INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

DOPPLER COMPENSATION FOR LEO SATELLITE COMMUNICATION SYSTEMS

By

GEORGE PATRICK AH-THEW

B. Eng. (McMaster University)

A Thesis Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOPPLER COMPENSATION FOR LEO SATELLITE COMMUNICATION SYSTEMS

To My Parents

DOCTOR OF PHILOSOPHY (1998) McMASTER UNIVERSITY

(Electrical and Computer Engineering) Hamilton, Ontario

TITLE: **Doppler Compensation for LEO Satellite Communication Systems**

AUTHOR: G. P. Ah-Thew

B. Eng. (McMaster University)

SUPERVISOR: Dr. C. R. Carter

NUMBER OF PAGES: xxxii, 271

ABSTRACT

This thesis investigates the carrier synchronization problem due to Doppler shift encountered in low data rate, low earth orbit (LEO) satellite communication systems. In particular, two new techniques are developed.

The first technique which is referred to as the novel technique, is based on the idea of applying a linearly decreasing frequency sweep to the received signal as a method to reduce the Doppler shift to tolerable levels, whereby conventional carrier synchronization may be employed. This method has shown to be better suited for satellites at larger cross track angles. However, for direct overhead satellite pass, improved Doppler compensation is achieved by utilizing the modified novel technique. Simulations have been carried out to explore the performance of the novel technique and results have shown that it is capable of reducing Doppler shift by at least a factor of two and performs better for satellites at larger cross track angles.

The second technique of Doppler compensation is based on Maximum Likelihood (ML) estimation of multiple parameters. This algorithm makes use of a model developed for the phase error introduced due to Doppler frequency shift and Doppler rate, and the two

implementable solutions investigated are; (1) the grid search, and (2) the ML quadriphase shift keying (QPSK) Doppler compensator. These solutions have been realized by utilizing approximations for low signal to noise ratio (SNR) conditions. The grid search is a non tracking technique and may be used for coarse Doppler compensation. Unfortunately, simulation results have shown that it is computationally intensive.

On another note, a block diagram consisting of a phase rotator, matched filtering and two feedback loops has been developed to represent the ML QPSK Doppler compensator. Three digital implementations of this structure have been analysed, namely; steepest descent, optimized steepest descent and Newton Raphson method. The first two, have two modes of operation; conventional feedback and adaptive feedback. Simulations have shown that the adaptive optimized steepest descent has the best performance, that is, the shortest lockup time. Comparisons to previous work in this area, have shown that the proposed Doppler compensator has superior performance.

ACKNOWLEDGEMENT

The author wishes to acknowledge his sincere appreciation to Dr. C. R. Carter for his constant encouragement, continued assistance, inspiring guidance and enthusiasm, and financial support throughout the course of this work.

Gratitude is also due to my supervisory committee members, Drs. P. Yip and and D. Conn for their valuable advice and wisdom.

Over and above, he would like to express his deep appreciation to his family, including his parents, sister and brother for their love and support during his study.

Finally, he is obliged to the Canadian Commonwealth for sponsoring him throughout his undergraduate and graduate studies at McMaster University.

GLOSSARY

List of Symbols

A	Signal amplitude
a_k	Data symbols
A_z	Azimuth angle
B_{IF}	Intermediate Frequency Bandwidth
b_k	Data symbols
c	Speed of light
$\cosh(x)$	Hyperbolic cosine of x
C/I	Carrier to Interference ratio
c_k	Complex data symbols
C/N	Carrier to Noise ratio
C/N_o	Carrier to Noise spectral density ratio

$(C/N_o)_r$	Received Carrier to Noise spectral density ratio
$(C/N_o)_R$	Required Carrier to Noise spectral density ratio
$E[x]$	Expectation of x
E_b/N_o	Bit Energy to Noise spectral density ratio
$EIRP$	Equivalent Isotropic Radiated Power
E_s	Symbol energy
$E[\epsilon_d(n)]$	Expectation of the frequency loop error
$E[\epsilon_r(n)]$	Expectation of the frequency rate loop error
f_c	Carrier frequency
f_d	Doppler frequency shift
$\hat{f}_d$	Estimate of f_d
$\bar{f}_d$	Grid search estimate of f_d
f_D	Actual Doppler frequency shift
$f_i(t)$	Dynamic or instantaneous frequency shift
f_r	Doppler rate
$\hat{f}_r$	Estimate of f_r

$\tilde{f}_r$	Grid search estimate of f_r
f_R	Actual Doppler rate
G_R	Receiver antenna gain
G_T	Transmitter antenna gain
G/T_s	Receiver gain temperature ratio
h	Altitude of satellite
$\underline{H}$	Hessian matrix
h_{GEO}	Geostationary Earth Orbit altitude
h_{LEO}	Low Earth Orbit altitude
$h(t)$	Signal pulse shape
k	Boltzmann's Constant
K	Constraint length
L	Free space loss or longitude difference between subsatellite point and user terminal
ln cosh(x)	Natural log of hyperbolic cosine of x
$L(\bar{r}(T_o)\|\underline{\alpha})$	Log-likelihood function
$LPF[\ \bullet\ ]$	Low Pass Filter

l_s	Latitude of subsatellite point
l_t	Latitude of user terminal
M	M-ary or over-sample factor
$m(t)$	Data signal
$\tilde{m}(t)$	Complex envelope of $m(t)$
$m_c(t)$	In-phase channel signal
$\tilde{m}_c(t)$	Complex envelope of $m_c(t)$
$m_s(t)$	Quadrature channel signal
$\tilde{m}_s(t)$	Complex envelope of $m_s(t)$
N	Total number of symbols observed
N_o	Noise spectral density
$N \times N$	Noise × Noise
$n(t)$	Noise
$\tilde{n}(t)$	Complex envelope of $n(t)$
$n_c(t)$	In-phase component of $n(t)$
$\tilde{n}_c(t)$	Complex envelope of $n_c(t)$
$n_s(t)$	Quadrature component of $n(t)$

$\tilde{n}_s(t)$	Complex envelope of $n_s(t)$
P	Location of satellite
$P(x)$	A priori probability of x
P_c	Signal power level
$p(r)$	Probability of $r(t)$
$p(r\|\underline{\alpha})$	Conditional probability density function of $r(t)$ given $\underline{\alpha}$
P_R	Power at the receiver input
P_T	Transmitted power
R_b	Uncoded data rate
R_c	Coded data rate
R_e	Mean earth radius
R_s	Symbol Rate
$r(t)$	Received signal
$\tilde{r}(t)$	Complex envelope of $r(t)$
s	Slant range
S	Subsatellite point
$S\times N$	Signal × Noise

$sgn(x)$	Signum of x
$SGN[\ \cdot\ ]$	Signum function
SNR	Signal to Noise ratio
$S \times S$	Signal × Signal
$s(t)$	Transmitted signal
$\tilde{s}(t)$	Complex envelope of $s(t)$
$\tilde{s}^*(t)$	Conjugate of the complex envelope of $s(t)$
$s(t\|\underline{\alpha})$	Trial signal
t	Time
T	Sampling period ($T_s = MT$) or location of user terminal
$\tanh(x)$	Hyperbolic tangent of x
T_b	Bit duration
T_o	Observation period (estimation period equal to $NT_s = M \times N \times T$)
T_s	System noise temperature or symbol period
$U(kT_s)$	Refer to (5.30)
v	Velocity of satellite
v_d	Relative radial velocity of satellite
α	Rolloff factor

$\underline{\alpha}$	Vector containing estimation parameters
α_{opt}	Optimum step gain
γ	Elevation or grazing angle
Δf	Doppler frequency shift
$\lvert\Delta f\rvert$	Tracking range
Δf_d	Change in Doppler frequency
$\Delta \bar{f}_d$	Change in $\bar{f}_d$
Δf_r	Change in Doppler rate
$\Delta(t)$	Normalized Doppler shift
ΔT	User terminal clock time base delay
δ	Satellite beamwidth, tilt or nadir angle
$\delta(x)$	Unit impulse
ϵ	Feedback loop error
$\underline{\epsilon}(k)$	Vector containing $\epsilon_d(k)$ and $\epsilon_r(k)$
$\epsilon_d(k)$	Frequency loop error
$\epsilon_r(k)$	Frequency rate loop error
$\hat{\theta}$	Estimate of the modulating phase

θ_ϵ	Phase error
θ_i	Modulating phase
θ_o	Arbitrary carrier phase
$\Lambda(r\|\underline{\alpha})$	Likelihood function
λ	Carrier wavelength
$\mu(t)$	Total carrier phase shift
σ_n^2	Variance of $n(t)$
$\sigma_{\tilde{n}}^2$	Variance of $\tilde{n}(t)$
$\sigma_{n_c}^2$	Variance of $n_c(t)$
$\sigma_{n_s}^2$	Variance of $n_s(t)$
τ	Time delay
Φ	Satellite coverage angle
Φ_{max}	Maximum satellite coverage angle
ϕ_k	Phase symbol
$\Psi(t)$	Phase error introduced by Doppler frequency shift and Doppler rate
$\hat{\Psi}(t)$	Estimate of the phase error $\Psi(t)$

ω_c	Carrier angular frequency
$\frac{\partial U(kT_s)}{\partial f_d}$	Gradient of $U(kT_s)$ with respect to f_d
$\frac{\partial U(kT_s)}{\partial f_r}$	Gradient of $U(kT_s)$ with respect to f_r

List of Acronyms

AFC	Automatic Frequency Control
AM	Amplitude Modulation
ASEAN	Association of South-East Asian Nations
ASS	Amateur Satellite Service
ATS	Application Technology Satellite
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
C-band	4 - 8 GHz
CCIR	International Radio Consultative Committee
CDMA	Code Division Multiple Access
COMETS	Communication and Broadcasting Engineering Test Satellite

DA	Data Aided or Demand Assignment
DAMA	Demand Assignment Multiple Access
DBS	Direct Broadcast Satellite
DC	Direct Current
DPSK	Differentially Phase Shift Keying
DSB-SC	Double Side Band Suppressed Carrier
DSB-SC	Double Side Band Amplitude Modulation
DSI	Digital Speech Interpolation
ELT	Emergency Location Transmitter
EPIRP	Emergency Position Indication Radio Beacon
ETS	Engineering Test Satellite
FACC	Ford Aerospace and Communication Corporation
FD	Frequency Detector
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FMF	Frequency Matched Filter
FRMF	Frequency Rate Matched Filter
GEO	Geostationary Earth Orbit
GPS	Global Positioning System
GSM	Global System for Mobile communications
HDTV	High Definition Television

HEO	Highly Elliptical Orbit
HF	High Frequency (3 - 30 MHz)
HPA	High Power Amplifier
INSAT	Indian National Satellite
ISI	Inter Symbol Interference
ITU	International Telecommunication Union
ITU-R	Radiocommunication section of the ITU
JCSAT	Japan Communications Satellite Company Inc.
JDC	Japan Digital Cellular
K-band	18 - 27 GHz
Ka-band	27 - 40 GHz
Ku-band	12 - 18 GHz
L-band	1 - 2 GHz
LCD	Liquid Crystal Display
LEO	Low Earth Orbit
LHCP	Left Hand Circular Polarization
LOS	Line Of Sight
MEO	Medium Earth Orbit
MF	Medium Frequency (0.3 - 3 MHz)
ML	Maximum Likelihood
MPSK	Multiple Phase Shift Keying

MSS	Mobile Satellite Service
NADC	North American Digital Cellular
NASA	National Aeronautics and Space Administration
NRZ	Non Return to Zero
NTT	Nippon Telegraph and Telephone Public Corporation
OBP	On Board Processing
PA	Pre-Assigned
PCS	Personal Communication Services
PCN	Personal Communication Networks
PD	Phase Detector
PLL	Phase Lock Loop
PM	Phase Modulation
PSTN	Public Switched Telephone Network
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RDSS	Radio Determination Satellite Service
RHCP	Right Hand Circular Polarization
RF	Radio Frequency
SAR	Search And Rescue
SARSAT	Search And Rescue Satellite Aided and Tracking
SB	Start of Burst

S-band	2 - 4 GHz
SCC	Space Communications Corporation
SCPC	Single Channel Per Carrier
SDMA	Space Division Multiple Access
SMF	Signal Matched Filter
STES	Small Transportable Earth Station
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TT&C	Telemetry Tracking and Command
TVRO	Television Receive Only
TWTA	Travelling Wave Tube Amplifier
UHF	Ultra High Frequency (30 - 300 MHz)
VAFB	Venderberg Air Force Base
VCO	Voltage Controlled Oscillator
VHF	Very High Frequency (300 - 1000 MHz)
VLSI	Very Large Scale Integration
VSAT	Very Small Aperture Terminal
WHO	World Health Organization

CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

Introduction

1.1 Global Communication Systems

Recent growth in mobile communications has been dramatic and is expected to continue for many years to come. It has been predicted that the number of cellular users worldwide will exceed 100 million by the year 2001. Networks providing the cellular services are however centered on the main population areas of the globe. As a matter of fact, remote areas of some industrialized countries, are now and are likely to remain poorly served. Vast areas of the globe will still not be covered by terrestrial cellular systems. At the present time, it is difficult to move from city to city with a single user terminal. For example in Canada, digital personal communication services (PCS) at 1.3 GHz are offered by four different networks and they are incompatible to each other. Clearnet and Bell Mobility are spread spectrum, code division multiple access (CDMA) systems, while Cantel is time division multiple access (TDMA) and Microcell is global system for mobile

communications (GSM). Conversely, in Europe the widespread of GSM is a good indicator of the perceived desire by users to be able to travel widely with a single mobile user terminal. In addition, these systems are the winning choice for medium to high traffic density areas such as urban and suburban environments.

Satellite communications, in this scenario, is well suited to filling the gaps in coverage area left by the terrestrial systems. This also includes regions where terrestrial networks are not competitive for thin route applications or under developed areas. They are unlikely to be able to replace the terrestrial communication system because of their relative difficulty in providing substantial frequency reuse and subsequent inability to supply adequate capacity to meet predicted user demand in highly populated areas. Besides, this approach is in line with the role played by satellite systems in the personal communication world: that is, complementing the terrestrial technologies and exploiting the relevant capabilities whenever possible and convenient. Satellites are ideal candidates for providing thin-route communications to a limited number of users. The main reason why they have not been used in this way in the past is because of the high cost of technology required to deliver voice communications to a hand held user terminal. Recent technological advances in both satellite and terrestrial cellular communications technology are now leading to the realization of the dream of a global mobile personal communication system within the next decade. The terrestrial cellular system has already proven its capability, but the proposed satellite systems are either in their testing phase or are still at the drawing boards. It must also be noted that progress in other areas have made remarkable contributions such as digital voice processing

and component miniaturization.

