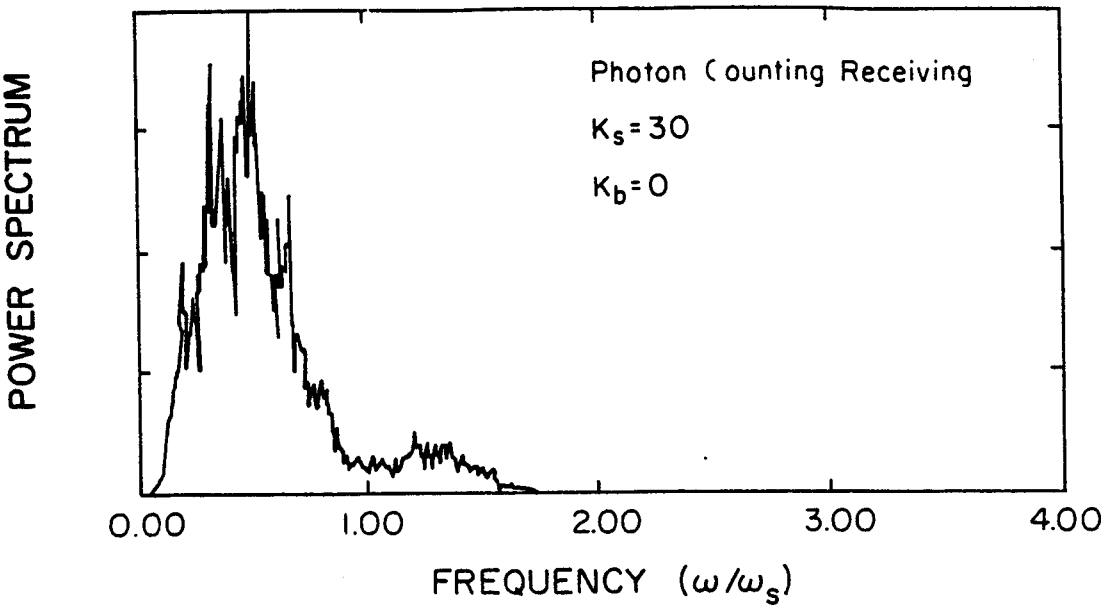
$$s(t) = G^2\frac{\lambda_s^2}{M}\sum_{k=-\infty}^{\infty}[p(t)*h(t-kT_s)]^2 \ , \tag{6.19}$$

and the noise is simply the part of the preprocessor output that is not periodic,

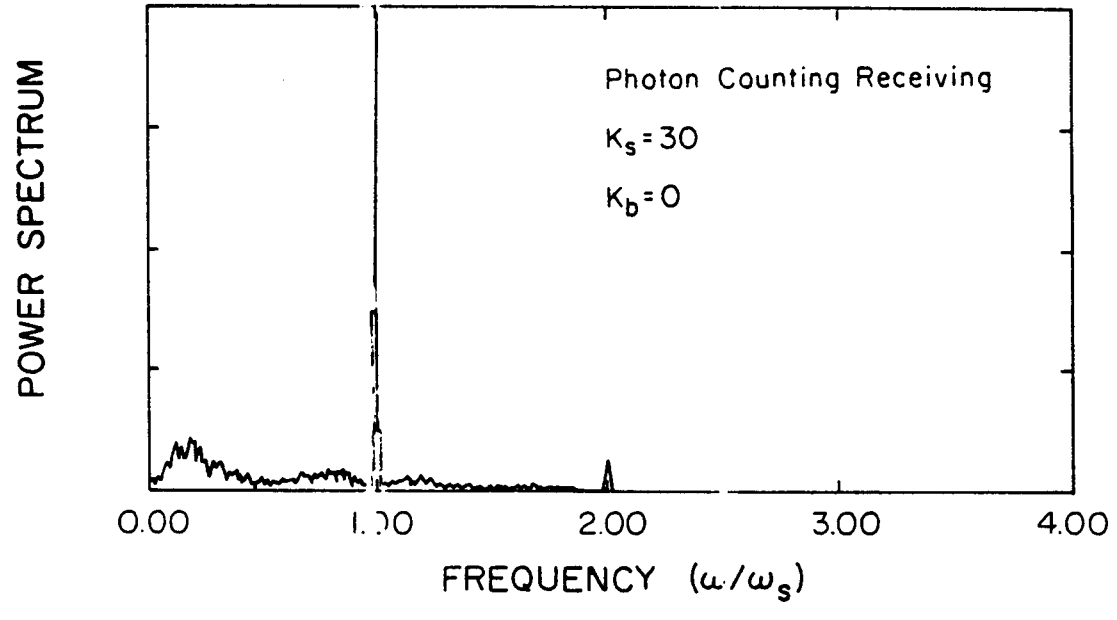
$$n(t) = i_2(t)-s(t) \ . \tag{6.20}$$

Figure 6.5 shows the power spectra of a computer simulated photodetector output before and after preprocessing. The simulation was carried out assuming that a photon counting detector ($F=1$) was used to detect the optical radiation. Note that the frequency component at slot frequency is strongly enhanced by the preprocessor. The analytic expression for the power spectrum of the preprocessed signal is quite complicated. However, the expression can be simplified considerably if it is assumed that $H(\omega)$ blocks the dc component of the signal, and that both $P(\omega)$ and $H(\omega)$, the Fourier transforms of $p(t)$ and $h(t)$, are slowly varying functions of frequency compared to $R(\omega)$, defined in Eq. (6.14). The first assumption is made because the dc component of the detector output contains no timing information and will only contribute noise to the squared signal. The second assumptions holds for higher-order PPMs (i.e., for M large). Because, in general, $h(t)$ and $p(t)$ are pulses of width comparable to T_s, their Fourier transforms will have supports on the order of $\omega_s=2\pi/T_s$. The support for $R(\omega)$, on the other hand, is on the order of ω_s/M, where M is the PPM order. Consequently, for higher-order PPMs, the assumption that $P(\omega)$ and $H(\omega)$ are slowly varying compared to $R(\omega)$ usually holds. For high signal counts, the power spectrum of the preprocessed signal $i_2(t)$ near the slot frequency can be approximated under these assumptions by (Appendix B) [169]

$$S_2(\omega) = S_s(\omega)+S_n(\omega) \ , \tag{6.21}$$



(a) Before



(b) After

Fig.6.5. Power spectra of the simulated photodetector output for a 4-ary PPM (a) before and (b) after nonlinear preprocessing. The preprocessing filter impulse response $h(t)$ used in this simulation was

$$h(t) = \begin{cases} \sin(\omega_s t) & 0 \le t \le T_s \\ 0 & \text{elsewhere} \end{cases}.$$

where

$$S_s(\omega) = \frac{\lambda_S^4 G^4}{T_w^2} \frac{1}{2\pi} |P(\omega)H(\omega) * P(\omega)H(\omega)|^2 \delta(\omega - \omega_s) \tag{6.22}$$

is the spectral component corresponding to the first harmonic of $s(t)$, and

$$S_n(\omega) = 4\frac{\lambda_S^3 G^4 F}{T_w} \frac{1}{(2\pi)^2} \int P(\omega')P^*(\omega'')P^*(\omega'-\omega'')H(\omega)H(\omega-\omega')H^*(\omega'')H^*(\omega-\omega'')d\omega' d\omega'' \tag{6.23}$$

is the power spectrum of the random, non periodic component, $n(t)$. The presence of a spectral peak at $\omega=\omega_s$ implies that the output of the preprocessor contains a sinusoidal component with frequency $\omega=\omega_s$. The amplitude of this sinusoidal component can be given by

$$A = \frac{2\lambda_S^2 G^2}{T_w} \left| \frac{1}{2\pi} \int P(\omega')H(\omega')P(\omega_s-\omega')H(\omega_s-\omega')d\omega' \right| , \tag{6.24}$$

where $\omega_s=2\pi/T_s$ is the slot frequency.

The variance of the timing error for a PLL driven by a sinusoidal signal embedded in an additive white noise can be approximated by

$$\sigma_\varepsilon^2 = \frac{1}{(2\pi)^2} \frac{1}{\rho} = \frac{2B_L S_n(\omega_s)}{(2\pi)^2 A^2} , \tag{6.25}$$

where $\varepsilon=\tau/T_s$ is the normalized timing error, $S_n(\omega)$ is the power spectral density of the additive noise, A is the amplitude of the input sinusoid, and B_L is the closed-loop bandwidth of the tracking loop. The parameter ρ in Eq. (6.25) can be interpreted as the ratio of signal power and the amount of noise power admitted by the loop with bandwidth B_L. Equation (6.25) was derived for a sinusoidal signal in an AWGN. The shot noise at the output of the squarer, however, is non-stationary. For non-stationary noise, the phase error variance of the PLL is in general a function of time. Nevertheless, it can be shown [170] that for a sufficiently narrow loop bandwidth, phase error is stationary, and the phase error variance for the receiver in Figure 6.4 is indeed given by Eq. (6.25).

By substituting the expression for noise power spectrum from (6.23) and the expression for signal amplitude from (6.24) into (6.25), the following expression for the variance of timing error is obtained:

$$\sigma_\varepsilon^2 = \frac{\gamma F M B_L T_s}{(2\pi)^2 K_S} , \tag{6.26}$$

$$\gamma = \left[\frac{2T_s \int P(\omega')P^*(\omega'')P^*(\omega'-\omega'')H(\omega')H(\omega-\omega')H^*(\omega'')H(\omega-\omega'')d\omega' d\omega''}{|P(\omega)H(\omega)*P(\omega)H(\omega)|^2} \right]_{\omega=\omega_s}, \quad (6.27)$$

where K_s is the signal count per word, M is the order of the PPM, B_L is the equivalent closed-loop bandwidth, and γ is a dimensionless parameter which depends only on the pulse shape and the preprocessing filter transfer function. Note that because $f_s=1/T_s$, the parameter $B_L T_s$ is actually a ratio of the loop bandwidth and the slot frequency. The values of γ for some choices of preprocessing filter transfer functions are listed in Table 6.1. For a given pulse shape, the transfer function of the preprocessing filter can be chosen to minimize the value of γ. One choice is to model $H(\omega)$ as an ideal low pass differentiator with bandwidth B_0. Figure 6.6 is a plot of the value of γ versus the bandwidth of the differentiator for the case where p(t) is a square pulse of width T_s. It is shown that the value of γ is minimized ($\gamma_{min}\approx 3.2$) for $B_0 T_s \approx 1.3$. Also shown in Figure 6.6 is the value of γ evaluated using the preprocessing filter which consists of a low pass Gaussian filter with RMS bandwidth B_0, followed by an ideal differentiator. The minimum values of γ obtained using both preprocessing filters are similar. Equation (6.26) shows that the variance of the phase error is inversely proportional to the signal power (i.e., signal photocount K_S) and is proportional to the number of time slots M and the loop bandwidth. Compared to an optimal, decision-driven, early-late gate timing estimator, the PLL requires γ times more signal power (i.e., 5 dB for $\gamma = 3.2$) to achieve comparable timing performance. However, the performance of the PLL does not depend critically on the assumption of pulse shape or the decision process. Furthermore, the important system performance factor is bit error rate (BER), not the timing variance. The system BER, which is calculated in Section 6.3, shows that the actual power penalty for using a suboptimal PLL synchronizer is small (less than 0.1 dB) for a sufficiently narrow loop bandwidth ($B_L T_s < 10^{-3}$) even when a conservative value of 10 is assumed for γ.

It should be noted that the loop bandwidth B_L actually increases with increasing signal amplitude. In order to accommodate a wide dynamic range of input signals, it is the usual practice to precede the PLL by either an automatic gain control (AGC) circuit or a limiter. The effect of the AGC is to dynamically scale the input signal so that the signal amplitude and, consequently the loop bandwidth, remain essentially constant. By using an AGC or limiter, the performance of the PLL will improve inversely with the signal strength. Equation (6.26) was verified by a computer simulation [68], [169] for direct detection PPM systems using photon counting detectors ($F=1$). The results, which are summarized in Figures 6.7 and 6.8, are in excellent

Table 6.1 Values of γ for various choices of preprocessing filter.

Filter	γ
$h(t) = \begin{cases} \sin(2\pi t/T_s) & 0 \le t \le T_s \\ 0 & \text{elsewhere} \end{cases}$	3.58
$H(\omega) = \begin{matrix} j\omega & \lvert\omega\rvert \le 2\pi/T_s \\ 0 & \text{elsewhere} \end{matrix}$	6.97
$H(\omega) = j\omega e^{-\omega^2/\omega_s^2}$, $\omega_s = 2\pi/T_s$	3.70

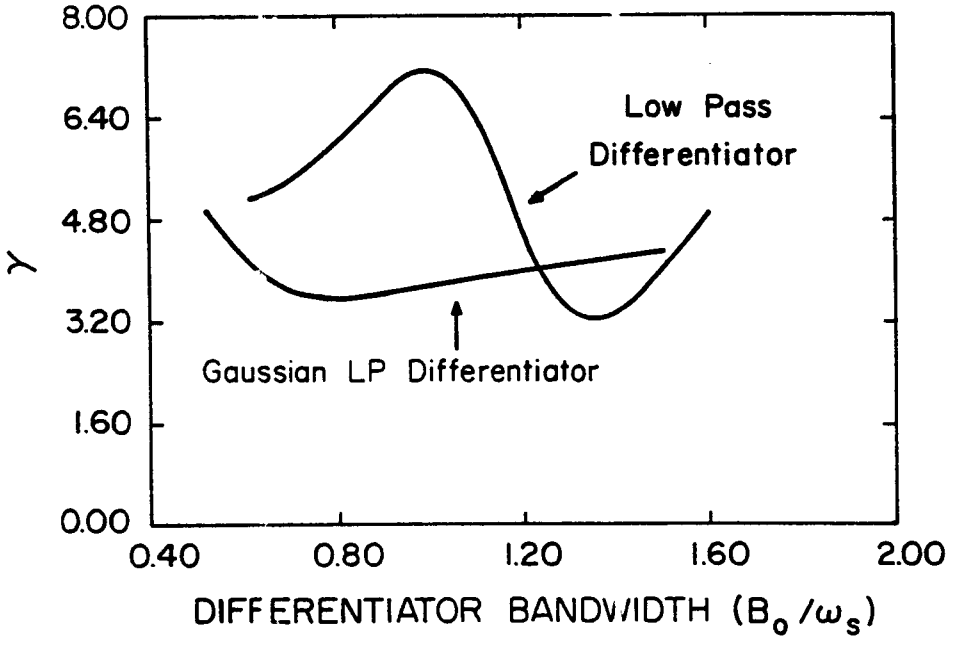


Fig.6.6. The value of γ as a function of the differentiator bandwidth for PPM systems with rectangular pulses of width T_s. $H(\omega)$ is modeled as an ideal low-pass differentiator, where B_0 is the cutoff frequency, and as a Gaussian low-pass filter, where B_0 is the RMS bandwidth, followed by a differentiator.

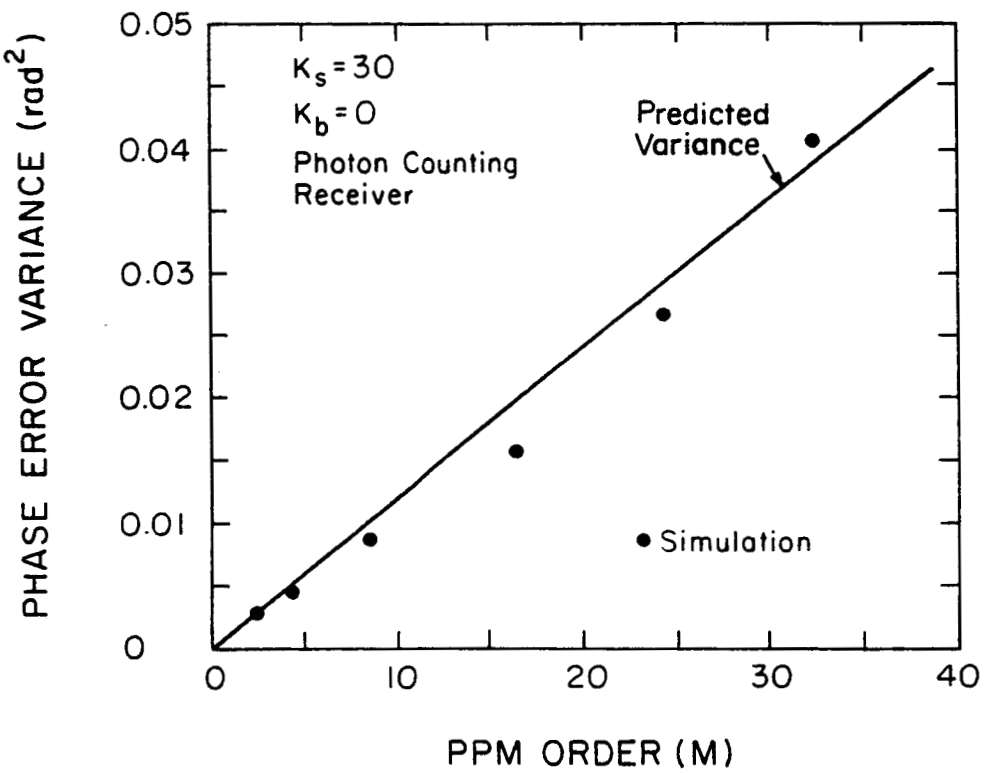


Fig.6.7. Computer simulated phase error variance versus the order of PPM for an M-ary PPM system with $B_L T_s$=0.01 and γ=3.6.

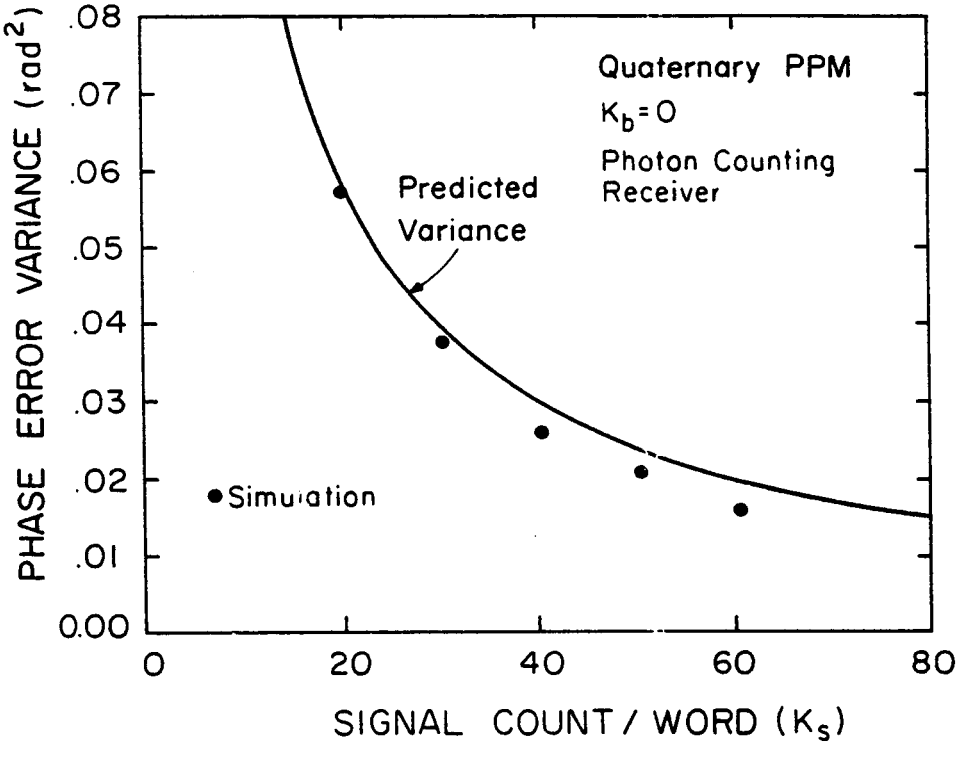


Fig.6.8. Computer simulated phase error variance versus the signal count rate for a 4-ary PPM receiver with $B_L T_s$=1/12 and γ=3.6. The solid curve is the theoretical variance predicted.

agreement with the theory. In particular, the phase error variance for binary PPM systems is found to be in good agreement with the theoretical predictions even though the derivation leading to Eq. (6.26) required M to be large.

6.2 Synchronization of Heterodyne Systems

Unlike the direct detection systems where the shot noise follows the Poisson statistics, the noise in a heterodyne receiver is predominately Gaussian distributed. The problems of timing synchronization and carrier recovery for a heterodyne system are therefore similar to that of an AWGN channel.

6.2.1 Synchronization of Coherent Heterodyne Channels

Synchronization of a coherent heterodyne channel consists of several steps. First, the receiver must establish the reference carrier synchronization by phase-locking onto the incoming carrier modulated signal. The problems of carrier synchronization and its impact on the performance of the heterodyne system have been analyzed in Chapter 4, and will not be repeated here. After a coherent carrier reference is established, symbol timing synchronization can then be achieved by phase-locking onto the demodulated baseband signal.

Several methods can be used to establish the symbol timing synchronization for a coherent heterodyne channel [28]-[31]. Figures 6.9(a)-(d) show the structures of some commonly used symbol synchronizers for the coherent BPSK system. The ML timing estimator shown in Figure6.9(a) offers the optimal timing performance, but is impractical to use because of its complexity [28]. The early-late gate loop is an approximation of the ML tracking loop. The received signal is integrated over the early and late gate period, and nonlinearly (absolute value) conditioned to make the integrator outputs bit independent. The channels are then subtracted and the difference is fed into a timing oscillator. The structure of an early-late gate tracking loop is shown in Figure6.9(b). Symbol timing synchronization can also be achieved using the data transition tracking loop (DTTL) shown in Figure6.9(c). The in-phase branch determines the polarity of the bit transition while the quadrature phase branch obtains a measure of the synchronization error. Finally, like the direct detection PPM system, nonlinear phase-locked loops can be used to track the symbol timing of an AWGN channel. Examples of the nonlinear PLL for synchronizing the BPSK system are shown in Figure6.9(d).