1.2 Third World Domestic Satellite System Survey

In general LEO satellite systems can provide a limited capacity with respect to terrestrial networks, nevertheless they are particularly suited to cover large surface areas offering thin route services, since it is either not feasible or economical to implement a cellular network. This is an important issue for third world countries since they are now placing the infrastructure to connect to the information superhighway. It is generally acknowledged that universal access to the information superhighway and information exchange, will rapidly enhance economic growth of all nations, including developed countries. Terrestrial networks have not penetrated developing countries to a large extent, since it is costly to implement such systems as a thin route service for large coverage areas. Can satellite systems be used as a possible solution to propel such nations with under developed telecommunications infrastructure, into the abundant sources of services available via the information superhighway? One way to find out, is to conduct a survey.

Appendix A presents a survey of domestic satellite communication systems for islands and developing countries. Elements pertaining to communications techniques as well as service applications are addressed. In particular, features such as service requirements, geographical location, and availability of satellite slots which are related to custom designing a thin-route domestic communication satellite system are investigated. Systems for India,

Peru, South Africa, Indonesia, Japan, and Seychelles have been examined owing to their inherent differences which make them candidates for domestic satellite communications. Some of the major trends and advances related to the type of multiple access and the utilization of higher equivalent isotropic radiated power (EIRP), data rates and frequency band such as K, millimetre wave and optical band are treated with greater detail since they are of utmost importance in today's strategic planning and design on domestic satellite systems. In addition, satellite slots, interference and rain fade counter measures in the case of tropical countries are also investigated.

For island countries, the bulk of domestic communications is accomplished over the high frequency (HF)/ ultra high frequency (UHF)/ very high frequency (VHF) bands. This system has many shortcomings including poor geographical coverage, uncertain ionospheric conditions and a limited number of users, where the latter is determined by available bandwidth and interference introduced by other users [1]. UHF and VHF direct-line-of sight radio paths between transmitters and receivers are reduced due to multipath over areas of water while, HF systems are limited due to ionospheric conditions [2].

Conversely, INTELSAT and INMARSAT are mostly responsible for all the international communication services where the former is a global satellite system composed of more than 20 satellites linking over 180 countries [3]. To date, developing countries constitute the bulk of the membership of INTELSAT [1]. Other consortia include: EUTELSAT providing telecommunications services to European countries; INTERSPUTNIK servicing Russia and its allies; ARABSAT linking the Arab region;

PALAPA connecting the Association of South-East Asian Nations (ASEAN); and, NORSAT uniting the Nordic countries [4].

Looking back a few decades, one can conclude that most developing countries and island nations use satellite technology for international communication purposes instead of domestic [1]. However, now the trend is the use of this technology for domestic and regional communication purposes since leasing a complete transponder or a fraction of a transponder for international and domestic purposes is expensive and often does not economically satisfy the telecommunication requirements. The growth of the telecommunications industry has forced many countries to upgrade their existing domestic communications system introducing services in new frequency bands and transmitting media, such as a fibre optic cable systems, and aiming towards full digitalization of their telecommunications networks [1, 5]. Unfortunately, in many cases, an undersea cable system connecting all the islands of an archipelago is unjustifiable since the cost of introducing such a system is high, and the time required for implementation is long. Furthermore, for these developing countries, underwater maintenance and repair is difficult due to lack of deep sea equipments. A few countries have implemented the use of copper submarine cables and are now faced with technical problems in regards to the transmission of wideband signals [5-6].

A satellite based system provides many advantages in point-to-point and point-to-multipoint communication. NAVSTAR and Meteorological satellites offer a multitude of services such as velocity, location, time, weather information, and search and rescue satellite aided tracking (SARSAT), to the entire globe. As a means of lowering cost, domestic

satellite systems should be designed to provide a multi-purpose service; that is, offer similar services in addition to the previously mentioned communication services. Thus far, only a handful of countries have implemented multi-purpose domestic satellite systems.

1.3 LEO Satellite Systems

Various options exist for satellite communications, and are well documented in the literature. Geostationary satellites are most often used for commercial communication applications and can be used for personal communication networks (PCN). But the satellite will require higher power to compensate for the capabilities of a small user terminal, hence increasing both their cost and complexity. This system suffers from the high cost of launching the satellites, relatively long transmission delays which are unacceptable to today's users after getting accustomed to terrestrial cellular services. Furthermore, the PCN service deteriorates in some situations where low elevation angles at high latitudes are encountered. Satellites in non-geostationary orbits can overcome this problems. We are thus referring to LEO, medium earth orbit, (MEO) and highly elliptical orbit (HEO) satellite systems employing either polar, inclined or equatorial orbits. Each has its own distinctive combination of transmission delay, satellite power requirements, and impact of the mobile user terminal design. Moreover, a high elevation angle may be obtained for a particular geographical area by selecting the right orbits since satellites are no longer fixed (e.g. as in geostationary earth orbit (GEO)) to the equatorial plane. Furthermore, each system has its

own approach in order to provide reliable communication services. For example, for a system to ensure continuous worldwide coverage or service to a single location, the number of satellites required increases as the orbit height decreases. However, this is offset by the lower launch cost per satellite and shorter propagation delays at lower altitudes.

Radio amateurs were the first to design and develop low cost communication satellite systems operating in low earth orbits and therefore, the first users of PCS provided by LEO satellite systems. In addition, this technology first demonstrated in amateur satellite service (ASS) has already been transferred to the commercial sector.

1.4 Scope and Structure of the Thesis

This thesis is organized as follows. We start our research by performing a survey on domestic satellite communication systems for islands and developing countries, provided in Appendix A. In Chapter 2, we examine the LEO satellite system, called Iridium developed by Motorola, which is based on a constellation offering worldwide coverage of mobile personal communication services. Iridium was chosen because it is primarily a thin-route voice communication system and was considered to be the most promising, challenging, and feasible system. All major aspects relevant to the space segment and to the ground segment are outlined. Of importance is the realization that of all satellite systems, Iridium offers minimum time delay in point-to-point communications. However, there are several problems. These are: the large Doppler shifts in uplink and downlink transmissions; the

significant change in Doppler shift when the user and satellite are at minimum range, and the problem of overhead satellite coverage of a user providing service for less than 9 minutes.

Chapter 3 investigates techniques of reducing Doppler shift on the uplink. For conciseness, a model is developed to generate Doppler frequency shift and Doppler rate curves for a user located anywhere on earth and for satellites at any cross track angle. A novel technique is proposed as a possible solution to compensate for Doppler shift on the uplink for the general case. At the conclusion of Chapter 3, a modified version of the novel technique is suggested for the overhead pass satellite case.

Chapter 4 is concerned with the problem of carrier synchronization. There are of course a number of strategies for obtaining carrier synchronization. First we examine conventional methods of carrier synchronization of QPSK signals. Then, we explore current methods employed to compensate for Doppler shift. However, these methods only consider constant Doppler shifts and do not address changing Doppler shifts as encountered in typical LEO satellite communication systems. As we will see, the QPSK acquisition range fundamental limit is derived and the short comings of current carrier synchronization techniques are discussed. We end the chapter by selecting the ML technique as our strategy to solve the Doppler shift problem in the following chapter.

In Chapter 5, we formulate the problem by developing a model for the phase error introduced by Doppler frequency shift and Doppler rate. Employing the ML technique and some reasonable assumptions, two possible practical solutions are obtained namely; (1) the grid search, and (2) the ML QPSK Doppler compensator, where the latter is based on

matched filtering. Towards the end, we present the structure of a ML QPSK Doppler compensator and its digital implementation.

Chapter 6 documents computer simulation results based on the theory of Chapter 5. The grid search and s-curve (Discriminator Characteristic) are simulated for various conditions. Closed form s-curve representation is derived for zero Doppler rate and an error analysis at equilibrium is discussed. Based on the ML QPSK Doppler compensator, three methods of estimating the Doppler shift are developed. Computer simulations are conducted to test and compare the effectiveness of these methods and the adaptive optimized steepest descent method is found to be optimum. Its performance (lockup time) is then compared to the existing carrier synchronizer.

The final chapter provides contributions of this thesis, conclusions, and direction for future research work.

CHAPTER 2

Iridium LEO Satellite Communications System

2.1 Introduction

This chapter examines a new LEO satellite communication system called Iridium, which is a worldwide digital satellite-based communication system developed by Motorola. Particular attention is paid to its creative frequency reuse technique, the satellite constellation, the satellite design, the ground infrastructure (gateway and user unit network), the multiple access, modulation employed, frequency band utilization, channel coding and the Doppler effect problem which is inherent in LEO systems. A few other areas such as radio frequency (RF) interference avoidance techniques, type of polarization used, proper selection of transmitting frequency, propagation delay, call routing procedures, services

offered and overall system's reliability will also be described.

The Iridium worldwide, digital, satellite-based PCS is composed of a constellation of 66 smart LEO satellites all networked together, as shown in Figure 2.1. The system obtained its name from its originally proposed constellation of 77 satellites, and its analogy to the chemical element Iridium with atomic number 77 [7-9]. Proposed in 1990, Iridium is scheduled for operation by late 1998 and the Motorola authorities have no intention of changing the project name to Dysprosium (atomic number 66). The first 5 satellites were launched by McDonnel Douglas Delta II, on May 5th 1997 at Venderberg Air Force Base (VAFB) in California.

The most important feature of this system is that it is capable of providing personal telephone service between any two points on earth, and within 30.48 km above sea level at any time of the day or night [10-11]. It is not intended to replace or substitute the cellular telephone service, but rather extend the radio-telephone coverage to the entire world. It has unprecedented spectrum efficiency, unlimited service flexibility, and worldwide coverage.

Iridium services will include data transfer, cellular-like voice, facsimile, paging, messaging services, and radio determination satellite service (RDSS). The estimate voice cost per minute is $3.00, which is about 3-10 times, today's cellular rate. It is anticipated that in the near future, Iridium will replace the HF communication systems which are extensively used in sparsely populated areas and the usage fee might drop to about $2.00 per minute [12-15].

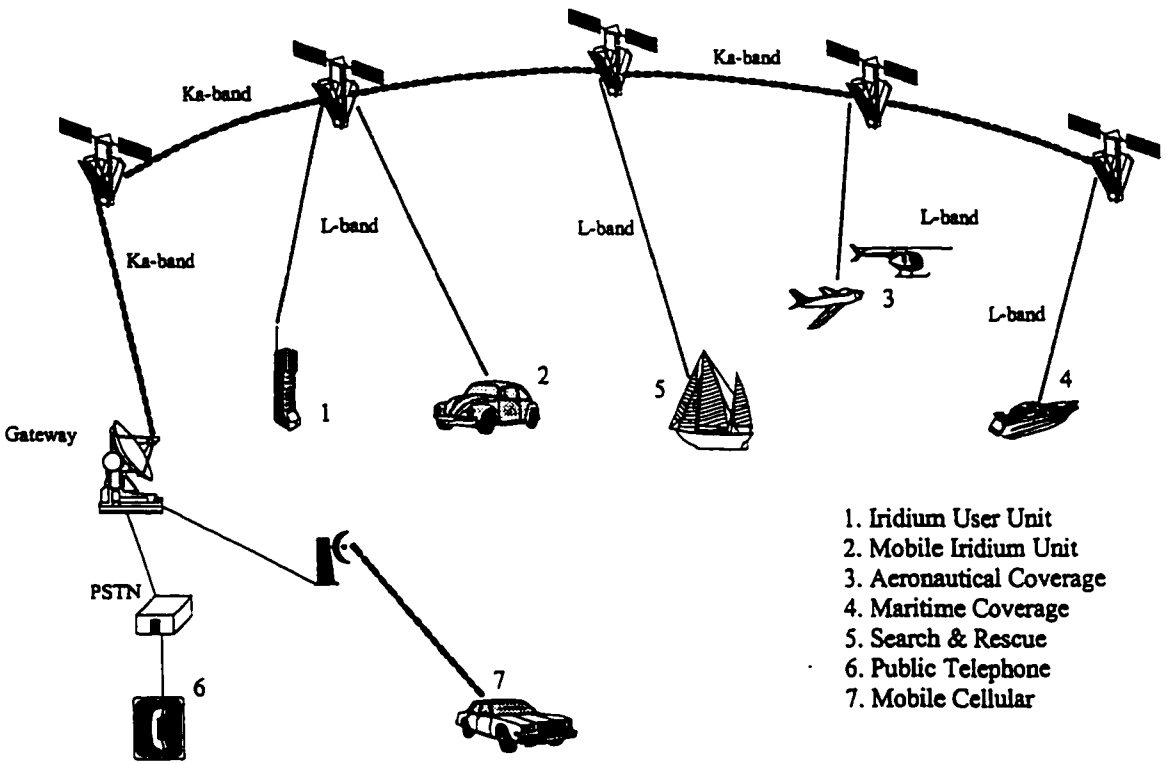


Figure 2.1: Iridium LEO satellite network

Some of the material presented in this chapter has not been published by Motorola, due to fact that the Iridium system is still undergoing thorough field tests, and the final design is being continually updated to further enhance the overall system. Motorola has published only a limited number of papers without revealing in depth technical information about the system. Iridium is now facing competition from several other proposed satellite communication systems such as Globalstar (Loral and Qualcomm), Odyssey (TRW), Orbcomm (Orbital Sciences Corporation and Teleglobe) and Teledesic (Teledesic Corporation).

2.2 Space Segment

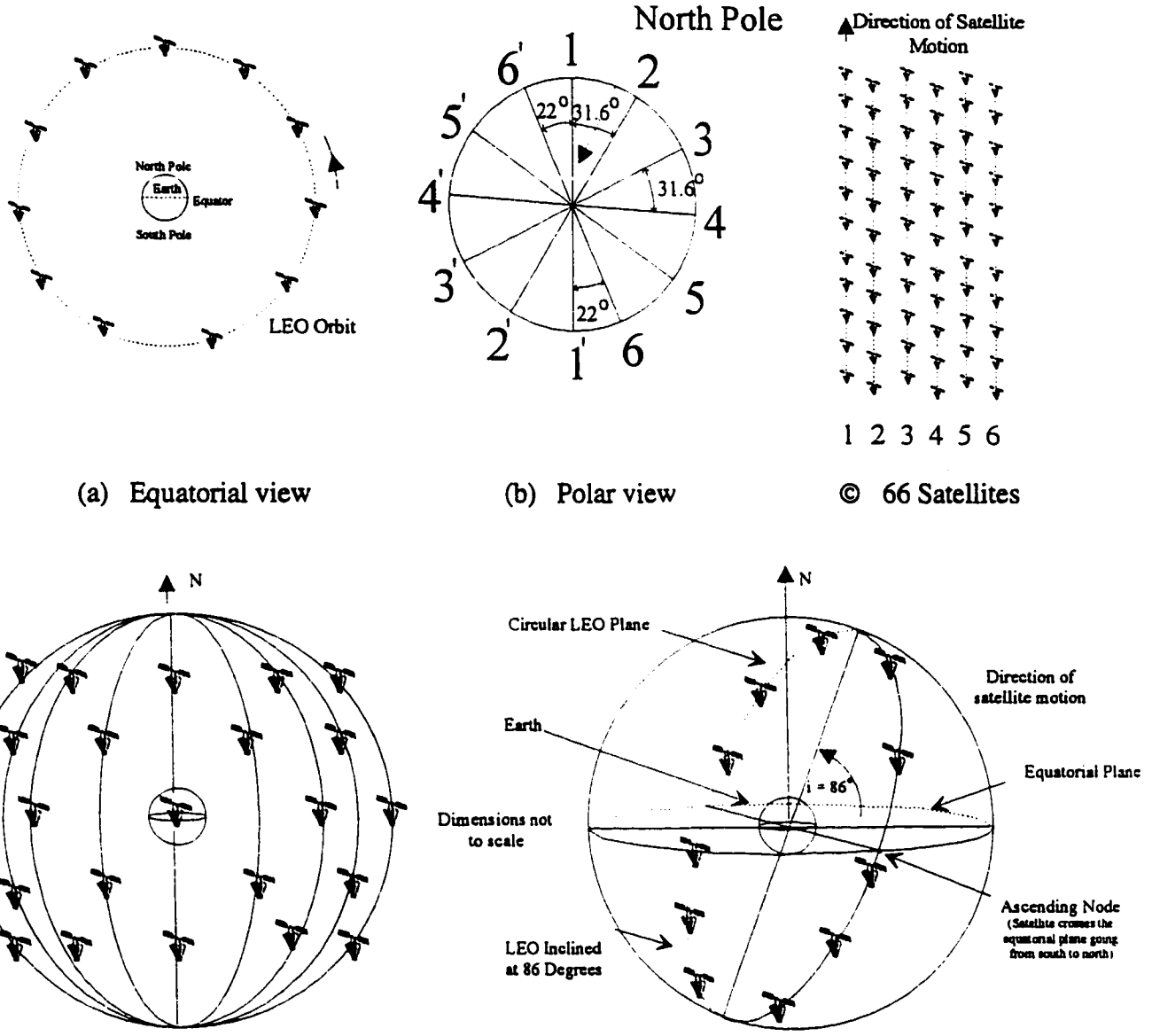
Full continuous earth coverage is a major driving factor in the design of Iridium. As the altitude of the satellites increase, fewer satellites are needed, since each has a wider coverage area. As a matter of fact, at geostationary orbit only three satellites are needed to provide global coverage, excluding the polar regions. Some systems employ multifold coverage as in the case of NAVSTAR global positioning system (GPS) which requires that every user be able to observe four satellites simultaneously. Hence, such a system requires a four-fold or quadruple coverage satellite constellation. Conversely, Iridium requires only a single-fold coverage constellation since at least one satellite is in view at any point in time, making use of the Rosettes Constellations [11, 16-21].

2.2.1 Constellation

The space segment consists of a constellation of 66 digitally networked, smart LEO satellites. There are 6 optimally-phased circular polar orbits with 11 satellites in each. A schematic of the constellation is shown in Figure 2.2.

The orbital period is 100.4 minutes with each satellite being in view to a stationary user on earth for about 10 minutes. The altitude of the satellites, 780.6 km, was chosen for a variety of reasons including: orbital lifetime; launch costs; radiation effect on the semiconductor components; and, communication parameters including guaranteed grazing

angles to the user units of at least 8° [15].



(a) Equatorial view (b) Polar view © 66 Satellites

(d) Polar constellation (e) 86° Circular inclined LEO

Figure 2.2: Satellite constellation

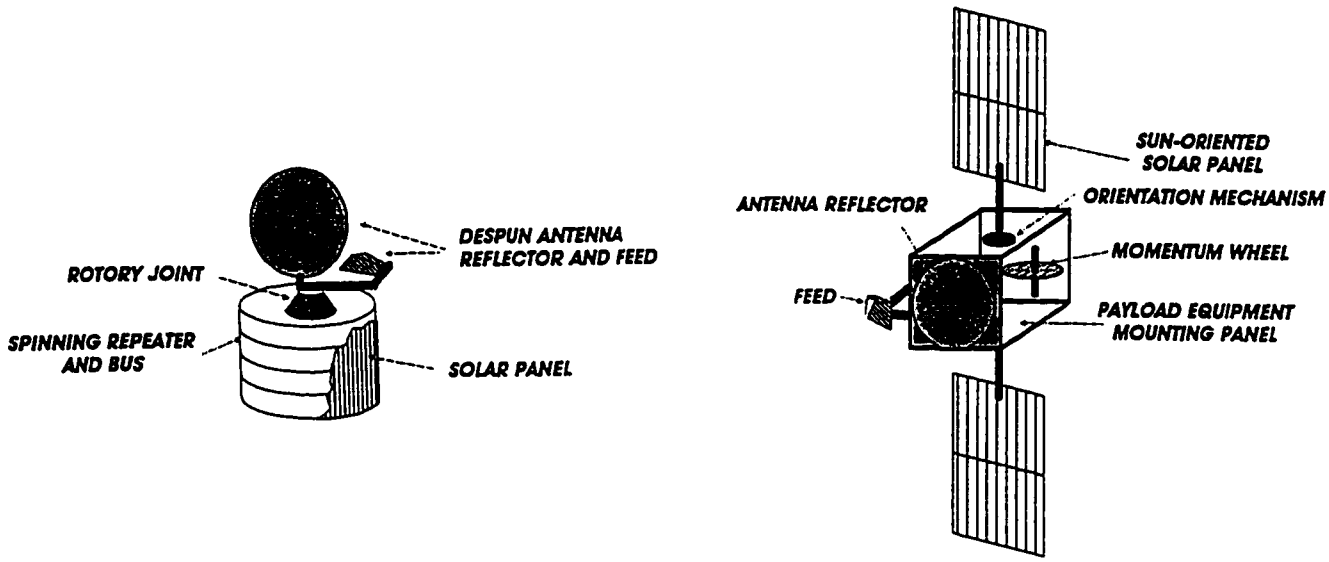
All the satellites travel in the same direction, which means that they co-rotate towards the north pole on one side of the earth and come down towards the south pole on the other

side. The 11 satellites in each orbit plane are nearly evenly spaced as shown in the equatorial view of one of the orbits in Figure 2.2(a). Figure 2.2(b) depicts the view of the orbits from the North Pole. Satellites in the semi-orbits 1-6' and 6-1' are moving in the opposite directions (counter-rotating). The area between the 1st and 6th plane is referred to as counter-rotating seam. In order to achieve an efficient satellite coverage in the counter-rotating seam, orbital planes 1 and 6 are separated by 22°. The rest of the adjacent planes are equally separated by 31.6° (2x22° + 10x31.6° = 360°). Figure 2.2(c) shows each satellite's phase relationship with respect to one another. Satellites in planes 1, 3, and 5 are in-phase with one another while those in planes 2, 4, and 6 are in-phase with each other but halfway out-of-phase with those in 1, 3, and 5. Hence, they are referred to as being optimally-phased and provide an efficient global coverage as shown in Figure 2.2(d). From a practical point of view, to avoid collisions at the poles, and to simplify station keeping, minor corrections to the above described constellation had to be made. Therefore, the orbits are inclined at 86° to the equator and one of the orbits is illustrated in Figure 2.2(e). Note that a satellite in an inclined orbit as mentioned above, sees a slightly different part of the earth's surface with each pass over the earth.

2.2.2 Satellite

There are two methodologies for achieving set attitude control. The two classes are the spin stabilized and the three-axis or sometimes referred to as body stabilized satellites.

A typical spin stabilized and three-axis stabilized satellite are shown in Figure 2.3.



(a) Spin stabilization (B) Body stabilization or 3-axis stabilization

Figure 2.3: Spin stabilized and three-axis (or body) stabilized satellites

In a spin stabilized system, the body which is usually in the shape of a drum, is spun about the pitch axis and with the help of the rotary joint, the antenna is spun in the opposite direction so that the antenna remains pointing at the earth. Furthermore, in a spin stabilized satellite system, the solar panels are mounted on the body. For the three-axis stabilized system, the body of the satellite is maintained in a fixed position relative to the earth. This is achieved by means of momentum wheels which act like gyros. In most cases, the solar panels are mounted on flat panels and are extended from the main body of the satellite [4, 22].

The Iridium satellites have a triangular shaped structure and are three-axis stabilized or body-stabilized satellites, as shown in Figure 2.4. Each satellite is about one metre in diameter, two metres tall and weighs 700 kg, thus allowing the use of flexible, and inexpensive launch vehicles such as Ariane, Atlas, Delta, Proton, Pegasus and Taurus [16].

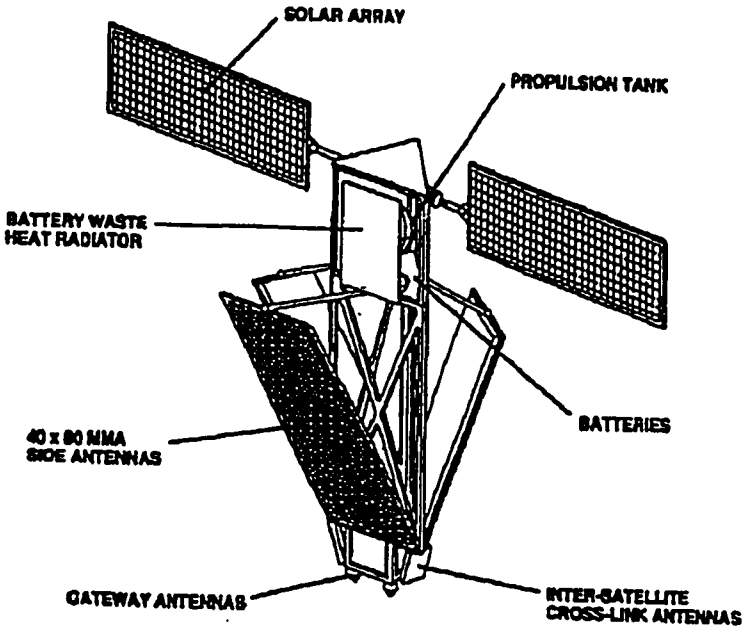


Figure 2.4: Iridium satellite with 3 phased array antennas

This is twice the weight of the original design. The low cost launch is required to maintain the constellation after it is put into operation, and a replacement satellite will be launched on average once every month. The expected orbital lifetime of a satellite is about 5 to 8 years.

Every satellite has the capability to generate 48 hexagonally shaped, interlaced spot beams (cells) by using state-of-the-art L-band (1610-1626.5 MHz) phased array antennas each producing 16 beams [15, 23]. The phased array antennas are located on the sides of the satellite and consist of a total of 165 elements in 10 columns (5 columns with 16 elements and the other 5 columns with 17 elements) [24].

L-band was chosen because of its suitability for direct satellite-to-user links due to its relatively short wavelengths, and low propagation losses [25]. Furthermore, L-band hardware is quite economical. These beams are fixed relative to the satellite and travel at an orbital speed of 7460 m/s (26,900 km/hour).

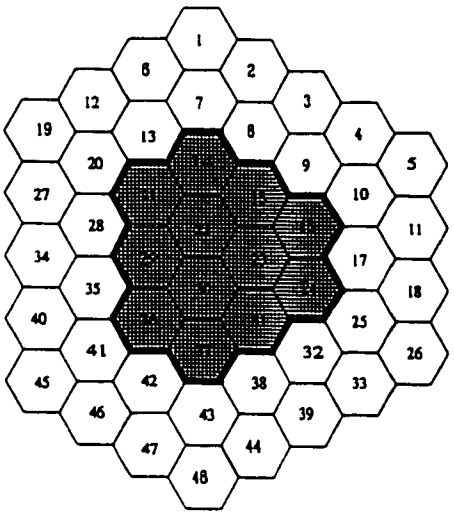


Figure 2.5: Iridium 48 cell pattern

Each spot beam creates a service area (cell) of about 602 km in diameter compared to cellular telephone cell which is roughly 1.6 km in diameter. The satellite antenna spot beam pattern (foot print) is shown in Figure 2.5, where the cells are numbered 1 through 48.

Since the Iridium is a direct line-of-sight (LOS) digital satellite communication system that operates in a cellular communications architecture, the cells created by the satellites allow the channels to be reused many times. A 12-cell frequency reuse pattern with cells designated A to L is utilized to provide part of the spectral efficiency in a similar way to terrestrial cellular telephone, and basically provides a buffer of three cells between any two cell using the same channel [26]. The basic 12-cell cluster is shown in Figure 2.6.