The performances of the various symbol synchronizers have been investigated extensively [28], [31]. It can be shown that the variance of the synchronization error for these loops generally have the following form:

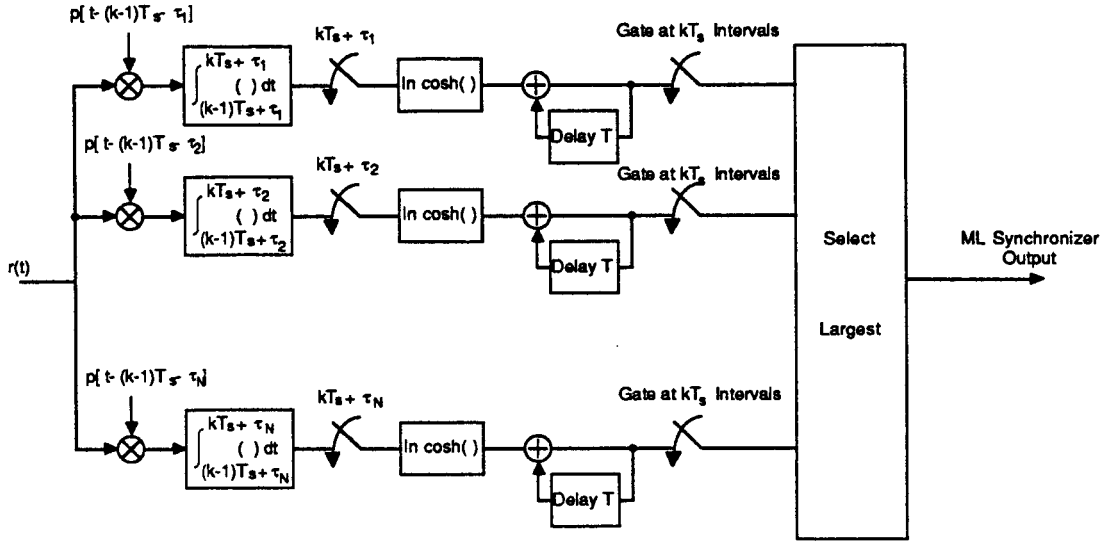
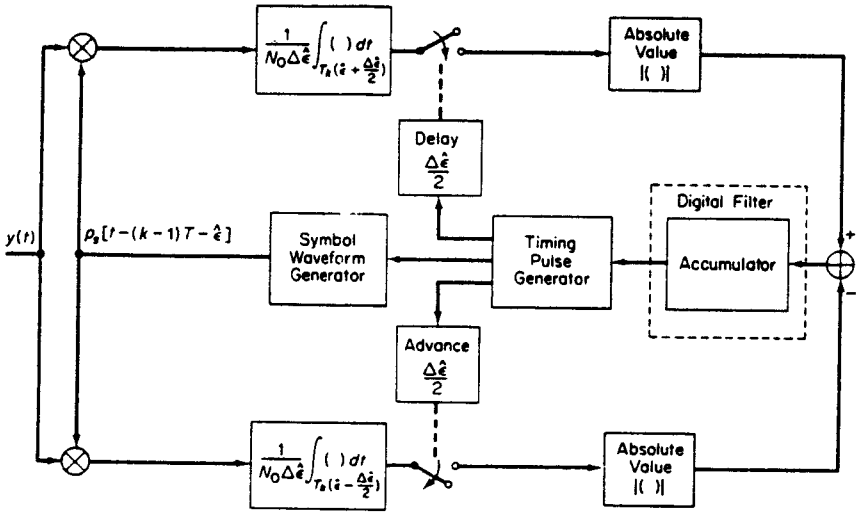


Fig.6.9(a). Maximum likelihood (ML) timing synchronizer for the BPSK system.



An early-late gate type of symbol synchronizer with absolute value type of nonlinearity

Fig.6.9(b). Early-late gate tracking loop.

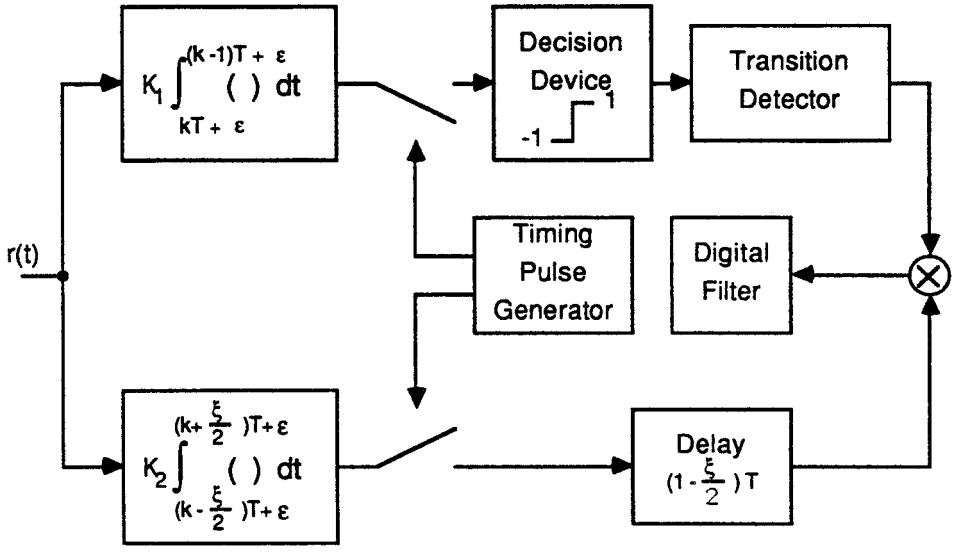
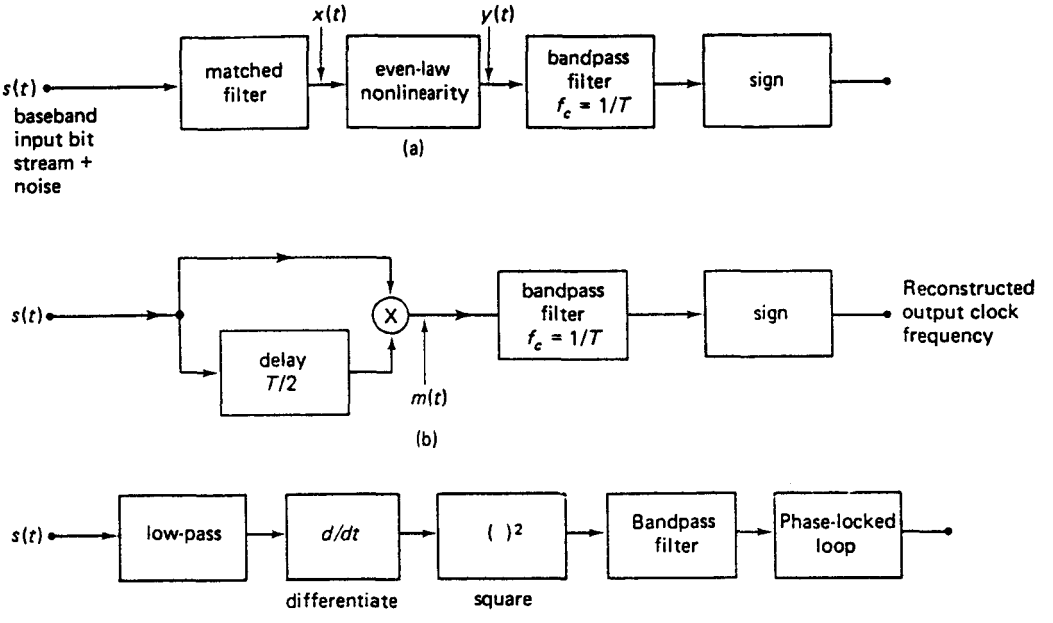


Fig. 6.9(c) The data-transition tracking loop.



Three types of nonlinear-filter bit synchronizers; (a) nonlinear filter using a matched filter and even-law nonlinearity, (b) delay-and-multiply synchronizer, and (c) differentiate-and-square synchronizer. A phase-locked loop can be employed in place of or in addition to the bandpass filter to provide better waveshape and narrower bandwidth.

Fig.6.9(d). Nonlinear phase-locked loops.

$$\sigma_{\varepsilon}^2 = \gamma \frac{B_L T_s}{\beta_c} \ , \tag{6.28}$$

where B_L is the bandwidth of the tracking loop, β_c is the coherent SNR, and γ is a dimensionless constant which depends on the loop structure and the pulse shape.

6.2.2 Synchronization of Noncoherent Heterodyne FSK Systems

Figure 6.10 shows a block diagram of a symbol synchronizer for the noncoherent binary FSK channel. During the acquisition period, a sequence of alternating frequencies is transmitted. This acquisition sequence is then envelope detected using two envelop detectors, each centered at one of the signal frequencies, over a T_s interval in quadrature with the primary detection interval (Figure 6.10). A maximum likelihood estimate of the normalized timing offset ε can then be derived from the output of these envelope detectors as

$$\varepsilon = \frac{\tau}{T_s} = \frac{M_1 - M_2}{2(M_1 + M_2)} \ , \tag{6.29}$$

where M_1 and M_2 denote the outputs of the sync envelope detectors. The conditional probability density function of this timing estimator is difficult to derive [28]. However, at high SNR, the output of the envelope detector can be approximated by a signal embedded in a Gaussian noise. In which case the estimator in Eq. (6.29) becomes unbiased and the variance of the estimator error can be easily calculated by

$$\sigma_{\varepsilon}^2 \simeq \frac{Var(M_1 - M_2)}{E^2[M_1 + M_2]} = \frac{1}{4\beta} \ , \tag{6.30}$$

where $\beta = A^2 T_s / 4N_s$ is the SNR. The timing error variance can be reduced by averaging the estimator output over a large number of code periods or, equivalently, by using a tracking loop with a narrow loop bandwidth. The analysis of the tracking loop follows closely the derivation leading to Eq. (6.8) [3]. It can be shown that, for a tracking loop with bandwidth B_L, the variance of the tracking error can be written as

$$\sigma_{\varepsilon}^2 = \frac{B_L T_s}{4\beta} \ . \tag{6.31}$$

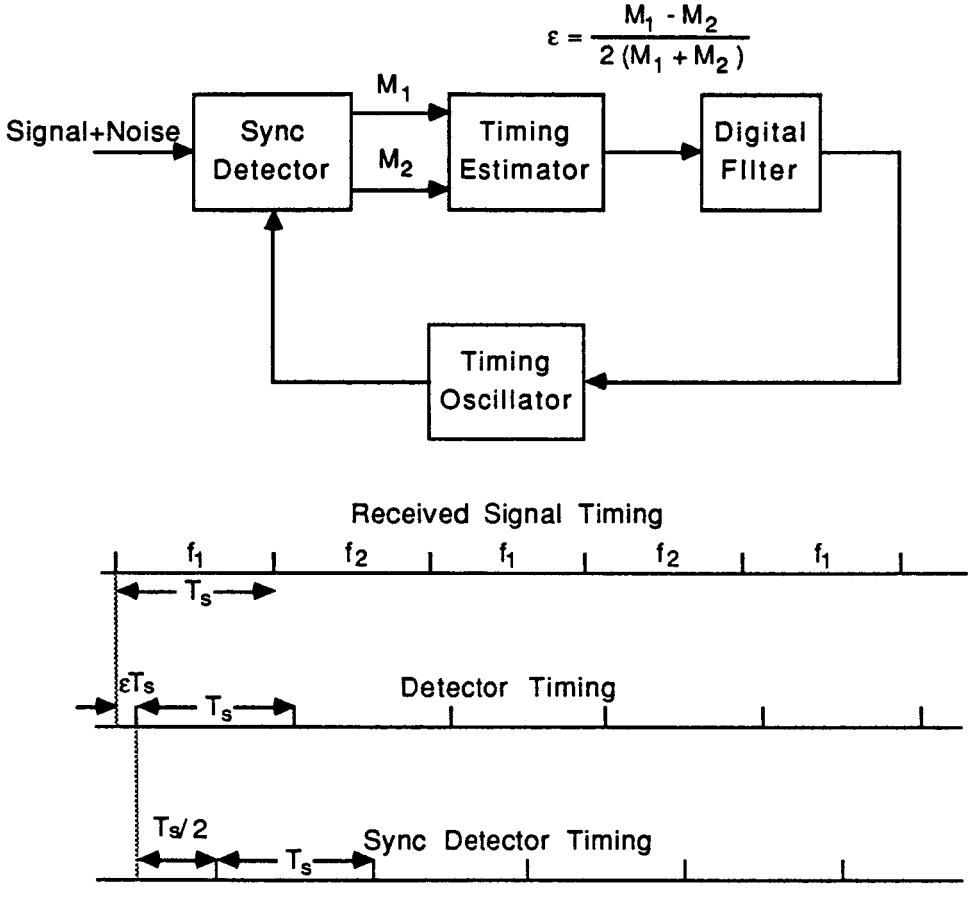


Fig.6.10 The structure and the timing diagram of the symbol synchronization loop for a noncoherent FSK system.

6.3 Error Performance in the Presence of Timing Error

6.3.1 Direct Detection PPM Systems

In the presence of a timing error, signal energy that was transmitted during a symbol interval may be interpreted as a strong background interference in the adjacent intervals. For a direct detection PPM system, such an intersymbol interference will result in a loss of SNR and, hence, the increase in the probability of decoding error. For systems employing the quantum limited, photon counting detectors, the probability of error in the presence of a normalized timing error ε can be approximated by [3]

$$P_E(K_S,K_B,\varepsilon)=\sum_{k=0}^{\infty}\frac{[(1-\varepsilon)K_S+K_B]^k}{k!}\left[\sum_{j=0}^{k}\frac{(\varepsilon K_S+K_B)^j}{j!}\right]\left[\sum_{j=0}^{k}\frac{(K_B)^j}{j!}\right]^{M-2}\cdot e^{-(K_S+MK_B)}\ , \tag{6.32}$$

where K_S, K_B are the expected photocount due to the signal and background in one slot period, and ε is the normalized timing offset. In practice, APDs are commonly used to detect the PPM encoded signal. In the presence of a fixed timing error ε, the probability of decoding error for systems using APDs can be written as

$$P_E(\varepsilon)=\frac{M}{2(M-1)}\left\{1-\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\cdot\left[\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_0}e^{-\frac{(y-\mu_0)^2}{2\sigma_0^2}}dy\right]^{(M-2)}\cdot\left[\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma'_0}e^{-\frac{(y-\mu'_0)^2}{2\sigma'^2_0}}dy\right]dx\right\} \tag{6.33}$$

where

$$\sigma_1^2=G^2F[(1-\varepsilon)K_s+K_b+I_bT_s/e]+I_sT_s/e+K_{th}\ ,$$

$$\mu_1=G[(1-\varepsilon)K_s+K_b+I_bT_s/e]+I_sT_s/e\ ,$$

$$\sigma_0^2=G^2F[K_b+I_bT_s/e]+I_sT_s/e+K_{th}\ ,$$

$$\mu_0=G[K_b+I_bT_s/e]+I_sT_s/e\ ,$$

$$\sigma_0'^2=G^2F[\varepsilon K_s+K_b+I_bT_s/e]+I_sT_s/e+K_{th}\ ,$$

$$\mu_{0'}=G[\varepsilon K_s+K_b+I_bT_s/e]+I_sT_s/e\ .$$

In writing Eq. (6.33), I_b, I_s are the gain dependent and gain independent dark currents, and K_{th} is the equivalent photocount rate of the thermal noise. At sufficiently high SNR, Eq. (6.33) can be approximated by the union bound:

$$P_E(\varepsilon) \le \frac{1}{2}\mathrm{erfc}\left[\frac{\mu_1-\mu_{0'}}{\sqrt{2(\sigma_1^2+\sigma_\delta^2)}}\right] + (M-2)\frac{1}{2}\mathrm{erfc}\left[\frac{\mu_1-\mu_0}{\sqrt{2(\sigma_1^2+\sigma_\delta^2)}}\right] \quad . \tag{6.34}$$

Note that for $M=2$, the union bound gives an exact evaluation of Eq. (6.33). In deriving Eqs. (6.32) and (6.33), we have neglected the end-effect at the word boundaries as they may contain energy spillover from adjacent words. For higher order PPMs ($M>4$) the assumption that the end-effect is negligible can be justified [3]. Furthermore, since APDs are more commonly used in direct detection PPM systems, and because the union bound proves to be an effective approximation of the PBE, Eq. (6.34) will be used to evaluate the PBE in the presence of a timing error.

For receivers employing dynamic phase synchronization circuits such as a PLL, the receiver timing error is randomly distributed. However, when the loop bandwidth B_L is much smaller than the word frequency, the timing error will be slowly varying so that it is approximately constant during a given word period. The unconditional error probability for the PPM system is therefore the expectation of (6.33) with respect to the distribution of ε. For a well-synchronized system, ε can be assumed to be Gaussian distributed, and the unconditional PBE is given by

$$P_E = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\int P_E(\varepsilon)e^{-\varepsilon^2/2\sigma_\varepsilon^2}d\varepsilon \quad . \tag{6.35}$$

Equation (6.35) was evaluated numerically and the resulting PBE is plotted in Figure 6.11 versus the expected signal count, K_s, for a 4-ary PPM system. The results show a severe degradation of receiver performance with increasing timing error variance, particularly when $\sigma_\varepsilon \approx 0.1$.

Numerical evaluation of PBE has been carried out for various signaling conditions and receiver loop bandwidths. The results are shown in Figures 6.12 through 6.14. Figure 6.12 is a plot of the PBE versus signal count for various loop bandwidths. Note that these results and the results plotted in Figures 6.13 and 6.14 were calculated by assuming a conservative value of 10 for γ. It was shown in Section 6.1.3 that $\gamma \approx 3$ for well-designed systems. For small loop bandwidths ($B_L T_s < 10^{-3}$), the performance of the PLL synchronized system is almost identical to that of the perfectly synchronized system. However, for large loop bandwidths ($B_L T_s > 10^{-2}$), higher signal levels are needed to compensate for the effects of synchronization errors. The increase in signal power needed to compensate for the imperfect synchronization (i.e., maintain a given BER) can be described equivalently in terms of decibel loss in signal power due to synchronization error. Figure

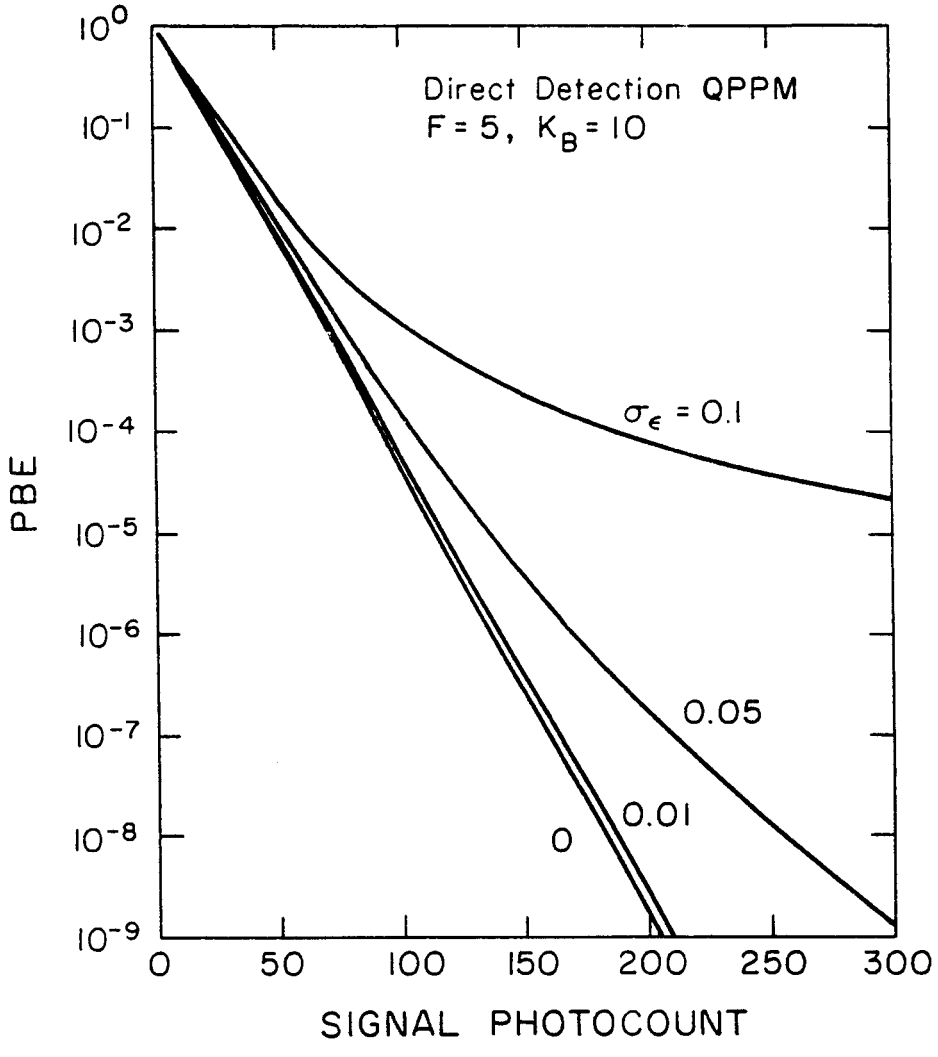


Fig.6.11. The probability of bit error (PBE) as a function of K_S, the expected signal photocount per word, for different values of the receiver timing error variance.