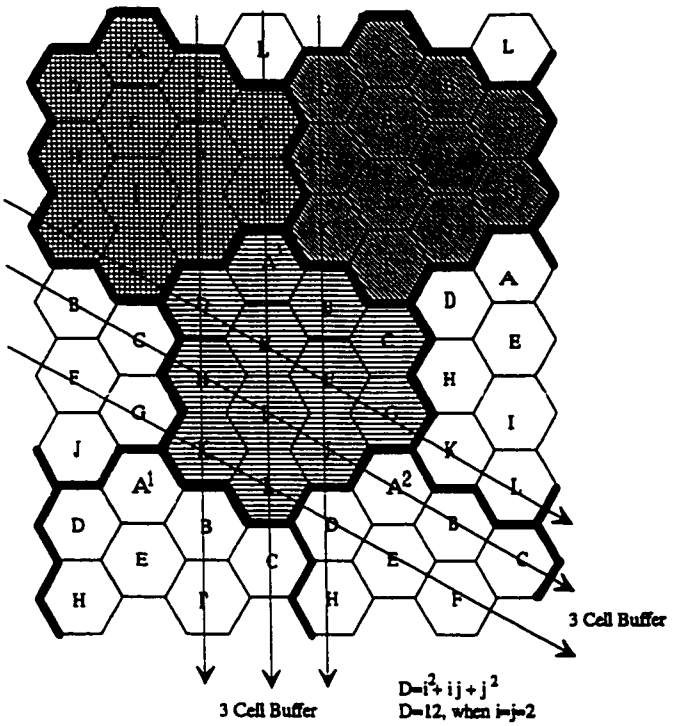


Figure 2.6: 12-cell cluster

In cellular telephone, the base stations are stationary and the users are mobile; hence, the handover of a user is to one of the six surrounding cells. Alternatively, because of the high speed movement of the Iridium cells, the users appear stationary on earth with respect to the cell. The user's position and the direction and speed of the cells are known. This simplifies the handover procedure, since it now becomes one of two cells and involves fewer handoff options, and better information to choose from. This corresponds to simplicity of the system. The Iridium looks like an upside down cellular implementation, where the satellites are the base station antennas. But, unlike the cellular telephones, Iridium cells move through the users, rather than the users moving through the cells.

Handovers between cells of a single satellite, which are also known as Intra-Satellite handovers typically occur every 60 to 90 seconds [14]. In addition, handovers between satellites, which are also known as inter-satellite handovers occurs every nine minutes. Refer

to Figure 2.7 for a global representation of the spot beam coverage.

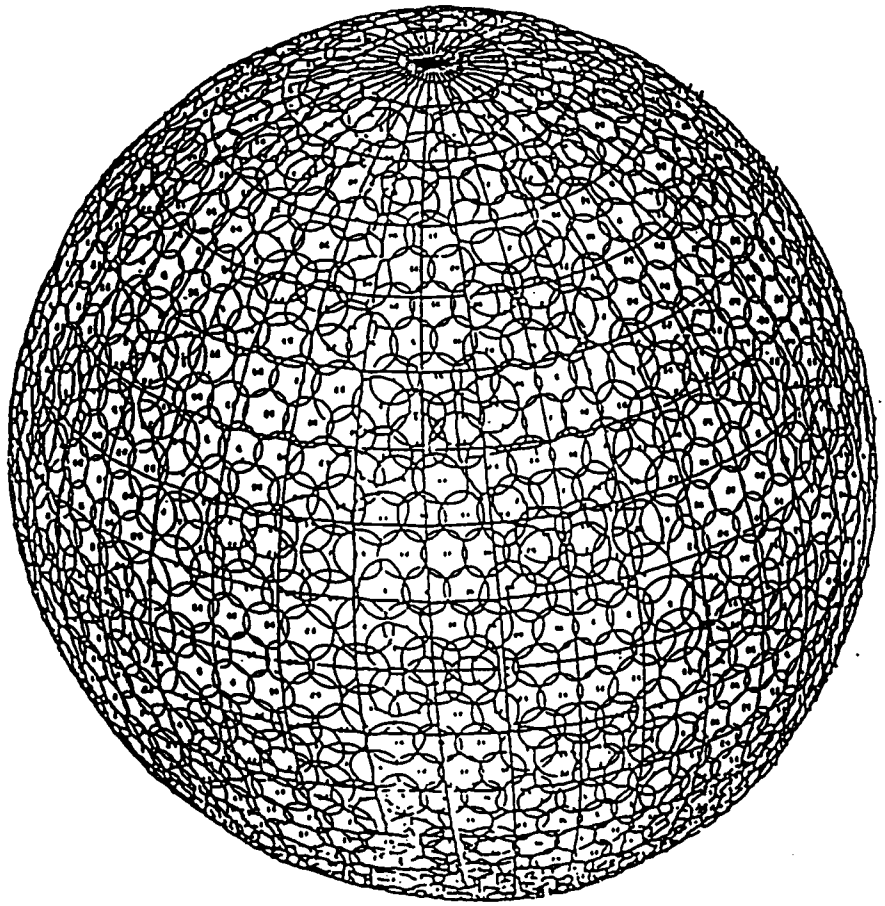


Figure 2.7: Global representation of spot beam coverage

It is obvious that the cell patterns will overlap little in the low latitude (near the equator), and very much in the high latitudes (near the poles). This is due mainly to polar satellite constellation and to a very small extent, the bulging of the earth at the equator (polar radius = 6355 km, and equatorial radius = 6377 km). This causes two problems which are not encountered in cellular telephone, namely;

1) Depending on the satellite's position in orbit (an individual cell will be on at the equator and at some high latitude in each orbit it will be turned off for a period of

time. At any one time approximately 2,150 (68%) of Iridium's 3168 cells are on), the beams must be turned on/off. Obviously, if operation is prohibited in some part of the world, more cells will be turned off [15].

2) The channel reuse pattern must be reset at different times during the orbital period to allow a 12-cell reuse pattern.

Implementing the 12-cell reuse pattern, the same channel can be reused 180 times (2,150 active cells /12-cell reuse pattern = 180). As the satellites approach the poles they handle more satellite-to-satellite data traffic since they are less burdened by the up and downlink activity. It is interesting to know that as the satellites approach the poles, the outer cells are progressively turned off for reasons mentioned previously, and in order to conserve power and control interference with satellites in adjacent planes. Furthermore, as each beam passes over a different geographic region, it has the ability to utilize only those frequencies, and provide only those services authorized in that region.

Replacement satellites can be placed into orbit in about 36 hours. The satellites will be travelling at about 7460 m/s. Each satellite on orbit would cost around $29 million. Due to the relatively low power requirements, the high power amplifiers (HPA) are only rated at 5 Watts peak power. This calls for a microwave solid state amplifier, which is linear compared to a travelling wave tube amplifier (TWTA), thus reducing AM-AM and AM-PM effects inherent in the latter [27].

2.3 Ground Segment

The ground segment consists of the gateways and the user units. The network of satellites which form the backbone of the Iridium allows a call between Iridium users to be routed without using any part of the terrestrial networks. Furthermore, if one of the terminals is not an Iridium terminal, the system will choose a route, through a gateway, which minimizes the usage of the terrestrial network. The completed ground segment of the system will cost around $3,370 million.

2.3.1 User units

Due to the satellite's low altitude and relatively high EIRP and small spot beams, there is a considerable reduction in the RF power requirements of the user units (use of small antennas instead of satellite dishes). The Iridium handheld unit will incorporate a low profile whip antenna (omnidirectional antenna, Quadrifliar Helix to be more precise) and will be designed to comply with the same size, weight and battery life constraints as terrestrial cellular units. Typical handheld units are shown in Figure 2.8.

Figure 2.8: Handheld units for the Iridium system

The user units are rated at 390 mW radiated power, well within the restrictive personal RF radiation limits recommended by World Health Organization (WHO). These portable handheld units can operate for 24 hours on a single recharge (23 hrs standby and 1 hr operation). Some units have an optional GPS capability which enhances the unit's timing, position, and warm-up capabilities; however, the system is not dependent on GPS for operation. A unit without the GPS option is functional within 1 minute from a "cold start". The user unit is capable of monitoring other channels apart of its own preassigned in assisting in the handover operations. Apart from digital voice communication, the handset is also capable of sending and receiving paging, fax, and data messages. This is achieved by using a standard communications port (data interface) located on the side of the unit to allow connectivity to a fax, printer or other data unit similar to features found on today's cellular phones. The estimated cost of a unit is about $3,500, but the price will eventually drop as more units are mass-produced. It is anticipated that dual-mode user units, which will operate both in local cellular and Iridium environment, will be manufactured [28]. The advantage

of such a unit is that the user will be able to operate the unit in cellular-mode when a cellular system is available and, in the Iridium-mode, which is slightly more costly, whenever a cellular system is not available. When a call is made, the dual mode unit will first try to use a cellular channel from the local terrestrial system either the analogue or digital service. The unit will only transmit to an overhead satellite if and only if a terrestrial cellular system is not available.

2.3.2 User Unit Overview

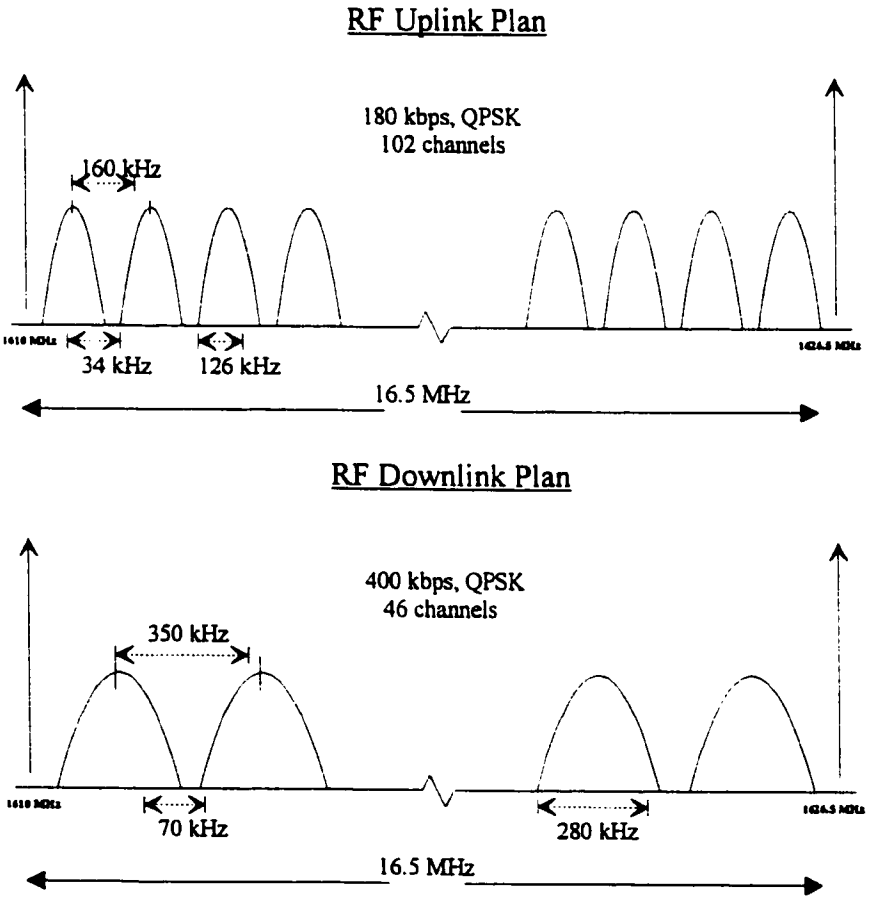


Figure 2.9: User unit RF uplink and downlink plan

The Iridium satellite system utilizes two portions of the frequency spectrum for establishing communication links. The uplink from user to satellite and downlink from satellite to user employs the L-band (1610-1626.5 MHz) corresponding to a total bandwidth of 16.5 MHz. Figure 2.9 shows the user unit RF uplink and downlink plan. The user unit uplink channel bandwidth is 128 kHz and coded data rate of 180 kbps while the downlink channel is 280 kHz at 400 kbps coded data rate. The maximum number of available uplink channels is 64 of which 55 are available for communications and 9 for control. On the other hand, the maximum available downlink channels is 29 of which 25 are used for communications and 4 for control. The four control channels are used specifically to provide the digital speech interpolation (DSI) control information, paging and synchronization of signals. Digital speech interpolation of 2.2:1 is employed on the downlink to achieve some spectral efficiency. There is an equivalent of 55 downlink channels due to the implementation of DSI of 2.2:1 on the 25 downlink channels (2.2x25 = 55). The frequency plan and specification of the user unit is given in Table 2.1.

	User Unit to Satellite	Satellite to User Unit
RF Bandwidth (MHz)	1610-1626.5 (L-band)	1610-1626.5 (L-band)
Number of Channels	102	46
Channel Bandwidth (kHz)	128	280
Channel Spacing (kHz)	160	350
Guard Bands (kHz)	34	70
Maximum Available Channels	64	29

Available Channels for Communications	55	25, with 2.2:1 DSI to obtain 55 Channels
Available Channels for Control	9	4
Coded Data Rate (kbps)	180	400
FEC Rate	¾	¾
Modulation	QPSK	QPSK
BER	10^{-2}	10^{-2}
Polarization	RHCP	RHCP

Table 2.1: Frequency plan and specification of user unit

The type of modulation employed is QPSK which has a constant envelope and generally suitable for satellite communications [29]. The multiple access of Iridium is a combination of TDMA, frequency division multiple access (FDMA), and space division multiple access (SDMA). Effectively, 55 FDMA channels are available, each being a TDMA carrier. SDMA is achieved due to the employment of non overlapping multiple spot beams, therefore, permitting channels to be reused.

Various techniques have been implemented to improve the performance of the system, including the commonly used raised-cosine filtering to reduce the spectral content of each channel and a special 4.8 kbps vocoder as the source encoder. Convolutional forward error correction (FEC) with rate ¾ and constraint length $K = 7$ channel encoding with Viterbi decoding is employed to achieve bit error rate (BER) of 10^{-2} with $E_b/N_o = 3.1$ dB for L-band, which is significantly lower than $E_b/N_o = 6.9$ dB for the Ka-band link with

BER of 10^{-7} cited on page 40.

The QPSK downlink signal has a coded data rate of 400 kbps, which corresponds to a symbol rate of 200 kBaud/s (400E3 (b/s)/2 (b/symbol) = 200 kBaud/s), and a Nyquist RF bandwidth of 200 kHz. This signal is bandlimited by means of raised-cosine filter with a roll-off factor of 40% and therefore the RF signal has a bandwidth of 280 kHz (200E3 x 1.4 = 280 kHz). The raised-cosine filtering provides a downlink efficiency of 1.43 bit/s/Hz (400/280 = 1.43 bit/s/Hz). The pulse shaped QPSK format was chosen to allow tighter effective channel spacings, combat multipath fading inherent in satellite communications, and shadowing due to vegetation [29]. The implementation of all the above digital techniques results in a higher quality voice communication system than achievable with analog systems [14]

Iridium is designed to be bi-directional or duplex system using one carrier per channel. This means that the satellite alternatively transmits and receives on the same frequency band and this is referred to as time division duplex (TDD) or ping pong.