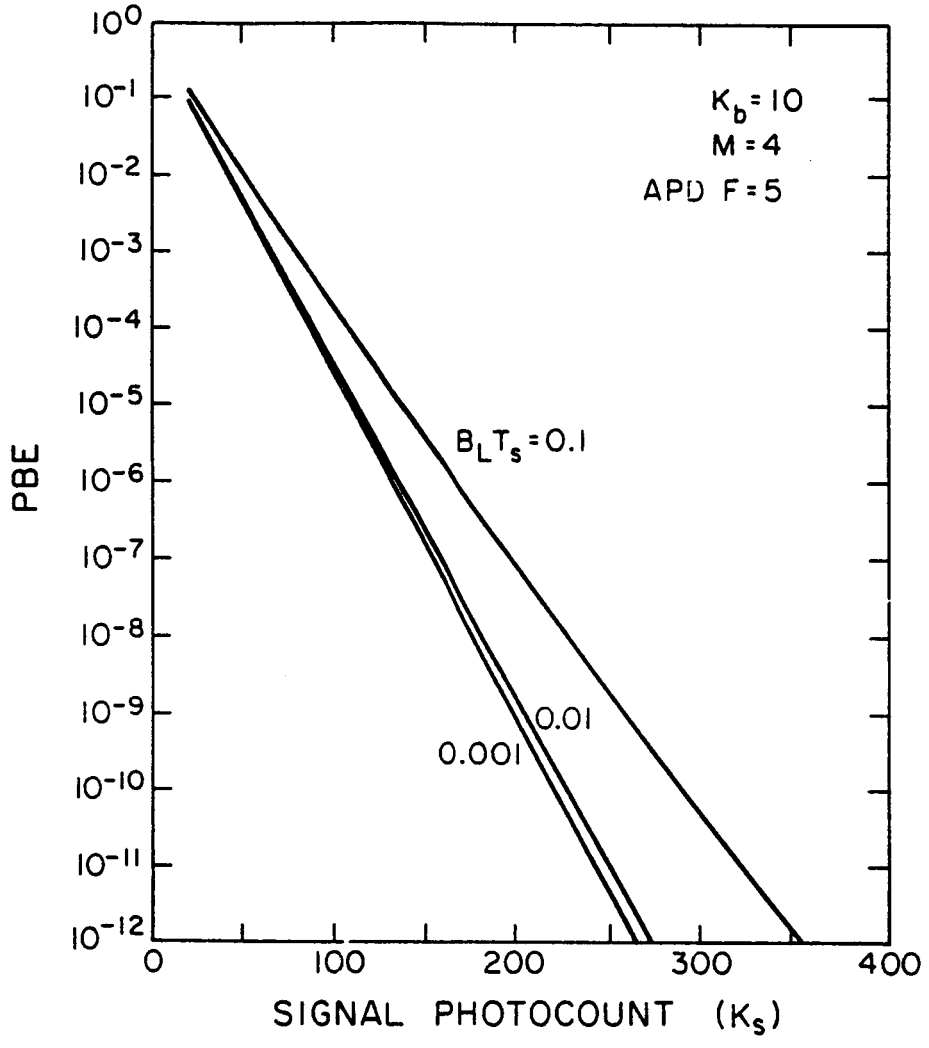


Fig.6.12. Probability of bit error (PBE) for a PLL synchronized 4-ary PPM system versus K_S, the expected signal photocount per word, for various loop bandwidths. The value of γ is set equal to 10 in this simulation.

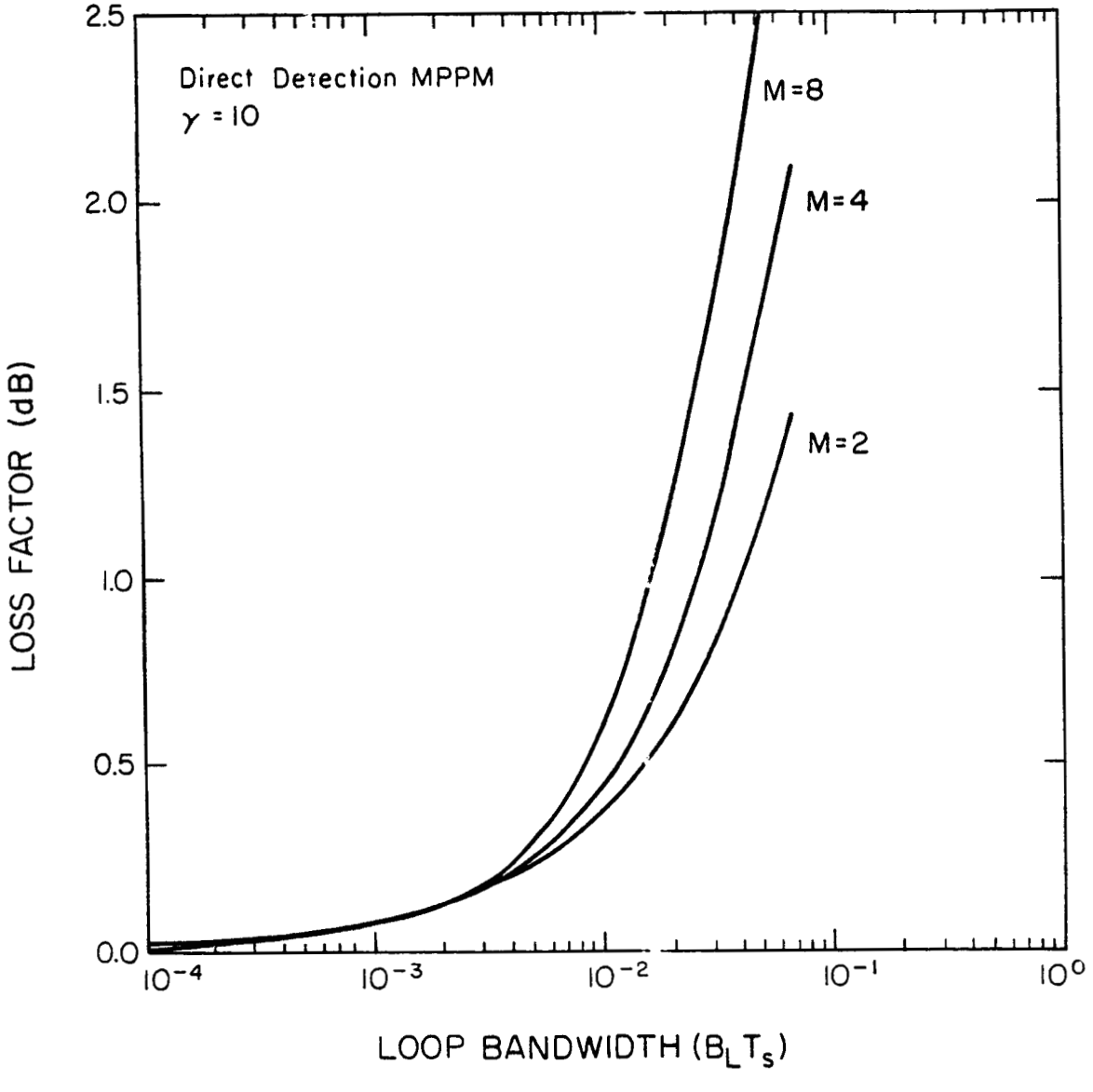


Fig.6.13. Power penalty versus the loop bandwidth for a PLL synchronized PPM system at $PBE=10^{-9}$, $K_B=1$, and $\gamma=10$.

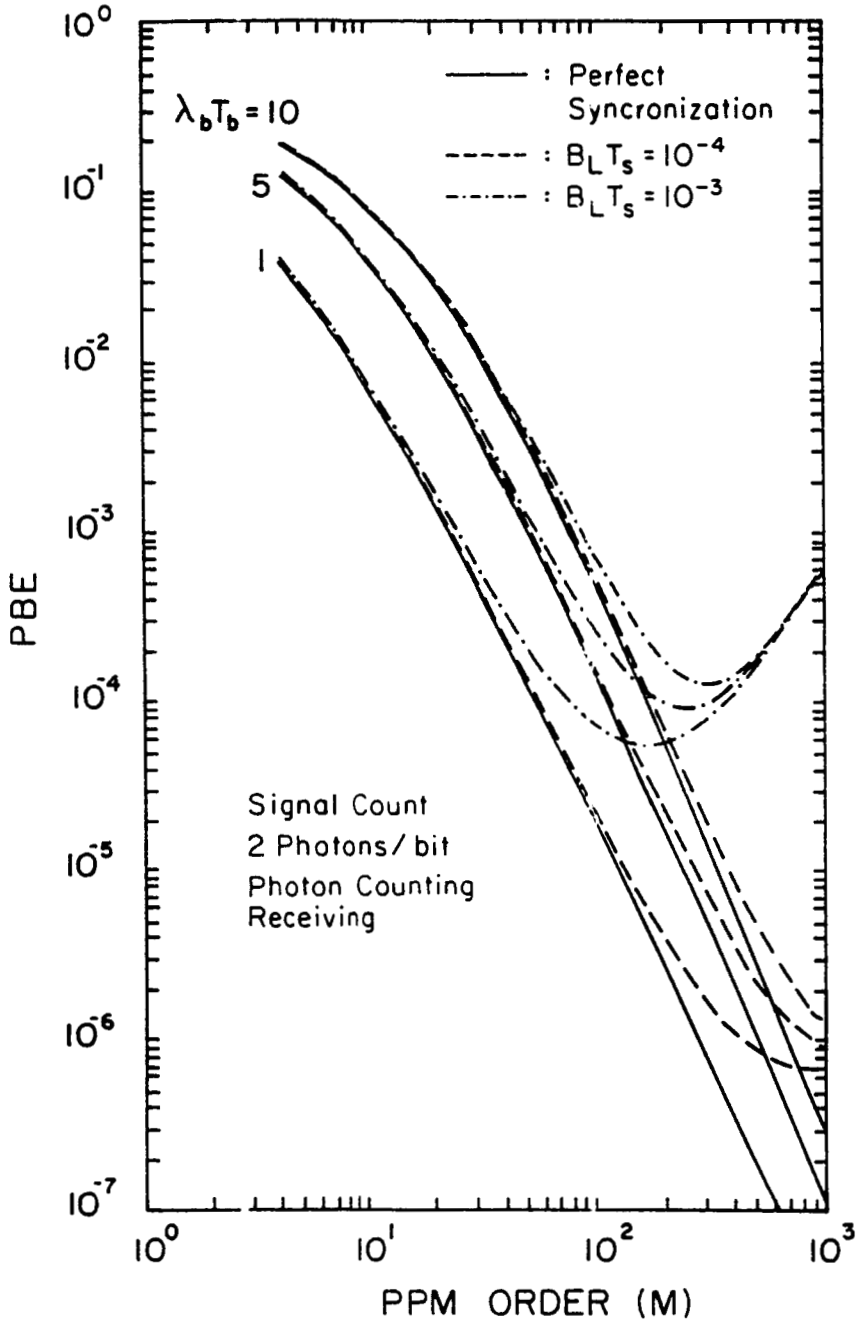


Fig.6.14. Probability of bit error (PBE) of a PLL synchronized PPM system versus M for various background levels and loop bandwidths. The signal photocount per bit is fixed at 2, and the equivalent data rate is kept fixed.

6.13 is a plot of the loss factor versus the loop bandwidth at a fixed PBE of 10^{-9} for several PPM systems. Notice that the M=2 curve is included because computer simulation of phase error variance agrees well with (6.26). It is seen from Figure 6.13 that for small loop bandwidths ($B_L T_s < 10^{-3}$), imperfect synchronization accounts for less than 0.1 dB loss in signal power, while at higher loop bandwidths ($B_L T_s \approx 0.1$), the loss can be significant. The loss factor curves plotted in Figure 6.13 are for a fixed PBE of 10^{-9}. However, it can be easily deduced from the linearity of curves shown in Figure 6.13 that similar losses can be expected for PBE ranges from 10^{-3} to 10^{-9}.

Because of the coding gain, the PBE of a perfectly synchronized PPM system decreases with increasing PPM order. For PLL synchronized PPM systems, however, the variation of PBE with PPM order is more complicated because the timing error is also a function of the PPM order. Figure 6.14 is a plot of the PBE versus PPM orders for various loop bandwidths and background count rates with the number of signal photons per bit and the equivalent bit period kept constant. The data show that for small loop bandwidths ($B_L T_s < 10^{-4}$) and for $M < 10$, the performance of the PLL synchronized system is almost indistinguishable from the perfectly synchronized system. Because the phase error variance increases with PPM order, the degradation in performance becomes more pronounced at larger values of M. At high enough values of M the synchronization error will dominate the coding gain so that eventually the PBE will become an increasing function of M. Consequently, an optimal order of PPM exists where the coding gain balances the synchronization error and the PBE is minimum. Figure 6.14 shows that for a PPM system using the photon counting detector and operating with a received signal level of 2 photons per bit, a background count rate of 1 photon per bit interval and $B_L T_s = 10^{-3}$, the PBE is minimum near $M \approx 100$. It should be noted that Figure6.14 was generated assuming that the transmitter laser has a perfect modulation extinction. In the presence of a finite extinction ratio, the detector SNR is limited by the residual signal in the background slots, and the curves shown in Figure6.14 must be modified. In fact, for a PPM system with perfect timing synchronization, the performance of the receiver improves with increasing PPM order until the limiting SNR is achieved, and further increase in the PPM order will only result in a deteriorated system performance. This optimum PPM order is typically on the order of M=10 for modulation extinction ratios between 1 and 10%. Consequently, the optimum PPM order due to synchronization error in Figure6.14 has very little impact on the design of the practical PPM channel.

6.3.2 Impact of Synchronization Error on Heterodyne Systems

Just as it was in the direct detection system, synchronization error affects the performance of the heterodyne channel by introducing inter-symbol interference at the decoder. For simplicity, the following analysis will assume that the carrier phase fluctuation can be ignored and that the heterodyne channel can be approximated by an AWGN. Furthermore, the analysis will be confined to the binary coherent PSK (CPSK) and the noncoherent FSK (NCFSK) systems. Detailed discussions on the impact of symbol synchronization error can be found in many communication texts [28]-[31].

Given a normalized timing error ε, the conditional PBE for a binary CPSK system with rectangular baseband pulses can be written as [28], [31]

$$P_E(\varepsilon) = \frac{1}{4}\mathrm{erfc}(\sqrt{\beta_c}) + \frac{1}{4}\mathrm{erfc}[\sqrt{\beta_c}(1-2|\varepsilon|)] \ , \qquad \text{(CPSK)} \quad (6.36)$$

where β_c is the coherent signal-to-noise ratio of the receiver in the absence of timing error. The conditional PBE can also be calculated for a noncoherent FSK (NCFSK) system as [28]

$$P_E(\varepsilon) = 1 - \int_0^{\infty}[1-Q(\sqrt{\rho}\varepsilon, y)][1-e^{-\frac{y^2}{2}}]^{M-2} y e^{-\frac{1}{2}[y^2+\rho(1-\varepsilon)^2]} I_0(\sqrt{\rho}(1-\varepsilon)y)\,dy \ , \qquad \text{(NCFSK)} \quad (6.37)$$

where ρ is the signal-to-noise ratio of the noncoherent receiver, and $Q(\alpha,\beta)$ is Marcum's Q function defined in Eq. (3.11). Equation (6.37) is in general very difficult to calculate. When the signal-to-noise ratio is sufficiently high, union bound can be used to approximate the error rate:

$$P_E(\varepsilon) \le \frac{1}{2}\left[1+Q[\sqrt{\rho}\varepsilon, \sqrt{\rho}(1-|\varepsilon|)] - Q[\sqrt{\rho}(1-|\varepsilon|), \sqrt{\rho}\varepsilon]\right] + \frac{(M-2)}{2} e^{-\rho(1-|\varepsilon|)/2} \ . \qquad \text{(NCFSK)} \quad (6.38)$$

The unconditional PBE can be evaluated by taking the expectation of the conditional PBE given by Eqs. (6.36) and (6.37) with respect to the distribution of the timing error. Again, for a well-synchronized system, the timing error can be assumed to be Gaussian distributed so that the unconditional PBE is given by Eq. (6.35) [28]. The random synchronization error affects the performance of the heterodyne channel in two respects. First, the synchronization error introduces intersymbol interference which reduces the SNR and increases the PBE. Second, given an RMS sync error, there is a finite probability that the normalized synchronization error will exceed 1/2. For this case, the output of the detector matched to the signal will be smaller than the output

of the detector matched to the adjacent symbol and, therefore, results in a large probability of error. At high SNR, the PBE due to the sync error can dominate the system error rate and result in an error rate "floor" which cannot be improved by increasing the SNR. Given an RMS sync error σ_ε, the error rate "floor" can be given by

$$PBE = 2\times\int_{1/2}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_\varepsilon}e^{-x^2/2\sigma_\varepsilon^2}dx = \mathrm{erfc}\left[\frac{1}{2\sqrt{2}\sigma_\varepsilon}\right] \quad . \tag{6.39}$$

The degradation of the performance of heterodyne CPSK and NCFSK channels due to the synchronization error can be seen in Figures 6.15(a) and 6.15(b) where the PBE is plotted against the receiver SNR. Note the appearance of error rate "floors" when σ_ε>0.1. The power penalties associated with the synchronization error for the heterodyne CPSK and NCFSK systems can be calculated as a function of the variance of the synchronization error. These power penalties are plotted in Figures 6.16(a) and 6.16(b) versus the RMS synchronization error at $PBE=10^{-3}$, 10^{-6}, and 10^{-9}, respectively. Note that, at small σ_ε, the power penalties are independent of the PBE. However, at large RMS sync errors, the error rate "floor" due to the synchronization error can exceed the desired PBE, and result in a rapid increase in the power penalty.

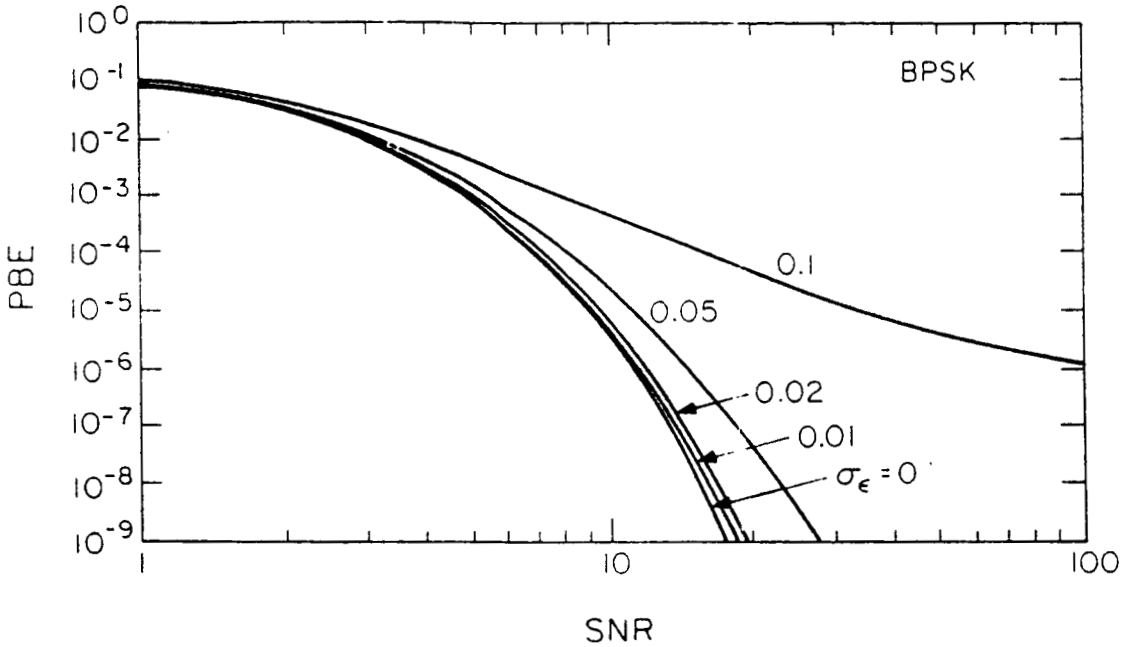
In practice, symbol synchronization of the heterodyne system is usually accomplished using a tracking loop such that the resulting RMS receiver timing error is inversely proportional to the SNR (cf. §6.2). In this case the unconditional PBE of the heterodyne channel can be represented as a function of the loop bandwidth and the signal power. Figures 6.17(a) and 6.17(b) are plots of the power penalty versus the synchronization loop bandwidth, assuming that the variance of the synchronization error can be related to the loop bandwidth and the SNR β by

$$\sigma_\varepsilon = \frac{\gamma}{(2\pi)^2}\frac{B_L T_s}{\beta} \quad . \tag{6.40}$$

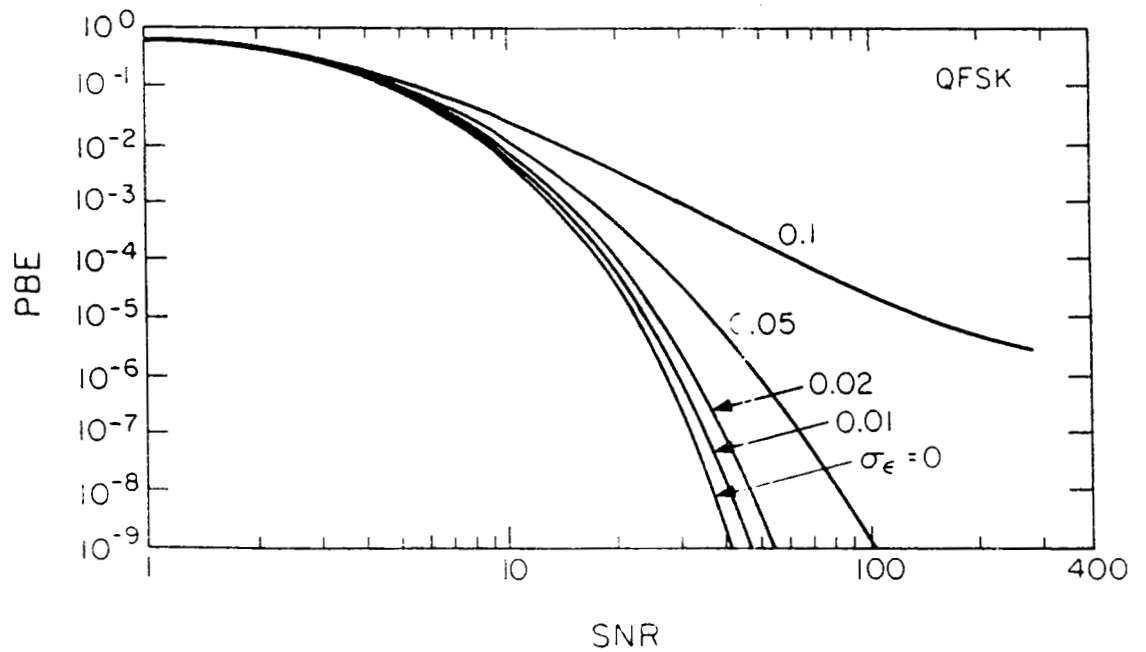
Note that, by tightening the loop bandwidth, the power penalty associated with the symbol synchronization can be minimized.