2.3.3 Doppler Shift and Doppler Rate

Iridium makes use of circular polarization to achieve the flexibility with alignment of antennas (due to relative motions of satellite-user unit, and satellite-gateway). Only right hand circular polarization (RHCP) is used, for reasons which will be explained in section 2.3.5. Linearly polarized signals are used for inter-satellite links (intra and inter orbital links)

since the satellites maintain a fixed position relative to each other.

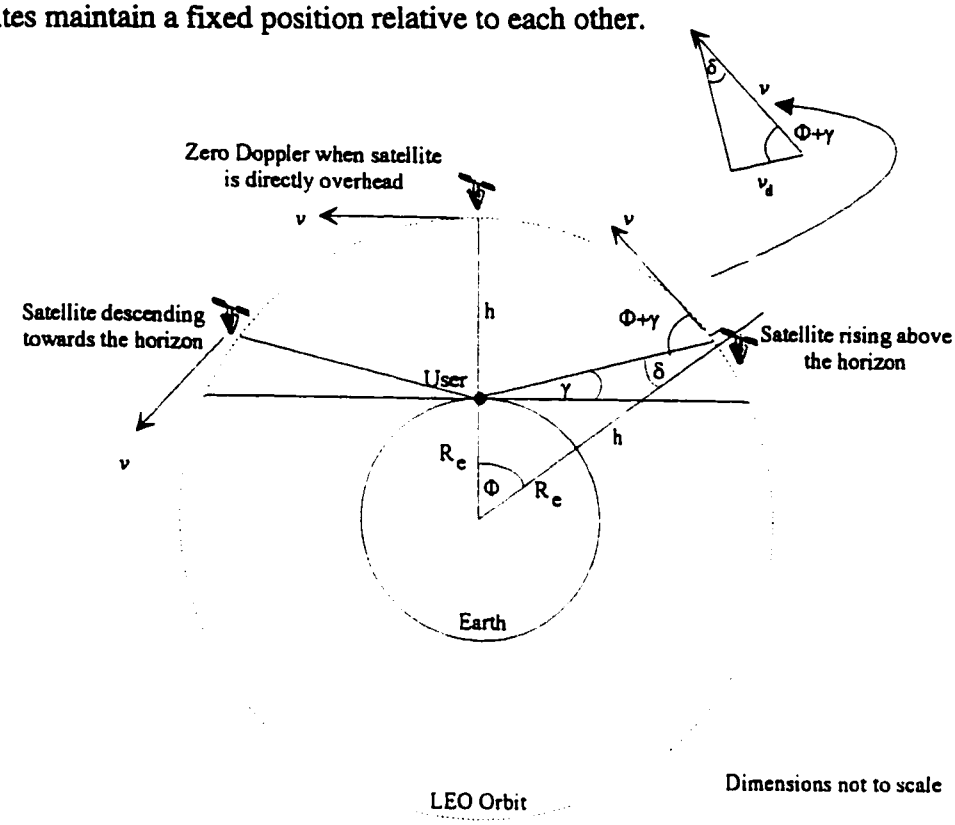


Figure 2.10: Calculation of Doppler frequency shift

The movement of the satellites produces a Doppler shift of the downlink transmissions as seen by the user unit and a corresponding shift of the uplink signals as seen by the satellite. The Doppler shift at any instant in time is directly proportional to the carrier frequency employed and the relative radial velocity between the LEO satellite and the user [30-31]. Due to the relative speed of the satellite with respect to the user, the Doppler effect is substantial and attains a maximum of 35.5 kHz. For illustration purposes Figure 2.10 shows how to calculate the maximum Doppler frequency shift. We make the assumptions that the satellite passes directly overhead, minimum elevation or grazing angle of $\gamma = 8°$ and a mean carrier frequency of 1618 MHz (about half way within the user L-band range). The Doppler

frequency is given by

$$f_d = \frac{v_d}{\lambda} = \frac{v_d f_c}{c} \tag{2.1}$$

where v_d is the relative radial velocity between the LEO satellite and the user, f_c is the carrier frequency and c is the speed of light. A positive Doppler frequency shift implies that the satellite is rising above the horizon whereas a negative Doppler frequency shift implies that the satellite is descending towards the horizon. Zero Doppler occurs when the satellite is directly overhead of the user (satellite at zenith) since there is no relative radial velocity.

From Figure 2.10 we obtain an expression in terms of R_e (mean radius of the earth), h (altitude of the satellite), and elevation angle γ and beamwidth, tilt or nadir angle δ.

$$\frac{R_e + h}{\sin(\gamma + 90^o)} = \frac{R_e}{\sin(\delta)} \tag{2.2}$$

The relative radial velocity between the satellite and user is given by

$$v_d = v \cos(\Phi + \gamma) = v \sin(\delta) \tag{2.3}$$

where Φ is the satellite coverage angle. Using (2.1), (2.2) and (2.3) and simplifying we find that the Doppler frequency shift is

$$f_d = \frac{v f_c R_e \cos(\gamma)}{c (R_e + h)} \tag{2.4}$$

Figure 2.11 shows the variation in Doppler frequency shift for an overhead pass from the instant the satellite is in view, rising above the horizon (minimum grazing angle) until it disappears.

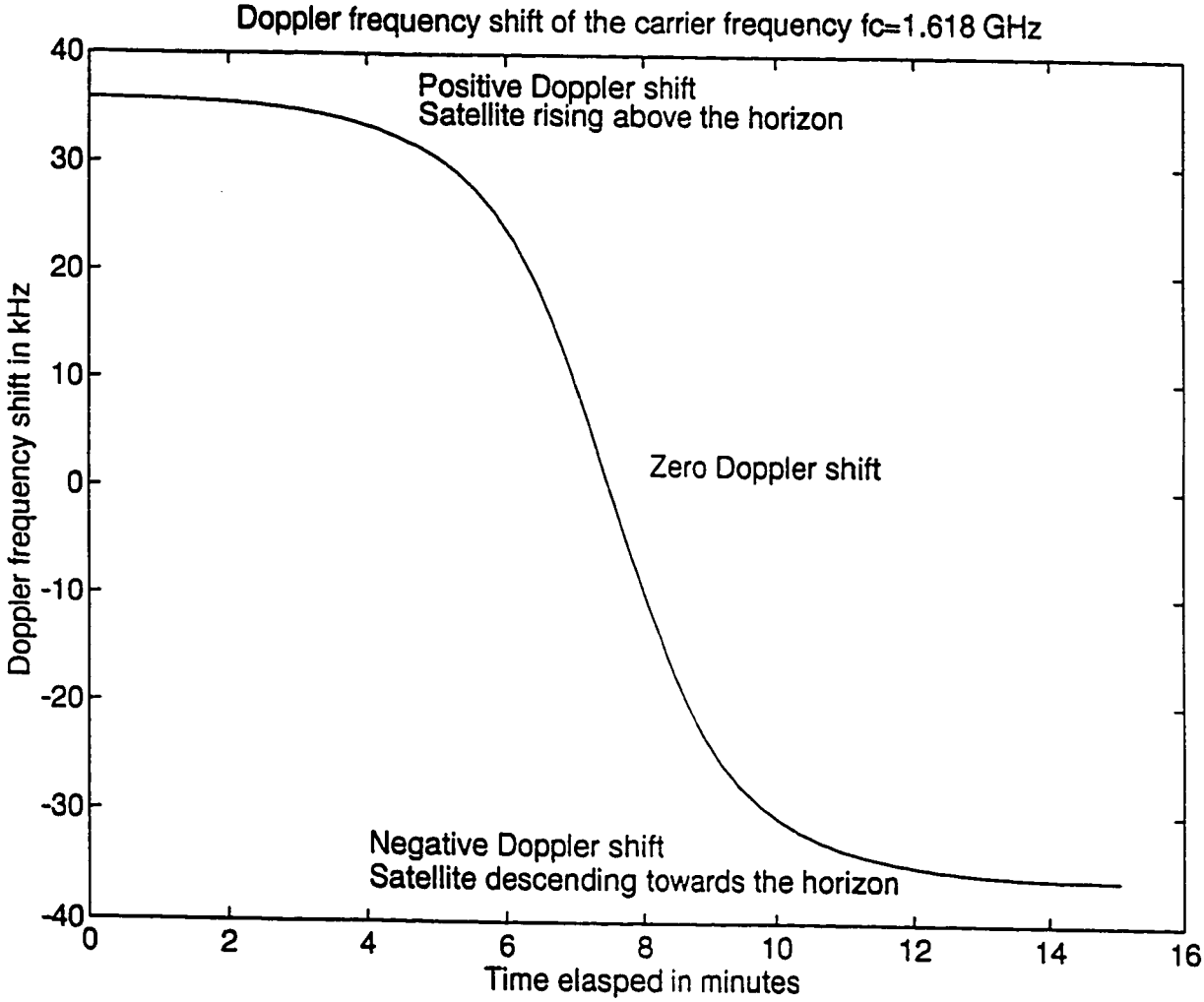


Figure 2.11: Doppler frequency shift of a typical Iridium satellite

By taking the derivative of the Doppler shift curve with respect to time, we obtain the Doppler rate curve, illustrated in Figure 2.12. The absolute maximum Doppler rate is around 340 Hz/s.

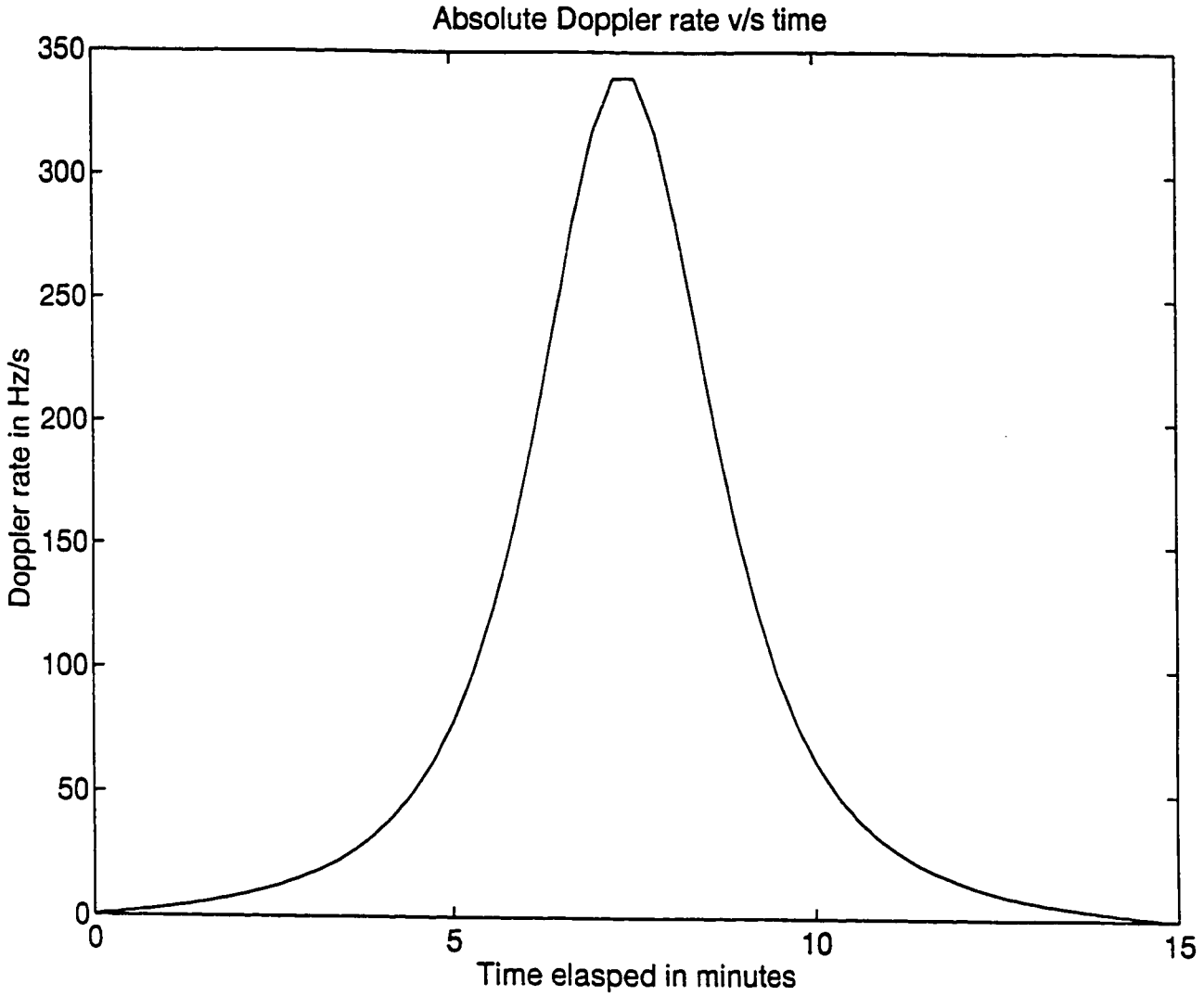


Figure 2.12: Absolute Doppler rate of a typical Iridium satellite

2.3.4 RF Interference Techniques

The obvious way to prevent RF interference would be not to use the frequencies that other satellite systems are utilizing. The Iridium satellites are capable of operating at six

different channels, each of 15 MHz in bandwidth, within the 100 MHz band. The channels may be used in such a way that its frequency is different for any pass where the satellite is pointing at a geostationary satellite (when it passes through a GEO satellite beam). Also, due to the relatively fixed position of all the geostationary satellites, the time duration of the actual impingement of an Iridium satellite beam is relatively short (2-4 seconds maximum). Alternatively, an Iridium satellite may link to another Iridium satellite (taking a longer route than necessary), to reduce the interference to geostationary satellites. This technique allows the Iridium system to route its traffic to one or more gateways to connect to the terrestrial telephone network providing a typical back-up procedure for a gateway failure.

2.3.5 RF Interference Avoidance to Other Systems

The frequency band utilized by radio astronomy is 1610.5-1613.5 MHz. Since the radio astronomy locations are permanent, the Iridium system is capable of avoiding the use of the frequencies in the above range whenever a cell passes over one of these locations. The Soviet navigational satellite system GLONASS uses L-band frequencies up to 1615.5 MHz. Harmful interference will be avoided by not using any Iridium frequency bands in the range of 1610-1616 MHz that are being utilized by GLONASS.

The Iridium makes use of the right hand circular polarization only, as a means of avoiding interference to GEOSTAR which is a CDMA mobile communication satellite system, utilizing left hand circular polarization (LHCP) in the same frequency band

(otherwise frequency reuse could have doubled the capacity).

2.3.6 Typical Receiver Downlink Budget

For simplicity, we assume the satellite is directly overhead and that the carrier frequency utilized f_c = 1618 MHz and downlink intermediate frequency bandwidth B_{IF} = 280 kHz. A typical downlink budget for the handset is calculated as follows.

The EIRP is given by

$$EIRP = P_T \times G_T$$
$$[EIRP] = [P_T] + [G_T] \qquad (2.5)$$

where [x] denotes 10log (x).

Satellite Transmitter Power	$[P_T]$ 5W	6.99 dBW
Antenna Gain / Line Loss	$[G_T]$	5.46 dBi
Satellite EIRP	$[EIRP]$	**12.45** dBW

The free-space loss is given by

$$L = \left(\frac{4\pi \ h_{LEO}}{\lambda} \right)^2 \qquad \textbf{(2.6)}$$

Free Space Loss	20log[(4×π×780.6E3)/λ]	154.47 dB
Polarization Loss*		3.00 dB
Mean Vegetation / Atmospheric Loss		0.80 dB

Total Propagation Loss	$[L]$	**158.27** dB

*Note that a 3 dB polarization is included to account for the fact that while circular polarization is radiated from the satellite, while the receiving antennas are vertically polarized.

The figure of merit of a receiver is given by the gain temperature ratio G/T_s.

Receiver Antenna Gain	$[G_T]$	3.01	dBi
System Noise Temperature	T_s	298.93	K

Receiver G/T_s	$[G/T_s]$	**-21.75**	dB/K

The carrier to noise power ratio (C/N) is given by

$$C/N = EIRP \times (1/L) \times (G_R/T_s) \times (1/(kB_{IF}))$$
$$[C/N] = [EIRP] - [L] + [G_R/T_s] - [k] - [B_{IF}] \quad \textbf{(2.7)}$$

where the Boltzmann's constant k = 1.38E-23 J/K can also be written as $[k]$ = -228.6 dB J/K. Conversely, the carrier to noise density ratio is expressed as

$$C/N_o = (C/N) \times B_{IF}$$
$$[C/N_o] = [C/N] + [B_{IF}] \quad \textbf{(2.8)}$$
$$[C/N_o] = [EIRP] - [L] + [G_R/T_S] - [k]$$

Boltzmann's Constant	$[k]$	-228.6	dB J/K

The carrier to noise density ratio (C/N_o) without interference is given by

C/N_o (without C/I)	$[C/N_o]$	61.03	dB

where C/I in the carrier to interference ratio. The overall carrier to noise density ratio in the presence of interference is given by

$$\left(\frac{C}{N}\right)_{received} = \left(\frac{1}{(C/N)} + \frac{1}{(C/I)}\right)^{-1} \quad \textbf{(2.9)}$$

C/I (Downlink)	$[C/I]$	18.00	dB

Received $(C/N_o)_r$	$[C/N_o]_r$	**60.73**	dBHz

The bit energy to noise density ratio is given by

$$[E_b/N_o] = [EIRP] - [L] + [G_R/T_s] - [k] - [R_b] \quad \textbf{(2.10)}$$

Where R_b is the uncoded data rate. Furthermore, the following relationship can be used to convert E_b/N_o to C/N_o.

$$C/N_o = (E_b/N_o) \times R_b$$
$$[C/N_o] = [E_b/N_o] + [R_b] \quad \textbf{(2.11)}$$

Required E_b/N_o (BER 10^{-2})	$[E_b/N_o]$	3.10	dB
Channel Data Rate (Coded)	R_c FEC rate 3/4	400	kbps
Channel Data Rate (Uncoded)	R_b	54.77	dBHz
QPSK Modulation		----------	
Required $(C/N_o)_R$	$[C/N_o]_R$	**57.87**	dBHz

The link margin can be calculated as follows

$$Link\ Margin = [C/N_o]_{received} - [C/N_o]_{required} \quad \textbf{(2.12)}$$

Subscriber Downlink Margin $[C/N_o]_t$ - $[C/N_o]_R$	**2.86**	dB

2.3.7 Gateways (Earth Stations)

A gateway is a ground station that handles call setup, caller location, controls the users access to the system, and collects necessary data to support the billing. The initial system configuration will utilize a minimum of 20 gateways which could be expanded to 250. Access to the system, and telemetry, tracking, and command (TT&C) will be co-ordinated through two Systems Control Centres, which may be co-located at gateway stations. The Master Control Centre will be located in Virginia near Washington, DC and the Backup Control Centre in Italy [9, 23]. These Control Centres will be located at high latitude to allow more satellites to be in view at any time, and to be in direct contact with the each satellite for longer time durations.

Each gateway employs a minimum of three 3.3 metre tracking dish antennas that are separated by up to 32 km as shown in Figure 2.13 [32].

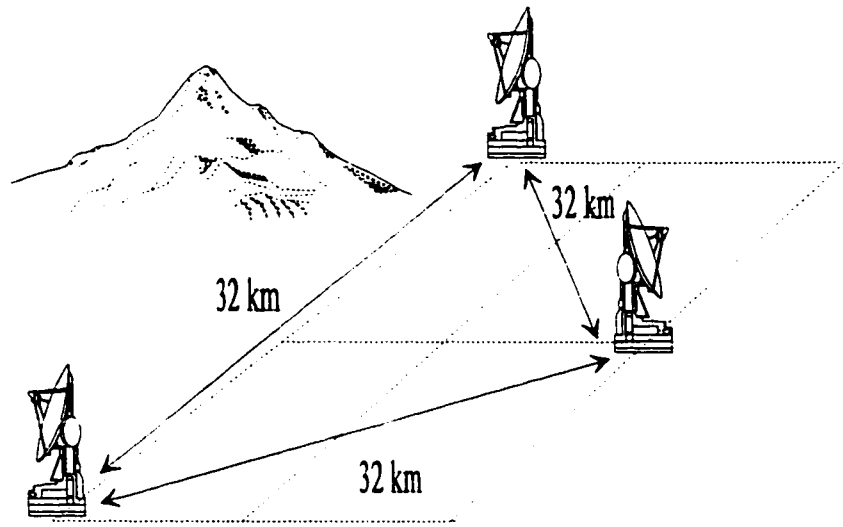


Figure 2.13: Geometrical layout of the gateways

Signal fading is a problem in satellite communications and is mainly due to rain fade on the received signal. This can be alleviated by employing space diversity which is a technique that makes use of two or more receive antennas at different locations as shown in Figure 2.14. The distance is based on the fact that it is rare to have a fade in all antennas at the same time; hence, switching to the antenna with the strongest received signal solves the fading problem.

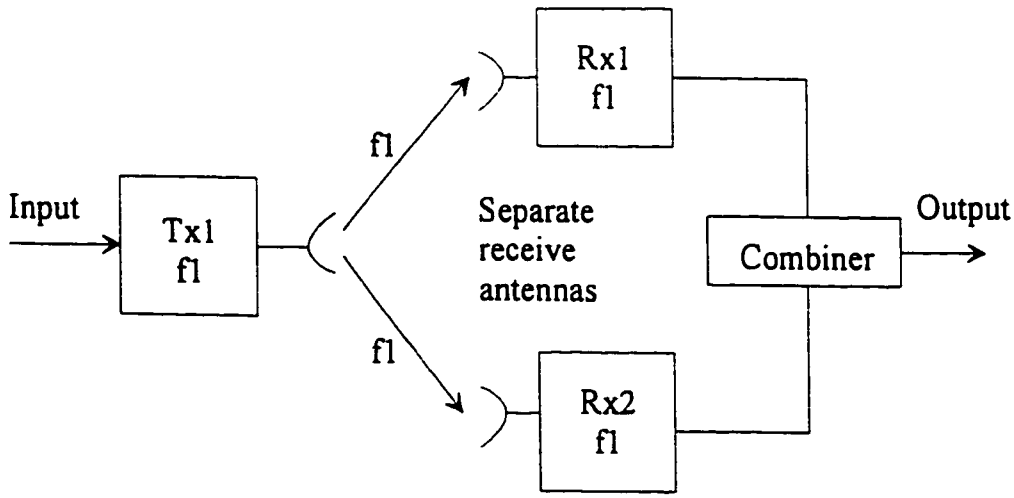


Figure 2.14: Space diversity using two receive antennas

To the current satellite, two antennas provide space diversity, which corresponds to a better link during fading. As the current satellite disappears over the horizon, the third antenna is used to provide a link to the next satellite that appears at a different location above the horizon. This technique allows for smooth handover. Finally, gateways with four antennas ensures a communication link even during maintenance activities.

The gateways are connected to the public switched telephone network (PSTN) which allows the connectivity between non-Iridium and Iridium users.

2.4 Satellite Links

Table 2.2 shows the frequency plan and specification of satellite and gateways.

	Intersatellite Link	Uplink	Downlink
RF Band (GHz)	22.55-23.55 (Ka)	27.5-30.0 (Ka)	18.8-20.2 (K)
Number of Channels	8	6	6
Channel Bandwidth (MHz)	25	15	15
Coded Data Rate (Mbps)	25	12.5	12.5
FEC Rate	½	½	½
Modulation	QPSK	QPSK	QPSK
BER	10^{-7}	10^{-7}	10^{-7}
Polarization	Vertical Polarization	RHCP	RHCP

Table 2.2: Frequency plan and specification of satellites and gateways

The 66 satellites are interconnected by Ka-band crosslinks which lie in the frequency range 22.55-23.55 GHz thereby forming a global network in space. Up to 8 channels are available for these links, each occupying a bandwidth of 25 MHz at coded data rate of 25 Mbps. Hence the total effective usable bandwidth is 200 MHz (8x25=200 MHz) from a total of 1 GHz (23.55-22.55=1 GHz). The uplink from gateway (ground station) to satellite utilizes the Ka-band (27.5-30 GHz). Up to 6 channels are available, each occupying a bandwidth of 15 MHz at coded data rate of 12.5 Mbps. The total usable uplink bandwidth is 90 MHz (6x15=90 MHz) from a total of 2.5 GHz (30-27.5=2.5 GHz). The downlink from satellite to gateway employs the K-band (18.8-20.2 GHz). A maximum of six 15 MHz bandwidth channels at 12.5 Mbps are available, giving a total usable bandwidth of 90 MHz from a total of 1.4 GHz (20.2-18.8=1.4 GHz). The satellite-to-satellite links and up and downlinks for the gateways employ QPSK which is also the most commonly used modulation in satellite communications. Convolutional channel encoding with FEC rate of ½ with constraint length K=7 and Viterbi decoding is used, thus allowing the links to achieve BER of 10^{-7} with E_b/N_o =6 .9 dB.

Figure 2.15 below shows the inter-satellite links that a satellite has to maintain to allow full connectivity of the Iridium space network.

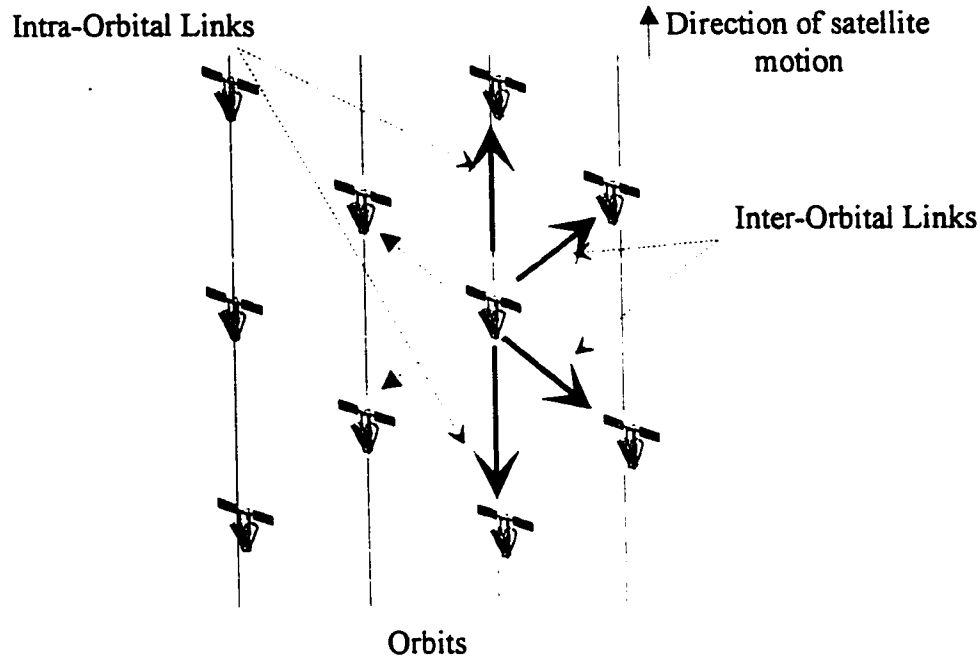


Figure 2.15: Intra-orbital and inter-orbital satellite Links

There are 11 satellites per orbit and each has to maintain forward and backward looking Ka-band links to two satellites in the same orbital plane. These links are referred to as intra-orbital links and are about 4033 km long. Since the satellites' positions in the same orbit are relatively fixed to one another, fixed wave guide slot antennas are used for these links. This also explains the reason for the use of vertical polarization for the intra-orbital links. Each satellite also maintains links to two of the nearest satellites in each adjacent planes, thereby making a total of up to 4 cross-links per satellite. Inter-satellite links to the adjacent orbits are referred to as inter-orbital links and may vary to a maximum distance of 4388 km. Mechanically steered wave guide slot antennas are employed for these links as the links are changing orientation dynamically.

Furthermore, up to four independent gateway links must be maintained per satellite.