6.4 Summary

The results in this chapter show that, given the detector output signal, the performance of the symbol synchronizer can be improved by reducing the bandwidth of the tracking loop and increasing the SNR. Even when suboptimal symbol synchronizers are used, the power penalty associated with the synchronization error can be minimized by reducing the bandwidth of the symbol tracking loop. It is seen from the analyses that,

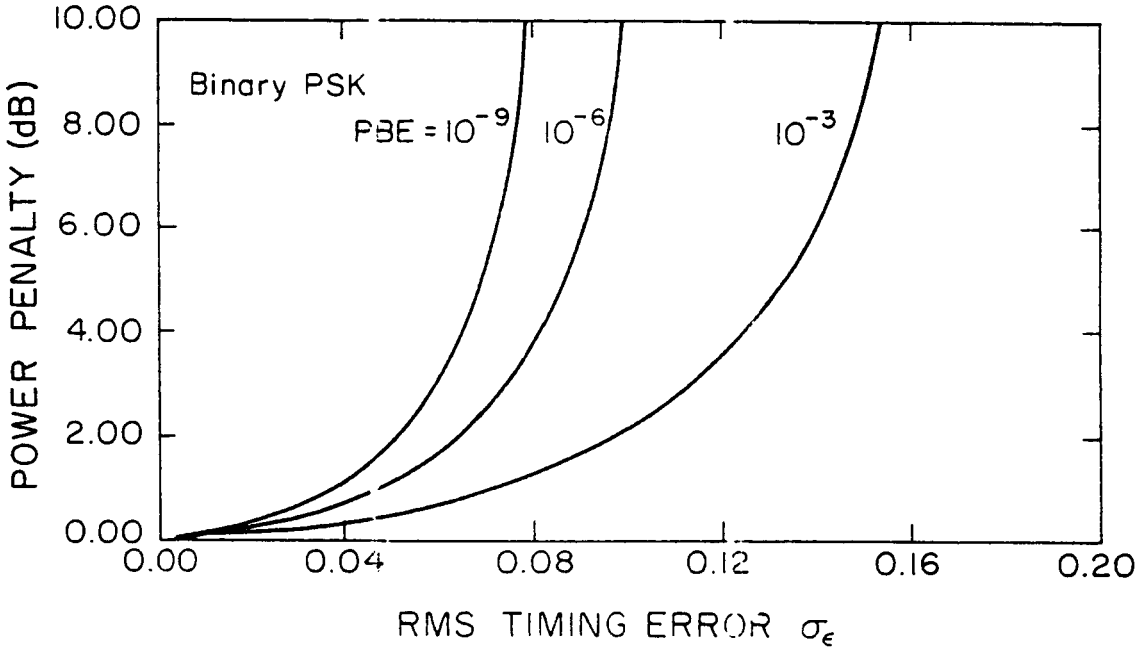


(a)

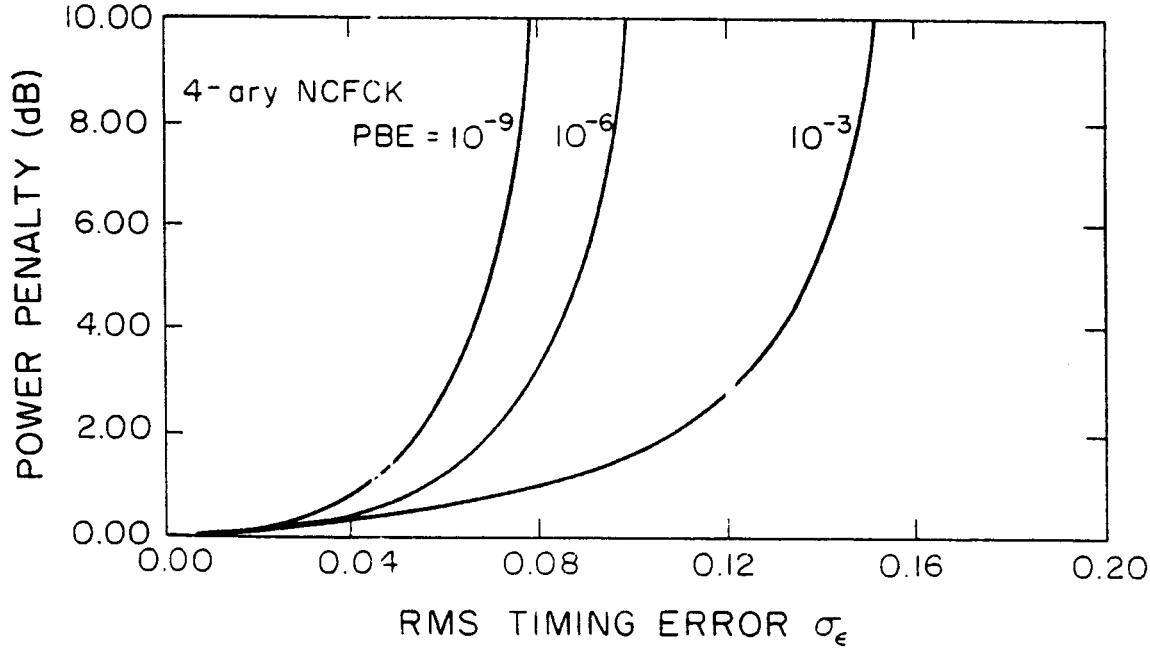


(b)

Fig.6.15. Probability of bit error (PBE) versus the SNR for (a) bin ary CPSK system and (b) 4-ary NCFSK system.

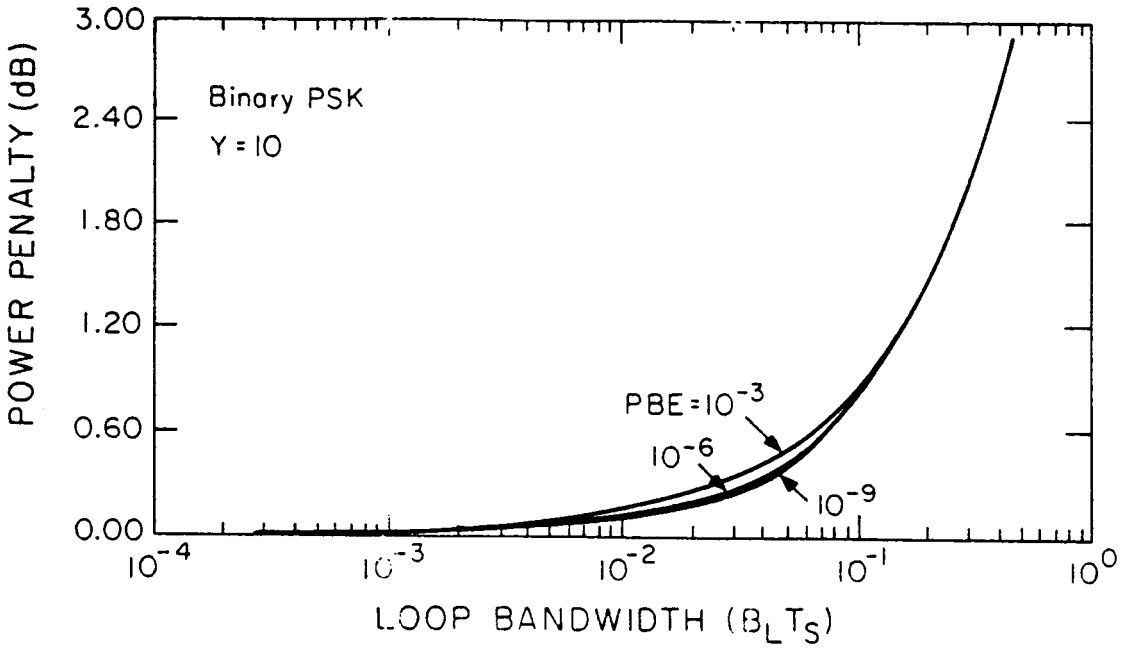


(a)

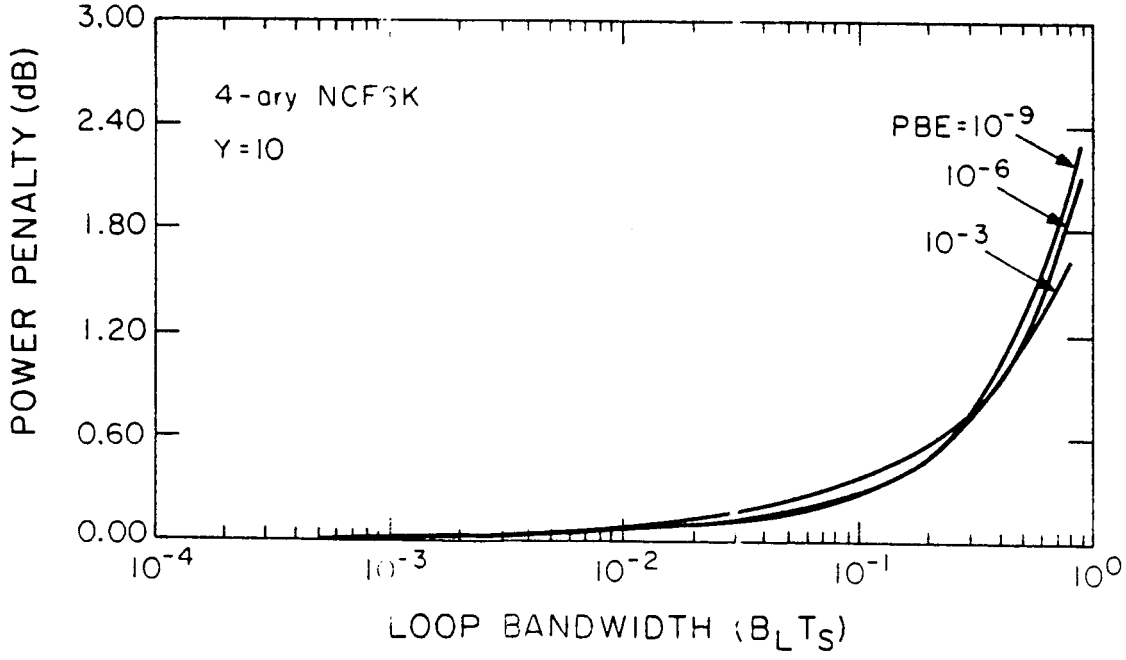


(b)

Fig.6.16. Power penalty due to the symbol synchronization error versus the RMS timing error for (a) binary CPSK and (b) 4-ary NCFSK systems at different error rates.



(a)



(b)

Fig.6.17. The power penalty versus the synchronization loop bandwidth for (a) binary CPSK and (b) 4-ary NCFSK systems at different error rates. The variance of the synchronization error is related to the SNR by Eq.(6.39).

with a loop bandwidth that is less than 0.1% of the slot frequency, the direct detection PPM system can be slot synchronized with less than 0.1dB of power penalty. It should be noted that for PPM decoding, word and frame synchronizations must also be achieved. Such a synchronization can be achieved using a coding technique after the slot synchronization is accomplished [165], [166]. For heterodyne systems, the analysis also shows that the power penalty due to symbol synchronization is small. In fact, it was seen in Chapter 4 that a more dominating source of error in the heterodyne system is due to the carrier synchronization and the carrier phase noise.

7. PERFORMANCE EVALUATION

Optical communication links can be implemented using either direct or heterodyne detection receivers [3]. The direct detection system [3]-[5] offers simplicity in design and, therefore, potentially higher reliability. In contrast, the heterodyne system [3], [6]-[8] provides higher detection sensitivity, better background noise rejection, and the potential of frequency and phase encodings. In principle, the heterodyne system has a 5-10 dB advantage in the detection sensitivity over the comparable direct detection channel [6]-[8]. In practice, however, the heterodyne channel is very sensitive to the LO tracking error [54] and the laser phase noise [44]-[53]. When both the LO tracking error and the carrier phase noise are considered, the performance of the heterodyne channel can be seriously deteriorated.

Both direct [1]-[5] and heterodyne [6]-[8] detection systems have been studied for ISL implementations. In this chapter, the design of typical direct and heterodyne detection ISLs will be studied and the performances of both systems in a realistic operating environment will be analyzed.

7.1 Design Specifications

Table 7.1 summarizes the mission requirements for an optical ISL. Because of the large demand projected for future space missions, a data rate which exceeds 200 Mbps with an averaged PBE less than 10^{-6} is desired. The mean-time-to-failure (MTTF) of the system should be better than 5 years. Based on the these mission requirements, detailed specifications on the ISL can be defined. Table 7.2 lists some of the parameters that are critical to the performance of the ISL. The choice of these parameters will be discussed in detail in following sections.

7.1.1 Modulation Format

Given the desired data rate (>200Mbps), modulation formats must be properly specified for both direct and heterodyne detection systems.

Direct detection PPM system For direct detection systems, pulse-position modulation (PPM) is used to modulate the transmitter laser. The PPM is chosen over the OOK scheme because of its coding gain and the noise immunity [3], [11]. It should be noted that because of the coding gain, the performance of the PPM system improves with increasing PPM order. In practice, however, it is not practical to use PPM with order

Table 7.1 ISL mission requirements.

LINK DISTANCE	z = 40,000 km (LEO-GEO) / 80,000 km (GEO-GEO)
DATA RATE	R = 200 Mbps
BIT ERROR RATE	$< 10^{-6}$
SYSTEM WEIGHT	50 kg
LIFETIME	>5 yr

Table 7.2 ISL Design Parameters.

MODULATION FORMAT

TRANSMITTER LASER
- POWER
- EXTINCTION RATIO
- MODULATION BANDWIDTH
- MULTIPLEX/DEMULTIPLEX SCHEMES
- LINEWIDTH

OPTICAL ANTENNA
- APERTURE DIAMETERS

TRANSMITTER AND RECEIVER OPTICS
- TRACKING SPLIT
- EFFICIENCY
- RECEIVER FIELD-OF-VIEW (FOV)
- RECEIVER BANDWIDTH

OPTICAL DETECTOR
- QUANTUM EFFICIENCY
- BANDWIDTH
- GAIN
- NOISE CHARACTERISTICS

which exceeds M=256 because, at higher order PPMs, the system complexity and the required transmitter and receiver bandwidths increase rapidly. Furthermore, in the presence of a finite modulation extinction ratio, the receiver signal-to-noise ratio (SNR) is limited by the leakage signal in the background slots (cf. Chap. 3) and an optimum PPM order exists which maximizes the system performance. For the analysis of direct detection PPM systems, it will be assumed that M=4.

Heterodyne NCFSK system For heterodyne systems, the noncoherent frequency-shift keying (NCFSK) scheme is chosen because of its superior performance in the presence of carrier phase noise [6]-[8]. It should be noted that although coherent heterodyne schemes offer an improved detection sensitivity over the NCFSK system, they are not suitable for ISL implementations because the signal power required for optical carrier synchronization is difficult to achieve in practice [152].

For the heterodyne NCFSK system, tone spacing between frequencies is an important parameter (cf. Chap. 4). The performance of the NCFSK system can be optimized if the different transmitted signal waveforms are orthogonal [28], [31]. This is possible if the tone spacing is an integer multiple of the symbol frequency, $f_s=1/T_s$ [28], [29]. In the presence of carrier phase noise, the system power penalty also depends on the tone spacing (cf. Chap. 4). The greater the tone spacing, the smaller the impact of carrier phase noise on the system performance [8], [36]. However, greater tone spacing implies larger system bandwidth and, in practice, it is not possible to increase the tone spacing indefinitely. For the analysis of heterodyne NCFSK systems, a tone spacing $v_d=f_s$ will be assumed.

Finally, since a fair comparison between direct and heterodyne systems can only be made if both systems have similar coding gains, the analysis of the heterodyne channel will be carried out for a 4-ary NCFSK system.

7.1.2 Transmitter Laser

One of the basic assumptions made at the beginning of the study is the availability of semiconductor lasers with sufficiently high power and reliability. The advantages of choosing a semiconductor laser include higher power efficiency and reliability. In addition, semiconductor lasers are smaller in size and can be more easily integrated.

Direct detection PPM system For use in a direct detection PPM system, laser power as high as several tens or hundreds of milliwatts will be needed. A frequency and polarization multiplexed system (cf. Fig. 2.1) can be used to provide the required transmitter power [102], [103]. This scheme is chosen because of the limited power output of single laser diodes and the fact that laser diode arrays currently available suffer from poor reliability and large far-field beam divergence. Other beam combining techniques such as grating [105] and waveguide combining [102], [103] are more difficult to implement and require a longer development cycle. The power output of the combined laser beam is assumed to be 200 mW [105], [171].

In addition to the power output, the selection of transmitter lasers for direct detection PPM systems also depends on the modulation bandwidth and the extinction ratio. The modulation bandwidth of laser diodes has been demonstrated to exceed 15GHz [91] and, for most practical systems, is well above the desired data rate. The modulation extinction ratio is defined as the ratio of power output when no signal is transmitted versus that when a signal pulse is sent [23]. In the presence of a finite extinction ratio, the detector background count is dominated by the residual optical signal which increases with increasing peak transmitter power. This finite extinction ratio can have a significant effect on the coding gain of the PPM channel because it limits the maximum achievable SNR (cf. Chap.3). Typical values of the extinction ratio are between 1 and 10%. For the analysis of the direct detection PPM system, a modulation extinction ratio of 5% will be assumed.

Heterodyne NCFSK system The wavelength and polarization combining technique provides sufficient power output for the direct detection PPM link. When implementing the heterodyne NCFSK system, however, lasers with narrow spectral linewidths are necessary. Laser diodes currently available with high power output generally have spectral linewidths on the order of 0.01-1nm or, equivalently, 10^9-10^{11}Hz, and provide little control over the lasing frequency [96]. Consequently, these lasers are not suitable for the heterodyne application. Instead, laser diodes with narrow spectral linewidth must be used. These lasers generally employ feedback to stabilize the lasing wavelength and suppress the carrier phase noise. Techniques such as distributed feedback (DFB) [112]-[115], distributed Bragg reflector (DBR) [116], or coupled cavity [109] can be used to achieve stable, single mode operation. However, laser diodes designed for narrow spectral linewidths generally have smaller output power. Fortunately, over the past few years, there has been a steady increase in the power output of frequency stabilized lasers. For the analysis of the heterodyne ISL, it will be assumed that, by 1990, frequency stabilized, narrow spectral linewidth, laser diodes will be available with an

average power output of 30 mW [171] and a spectral linewidth no greater than 5 MHz.

An alternative to using frequency stabilized laser diodes in heterodyne systems is to use the diode pumped solid state Nd:YAG laser. Previous efforts in developing the diode pumped Nd:YAG laser have been limited by the low efficiency of the pumping mechanism. Recently, however, by using a tightly focused end-pump geometry, pumping efficiency as high as 8% had been achieved which resulted in 80 mW CW power output with only 1W of electrical power into a single semiconductor laser array pump [85]. Compared to semiconductor diode lasers, the diode pumped Nd:YAG laser offers a much improved spectral purity and a potential increase in output power. However, a wideband external electro-optical modulator is needed to modulate the laser which introduces additional loss in the signal power and consumes more prime electric power. A diode pumped Nd:YAG laser is expected to produce a power output of approximately 100mW with good spectral purity.

7.1.3 Antenna Gain

The transmitter and receiver antenna gains or, equivalently, the transmitter and receiver aperture diameters, determine the amount of signal power that can be focused onto the receiver.

Transmitter Antenna Gain Ideally, when spatial tracking can be perfectly realized, it is desirable to choose a large transmitter antenna gain so that a higher signal intensity can be focused onto the receiver. In the presence of random pointing errors, however, transmitter antenna gain cannot be increased arbitrarily. In fact, given the RMS pointing error, an optimal transmitter antenna gain exists which minimizes the required signal power (cf. Chap. 5). When the pointing error at the transmitter can be modeled as a Rayleigh distributed random variable, the optimal antenna gain is inversely related to the square of the RMS pointing error at the transmitter so that the product $\gamma_T = G_{T_{opt}} \sigma_{\theta_T}^2$ is independent of the RMS pointing jitter. Table 7.3 summarizes the parameter γ_T for a direct detection 4-PPM system and a heterodyne noncoherent 4-FSK system at different values of the PBE. Given the RMS pointing error, the optimal transmitter antenna gain can be chosen according to Table 7.3.

Because the optimal transmitter antenna gain is inversely proportional to the square of the RMS pointing jitter, the optimal antenna gain can be very large for systems with small RMS pointing jitters. For such systems, a suboptimal design in which the transmitter antenna gain is smaller than the optimal value is usually

Table 7.3 The parameter $\gamma_T = G_{T_{opt}} \sigma_{\theta_T}^2$ at different values of the PBE.

PBE	$G_T \sigma_\theta^2$	
	Direct Detection QPPM †	Heterodyne QFSK
10^{-3}	0.42	0.48
10^{-6}	0.22	0.27
10^{-9}	0.15	0.19

† Derived with receiver parameters shown in Table 7.4.

systems, a suboptimal design in which the transmitter antenna gain is smaller than the optimal value is usually chosen. Because the antenna gain is not optimized, higher signal power is needed to maintain the system performance. However, since the transmitted beamwidth increases with decreasing transmitter antenna gain, such systems are less sensitive to the random pointing error. As a result, the power penalty due to random pointing error is much smaller for systems using the suboptimal design. Figures 5.14 and 5.15 show the power penalty versus the antenna gain for a 4-ary PPM system and a 4-ary NCFSK system, respectively. Note that when the antenna gain is much smaller than the optimal value, the power penalty is very small (<0.5 dB).

Receiver Antenna Gain For direct detection systems, the amount of signal power collected by the receiver is proportional to the area of the receiver aperture. For such systems, it is desirable to use a receiver with as large an aperture diameter as possible. In practice, however, a larger receiver aperture implies a larger system weight, a higher launch cost and a greater power demand for the steering system. Consequently, the size of the receiver aperture cannot be increased indefinitely. For the analysis of the direct detection ISL, a receiver aperture diameter of 30 cm will be assumed [171]. In contrast, a larger aperture diameter does not necessarily imply a greater SNR for a heterodyne receiver, because, in addition to the receiver antenna gain, the SNR of a heterodyne receiver also depends on the alignment between the signal and the LO. Given an RMS tracking error of the LO, an optimum receiver antenna gain can be found which optimizes the system performance (cf. Chap. 5). When the tracking error can be modeled as a Rayleigh distributed random variable, this optimum receiver antenna gain can be inversely related to the square of the LO tracking error. In fact, given the desired PBE and the modulation format, the parameter $\gamma_R = G_{R_{opt}} \sigma_{\theta_R}^2$ is equal to the parameter γ_T given in Table 7.3 and is independent of the RMS tracking error. For the analysis of the heterodyne ISL, it will be assumed that the diameter of the receiver aperture is chosen such that the required transmitter power is minimized.

7.1.4 Transmitter and Receiver Optics

In addition to the transmitter and receiver aperture diameters, other transmitter and receiver optical parameters that affect the performance of the ISL include the transmitter and receiver optical efficiencies, the tracking split, the receiver field-of-view (FOV), and the optical bandwidth. The efficiencies of the transmitter and receiver optics depend on the number of reflecting and refracting surfaces. For the analysis, optical efficiencies of 50% and 35% will be assumed for the transmitter and receiver optics, respectively. The

will be assumed to be 5%.