The up and down links for the gateways, and user units employ RHCP. More will be said about this choice of polarization. The satellites include base-band processors capable of 3 giga flops, which feasibly permits them to perform intelligent call routing procedures, thus the reason for referring to them as being "smart". Therefore they are not considered as "bent pipes" or conventional non-regenerative repeater satellites.

2.5 General System Overview

2.5.1 Services Offered

The Iridium system is designed to provide a global high grade voice and data communication service on a continuous basis. There are five basic types of services offered by the Iridium system.

1) High quality digital voice communication service using duplex (two-way) 4.8 kbps.
2) Data service using 2.4 kbps modem.
3) Paging service.
4) Facsimile service using two types of mobile facsimile units. A stand alone unit, and one used with the Iridium telephone.
5) RDSS service, and two-way messaging. The former service is available on all types of user units except some types of pagers, which includes geolocating (latitude, longitude and altitude) the user and time. The latter service is used to report the

user's position and relay (transmit and receive) short messages. Paging messages can also be left when voice linkage may not be possible, and the system does not guarantee building penetration.

The system is designed with 16 dB fade margin which permits communication in a variety of harsh fading environments such as foliage and inside a vehicle. The user unit will be able to store incoming faxes into memory for which a hard copy could be obtained later. Alternatively, the fax could be viewed by scrolling down the liquid crystal display (LCD) display screen on the user unit. The geolocation accuracy of the Iridium system is 1.8 km compared to GPS which has an accuracy of 16 m. Some user units will have a built-in GPS chip to enhance its accuracy.

All the above services will find applications in remote areas where it has not been feasible to provide reliable communication systems in the past due to either economic, infrastructure requirements, or size limitations. It is anticipated that if the usage charge, which will be set by the Iridium gateway operators, drops to about $2.00 per minute, then the Iridium system will definitely replace traditional HF services, and geostationary mobile satellite service such as INMARSAT, and MOBILESAT (Australian domestic mobile service provided by AUSSAT). The Iridium system is estimated to break even financially with minimum infrastructure, at about ½ million users worldwide. Currently there are about 30 million cellular users worldwide and it is predicted that this will exceed 100 million by the year 2001 and the total number of Iridium users worldwide will reach 6 million.

From Table 2.1 on page 18, we find there are 55 communication channels per cell. Each carrier (channel) is divided into 24 time slots using TDMA as shown in Figure 2.16.

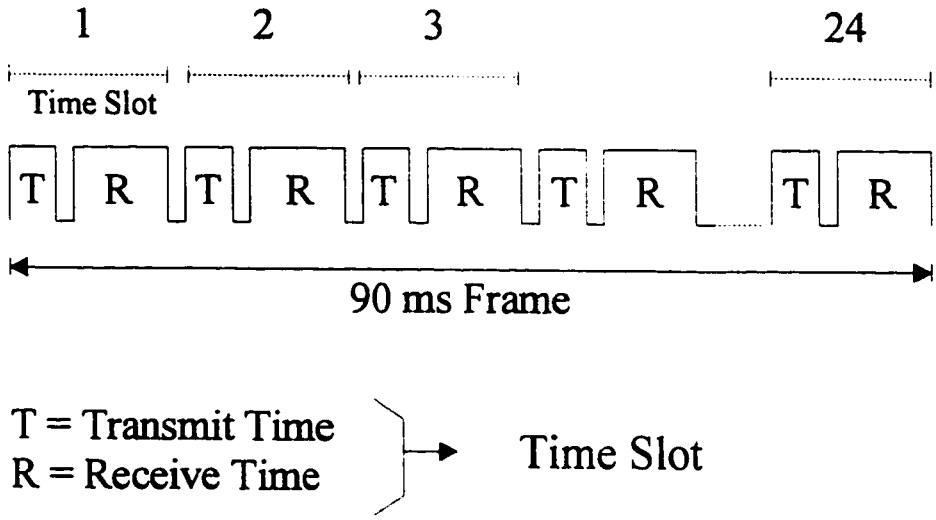


Figure 2.16: Iridium TDMA frame format

Groups of 24 consecutive time slots form TDMA frames, with duration 90 ms. A transmission and reception occupies one time slot position within a TDMA frame. Since Iridium employs a 12-cell frequency reuse pattern, on average each cell will be assigned 2 TDMA slots (24/12 = 2). There are 2150 active cells worldwide and Iridium has a worldwide capacity of 55x2x2150 = 236500 full duplex voice channels. This corresponds to 3583 full duplex voice channels per satellite (236500/66 = 3583).

2.5.2 Routing of Calls

The user unit has non-volatile memory in which invaluable information such as

registration number, telephone number, date, and time will be stored. The unit is referred to be in "standby mode" when it is turned on, and it locks onto the strongest satellite signalling channel. This allows the unit to geolocate itself. The normal mode of operation would be to either make a call or leave the unit in "standby mode", receive a page, then return the call.

2.5.3 How to Place a Call from an Iridium Unit

The Iridium user dials the appropriate number and depresses the send button to start the call setup mode. This enables the unit to send its own phone number, the number being called, and the user's registration number directly to the nearest overhead satellite. Figure 2.17 shows a simplified illustration of a typical call routing procedure to either an Iridium or non-Iridium user. The satellite then passes this setup information to the regional gateway. The regional gateway is the gateway closest to the user at the time the call is initiated. Then the regional gateway passes this information to the caller's home gateway to verify the caller's validity. The home gateway is the gateway where the user activated the user unit. After validation, the regional gateway determines the optimum (shortest) call routing, through the constellation of satellites, to the destination.

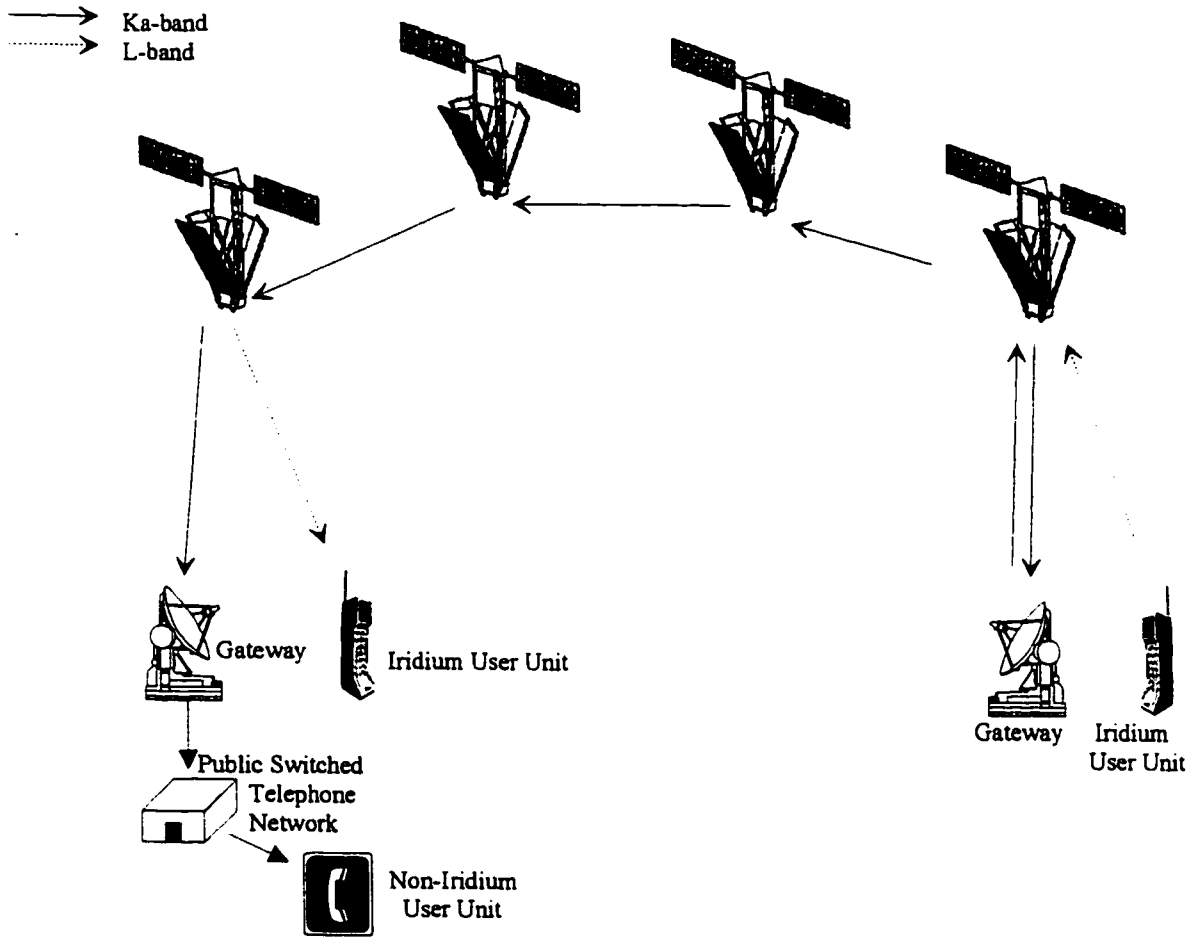


Figure 2.17: A simplified illustration of a typical call routing procedure

If the destination is a non-Iridium user (cellular or terrestrial telephone), the destination will be the closest gateway to the non-Iridium user. Conversely, if the destination is an Iridium user, its own home gateway must validate the caller, and provide his location. This is possible because a user location database at each home gateway is continuously being updated as the users change their location. Both the caller and destination setup information is sent to the regional gateway. This allows the regional gateway to provide optimum routing through the constellation of satellites. The call setup procedure usually takes no more than 30 seconds compared to 10 seconds in the cellular environment [10]. Once the call setup is

completed, the minimum one-way delay associated with Iridium is 10.4 ms (assuming 4 user unit-to/from-satellite links) compared to 270 ms using geostationary satellites (between two points within ⅓ of the earth's surface) as shown in Figure 2.18. However, the largest one-way propagation delay through the constellation of Iridium satellites is always less than 150 ms (assuming 4 user unit-to/from-satellite, and 10 satellite-to-satellite links).

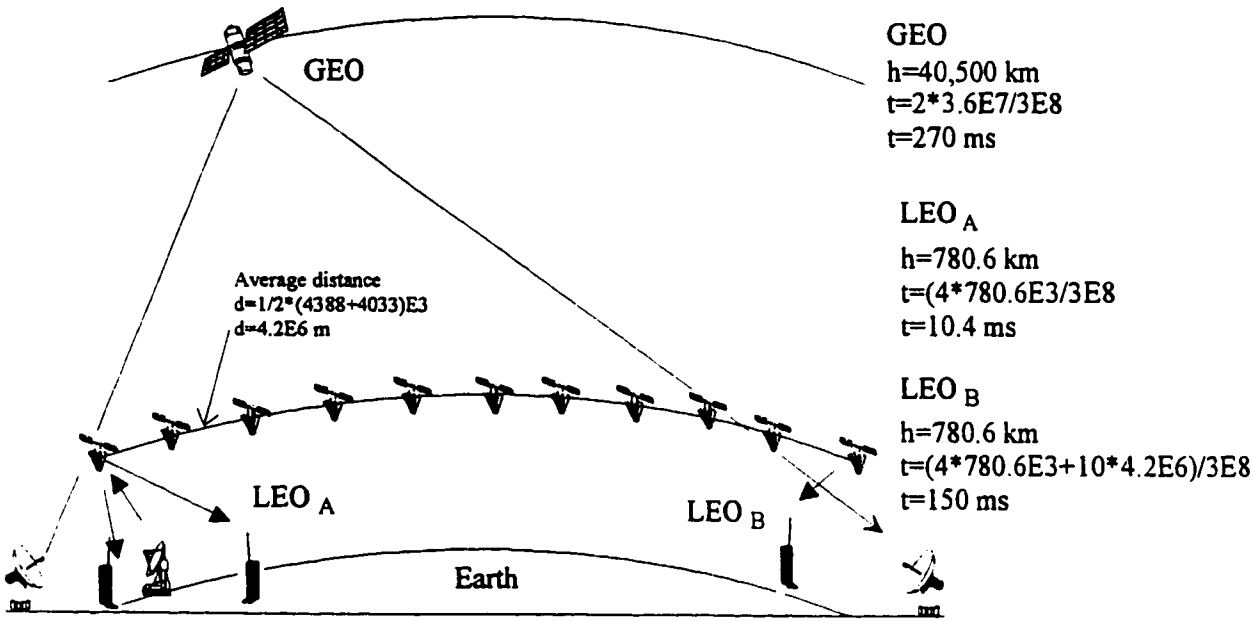


Figure 2.18: Propagation delays in satellite communications

The longest propagation delay over domestic terrestrial networks is around 25 ms and up to 100 ms for transoceanic cables. The delay in geostationary satellites is approximately 10 times more than longest delays in modern domestic terrestrial telephone circuits. The round-trip delay in geostationary satellite systems is about 540 ms, and adaptive echo cancellation

is required for echo control. Since the delay associated with the Iridium system is much less, there might be a need for an echo suppressor/canceller which could be implemented on a very large scale integration (VLSI) chip [33-35]. In the calculations above we have not accounted for the overhead in processing time for voice coding and decoding and other on board processing (OBP) [36].

2.5.4 Call Routing from a Non-Iridium User and Back-up Technique

A user from a public telephone can also place a call to an Iridium user. The call will be setup by entering the Iridium system through the regional gateway closest to the caller. The Iridium user's home gateway will validate whether the user is still current and provide information about its last known location. Then the setup process proceeds as previously discussed.

In an event of a gateway failure, the system will find an alternate route to complete the communication link via another gateway. Conversely, if a satellite fails, there are two possible solutions. The simplest remedy would be to reroute the data traffic to the nearest satellite to complete the link. This solution will provide a slightly longer route hence a longer one-way delay in call. The last option would be to replace the satellite in 36 hours. In this process, since the Iridium satellites are made of special materials, the faulty satellite is brought to a lower orbit, where it burns up on re-entry into the earth's atmosphere.