In addition to the signal power, the performance of the direct detection PPM system also depends on the amount of background power collected by the receiver. In order to reduce the amount of background radiation collected by the receiver, both the receiver FOV and the optical bandwidth should be carefully controlled. However, a small receiver FOV implies a more stringent tracking requirement, and a narrow band optical filter may have an unacceptably high loss. Furthermore, when the transmitter laser is frequency multiplexed, the receiver optical bandwidth should be wide enough to admit all wavelengths of the transmitter output. For the analysis, a receiver bandwidth of 2 nm and FOV of 0.1 mrad will be assumed.

7.1.5 Optical Detectors

Optical detectors with high bandwidth and low noise are required for ISL applications. The quantum efficiencies of semiconductor photodetectors are typically on the order of 50%-90%. *Pin* photodiodes with bandwidths as high as 20GHz have been reported in laboratory work [73]. However, because of the low detector gain, *pin* diodes are used primarily in heterodyne systems where the detector thermal noise is negligible compared to the LO shot noise. For direct detection ISLs, APDs with large internal gain are needed to overcome the detector thermal noise. Table 7.4 lists the typical characteristics of a silicon-APD based direct detection receiver. For the comparison, the ratio of effective ionization coefficients of 0.01 will be assumed. An effective ionization ratio as small as 0.007 can be achieved with selected commercial devices [171].

Given the amount of optical power incident on the detector surface, the gain of the APD is an important parameter that must be properly specified. The effect of detector thermal noise on the receiver SNR can be minimized by increasing the detector gain. However, larger detector gain leads to an increased excess noise factor F which in turn can reduce the receiver SNR. Given K_S=200 and T_s=5ns, the optimal detector gain can be evaluated from Eq. (3.24) to be approximately 150. The excess noise factor at the optimal detector gain is therefore F_{opt}≈3.5. Note that the optimal gain decreases with the 1/3 power of K_S. At large signal level the optimal gain is 1 and the optimal excess noise factor is equal to 1. Therefore, a design which optimizes the gain at K_S=200 will have a power penalty of 5.4 dB at K_S=∞.

Table 7.4 Typical Si-APD based direct detection receiver parameters.

Ratio of Ionization Coefficients	k_{eff}	0.01
Detector Gain	G	150
Excess Noise Factor	F	3.5
Detector Dark Current		
Gain Dependent	i_b	10^{-10}A
Gain Independent	i_s	10^{-8}A
Noise Equivalent Temperature	T_{eq}	400°K
Load Resistance	R_L	2000 Ω

7.1.6 Design Summary

The design specifications for the ISL are summarized in Table 7.5. Note that the optimal antenna gains can be calculated from the tabulated value of γ_T shown in Table 7.3. The background count rate at the direct detection receiver can be calculated from Eq. (2.35) to be

$$\lambda_B \approx \begin{cases} 5\times10^8/s & \text{cloud cover} \\ 5\times10^7/s & \text{no cloud cover} \end{cases} . \tag{7.1}$$

For the analysis of ISL performance, it will be assumed that RMS pointing and tracking jitters of 1 μrad can be achieved and that the average background count rate is $5\times10^8/s$.

Given the receiver parameters and the background photocount rate, the receiver sensitivity, which is defined as the received power needed to achieve a given PBE, can be calculated. Table 7.6 summarizes the receiver sensitivity at different values of the PBE and coding orders at a data rate of 200 Mbps.

7.2 Performance Evaluation

Given the link specifications, it is of interest to determine whether the mission requirements can be met within the design constraints. Furthermore, it is desirable to investigate the various trade-offs among the system performance, complexity, and component availability.

7.2.1 Performance of the Direct Detection PPM System

The link budgets for a 4-ary PPM link at 10^{-6} and 10^{-9} error rates are shown in Tables 7.7-7.8. These link budgets are calculated assuming that the RMS pointing error is equal to 1 μrad. Note that, by using a 200mW transmitter laser and a QPPM encoding scheme, a 200 Mbps channel can be implemented with 6.2 dB and 3.3 dB power margins, respectively. These link budgets show that the transmitter pointing error dominates the power allocation by limiting the transmitter antenna gain and imposing an additional power penalty. In contrast, the power penalty due to receiver synchronization error is relatively small (0.1 dB).

The maximum data rate, R_{max}, that the link can achieve is plotted in Fig. 7.1 versus the transmitter power P_L for a QPPM system with RMS transmitter pointing error of 1 μrad. A 3 dB power margin was included in the calculation of data rate. Note that at high laser powers, R_{max} increases linearly with increasing signal power. At high signal powers, the receiver SNR is determined by the number of signal photons

Table 7.5 ISL design summary.

	Direct Detection QPPM	Heterodyne Q-NCFSK
Modulation Format	4-ary PPM	4-ary FSK
Transmitter Laser		
Power	200 mW	25 mW
Extinction Ratio	5%	-
Modulation Bandwidth	>15 GHz	>15 GHz
Linewidth	-	3 MHz
Optical Antenna		
Transmitter Aperture Diameter	TBD†	TBD†
Receiver Aperture Diameter	30 cm	TBD†
Transmitter and Receiver Optics		
Transmitter Efficiency	50%	50%
Receiver Efficiency	35%	35%
Tracking Split	5%	5%
Receiver Optical Bandwidth	2 nm	-
Receiver Field-of-View	0.1 mrad	-
Optical Detector		
Quantum Efficiency	70%	70%
Gain	150	1
Excess Noise Factor	3.0	-
Gain Dependent Dark Current	10^{-10} A	-
Gain Independent Dark Current	10^{-8} A	-
Bandwidth	>10 GHz	>10 GHz
Noise Equivalent Temperature	400 °K	-
Load Resistance	2 KΩ	-

† To be determined by the PBE and RMS pointing/tracking errors.

Table 7.6 Receiver sensitivity at 200 Mbps.

M	PBE	Direct Detection PPM†		Heterodyne NCFSK	
		Number of Photons/Bit	Sensitivity	Number of Photons/Bit	Sensitivity
2	10^{-3}	58	-54.1 dBm	12	-61.0 dBm
	10^{-6}	111	-51.3 dBm	26	-57.6 dBm
	10^{-9}	162	-49.7 dBm	40	-55.7 dBm
4	10^{-3}	35	-56.4 dBm	7	-63.4 dBm
	10^{-6}	63	-53.8 dBm	14	-60.4 dBm
	10^{-9}	91	-52.2 dBm	21	-58.6 dBm
8	10^{-3}	29	-57.2 dBm	5	-64.7 dBm
	10^{-6}	50	-54.8 dBm	10	-61.9 dBm
	10^{-9}	71	-53.2 dBm	14	-60.4 dBm

† Derived with receiver parameters shown in Table 7.5.

Table 7.7 Link budget of a direct detection QPPM system at PBE=10^{-6}.

Modulation Format	4-PPM	
Data Rate	200 Mbps	
Laser Power	200 mW	23.0 dBm
Path Loss	40,000 km	-295.4 dB
Antenna Gain		
Transmitter†	12.6 cm	113.4 dB
Receiver	30 cm	120.9 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-6}	63 Photons/bit	-53.8 dBm
Pointing Loss		-1.8 dB
Temporal Tracking Loss		-0.1 dB
Link Margin		6.2 dB

† Optimum for σ_{θ_T} = 1 μrad.

Table 7.8 Link budget of a direct detection QPPM system at PBE=10^{-9}.

Modulation Format	4-PPM	
Data Rate	200 Mbps	
Laser Power	200 mW	23.0 dBm
Path Loss	40,000 km	-295.4 dB
Antenna Gain		
Transmitter†	10.6 cm	111.9 dB
Receiver	30 cm	120.9 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-9}	91 Photons/bit	-52.2 dBm
Pointing Loss		-1.6 dB
Temporal Tracking Loss		-0.1 dB
Link Margin		3.3 dB

† Optimum for σ_{θ_T} = 1 μrad.

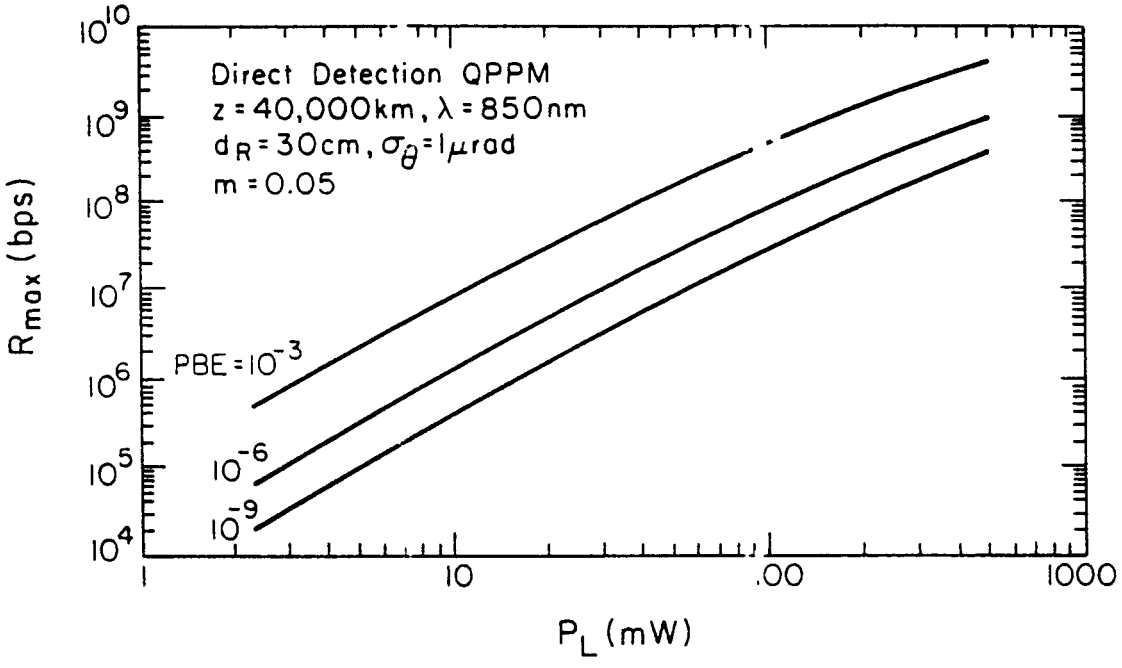


Fig.7.1. Maximum data rates R_{max} versus the transmitter power for a 4-ary PPM system with an RMS pointing error of 1 μrad. The data rates are calculated with a 3dB power margin.

collected at the receiver. Consequently, by fixing the SNR and hence the PBE, the maximum data rate increases linearly with increasing signal power. As the signal power decreases, the effect of background radiation becomes more prominent and the maximum data rate no longer depends linearly on the transmitter power.

The performance of the link can be improved by reducing the RMS pointing error at the transmitter. Since the optimal transmitter antenna gain is inversely related to the mean square pointing jitter, a 6 dB improvement in power margin can be realized if the RMS pointing jitter is reduced by a factor of two at the transmitter. However, higher cost and complexity of the optical system are needed to achieve a small RMS pointing jitter. In practice, RMS pointing jitter less than 0.2 μrad is difficult to achieve. Performance of the PPM system can also be improved by using higher order encoding schemes. In general, for a block-encoded system, $k=\log_2(M)$ bits can be transmitted for every symbol sent across the optical channel. Therefore, by increasing the order of PPM, a higher data rate can be achieved with a given signal power. However, at higher order PPMs, the system bandwidth and complexity increase rapidly with increasing PPM order. Furthermore, in the presence of a finite modulation extinction ratio, the maximum receiver SNR is limited by the residual signal in the background slots. Consequently, one cannot increase the PPM order indefinitely. Figure 7.2 is a plot of the R_{max} versus PPM order for an ISL with 200mW transmitter laser power at different values of the modulation extinction ratio m and the PBE. Note that for $0.01 \le m \le 0.1$, the performance improvement is negligible for $M > 10$.

7.2.2 Performance of the Heterodyne NCFSK System

Tables 7.9-7.10 display the link budget of a 4-ary heterodyne NCFSK channel with 1 μrad RMS pointing and tracking errors. A comparison of these link budgets with that of the direct detection PPM system shown in Tables 7.7-7.8 shows that the heterodyne 4-FSK system has an approximately 5-10 dB advantage in the detection sensitivity over the direct detection PPM system. However, this advantage is quickly offset by the smaller transmitter power, smaller receiver antenna gain, larger power penalty due to spatial tracking error, and power penalty due to carrier phase noise (cf. Table 4.4). It should be noted that the link budgets shown in Tables 7.9-7.10 do not take into account the excess intensity noise of the transmitter and the LO. At a data rate of 1 Gb/s using 4-ary FSK, the power penalty due to the intensity noise is approximately 1.5 dB [148].

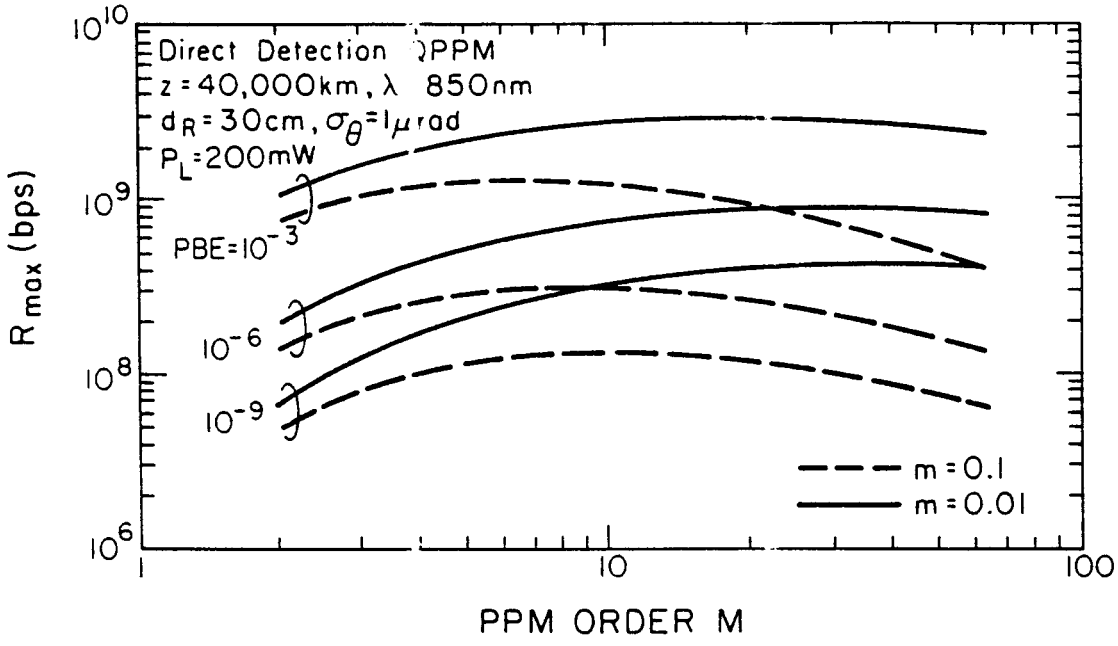


Fig.7.2. Maximum data rates R_{max} versus the PPM order for a direct detection PPM system at different values of the modulation extinction ratio and PBE. The transmitter power is assumed to be 200mW, and the data rates are calculated with a 3dB power margin.

Table 7.9 Link Budget of a heterodyne QFSK system at PBE=10^{-6}.

Modulation Format	4-FSK	
Data Rate	200 Mbps	
Laser Power	30 mW	14.8 dBm
Path Loss	40,000 km	-295.4 dB
Antenna Gain		
Transmitter†	14.1 cm	114.3 dB
Receiver†	14.1 cm	114.3 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-6}	14 Photons/bit	-60.4 dBm
Pointing and Tracking Loss		-4.0 dB
Temporal Tracking Loss		-0.1 dB
Phase Noise @ PBE=10^{-6} $\Delta_f T_s$=0.03, and $\nu_d T_s$=1		-0.8 dB
Link Margin		-4.1 dB

† Optimum for $\sigma_{\theta_T} = \sigma_{\theta_R} = 1$ μrad.

Table 7.10 Link Budget of a heterodyne QFSK system at PBE=10^{-9}.

Modulation Format	4-FSK	
Data Rate	200 Mbps	
Laser Power	30 mW	14.8 dBm
Path Loss	40,000 km	-295.4 dB
Antenna Gain		
Transmitter†	11.7 cm	112.7 dB
Receiver†	11.7 cm	112.7 dB
Optics Efficiency		
Transmitter	0.5	-3.0 dB
Receiver	0.35	-4.6 dB
Receiver Sensitivity @ 200Mbps and PBE=10^{-9}	21 Photons/bit	-58.6 dBm
Pointing and Tracking Loss		-3.8 dB
Temporal Tracking Loss		-0.1 dB
Phase Noise @ PBE=10^{-9} $\Delta_f T_s$=0.03, and $\nu_d T_s$=1		-3.0 dB
Link Margin		-11.1 dB

† Optimum for $\sigma_{\theta_T} = \sigma_{\theta_R} = 1$ μrad.

C-3

The fact that the heterodyne NCFSK system shows a negative power margin implies that the desired data rate (200 Mbps) cannot be achieved for links with 1 μrad pointing and tracking errors. In fact, the maximum data rate, R_{max}, that the ISL can support for the given link parameters is plotted in Fig. 7.3 versus the transmitter power. With a 3 dB link margin, the transmitter power needed to achieve 10^{-6} error rate is on the order of 100 mW. The performance can be improved by reducing the RMS pointing and tracking jitters. Since the required transmitter power is proportional to the product of mean square tracking and pointing jitters, a 12 dB improvement in power margin can be realized if the RMS pointing and tracking jitters can be reduced from 1 μrad to 0.5 μrad. In this case, the 200Mbps QFSK channel can be implemented with 0.9 and 7.9 dB of power margin at $PBE=10^{-9}$ and 10^{-6}, respectively. Performance of the heterodyne channel can also be improved by using higher order FSKs. However, as was the case with the direct detection PPM system, the transmitter and receiver bandwidths increase rapidly with increasing FSK order. Consequently, it is not possible to increase the encoding order indefinitely.

7.2.3 The ACTS-Shuttle Link

The link parameters for the proposed Advanced Communication Technology Satellite (ACTS)-shuttle link were supplied by NASA-Goddard and displayed in Table 7.11. Note that these parameters differ from those shown in Table 7.5 in that (a) a smaller laser power (70 mW average, 140 mW peak) is assumed for the direct detection link, (b) a static pointing loss (1-2 dB) is included in the link budget and, most significantly, (c) much smaller RMS pointing and tracking jitters are assumed. For such a small (0.21 μrad RMS) pointing jitter, the optimal antenna diameter is in excess of 50 cm. Consequently, the ACTS-shuttle link is designed using suboptimal antenna gains. Tables 7.12(a),(b) summarize the link budget for the direct detection PPM and heterodyne QFSK links calculated using the parameters shown in Table 7.11. Note that because the suboptimal design is chosen, the power penalty due to pointing and tracking errors is small (≈0.2-0.4 dB). As a result, the performance of the heterodyne QFSK system is superior to that of the direct detection PPM system. Furthermore, since the peak transmitter power is equal to twice the average power, the receiver SNR is optimized at $M=2$. Consequently, the binary PPM system has a higher power margin compared to the 4-ary PPM system.

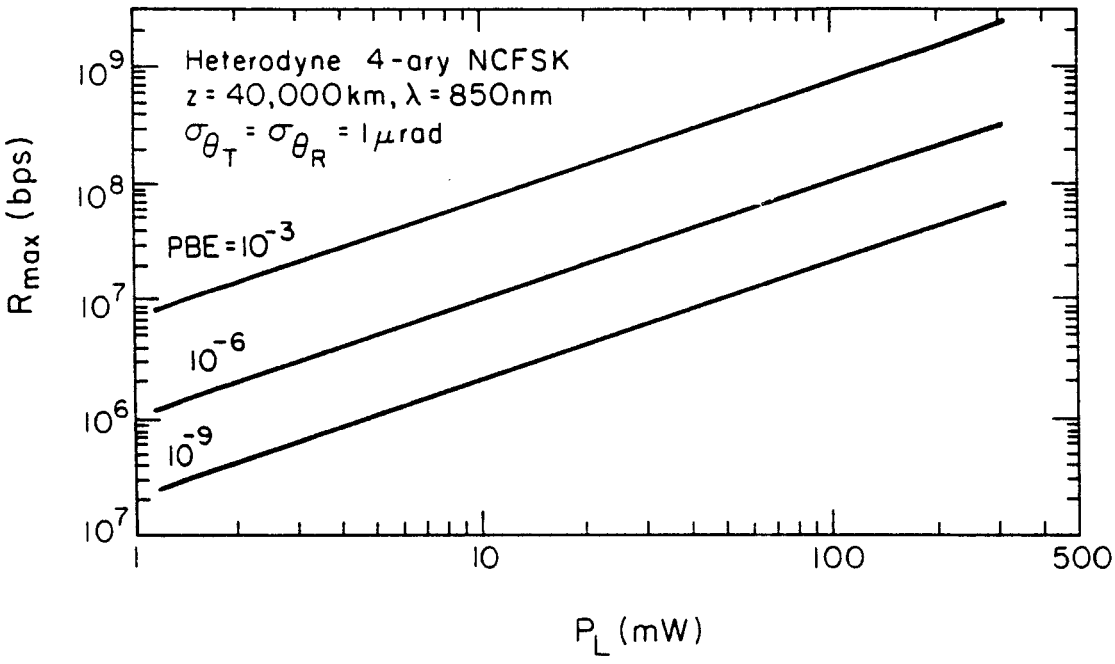


Fig.7.3. Maximum data rates R_{max} versus the transmitter power for a QFSK system with RMS pointing and tracking errors of 1 μrad. The data rates are calculated with a 3dB power margin.

Table 7.11 ACTS-shuttle link parameters.

	Direct Detection QPPM	Heterodyne Q-NCFSK
Data Rate	220 Mbps	220 Mbps
Modulation Format	QPPM, BPPM	QFSK
Transmitter Laser		
Power	70 mW average 140 mW peak	30 mW
Extinction Ratio	5%	-
Modulation Bandwidth	>15 GHz	>15 GHz
Linewidth	-	3 MHz
Optical Antenna Gain†		
Transmitter	113.7 dB	113.7 dB
Receiver‡	115.2 dB	113.7 dB
Transmitter and Receiver Optics		
Tracking Split	5%	5%
Receiver Optical Bandwidth	2 nm	-
Receiver Field-of-View	0.1 mrad	-
Optical Detector		
Quantum Efficiency	70%	70%
k_{eff}	0.007	-
Gain	250	1
Excess Noise Factor	3.75	-
Gain Dependent Dark Current	10^{-10} A	-
Gain Independent Dark Current	10^{-8} A	-
Bandwidth	>10 GHz	>10 GHz
Noise Equivalent Temperature	400 °K	-
Load Resistance	2 KΩ	-

† Includes optics losses.
‡ Assumes a 20 cm diamter aperture with 60% through-put.

Table 7.12(a) ACTS-shuttle link budget summary.

	Direct Detection QPPM	Heterodyne QFSK
Data Rate	220 Mbps	220 Mbps
Laser Power	15.4 dBm	14.8 dBm
Path Loss	-295.4 dB	-295.4 dB
Antenna Gain		
Transmitter†	113.7 dB	113.7 dB
Receiver†	115.2 dB	113.7 dB
Receiver Sensitivity	-53.0 dBm	-59.6 dBm
Pointing and Tracking Losses		
Static Tracking Error	-1.0 dB	-2.0 dB
Random Tracking Error	-0.2 dB	-0.4 dB
Temporal Tracking Loss	-0.1 dB	-0.1 dB
Phase Noise†	-	-0.8 dB
Link Margin	0.6 dB	3.1 dB

† Calculated at PBE=10^{-6}, $\nu_d T_s$=1, and $\Delta_f T_s$=0.03.

Table 7.12(b) ACTS-shuttle link budget summary.

	Direct Detection BPPM	Heterodyne QFSK
Data Rate	220 Mbps	220 Mbps
Laser Power	18.4 dBm	14.8 dBm
Path Loss	-295.4 dB	-295.4 dB
Antenna Gain		
Transmitter†	113.7 dB	113.7 dB
Receiver†	115.2 dB	113.7 dB
Receiver Sensitivity	-50.6 dBm	-59.6 dBm
Pointing and Tracking Losses		
Static Tracking Error	-1.0 dB	-2.0 dB
Random Tracking Error	-0.2 dB	-0.4 dB
Temporal Tracking Loss	-0.1 dB	-0.1 dB
Phase Noise†	-	-0.8 dB
Link Margin	1.2 dB	3.1 dB

† Calculated at PBE=10^{-6}, $\nu_d T_s$=1, and $\Delta_f T_s$=0.03.

7.3 Conclusions

The performance of the direct detection PPM ISL was evaluated and compared to that of the heterodyne NCFSK system. For systems with large (>0.5 μrad) RMS pointing jitters, the spatial tracking error dominates the link budget by limiting the maximum antenna gains and incurring additional losses. For systems with small pointing jitters, on the other hand, the antenna gains are usually determined by the cost and size constraints rather than performance consideration. When the antenna diameters are chosen to be much smaller than the optimal values, much smaller power penalties can be allocated for pointing and tracking errors. For the heterodyne system, power penalty due to the carrier phase noise is another major loss factor that must be carefully considered. For a 200 Mbps QFSK system with 100 MHz tone spacing and 3 MHz linewidth, the power penalty is approximately 0.8 dB at PBE = 10^{-6}. The power penalty due to carrier phase noise can be minimized by increasing the FSK tone spacing and reducing the IF linewidth. In contrast, the receiver synchronization error has a relatively minor effect (0.1 dB) on the link budget.

The performances of the direct and heterodyne detection ISLs can be compared by inspecting the link budgets shown in Tables 7.7-7.12. Note that, even though the heterodyne channel offers a better detection sensitivity, it is more sensitive to the spatial tracking error and the carrier phase noise. Therefore, for systems with large (≈1 μrad) pointing and tracking errors, the direct detection PPM system has a superior performance. On the other hand, for systems with small (<0.3 μrad) pointing and tracking jitters, such as in an ACTS-shuttle link, the heterodyne FSK system shows a larger power margin because of the higher sensitivity of the heterodyne receiver.

Finally, it should be noted that the comparison above did not take into account the possibility of performance improvement using error control codes (ECCs) [163]-[164]. By using a proper error control code, the desired PBE performance ($<10^{-6}$) can be achieved for systems with raw PBE on the order of 10^{-3}. However, since the ECC is expected to affect both direct detection PPM and the heterodyne FSK systems similarly, the conclusions derived from this study should not be affected significantly with the inclusion of ECCs.

APPENDIX A

APPROXIMATION OF THE DPSK ERROR RATE

An approximation to the DPSK error rate can be made by considering the alternative receiver structure shown in Figure A.1 [16]. For a given bit period, the phase of the received signal $r(t)$ is estimated and the output of the phase estimator is then differentially detected. When only white Gaussian noise is present at the estimator input, the error in estimating the phase angle can be shown to have the following distribution:

$$P(\Phi_n) = \frac{1}{2\pi} e^{-\gamma} + \sqrt{\gamma/\pi} e^{-\gamma \sin^2 \Phi_n} \cos\Phi_n \, [1 - \mathrm{erfc}(\sqrt{\gamma}\cos\Phi_n)] \;, -\pi \le \Phi_n \le \pi \;. \tag{A.1}$$

For a DPSK receiver, the output of the phase estimator during the k^{th} bit interval $\Phi(k)$ is delayed by one bit period and subtracted from the next estimator output. Since the additive noise over the two bit periods are uncorrelated, the resulted estimator error will be independent. Consequently, the probability distribution of the correlator output will be the self-convolution of Eq. (A.1). This calculation is in general very difficult. At sufficiently high signal-to-noise ratio, however, the error in estimating the phase angle is expected to be small. In which case $\sin\Phi_n \approx \Phi_n$, $\cos\Phi_n \approx 1$, $\mathrm{erfc}(\sqrt{\gamma}\cos\Phi_n) \approx 1$, and the probability distribution can be approximated by the sum of a constant and a Gaussian peak:

$$p(\Phi_n) \approx \frac{1}{2\pi} e^{-\gamma} + \sqrt{\gamma/\pi} e^{-\gamma\Phi^2} \;. \tag{A.2}$$

Using this approximation, the convolution can be calculated to be

$$p(\Psi) = \frac{1}{2\pi} e^{-\gamma} + \sqrt{\gamma/2\pi} e^{-\gamma\Psi^2/2} \;, \; -\pi \le \Psi_n \le \pi \tag{A.3}$$

where $\Psi(k) = \Phi_n(k) - \Phi_n(k-1)$ is the phase error at the correlator output. A decoding error occurs when the magnitude of Ψ exceeds $pi/2$. Using the probability distribution given by Eq. (A.3), the error probability can be written as

$$P_E = 2 \times \int_{\pi/2}^{\pi} p(\Psi) d\Psi \approx \frac{1}{2} e^{-\gamma} + \mathrm{erfc}(\sqrt{\pi^2\gamma/8}) \;. \tag{A.4}$$

At sufficiently high SNR, the difference between the PBE evaluated using Eq. (A.4) and the exact PBE is negligible.

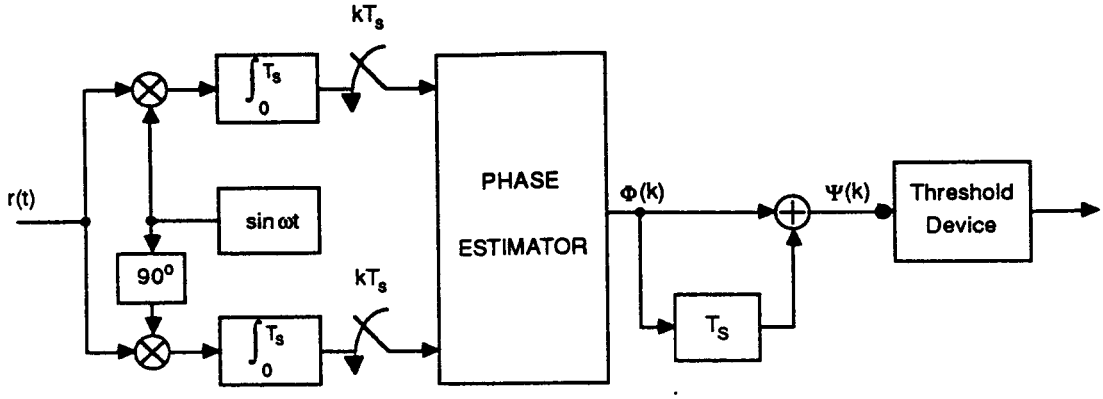


Fig.A1. Alternative implementation of the DPSK receiver.

When phase noise is present at the IF signal, the analysis becomes more complicated. Given a sample path of the carrier phase noise $\phi(t)$, the output of the integrator prior to the phase estimator can be written as

$$X(k) = \frac{1}{T_s}\int_0^{T_s} A\cos\phi(t)dt + N_c = A\cos\phi(\xi_k) + N_c \ ,$$
$$Y(k) = \frac{1}{T_s}\int_0^{T_s} A\sin\phi(t)dt + N_s = A\sin\phi(\xi'_k) + N_s \ , \tag{A.5}$$

where N_c and N_s are the in-phase and quadrature components of the shot noise, respectively, and ξ_k, ξ'_k are some values between $(kT_s,(k+1)T_s)$. In general $\xi_k \neq \xi'_k$. However, at small linewidth the carrier phase varies slowly so that $\phi(\xi_k)$ is approximately equal to $\phi(\xi'_k)$. In this case the output of the phase estimator can be written as

$$\Phi(k) = \Phi_n(k) + \Phi_p(k) \tag{A.6}$$

where $\Phi_p(k)=\phi(\xi_k)$ is the phase noise contribution to the phase estimator error. When the IF linewidth is small compared to the data rate, the phase noise angle $\Phi_p(k)$ can be given to a good approximation by the average carrier phase:

$$\Phi_p(k) \simeq \frac{1}{T_s}\int_{kT_s}^{(k+1)T_s} \phi(\tau)d\tau \ . \tag{A.7}$$

With the contribution of carrier phase noise to the phase estimator, the output of the phase correlator can be written as

$$\begin{aligned}\Psi(k) &= \Phi(k) - \Phi(k-1) = \Phi_n(k) + \Phi_p(k) - \Phi_n(k-1) - \Phi_p(k-1) \\ &= \Phi_n(k) - \Phi_n(k-1) + \frac{1}{T_s}\int_{kT_s}^{(k+1)T_s} \phi(\tau) - \phi(\tau - T_s)d\tau \ .\end{aligned} \tag{A.8}$$

Since the instantaneous distribution of Φ_p is Gaussian, the phase noise contribution to the correlator output is Gaussian with zero mean and variance

$$Var(\Phi_p(k) - \Phi_p(k-1)) = \frac{4\pi}{3}\Delta_f T_s \ . \tag{A.9}$$

Consequently, the distribution of $\Psi(k)$ can be given by the convolution of Eq. (A.3) with a Gaussian peak of variance $4\pi\Delta_f T_s/3$.

$$p(\Psi) = \frac{1}{2\pi} e^{-\gamma} + \frac{1}{\sqrt{2\pi}\sigma_\Psi} e^{-\Psi^2/2\sigma_\Psi^2} \tag{A.10}$$

where

$$\sigma_\Psi^2 = \frac{1}{\gamma} + \frac{4\pi}{3} \Delta_f T_s \ . \tag{A.11}$$

The probability of error, which is simply the probability that the magnitude of Ψ exceeds $\pi/2$, can therefore be approximated by

$$P_E \simeq \frac{1}{2} e^{-\gamma} + \frac{1}{2} \mathrm{erfc}(\pi/\sqrt{8}\sigma_\Phi) \ . \tag{A.12}$$

APPENDIX B

JOINT EXPECTATION OF THE PHOTODETECTOR OUTPUT

The joint characteristics function of the photodetector output $r(t)$ is defined as [57]

$$\Phi(\omega_1,\omega_2) = E\left[e^{-j[\omega_1 r(t)+\omega_2 r(\tau)]} \right] \tag{B.1}$$

where $j=\sqrt{-1}$, and $r(t)$ can be modeled as a filtered Poisson process [57]

$$r(t) = E\left[\sum_{\tau_j} G_j h(t-\tau_j) \right] \quad . \tag{B.2}$$

Note that by associating a random gain G_j with the j^{th} photon event, this particular model is suitable for modeling the output of the avalanche photodiode (APD) as well as the *pin* diode ($G \equiv 1$).

The joint characteristics function can be evaluated by substituting Eq. (B.2) into (B.1) and taking the expectation with respect to the photon arrival times t_k. The arrival time of the k^{th} photon can be shown to have the probability distribution

$$p_{t_k}(t_k) = \lambda(t_k)\left[\int_{t_0}^{T} \lambda(\xi) d\xi \right]^{-1} \tag{B.3}$$

where (t_0,T) is the observation period. The expectation can be evaluated by first condition on the number of photons detected N_T, then taking the expectation with respect to the random variable N_T. By substituting Eqs. (B.2) and (B.3) into (B.1) and using the fact that $\{t_k\}$ are independent and identically distributed, the conditional expectation can be written as

$$\begin{aligned} E\left[e^{-j[\omega_1 r(t)+\omega_2 r(\tau)]} \,|\, N_T=k \right] &= E_G\left[\int dt_1 dt_2 \cdots dt_k \prod_{v=1}^{k} \left\{ e^{-jG_v[\omega_1 h(t-\tau_v)+\omega_2 h(\tau-\tau_v)]} p_{t_v}(t_v) \right\} \right] \\ &= \prod_{v=1}^{k} \left\{ \int dt_v E_G\left[e^{-jG_v[\omega_1 h(t-\tau_v)+\omega_2 h(\tau-\tau_v)]} \right] \lambda(t_v) \right\} \cdot \left[\int_0^T \lambda(\xi) d\xi \right]^{-k} \\ &= \left\{ \int d\xi \Phi_G[\omega_1 h(t-\tau_v)+\omega_2 h(\tau-\tau_v)] \lambda(\xi) / \int_0^T \lambda(\xi) d\xi \right\}^k \end{aligned} \tag{B.4}$$

where Φ_G is the characteristics function of the detector gain G. In deriving Eq. (B.4), we have assumed that the detector gains $\{G_j\}$ are independent and identically distributed. The conditioning on the number of photons N_T can be removed by taking the expectation with respect to the random variable k. Using the fact

that N_T is Poisson distributed, the joint characteristics function of the photodetector output can be written as

$$\Phi(\omega_1,\omega_2) = \exp\left[\int_0^T \lambda(\xi)\left\{\Phi_G[\omega_1 h(t-\xi)+\omega_2 h(\tau-\xi)]-1\right\}d\xi\right] \quad . \tag{B.5}$$

Given the joint characteristics function in Eq. (B.5), the joint expectation of the photodetector output can be evaluated by differentiating Eq. (B.5) with respect to its arguments. Some of the joint moments are given by

$$E[r(t)r(\tau)] = <G^2>\int\lambda(\xi)h(t-\xi)h(\tau-\xi)d\xi + G^2\left[\int\lambda(\xi)h(t-\xi)d\xi\right]\left[\lambda(\zeta)h(t-\zeta)d\zeta\right] \tag{B.6}$$

$$\begin{aligned}
E[r^2(t)r^2(\tau)] = {} & <G^4>\int\lambda(\xi)h^2(t-\xi)h^2(\tau-\xi)d\xi \\
& + 2<G^3>G\left[\int\lambda(\xi)h^2(t-\xi)h(\tau-\xi)d\xi\right]\left[\int\lambda(\eta)h(\tau-\eta)d\eta\right] \\
& + 2<G^3>G\left[\int\lambda(\xi)h(t-\xi)d\xi\right]\left[\int\lambda(\eta)h^2(\tau-\eta)h(t-\eta)d\eta\right] \\
& + 2<G^2><G^2>\left[\int\lambda(\xi)h(t-\xi)h(\tau-\xi)d\xi\right]^2 \\
& + <G^2><G^2>\left[\int\lambda(\xi)h(t-\xi)d\xi\right]\left[\int\lambda(\eta)h(\tau-\eta)d\eta\right] \\
& + <G^2>G^2\left[\int\lambda(\xi)h(t-\xi)d\xi\right]^2\left[\int\lambda(\eta)h^2(\tau-\eta)d\eta\right] \\
& + <G^2>G^2\left[\int\lambda(\xi)h^2(t-\xi)d\xi\right]\left\{\int\lambda(\eta)h(\tau-\eta)d\eta\right]^2 \\
& + 4<G^2>G^2\left[\int\lambda(\xi)h(t-\xi)d\xi\right]\left[\int\lambda(\eta)h(t-\eta)h(\tau-\eta)d\eta\right]\left[\int\lambda(\zeta)h(\tau-\zeta)d\zeta\right] \\
& + G^4\left[\int\lambda(\xi)h(t-\xi)d\xi\right]^2\left[\int\lambda(\eta)h(\tau-\eta)d\eta\right]^2 \quad .
\end{aligned} \tag{B.7}$$

APPENDIX C

POWER SPECTRUM OF THE PHOTODETECTOR OUTPUT

C.1 Power Spectrum of the APD output

For a nonstationary process such as the photodetector output, the power spectrum can be defined as (Eq. (6.10))

$$S_r(\omega) = \lim_{T\to\infty} \frac{1}{2T} E_\lambda \left[\int_{-T}^{T}\int_{-T}^{T} E\,[r(t)r(\tau)\,|\,\lambda(t)] e^{-i\omega(t-\tau)} dt d\tau \right] \quad . \tag{C.1}$$

The conditional expectation in Eq. (C.1) is taken with respect to the shot noise process $r(t)$, conditioned on the detector photocount rate $\lambda(t)$. This expectation can be taken by differentiating the joint characteristics function $\Phi(\omega_1,\omega_2)$, and the result is given by Eq. (B.6). By substituting Eq. (B.6) into Eq. (C.1) and taking the Fourier transform, the power spectrum of the photodetector output can be written as

$$S_r(\omega) = \lim_{T\to\infty} \frac{1}{2T} G^2 H(\omega) E\,[F\,\Lambda_T(0) + |\,\Lambda_T(\omega)\,|^2] \tag{C.2}$$

where $H(\omega)$ is the Fourier transform of the detector impulse response $h(t)$, G, F are the detector gain and the excess noise, and

$$\Lambda_T(\omega) = \int_{-T}^{T} \lambda(t) e^{-i\omega t} dt \tag{C.3}$$

is the Fourier transform of $\lambda(t)$. For the PPM signaling scheme, $\Lambda_T(\omega)$ is given by

$$\Lambda_T(\omega) = 2T\left(\frac{\sin\omega T}{\omega T}\right)\lambda_B + \lambda_S P(\omega) \sum_k e^{-i\omega(kT_w + d_k T_s)} \tag{C.4}$$

where $P(\omega)$ is the Fourier transform of the pulse shape $p(t)$, and the summation is over the codewords in the interval $\{-T, T\}$. By using the summation rule

$$\sum_{i,j} = \sum_{i=j} + \sum_{i\neq j} \quad , \tag{C.5}$$

and the fact that $\{d_k\}$ are independent for nonoverlapping codewords, the expectation in Eq. (C.2) can be evaluated. Note that by using the fact

$$\lim_{T\to\infty} 2T\left(\frac{\sin\omega T}{\omega T}\right) = 2\pi\delta(\omega) \quad , \tag{C.6}$$

the dependence in the boundary T can be removed, and the resulting power spectrum is given by

$$S_r(\omega) = G^2|H(\omega)|^2\left[F\left(\lambda_B+\frac{\lambda_S}{M}\right) +2\pi\delta(\omega)\left(\lambda_B^2+\frac{2\lambda_B\lambda_S}{M}\right)\right.$$

$$+\frac{\lambda_S^2}{T_w}|P(\omega)|^2(1-|R(\omega)|^2) \qquad \text{(C.7)}$$

$$\left.+\frac{2\pi}{T_s}\lambda_S^2|P(\omega)|^2\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{2\pi k}{T_s}\right)\right] ,$$

where

$$R(\omega)=E[\exp(-jC_k\omega T_s)] = \frac{1}{M}\sum_{k=0}^{M-1}e^{-j\omega kT_s} . \qquad \text{(C.8)}$$

C.2 Power Spectrum of the Preprocessor Output

For a square-law nonlinear preprocessor, the power spectrum of the preprocessor output can be calculated by substituting Eq. (B.7) into Eq. (C.1) and taking the expectation with respect to the photocount rate $\lambda(t)$. Note that in the limit of a large photocount, only terms involving λ_3 and λ^4 are significant. The expression for the power spectrum can therefore be simplified by dropping terms corresponding to lower order λ's. By carrying out the Fourier transform on the remaining terms, the power spectrum of the preprocessor output can be written as

$$S_2(\omega) = \lim_{T\to\infty}\frac{1}{T}E\Big[2G^4F\,\Lambda_T(\omega)H_2(\omega)[\Lambda_T(\omega)H(\omega)*\Lambda_T(\omega)H(\omega)]^*$$

$$+ 2G^4F\,\Lambda_T^*(\omega)H_2^*(\omega)[\Lambda_T(\omega)H(\omega)*\Lambda_T(\omega)H(\omega)]$$

$$+ 4G^4F\int\Lambda_T(\omega')\Lambda_T^*(\omega'')\Lambda_T^*(\omega'-\omega'')H(\omega')H(\omega-\omega')H^*(\omega'')H^*(\omega-\omega'')d\omega'd\omega'' \qquad \text{(C.9)}$$

$$+ G^4|\Lambda_T(\omega)H(\omega)*\Lambda_T(\omega)H(\omega)|^2\Big]$$

where

$$H_2(\omega) = \int h^2(t)e^{-i\omega t}dt . \qquad \text{(C.10)}$$

The expectation is taken with respect to the data stream $\{d_k\}$. By substituting Eq. (C.5) into Eq. (C.10) and taking the expectation with respect to the data stream $\{d_k\}$, the power spectrum of the preprocessed signal can be obtained. Note that by taking expectation with respect to the third and fourth powers of λ, it is necessary to

use the sum rules

$$\sum_{k\neq l\neq m} = \sum_{k,l,m} - \sum_{k=l,m} - \sum_{k=m,l} - \sum_{l=m,k} + 2\sum_{k=l=m} \tag{C.11}$$

and

$$\begin{aligned}\sum_{i\neq j\neq k\neq l} = &\sum_{i,j,k,l} - \sum_{i=j,k,l} - \sum_{i=k,j,l} - \sum_{i=l,j,k} - \sum_{j=k,i,l} - \sum_{j=l,i,k} - \sum_{k=l,i,j} + \sum_{i=j,k=l} + \sum_{i=l,j=k} + \sum_{i=k,j=l} \\ &+2\sum_{i=j=k,l} + 2\sum_{i=j=l,k} + 2\sum_{i=l=k,j} + 2\sum_{j=k=l,i} - 6\sum_{i=j=k=l} .\end{aligned} \tag{C.12}$$

The results, after substitution, are very complicated. With some very simple assumptions, however, the expressions can be simplified considerably. First, it will be assumed that the preprocessor filter blocks the dc component of the signal, which is due largely to the uniform background count rate and contains no timing information. Next, note that after taking the expectation with respect to d_k, the expression of the power spectrum contains the integral of the following form:

$$I = \int |R(\omega')G(\omega,\omega')d\omega' \tag{C.13}$$

where $R(\omega)$ is given by Eq. (C.4), and $G(\omega,\omega')$ is some function of $H(\omega)$ and $P(\omega)$. A plot of the function $R(\omega)$ shows that the magnitude of $R(\omega)$ is significant only in the vicinity of $\omega=2\pi k/T_s$. Therefore, if the function $G(\omega)$ is slowly varying over the region where R is appreciable, we may approximate the function under the integral sign by a train of delta functions,

$$|R(\omega)|^2 \simeq \frac{2\pi}{T_s}\sum_k \delta(\omega-2\pi k/T_s) \tag{C.14}$$

where

$$\frac{T_w}{2\pi} = \int_0^{2\pi/T_s} |R(\omega)|^2 d\omega \tag{C.15}$$

is the area under each peak of $R(\omega)$. The width of each peak of $|R(\omega)|^2$ is on the order of $2\pi/MT_s$, where M is the order of the PPM. For most cases of interest, $H(\omega)$ and $P(\omega)$ are slowly varying with respect to $|R(\omega)|^2$ such that this approximation holds.

By using the above assumptions, the power spectrum of the photodetector output can be approximated. The result, after lengthy derivation, is given by

$$S_2(\omega) = \frac{\lambda_s^4 G^4}{T_w^2} \frac{1}{2\pi} |P(\omega)H(\omega) * P(\omega)H(\omega)|^2 \delta(\omega - \omega_s) \qquad \text{(C.16)}$$

$$+ 4\frac{\lambda_s^3 G^4 F}{T_w} \frac{1}{(2\pi)^2} \int P(\omega')P^*(\omega'')P^*(\omega' - \omega'')H(\omega)H(\omega - \omega')H^*(\omega'')H^*(\omega - \omega'')d\omega' d\omega'' \ .$$

LIST OF REFERENCES

[1] Ball Aerospace Systems Division, "Tracking and data acquisition system laser intersatellite communications study," F86-04, Aug. 1986; prepared under contract NAS5-29128 for National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD.

[2] M. Siegel and G. Hoshal, "Planning document for intersatellite link technology and experiment package," Department of Electrical Engineering and Systems Science, Rep. MSU ENGR-82-013, April 1982; prepared under Grant NAS5- 26445 for National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD.

[3] R. M. Gagliardi and S. Karp, *Optical Communications*. New York: Wiley, 1975.

[4] L. L. Jeromin and V. W. S Chan, "Performance of a coded optical communication system using an APD direct detection receiver," IEEE GLOBECOM'83, Nov. 1983.

[5] G. S. Mecherle, "Maximized data rate capability for optical communication using semiconductor devices with pulse position modulation," Ph.D dissertation, University of Southern California, Los Angeles, CA, May 1986.

[6] J. E. Kaufmann and L. L. Jeromin, "Optical heterodyne intersatellite links using semiconductor lasers," IEEE GLOBECOM'84, Atlanta, GA, Nov. 27-29, 1984.

[7] V. W. S. Chan, L. L. Jeromin, and J. E. Kaufmann, "Heterodyne lasercom systems using GaAs lasers for ISL applications," IEEE ICC'83, Boston, MA, June 19-22, 1983.

[8] L. L. Jeromin and V. W. S. Chan, "Modulation design for heterodyne optical communication system," IEEE GLOBECOM'83, San Diego, CA, Nov.28-Dec.1, 1983.

[9] S. D. Personick, "Receiver design for digital fiber optic communication systems. I," *Bell Syst. Tech. J.*, vol. 52, p. 843, July 1973.

[10] G. Keiser, *Optical Fiber Communications*. New York: McGraw-Hill, 1983.

[11] R. Gagliardi and S. Karp, "M-ary Poisson detection and optical communications," *IEEE Trans. Commun. Tech.*, vol. CT-17, no. 2, p. 208, April 1969.

[12] S. Karp, R. Gagliardi and E. O'Neill, "Communication theory for the free space optical channel," *Proc. IEEE*, vol. 58, no. 10, p.1611, Oct. 1970.

[13] G. S. Mecherle, "Impact of laser diode performance on data rate capability of PPM optical communication," IEEE MILCOM'85, Boston, MA, Oct.20-23, 1985.

[14] S. Karp and R. Gagliardi, "The design of PPM optical communication systems," *IEEE Trans. Commun. Tech.*, vol. CT-17, p. 670, Dec. 1969.

[15] J. R. Pierce, "Optical channels: practical limits with photon counting," *IEEE Trans. Commun.*, vol. COM-26, no. 12, pp. 1819-1821, Dec. 1978.

[16] R. J. McEliece, "Practical codes for photon communication," *IEEE Trans. Info. Theory*, vol. IT-27, no. 4, pp. 393-397, July 1981.

[17] W. K. Marshall, "A PPM modulator and demodulator for the 2.5bit/detected photon demonstration," *TDA Progress Report 42-68*, Jet Propulsion Laboratory, Pasadena, CA, Nov. 1981.

[18] J. R. Lesh, J. Katz, H. H. Tan and D. Zwillinger, "2.5 bit/detected photon demonstration program: description, analysis, and phase I results," *TDA Progress Report 42-66*, pp. 115-132, Jet Propulsion Laboratory, Pasadena, CA, Oct. 1981.

[19] J. Katz, "2.5 bit/detected photon demonstration program: phase II and III experimental results," *TDA Progress Report 42-70*, pp. 95-104, Jet Propulsion Laboratory, Pasadena, CA, June 1982.

[20] J. Katz, "High power semiconductor laser for deep space communications," *TDA Progress Report 42-63*, pp. 40-50, Jet Propulsion Laboratory, Pasadena, CA, June 1981.

[21] S. R. Forrest, "Optical Detectors: three contenders," *IEEE Spectrum*, vol. 23, no. 5, p.77, May 1986.

[22] G. E. Stillman and C. M. Wolfe, "Avalanche photodiodes," in *Semiconductors and Semimetals, Vol.12*. R. K. Willardson and A. C. Beer Ed. New York: Academic Press, 1977.

[23] J. A. Abshire, "Performance of OOK and low-order PPM modulations in optical communication when using APD-based receivers," *IEEE Trans. Commun.*, vol. COM-32, p.1140, Oct. 1984.

[24] R. M. Gagliardi and G. Prati, "On Gaussian error probabilities in optical receivers," *IEEE Trans. Commun.*, vol. COM-28, pp. 1742-1747, Sept. 1980.

[25] N. Sorenson and R. Gagliardi, "Performance of optical receivers with avalanche photodetection," *IEEE Trans. Commun.*, vol. COM-27, pp. 1315-1321, Sep.1979.

[26] L. Mandel and E. Wolf, "Optimum conditions for heterodyne detection of light," *Journ. Opt. Soc. Amer.*, vol. 65, no. 4, pp. 413-420, April 1975.

[27] D. Fink, "Coherent detection signal-to-noise," *Appl. Opt.*, vol. 14, no. 3, pp. 689-690, March 1975.

[28] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

[29] A. J. Viterbi, *Principles of Coherent Communications*. New York: McGraw-Hill, 1966.

[30] M. Schwartz, W. R. Bennet and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[31] J. J. Spilker, *Digital Communication by Satellite*. Englewood-Cliffs, NJ: Prentice-Hall, 1977.

[32] Y. Yamamoto, "Receiver performance evaluation of various digital optical modulation-demodulation systems in the 0.5- 10μm wavelength region," *IEEE J. Quantum Electron.*, vol. QE-16, no. 11, pp. 1251-1259, Nov. 1980.

[33] T. Okoshi, "Recent progress in heterodyne/coherent optical- fiber communications," *IEEE J. Lightwave Tech.*, vol. LT-2, no. 4, pp. 341-345, Aug. 1984.

[34] T. Okoshi, "Heterodyne and coherent optical fiber communications: Recent progress," *IEEE Trans. Microwave Tech.*, vol. MTT-30, no. 8, pp. 1138-1148, Aug. 1982.

[35] S. Saito, Y. Yamamoto, and T. Kimura, "Optical heterodyne detection of directly frequency modulated semiconductor laser signals," *Electron. Lett.*, vol. 16, pp. 826-827, Oct. 1980.

[36] L. L. Jeromin and V. W. S. Chan, "M-ary FSK performance for coherent optical communication systems using semiconductor lasers," *IEEE Trans. Commun.*, vol. COM-34, no. 4, pp. 375-381, April 1986.

[37] Y. Yamamoto and T. Kimura, "Coherent optical fiber transmission systems," *IEEE J. Quantum Electron.*, vol. QE-17, no. 6, pp. 919-935, June 1981.

[38] M. W. Fleming and A. Mooradian, "Fundamental line broadening of single-mode GaAlAs diode lasers," *Appl. Phys. Lett.*, vol. 38, p.511, 1981.

[39] B. Daino, P. Spano, M. Tamburrini, and S. Piazzolla, "Phase noise and spectral line shape in semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-19, no. 3, pp. 266-270, March 1983.

[40] Y. Yamamoto, "AM and FM quantum noise in semiconductor lasers, part 1: Theoretical analysis," *IEEE J. Quantum Electron.*, vol. QE-19, no. 1, pp. 34-46, Jan. 1983.

[41] Y. Yamamoto, S. Saito, and T. Mukai, "AM and FM quantum noise in semiconductor lasers, part II: Comparison of theoretical and experimental results for AlGaAs lasers," *IEEE J. Quantum Electron.*, vol. QE-19, no. 1, pp. 47-58, Jan. 1983.

[42] C. H. Henry, "Theory of the linewidth of semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-18, p.259, 1982.

[43] F. G. Walther and J. E. Kaufmann, "Characterization of GaAlAs laser diode frequency noise," *Proceedings of OFC'83*, New Orleans, LA, Feb. 1983.

[44] M. Tamburrini, P. Spano, and S. Piazzolla, "Influence of semiconductor-laser phase noise on coherent optical communication systems," *Opt. Lett.*, vol. 8, no. 3, pp. 174-176, March 1983.

[45] J. Franz, "Evaluation of the probability density function and bit error rate in coherent optical transmission systems including laser phase noise and additive Gaussian noise," *J. Opt. Commun.*, vol. 6, pp. 51-57, 1865.

[46] J. Franz, C. Rapp, and G. Soder, "Influence of baseband filtering on laser phase noise in coherent optical transmission systems," *J. Optical Commun.*, vol. 7, pp. 15-20, 1986.

[47] K. Kikuchi, T. Okoshi, M. Nagamatsu, and N. Henmi, "Degradation of bit-error rate in coherent optical communications due to spectral spread of the transmitter and local oscillator," *IEEE J. Lightwave Tech.*, vol. LT-2, no. 6, pp. 1024-1033, Dec. 1984.

[48] S. Saito, Y. Yamamoto, and T. Kimura, "S/N and error rate evaluation for an optical FSK-heterodyne detection system using semiconductor lasers," *IEEE J. Quantum Electron.*, vol. QE-19, no. 2, pp. 180-

193, Feb. 1983.

[49] F. Favre and D. Le Guen, "Effect of semiconductor laser phase noise on BER performance in an optical DPSK heterodyne-type experiment," *Electron. Lett.*, vol. 18, pp. 964-965, Oct. 1982.

[50] A. C. Van Bochove, J. F. Jacobs, and H. T. Nijnuis, "Influence of (semiconductor) laser linewidth on the error- rate floor in dual-filter optical FSK receivers," *Electron. Lett.*, vol. 21, pp. 280-283, March 1985.

[51] L. G. Kazovsky, "Decision driven phase-locked loop for optical homodyne receivers: performance analysis and laser linewidth requirements," *IEEE J. Lightwave Tech.*, vol. LT-3, no. 6, pp. 1238-1247, Dec. 1985.

[52] I. Garrett and G. Jacobsen, "Theoretical analysis of heterodyne optical receivers for transmission systems using (semiconductor) lasers with nonnegligible linewidth," *IEEE J. Lightwave Tech.*, vol. LT-4, p.323, March 1986.

[53] L. G. Kazovsky, "Performance analysis and laser linewidth requirement for optical PSK heterodyne communication systems," *IEEE J. Lightwave Tech.*, vol. LT-4, no. 4, pp. 415-425, April 1986.

[54] E. A. Swanson and V. W. S. Chan, "Heterodyne spatial tracking systems for optical space communication," *IEEE Trans. Commun.*, vol. COM-34, no. 2, pp. 118-126, Feb. 1986.

[55] P. Van Hove and V. Chan, "Spatial acquisition algorithm and systems for optical ISL," IEEE ICC'83, Boston MA, June 19-22, 1983.

[56] J. D. Barry, K. J. Held, G. S. Mecherle, and A. J. Einhorn, "Spacecraft system study: A study to define the impact of laser communication systems on their host spacecraft," FR84-75-717, Prepared for NASA Goddard Space Flight Center, Greenbelt, MD, April 1984.

[57] K. J. Held, and J. D. Barry, "Precision optical pointing and tracking from spacecraft with vibrational noise," *Proc. SPIE*, vol. 616, pp. 160-173, Jan. 1986.

[58] E. Sein, J. F. Clervoy, M. Lequime,B. Moreau, and J. L. Hibon, "Acquisition and fine-pointing control for a 400-Mbps link between a low-earth orbiter and a geosynchronous satellite," *Proc. SPIE*, vol. 616, pp. 141-159, Jan. 1986.

[59] L. M. Germann, R. Nelson, "Pointing, acquisition, and tracking subsystem for space-based laser communications," *Proc. SPIE*, vol. 616, pp. 118-128, Jan. 1986.

[60] J. D. Barry and G. S. Mecherle, "Beam pointing error as a significant design parameter for satellite-borne free space optical communication systems," *Optical Engineering*, vol. 24, no. 6, pp. 1049-1054, Nov. 1985.

[61] J. D. Barry and G. S. Mecherle, "Communication channel burst errors induced by Gaussian distributed mispointing," *Proc. SPIE*, vol. 616, pp. 137-140, 1986.

[62] P. W. Scott, and P. W. Young, "Impact of temporal fluctuations of signal-to-noise ratio (burst error) on free-space laser communications system design," *Proc. SPIE*, vol. 616, pp. 174-182, Jan. 1986.

[63] G. A. Koepf, R. Peters, and R. G. Marchalek, "Analysis of burst error occurrence on optical intersatellite link (ISL) design," *Proc. SPIE*, vol. 616, pp. 129-136, Jan. 1986.

[64] W. C. Lindsey, *Synchronization Systems in Communication and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1973.

[65] W. K. Marshall, "Detection and symbol synchronization for multiple-bit per photon optical communications," IEEE MILCOM'84, Los Angeles, CA, Oct. 21-28, 1984.

[66] C. C. Chen and C. S. Gardner, "Performance of PLL synchronized optical PPM communication systems," *IEEE Trans. Commun.*, vol. COM-34, Oct. 1986.

[67] U. Mengali and E. Pezzani, "Tracking properties of phase locked loops in optical communication systems," *IEEE Trans. Commun.*, vol. COM-26, pp. 1811-1818, Dec. 1978.

[68] D. L. Lafaw and C. S. Gardner, "Timing performance of phase-locked loops in optical pulse-position modulation communication systems," Electro-Optic Systems Lab., Univ. of Illinois, Urbana, IL, Rep. EOSL 884-001, Aug. 1984.

[69] G. E. Stillman, "Optoelectronics," in *Reference Data for Engineers: Radio, Electronics, Computer and Communications, 7th Ed.* Indianapolis, IN: SAMS, 1985.

[70] S. M. Sze, *Physics of Semiconductor Devices.* New York: Wiley, 1969.

[71] T. P. Pearsall and M. A. Pollack, "Compound semiconductor photodiodes," in *Lightwave Communications Technology, Part D*, (Semiconductors and Semimetals, vol. 22). W. T. Tsang, Ed. Orlando, FL: Academic Press, 1985.

[72] D. P. Schinke, R. G. Smith, and A. R. Hartman, "Photodetectors," in *Semiconductor Devices for Optical Communication* (Topics in Applied Physics, vol. 39). H. Kressel, Ed. Berlin, Heidelberg, New York: Springer-Verlag, 1982.

[73] V. Diadiuk and S. Groves, "Double heterostructure InGaAsP/InP *pin* photodetectors," *Solid State Electron.*, vol. 29, no. 2, p.229, Feb. 1986.

[74] D. L. Snyder, *Random Point Processes.* New York: Wiley, 1975.

[75] S. D. Personick, "Statistics of a general class of avalanche detectors with applications to optical communications," *Bell Syst. Tech. J.*, vol. 50, no. 10, pp. 3075-3095, Dec. 1971.

[76] R. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Trans. Electron Devices*, vol. ED-19, pp. 703-712, June 1972.

[77] J. Conradi, "The distribution of gains in uniformly multiplying avalanche photodiodes: Experimental," *IEEE Trans. Electron Devices*, vol. ED-19, pp. 713-718, June 1972.

[78] J. Muller, "Thin silicon film *p-i-n* photodiodes with internal reflection," *IEEE Trans. Electron Devices*, vol. ED-25, p.247, 1978.

[79] J. C. Campell, "High gain-bandwidth product avalanche photodiodes for multigigabit data rate," OSA Annual Meeting, Seattle, WA, Oct. 19-25, 1986.

[80] T. C. McGill and T. J. Watson, "HgTe-CdTe superlattice infrared detectors," OSA Annual Meeting, Seattle, WA, Oct. 19-25, 1986.

[81] J. H. McElroy, "CO_2 laser communication systems for near-earth space applications," *Proc. IEEE*, vol. 65, no. 2, pp. 221-257, Feb. 1977.

[82] W. Reiland, W. Englisch, and M. Endemann, "Optical intersatellite communication links: State of CO_2 laser technology," *Proc. SPIE*, vol. 616, pp. 69-76, Jan. 1986.

[83] M. M. Ross, "Space optical communications with the Nd:YAG laser," *Proc. IEEE*, vol. 66, no. 3, p.322, March 1978.

[84] M. M. Ross, "YAG laser operation by semiconductor laser pumping," *Proc. IRE*, vol. 56, pp. 196-197, 1968.

[85] D. L. Sipes, "A highly efficient neodymium:yttrium aluminum garnet laser end pumped by a semiconductor laser array," *Appl. Phys. Lett.*, vol. 47(2), pp. 74-76, July 1985.

[86] H. Kressel, M. Ettenberg, J. P. Wittke, and I. Ladany, "Laser diodes and LEDs for fiber optical communication," in *Semiconductor Devices for Optical Communication* (Topics in Applied Physics, vol.39). Berlin, Heidelberg, New York: Springer-Verlag, 1982.

[87] T. Murakami et al., "High-power AlGaAs laser with a thin tapered-thickness active layer," *Electron. Lett.*, vol. 22, pp. 217-218, 1986.

[88] U. J. Ungar, N. Bar-Chaim, and I. Ury, "High power GaAlAs window laser," *Electron. Lett.*, vol. 22, no. 5, p.279, 1986.

[89] W. T. Tsang, R. A. Logan, and M. Ilegems, "High-power fundamental-transverse-mode strip buried heterostructure lasers with linear light-current characteristics," *Appl. Phys. Lett.*, vol. 32, p.311, 1978.

[90] K. Takahashi, K. Ikeda, J. Ohsawa, and W. Susaki, "High efficiency and high power AlGaAs/GaAs lasers," *IEEE J. Quantum Electron.*, vol. QE-19, no. 6, pp. 1002-1008, June 1983.

[91] C. B. Su, et al., "15GHz direct modulation bandwidth of vapor phase regrown 1.3μm InGaAsP buried heterostructure lasers under cw operations at room temperature," *Electron. Lett.*, vol. 21, p.577, 1985.

[92] N. Holonyak, Jr., "Quantum-well semiconductor lasers (review)," *Sov. Phys. Semicond.*, vol. 19(9), pp. 943-958, Sep.1985.

[93] N. Holonyak, Jr., R. M. Kolbas, R. D. Dupuis, and P. D. Dapkus, "Quantum-well heterostructure lasers," *IEEE J. Quantum Electron.*, vol. QE-16, no. 2, pp. 170-185, Feb. 1980.

[94] G. L. Harnagel, P. S. Cross, D. R. Scifres and D. P. Worland, "11W quasi cw monolithic laser diode arrays," *Electron. Lett.*, vol. 21, no. 5, p.231, 1986.

[95] G. L. Harnagel, D. R. Scifres, H. H. Kung, D. F. Welch and P. S. Cross, "Five watt continuous wave AlGaAs laser diodes," *Electron. Lett.*, vol. 22, no. 11, p.605, 1986.

[96] L. Goldberg and J. F. Weller, "Injection-locked operation of a 20-element coupled-stripe laser array," *Electron. Lett.*, vol. 22, no. 12, pp. 858-859, July 1986.

[97] J. E. A. Whiteaway, "Zero-order supermode discrimination in semiconductor laser arrays," *Electron. Lett.*, vol. 22, no. 10, pp. 560-562, May 1986.

[98] J. J. Yang and M. Jansen, "Single-lobed emission from phase-locked array lasers," *Electron. Lett.*, vol. 22, no. 1, pp. 2-4, Jan. 1986.

[99] J. R. Leger, G. J. Swanson, and W. B. Veldkamp, "Coherent beam addition of GaAlAs lasers by binary phase gratings," *Appl. Phys. Lett.*, vol. 48, no. 14, pp. 888-890, April 1986.

[100] K. L. Chen and W. Wang, "Single-lobe symmetric coupled laser arrays," *Electron. Lett*, vol. 21, pp. 347-349, 1985.

[101] K. L. Chen and S. Wang, "Analysis of symmetric Y-junction laser arrays with uniform near-field distribution," *Electron. Lett.*, vol. 22, no. 12, pp. 644-645, June 1986.

[102] G. S. Mecherle, "Laser diode combining for free space optical communication," *Proc. SPIE*, vol. 616, pp. 281-291, Jan. 1986.

[103] D. L. Begley and W. L. Casey, "Selection of laser diode beam combining techniques for free space communication," *Proc. SPIE*, vol. 616, pp. 276-280, Jan. 1986.

[104] K. Nosu, "Multireflection optical multi/demultiplexer using interference filters," *Electron. Lett.*, vol. 15, no. 4, pp. 414-1415, July 1979.

[105] P. O. Minnot and J. B. Abshire, "Grating rhomb diode laser power combiner," *Proceedings of OE/LASE'87*, Los Angeles, CA, Jan. 11-16, 1987.

[106] N. Chinone and m. Nakamura, "Mode-stabilized semiconductor lasers for 0.7-0.8μm and 1.1-1.6μm regions," in *Lightwave Communication Technology, Part C* (Semiconductors and Semimetals, vol. 22), W. T. Tsang, Ed. Orlando, FL: Academic Press, 1985.

[107] H. Soda, Y. Motegi, and K. Iga, "GaInAsP/InP surface emitting injection lasers with short cavity length," *IEEE J. Quantum Electron.*, vol. QE-19, pp. 1035-1041, June 1983.

[108] L. A. Coldren, K. Furuga, B. I. Miller, and J. A. Rentschler, "Etched mirror and groove-coupled GaInAsP/InP laser devices for integrated optics," *IEEE J. Quantum Electron.*, vol. QE-18, pp. 1679-1687, 1982.

[109] W. T. Tsang, "The cleaved-couple-cavity laser," in *Lightwave Communication Technology, Part B* (Semiconductors and Semimetals, vol. 22), W. T. Tsang, Ed. Orlando, FL: Academic Press, 1985.

[110] C. Lin and F. Mengel, "Reduction of frequency chirping and dynamic linewidth in high-speed directly modulated semiconductor lasers by injection locking," *Electron. Lett.*, vol. 20, pp. 1073-1075, Oct. 1984.

[111] R. Wyatt, D. W. Smith, and K. H. Cameron, "Megahertz linewidth from a 1.5μm semiconductor laser with HeNe laser injection," *Electron. Lett.*, vol. 18, p.179, 1982.

[112] H. C. Casey, Jr., S. Somekh, and M. Ilegems, "Room-temperature operation of low-threshold separate-confinement heterostructure injection laser with distributed feedback," *Appl. Phys. Lett.*, vol. 27, p.142, 1975.

[113] Y. Suematsu, K. Kishino, S. Arai, and F. Koyama, "Dynamic single mode semiconductor lasers with a distributed reflector," in *Lightwave Communication Technology, Part B* (Semiconductors and

Semimetals, vol. 22), W. T. Tsang, Ed. Orlando, FL: Academic Press, 1985.

[114] K. Aiki, M. Nakamura, T. Kuroda, J. Umeda, R. Ito, N. Chinone, and M. Maeda, "GaAs-GaAlAs distributed-feedback diode lasers with separate optical and carrier confinement," *Appl. Phys. Lett.*, vol. 27, p. 145, 1975.

[115] F. K. Reinhart, R. A. Logan, and C. V. Shenk, "GaAs-$Al_xGa_{1-x}As$ injection lasers with distributed Bragg reflectors," *Appl. Phys. Lett.*, vol. 27, p. 45, 1975.

[116] P. Meissner, "Investigation of linewidth and sidemode suppression for a DBR laser with integrated passive waveguide," *Electron. Lett.*, vol. 22, no. 6, p. 340, 1986.

[117] A. Chraplyvy et al., "Simple narrow linewidth 1.5μm InGaAsP DFB external cavity laser," *Electron. Lett.*, vol. 22, no. 2, pp. 88-90, 1986.

[118] R. Wyatt and W. J. Devlin, "10 KHz linewidth 1.5μm InGaAsP external cavity laser with 55 nm tuning range," *Electron. Lett.*, vol. 19, p. 110, 1983.

[119] M. R. Mathews, K. H. Cameron, K. H. Wyatt and W. J. Devlin, "Packaged frequency-stable tunable 20KHz linewidth 1.3μm InGaAsP external cavity laser," *Electron. Lett.*, vol. 21, no. 2, pp. 113-115, 1985.

[120] B. J. Klein and J. J. Degnan, "Optical antenna gain 1. Transmitting antennas," *Appl. Opt.*, vol. 13, p.2137, Sept. 1974.

[121] B. J. Klein and J. J. Degnan, "Optical antenna gain. 3: Effect of secondary element support struts on transmitter gain," *Appl. Opt.*, vol. 15, no. 4, p. 977, April 1976.

[122] J. J. Degnan and B. J. Klein, "Optical antenna gain. 2: Receiving antennas," *Appl. Opt.*, vol. 13, no. 10, pp. 2397-2401, Oct. 1974.

[123] G. Abbas and V. Chan, "Optimal design and performance of a dual detector optical heterodyne receiver for local oscillator noise suppression," IEEE GLOBECOM'83, San Diego, CA, Nov. 28-Dec.1, 1983.

[124] W. K. Pratt, *Laser Communications Systems*. New York: Wiley, 1969.

[125] C. S. Gardner and R. L. Gallawa, "Optical communications," in *Reference Data for Engineers: Radio, Electronics, Computer, and Communications, 7th Ed.* Indianapolis, IN: SAMS, 1985.

[126] G. Arnold, P. Russer, and K. Petermann, "Modulation of Laser Diodes," in *Semiconductor Devices for Optical Communication* (Topics in Applied Physics, Vol.39). H. Kressel, Ed. Berlin, Heidelberg, New York: Springer-Verlag, 1982.

[127] D. Hall, A. Yariv, and E. Carmie, "Observation of propagation cutoff and its control in thin optical waveguides," *Appl. Phys. Lett.*, vol. 17, p.127, 1970.

[128] M. Kawabe, S. Hirata, and S. Namba, "Ridge waveguides and electro-optical switches in $LiNbO_3$ fabricated by ion bombardment-enhanced etching," *IEEE Trans. Circuit Syst.*, vol. CAS-26, p.1109, 1979.

[129] I. P. Kaminow, L. W. Stultz, E. H. Turner, "Efficient strip-waveguide modulator," *Appl. Phys. Lett.*, vol. 27, p. 555, 1975.

[130] J. I. Pankove, *Optical Processes in Semiconductors*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[131] F. K. Reinhart, "Electroabsorption in $Al_yGa_{1-y}As$- $Al_xGa_{1-x}As$ double heterostructures," *Appl. Phys. Lett.*, vol. 22, p.372, 1973.

[132] J. C. Campbell, J. C. De Winter, M. A. Pollack, and R. E. Nahory, "Buried heterojunction electroabsorption modulator," OSA Topical Meeting on Integrated and Guided Wave Optics, Salt Lake City, UT, Jan. 16-18, 1978.

[133] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[134] M. Nakamura, K. Aiki, N. Chinone, R. Ito, and J. Umeda, "Longitudinal mode behavior of mode stabilized AlGaAs injection laser," *J. Appl. Phys.*, vol. 49, pp. 4644-4648, Sep.1978.

[135] T. Isobe and M. Tokida, "Power amplification of FM and PM signals with synchronized IMPATT oscillators," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 906-911, Nov. 1970.

[136] C. F. Buhrer, D. Baird, and E. M. Conwell, "Optical frequency shifting by electrooptic effect," *Appl. Phys. Lett.*, vol. 5, pp. 29-31, July 1964.

[137] W. J. Thaler, "Frequency modulation of He-Ne laser beam via ultrasonic waves in quartz," *Appl. Phys. Lett.*, vol. 5, pp. 29-31, July 1964.

[138] F. K. Reinhart and R. A. Logan, "Electrooptic frequency and polarization modulated injection laser," *Appl. Phys. Lett.*, vol. 36, pp. 954-957, June 1980.

[139] J. E. Ripper, G. W. Pratt Jr. and C. G. Whitney, "Direct frequency modulation of a semiconductor laser by ultrasonic waves," *IEEE J. Quantum Electron.*, vol. QE-2, pp. 603-605, Sep.1966.

[140] J. M. Osterwalder and R. J. Rickett, "Frequency modulation of GaAlAs injection laser at microwave frequency rates," *IEEE J. Quantum Electron.*, vol. QE-16, pp. 250-252, March 1980.

[141] D. Welford and S. B. Alexander, "GaAlAs semiconductor diode laser 4-ary frequency shift key modulation at 100 Mb/s," MIT Lincoln Laboratory, Lexington, MA, 1984.

[142] S. Namba, "Electro-optical effect of zincblende," *J. Opt. Soc. Amer.*, vol. 51, p.7 6, 1961.

[143] I. P. Kaminow, J. R. Carruthers, E. H. Turner, and L. W. Stultz, "Thin film $LiNb_3$ electro-optic light modulator," *Appl. Phys. Lett.*, vol. 22, p.240, 1973.

[144] I. P. Kaminow, V. Ramaswamy, R. V. Schmidt, and F. H. Turner, "Lithium niobate ridge waveguide modulator," *Appl. Phys. Lett.*, vol. 24, p.622, 1974.

[145] H. D. Chadwick and J. C. Springett, "The design of a low data rate MFSK communication system," *IEEE Trans. Commun. Tech.*, vol. COM-18, p.740, June 1970.

[146] S. Stein and J. J. Jones, *Modern Communication Principles*. New York: McGraw-Hill, 1967.

[147] S. Yee, V. Chan, and G. Abbas, "Intensity noise statistics of GaAlAs lasers," OFC '83, New Orleans, LA, Feb. 28-March 2, 1983.

[148] L. L. Jeromin and D. Welford, "The effect of spurious intensity modulation in semiconductor laser diode lasers on the performance of optical heterodyne frequency shift-keyed communication systems," *IEEE J. Lightwave Tech.*, vol. LT-4, no. 6, pp. 590-595, June 1986.

[149] S. Saito and Y. Yamamoto, "Direct observation of Lorentzian lineshape of semiconductor laser and lindwidth reduction with external grating feedback," *Electron. Lett.*, vol. 17, pp. 325-327, April 1981.

[150] C. Georghiades and D. L. Snyder, "A proposed receiver structure for optical communication systems that employ heterodyne detection and a semiconductor laser as a local oscillator,," *IEEE Trans. Commun.*, vol. COM-33, no. 4, pp. 382-384, April 1985.

[151] C. N. Georghiades and D. L. Snyder, "Receiver performance for heterodyne optical communication," *IEEE Trans. Commun.*, vol. COM-34, no. 11, pp. 1096-1104, Nov. 1986.

[152] S. Kobayashi and T. Kimura, "Injection locking characteristics of an AlGaAs semiconductor laser," *IEEE J. Quantum Electron.*, vol. QE-16, no. 9, pp. 915-917, Sep. 1980.

[153] B. S. Glance, "Minimum required power for carrier recovery at optical frequencies," *IEEE J. Lightwave Tech.*, vol. LT-4, no. 3, pp. 249-255, March 1986.

[154] D. J. Malyon, D. W. Smith, and R. Wyatt, "Semiconductor laser homodyne optical phase-locked loop," *Electron. Lett.*, vol. 22, no. 8, p.421, Feb. 1986.

[155] H. Philipp, A. Scholtz, E. Bonek, and W. R. Leeb, "Costas loop experiments for a 10.6 μm communication receiver," *IEEE Trans. Commun.*, vol. COM-31, no. 8, pp. 1000- 1002, Aug. 1983.

[156] L. G. Kazovsky, "Balanced phase-locked loops for optical homodyne receivers: performance analysis, design considerations and laser linewidth requirements," *IEEE J. Lightwave Tech.*, vol. LT-4, no. 2, p.182, Feb. 1986.

[157] J. F. Hayes and W. C. Lindsey, "Power allocation - rapidly varying phase error," *IEEE Trans. Commun. Tech.*, vol. CT-17, no. 2, pp. 323-326, April 1969.

[158] L. G. Kazovsky, "Impact of laser phase noise on optical heterodyne receivers," presented in Conference on Laser and Electro-Optics (CLEO'85), Baltimore, MD, May 21-24, 1985.

[159] L. D. Vilesov and V. N. Veis, "Estimate of the angular position of a source of light when received by an array of photodetectors," *Radio Electron. Commun. Syst. (USSR)*, vol. 26, no. 8, pp. 85-87, 1983.

[160] K. Winnick, "Cramer-Rao lower bound on the performance of charge-coupled device optical position estimator," *J. Opt. Soc. Amer. A*, series 2, vol. 3, no. 11, pp. 1809-1815, Nov. 1986.

[161] L. G. Kazovsky, "Theory of tracking accuracy of laser systems," *Optical Eng.*, vol. 22, no. 3, pp. 339-347, 1983.

[162] I. M. Templyakov, "Acquisition and tracking of laser beams in space communications," *Acta Astronautica*, vol. 7, pp. 341-355, 1980.

[163] W. W. Peterson and E. J. Weldon, Jr., *Error Correcting Codes*. Cambridge, MA: MIT Press, 1970.

[164] R. J. McEliece, *The Theory of Information and Coding* (Encyclopedia of Mathematics and Applications), Reading, MA: Addison-Wesley, 1977.

[165] C. N. Georghiades, "Joint baud and frame synchronization in direct detection optical communications," *IEEE Trans. Commun.*, vol. COM-33, no. 4, pp. 357-360, April 1985.

[166] C. N. Georghiades and D. L. Snyder, "Locating data frames in direct detection optical communication systems," *IEEE Trans. Commun.*, vol. COM-32, pp. 118-123, Feb. 1984.

[167] R. Gagliardi and M. Haney, "Optical synchronization phase-locking with shot noise process," USCEE Rep.396, Aug. 1970.

[168] R. H. Forrester and D. L. Snyder, "Phase tracking performance of direct-detection optical receivers," *IEEE Trans. Commun.*, vol. COM-21, pp. 1037-1039, Sept. 1973.

[169] C. C. Chen and C. S. Gardner, "Phase-locked loop synchronization for direct detection M-ary PPM communication systems," Electro-Optics Systems Lab., Univ. of Illinois, Urbana IL, Rep. EOSL85-003 (UILU-ENG-85-2549), May 1985.

[170] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, Jan. 1975.

[171] J. B. Abshire, private communications.

CUMULATIVE LIST OF RADIO RESEARCH LABORTORY

AND ELECTRO-OPTIC SYSTEMS LABORTORY REPORTS

PREPARED UNDER NASA GRANT NSG-5049

RRL Rep. No. 469 - Gardner, C. S. and N. N. Rao (December 1975), The Effects of Random Path Fluctuations on the Accuracy of Laser Ranging Systems.

RRL Rep. No. 471 - Zanter, D. L., C. S. Gardner and N. N. Rao (January 1976), The Effects of Atmospheric Refraction on The Accuracy of Laser Ranging Systems.

RRL Rep. No. 477 - Gardner, C. S. and J. R. Rowlett (November 1976), Atmospheric Refraction Errors in Laser Ranging Data.

RRL Rep. No. 478 - Hendrickson, B. E. and C. S. Gardner (December 1976), Correction of Laser Ranging Data for the Effects of Horizontal Refractivity Gradients.

RRL Rep. No. 481 - Gardner, C. S. (February 1977), Statistics of the Residual Refraction Errors in Laser Ranging Data.

RRL Rep. No. 486 - Gardner, C. S. (July 1977), Comparison Between the Refraction Error Covariance Model and Ray Tracing.

RRL Rep. No. 488 - Gardner, C. S. (September 1977), Speckle Noise in Satellite Based Lidar Systems.

RRL Rep. No. 495 - Gardner, C. S. and G. S. Mecherle (April 1978), Speckle Noise in Direct-Detection Lidar Systems.

RRL Rep. No. 496 - Gardner, C. S. and A. M. Saleh (October 1978), Speckle Noise in Differential Absorption Lidar Systems.

RRL Rep. No. 499 - Gardner, C. S. (January 1979), A Technique for Remotely Measuring Surface Pressure from a Satellite Using a Multicolor Laser Ranging System.

RRL Rep. No. 502 - Palluch, E., J. D. Shelton and C. S. Gardner (May 1979), Operating Manual for the RRL 8 Channel Data Logger.

RRL Rep. No. 505 - Gardner, C. S. and R. Axford, Jr. (March 1980), Regression Models for Multicolor Satellite Laser Ranging.

RRL Rep. No. 510 - Gardner, C. S. (April 1981), Analysis of Target Signatures for Laser Altimeters.

RRL Rep. No. 511 - Gardner, C. S. (June 1981), Atmospheric Refraction Effects in Air Borne Laser Ranging.

RRL Rep. No. 514 - Tsai, B. and C. S. Gardner (December 1981), Remote Sensing of Sea State by Laser Altimeters.

RRL Rep. No. 518 - Gardner, C. S. (August 1982), Optical Communications.

RRL Rep. No. 519 - Im, K. E. and C. S. Gardner (September 1982), Atmospheric Effects on Baseline Error in Satellite Laser Ranging Systems.

RRL Rep. No. 526 - Im, K. E., B. M. Tsai and C. S. Gardner (September 1983), Analysis of Short Pulse Laser Altimetry Data Obtained over Horizontal Path.

RRL Rep. No. 527 - Tsai, B. M. and C. S. Gardner (March 1984), Theoretical and Experimental Analysis of Laser Altimeters for Barometric Measurements Over the Ocean.

EOSL Rep. No. 84-001 - Lafaw, D. A. and C. S. Gardner (August 1984), Timing Performance of Phase-Locked Loops in Optical Pulse Position Modulation Communication Systems.

EOSL Rep. No. 85-002 - Im, K. E. and C. S. Gardner (April 1985), Estimation of the Differential Pulse Propagation Times in Two-Color Laser Ranging Systems.

EOSL Rep. No. 85-003 - Chen, C. C. and C. S. Gardner (May 1985), Phase-Locked Loop Synchronization for Direct Detection Optical PPM Communication Systems.

EOSL Rep. No. 85-006 - Im, K. E. and C. S. Gardner (August 1985), Theoretical and Experimental Analysis of the Performance of Two-Color Laser Ranging Systems.

EOSL Rep. No. 87-002 - Chen, C. C. and C. S. Gardner (March 1987), Comparison of Direct and Heterodyne Detection Optical Intersatellite Communication Links.

PAPERS PUBLISHED

C. S. Gardner, "Effects of Random Path Fluctuations on the Accuracy of Laser Ranging Data," Applied Optics, 15, 2539-2545, October 1976.

C. S. Gardner, "Effects of Horizontal Refractivity Gradients on the Accuracy of Laser Ranging to Satellites," Radio Science, 11, 1037-1044, December 1976.

C. S. Gardner, "Correction of Laser Tracking Data for the Effects of Horizontal Refractivity Gradients," Applied Optics, 16, 2427-2432, September 1977.

C. S. Gardner, R. Rowlett and B. E. Hendrickson, "Ray Tracing Evaluation of a Technique for Correcting the Refraction Errors in Satellite Tracking Data," Applied Optics, 17, 3143-3145, October 1978.

C. S. Gardner, "Technique for Remotely Measuring Surface Pressure from a Satellite Using a Multicolor Laser Ranging System," Applied Optics, 18, 3184-3189, September 1979.

C. S. Gardner, "Target Signatures for Laser Altimeters: An Analysis," Applied Optics, 21, 448-453, February 1982.

B. M. Tsai and C. S. Gardner, "Remote Sensing of Sea State Using Laser Altimeters," Applied Optics, 21, 3932-3940, November 1982.

C. S. Gardner, B. M. Tsai and J. B. Abshire, "Remote Sensing of Atmospheric Pressure and Sea State from Satellites Using Short-Pulse Multicolor Laser Altimeters," Proceedings of NATO-AGARD Symposium on Propagation Factors Affecting Remote Sensing by Radio Waves, 345, (46-1)-(46-11), Oberammergau, FRG, May 24-28, 1983.

C. S. Gardner, B. M. Tsai and K. E. Im, "Multicolor Laser Altimeters for Barometric Measurements over the Ocean: Theoretical," Applied Optics, 22, 2571-2577, September 1, 1983.

C. S. Gardner and J. B. Abshire, "Atmospheric refraction and target speckle effects on the accuracy of laser ranging systems," Proc. Int. Conf. on Laser Ranging Instrumentation, 1, 29-41, Royal Greenwich Observatory, Hearstmonceux, UK, September 24-28, 1984 (invited paper).

B. M. Tsai and C. S. Gardner, "Time-Resolved Speckle Effects on the Estimation of Laser Pulse Arrival Times," J. Opt. Soc. Amer. A., 2, 649-656, May 1985.

J. B. Abshire and C. S. Gardner, "Atmospheric Refractivity Corrections for Satellite Laser Ranging," IEEE Trans. Geosci. Remote Sensing, GE-2, 414-425, July 1985.

C. S. Gardner, "Remote Sensing of Atmospheric Pressure and Sea State Using Laser Altimetry," Proc. 1985 Int. Geosci. Remote Sensing Symps., 1, 199-206, Amherst, MA, October 7-9, 1985.

K. E. Im and C. S. Gardner, "Estimation of Differential Pulse Propagation Times in Two-Color Laser Ranging Systems," J. Opt. Soc. Amer. A., 3, 143-156, Jan. 1986.

C. C. Chen and C. S. Gardner, "Performance of Phase Locked Loop Synchronized Optical PPM Communication Systems," IEEE Trans. Comm., COM-34, 988-994, Oct. 1986.

C. C. Chen and C. S. Gardner, "Loss Factors Associated with Spatial and Temporal Tracking Error in Intersatellite PPM Communication Links," Proc. IEEE Global Telescom Conf., 3, 1392-1397, Houston, TX, Dec. 1-4, 1986.

C. C. Chen and C. S. Gardner, "Impact of Random Pointing and Tracking Errors on the Design of Coherent and Incoherent Optical Intersatellite Communication Links," IEEE Trans. Comm., to be published, 1987.

K. E. Im, C. S. Gardner, J. B. Abshire and J. F. McGarry, "Experimental evaluation of the performance of pulsed two-color laser ranging systems," J. Opt. Soc. Amer. A., to be published, 1987.