2.5.5 Balancing the System Cost

At higher altitudes and in Van Allen radiation belts, semiconductors must withstand harsh ultra-violet radiation. At GEO radiation-tolerant semiconductors are used and they are expensive. The altitude was chosen to be below 1,100 km (just below the first Van Allen belt) because the radiation environment would drive up the hardware cost. On the other hand, altitude below 375 km would obviously cause excessive fuel requirements for station keeping, due to the increase in atmospheric drag and as a consequence lessen the orbit lifetime. Last but not the least, LEO was chosen because of its primary advantage in reducing path loss. Refer to Figure 2.19 for transponder power comparison of GEO and LEO satellite systems.

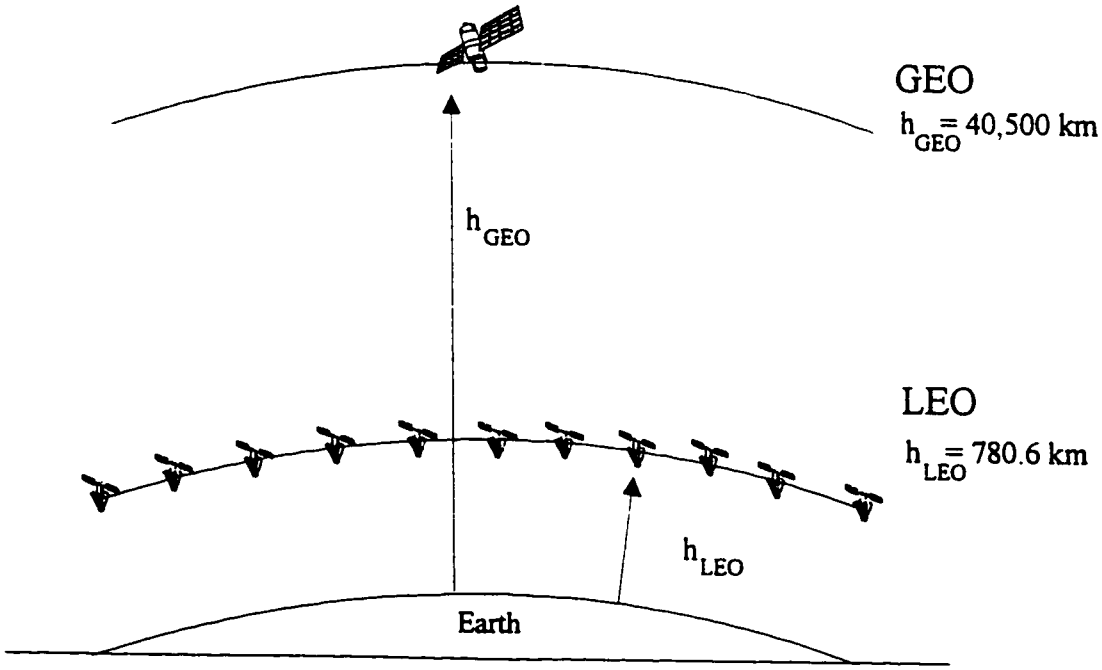


Figure 2.19: GEO and LEO transponder power comparison

The power at the receiver input is given by the relationship

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi h}\right)^2 \quad \textbf{(2.13)}$$

where P_T is the transmitted power, G_T is the transponder antenna gain, G_R the receiver antenna gain, λ the carrier wavelength and h the altitude of the satellite. For similar received signal power, transponder and receiver antenna gain, and carrier wavelength, we obtain a relationship comparing the transponder power required for GEO and LEO satellite systems (assuming the satellites are directly overhead).

$$\frac{P_T(GEO)}{P_T(LEO)} = \left(\frac{h_{GEO}}{h_{LEO}}\right)^2 = 2692 \quad \textbf{(2.14)}$$

Therefore, the power required for an Iridium satellite transponder is more than 2000 times (34 dB) less than geostationary systems [37]. This also corresponds in an overall reduction in the size of the satellite even taking into account the extra payload for the base-band processing and signal processing to maintain the LEO time-varying geometry constellation.

The three phase array antennas each producing 16 beams are fixed to the sides of the satellite and do not need to be deployed in space like other types of antennas. Most phased array antennas are used in situations when the antenna beam can be electronically steered by using beam forming techniques. But the phased arrays antennas on board the Iridium satellites are used to generate 48 contiguous narrow beams per satellite. The beams are not electronically steerable, thus keeping the satellite antennas less complex. For all the reasons

mentioned above, this translates into minimizing the cost of satellite production (assembly line production), launch cost, and allows feasible upgrades to increase the system's capacity, and to implement improvements resulting from technological advancements [7, 23].

2.5.6 Worldwide Roaming

In a cellular telephone system, when a user is in a remote city, calls cannot be initiated or received unless the user is in roam mode. Roaming is a term used, when a user in a remote service area can initiate or receive a call and the service will be billed via the caller's hometown cellular company. Even today, roaming using a cellular telephone is not possible at the European level due to the various cellular standards that those countries adopt. At the moment there are about 15 different standards for analog and digital cellular systems [38]. In Europe, these systems are being phased out by the introduction of the GSM, which is a digital system. However the digital systems in North America and Japan are the North American digital cellular (NADC) and Japan digital cellular (JDC), respectively. This solves the roaming problem at the continent level but worldwide roaming is still not possible. Since the Iridium satellite system is a worldwide communication system, worldwide roaming will be possible.

2.5.7 Other Proposed Systems Competing with Iridium

	Constellation	Ellipsat	Loral Qualcomm	I-CO	Iridium Inc.	TRW	GSC Teleglobe	Teledesic Corp.
Satellite System	Aries	Ellipso	Globalstar	Inmarsat-P	Iridium	Odyssey	OrbComm	Teledesic
Orbit Class	LEO	HEO + MEO	LEO	MEO	LEO	MEO	LEO	LEO
Altitude (km)	1018	[7846x520]+[8040]	1398	10,335	780	10,373	775	700
Number of Satellites	48	16	48 8 Spare	10 2 Spare	66 6 Spare	12 3 Spare	26 10 Spare	840 84 Spare
Inclination (°)	90	[116.5] + [0]	52	45	86	55	[45] + [90]	98.2
Number of Planes	6	[2] + [1]	8	2	6	3	[3] + [2]	21
Satellites / Plane	8	[5] + [6]	6	5	11	4	[8] + [1]	40
Multiple Access Technique	CDMA	CDMA	CDMA	TDMA	TDMA/FDMA/ TDD	CDMA	N/A	FDMA (Uplink) TDMA (Downlink)
Coverage	Global	North of 50° South Latitude	Within ± 70° Latitude	Global	Global	Major Land Masses	Within ± 65° Latitude	Global
Type of Satellite	Bent Pipe	Bent Pipe	Bent Pipe	Bent Pipe	Processing	Bent Pipe	Bent Pipe (S&F)	Processing
Satellite Mass (kg)	520	650	450	1925	700	2207	43	771
Beams / Satellite	7	61	16	163	48	37	N/A	64x9=576
System Cost ($ Billion)	0.29	0.7	1.7	2.6	3.37	1.8	< 0.25	9
User Terminal ($)	1,500	1,000	750	Few Hundred	3,000	550	100 -300	N/A
Call Rates ($)	0.50	0.50	0.60	2.00	3.00	0.65	N/A	0.04
Service Types	Voice Data Fax	Voice Data Fax Paging Messaging Position Location	Voice Data Fax Paging	Voice Data Fax Paging	Voice Data Fax Paging Messaging Position Location	Voice Data Fax Paging Messaging Position Location	Data Fax Paging	Voice Data Fax Paging Video
Data (kbps)	2.4	0.3 - 9.6	9.6	2.4	2.4	9.6	2.4 Uplink 4.8 Downlink	16 - 2048
Voice (kbps)	4.8	4.8	2.4 / 4.8 / 9.6	4.8	4.8	4.8	No Voice Service	16
Uplink Frequency Band	L - Band	L - Band	L - Band	S - Band	L - Band	L - Band	VHF - Band	Ka - Band
Downlink Frequency Band	S - Band	S - Band	S - Band	L - Band	L - Band	S - Band	VHF - Band	Ka - Band

Table 2.3: Proposed satellite systems

In addition to the four proposed LEO satellite systems mentioned which are competing with Iridium, there are other proposed MEO and HEO satellite systems as well. Table 2.3 shows some of their characteristics.

2.6 Conclusions

In this chapter a detailed analysis of the Iridium satellite system is presented. Iridium makes efficient use of the limited and valuable frequency spectrum by implementing advances from both cellular and satellite communications. The major contributions for this efficiency are the use of digital speech interpolation, time division duplex, multiple spot beams and 12-cell frequency reuse which allows the implementation of a multiple access which is a combination TDMA, FDMA, and SDMA. Various other technical aspects of the system have been found to be particularly interesting, namely the RF interference avoidance techniques utilized, type of beam polarization, type of digital modulation, call routing procedure, and system backup procedures.

It appears quite likely that the penetration of Iridium will be more pronounced in; (1) many developing countries with no mobile telephone service, (2) sparsely populated areas around the world for thin-route communications, and where the infrastructure for mobile and/or non-mobile terrestrial telephone is too expensive, (3) areas where a more reliable long distance communication service replaces the HF terrestrial systems, and aeronautic applications, and (4) areas where a better search and rescue (SAR) service is required (for off-shore fishing countries, maritime, forest, and mountain related activities areas). In the case of SAR, Iridium will be able to provide more information than currently used emergency position indication radio beacons (EPIRB), which utilizes SARSAT.

In the third world countries, especially those with telephone density of less than 5

telephones per 100 people, the simplest and most economical infrastructure required to provide the service would be an Iridium terminal in a phone booth which could be solar powered [7, 9]. On the other hand, in developed countries, the Iridium system will bring instant communication service to everybody's fingertips at a reasonable price. If the user unit, and usage fee becomes more competitive than the cellular telephone rate, Iridium will eventually get a bigger share of the market.

Since the Iridium satellites are relatively inexpensive, it would be feasible to upgrade and expand the system whenever satellites are replaced. Clearly for developing countries to reach their full potential, they must have access to advance communication capabilities. Laying of copper wires and/or optical fibre is fine for developed countries but the cost involved and most importantly, the time required to bring developing countries to the same capabilities prevent them from competing on an equal economic basis. A partial solution to this problem is global wireless interconnection [39-43]. Iridium happens to provide an instant PCS where narrow-band (Narrow-band is defined as $\leq$ 33.6 kbps) communication is concerned.

Even though the Iridium satellite system has some very attractive features for today's needs, there exist technical issues peculiar to the system. Iridium has a fundamental system propagation delay of 5.2 ms which constitutes one user-satellite up and downlink. In addition, an SNR improvement of over 2000 times over GEO systems, is another advantageous feature inherent to LEO systems, such as Iridium. However, it has been shown that the typical Doppler frequency shifts in the range of 35.5 kHz and absolute Doppler rate

or 340 Hz/s are expected for the L-band QPSK signal. Furthermore, due to the dynamic constellation of the satellites, the coverage time of each satellite beam footprint is 10 minutes while that for a single cell is between 60-90 seconds. Therefore, the system requires frequent handovers.

One of the most important issues is the effect of Doppler shift, whose value and rate are large compared GEO satellite systems. In the following Chapters we investigate the problems encountered due to the presence of Doppler shift and techniques for Doppler correction to achieve fast lock-ups, which is necessary to make the system efficient.

CHAPTER 3

Novel Technique for Doppler Compensation

In this Chapter, we establish the relationship between Doppler frequency shift and Doppler rate, and satellite cross track angle. Furthermore, we introduce the novel technique of Doppler compensation as well as a refined version for the overhead pass case. The novel technique has the effect of reducing Doppler frequency shift to tolerable levels whereby conventional methods can be utilized to achieve carrier synchronization. The overall gain is a considerable reduction in lock-up time of the Doppler shifted signal.

3.1 Background

In Chapter 2, the Doppler shift and Doppler rate curves were obtained from a very

basic simulation model consisting of a user at the equator and a polar LEO satellite which passes directly overhead. The cross track angle of a satellite is the angle subtended between the orbital planes of the satellite in question and the orbit of an over head passing satellite. In other words, an over head passing satellite has a cross track angle of zero. This system was shown to provide maximum duration period for viewing the satellite, within the allowed elevation angles, and also produce the highest Doppler shift and Doppler rate. In the real world, this is not always the case. Often the satellite passes by the user at a non-zero cross track angle. This section is devoted to deriving the expressions relating the Doppler frequency shift and Doppler rate to the satellite cross track angles and user position.

3.1.1 Earth Coverage by Satellite at Altitude h

Iridium provides single-satellite coverage with a minimum elevation or grazing angle $\gamma = 8°$. The geometry of the earth coverage by an LEO satellite is shown in Figure 3.1.

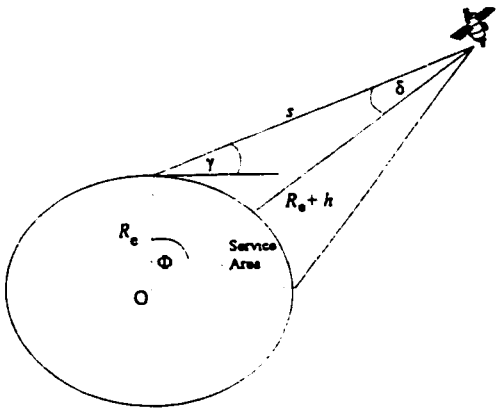


Figure 3.1: Earth coverage by an LEO satellite at altitude h

The beamwidth or tilt/nadir angle of the satellite is given by [4, 44-49]

$$\delta = \arcsin\left[\frac{R_e}{(R_e + h)}\cos(\gamma)\right] \tag{3.1}$$

The slant range is given by [46-47, 49-50]

$$s = \sqrt{R_e^2 + (R_e + h)^2 - 2R_e(R_e + h)\cos(\Phi)} \tag{3.2}$$

Since $\delta = 90^o - (\Phi + \gamma)$, the maximum satellite coverage angle is given by

$$\Phi_{max} = \arccos\left[\frac{R_e}{(R_e + h)}\cos(\gamma_{min})\right] - \gamma_{min} \tag{3.3}$$

Using $R_e = 6366$ km, $h = 780.6$ km and $\gamma = 8°$, we get $\delta = 61.90°$, $s = 2480$ km and $\Phi_{max} = 20.10°$.

3.1.2 Geometry of LEO Satellite

To provide insight on the extent of Doppler shift and Doppler rate consider the geometry given in Figure 3.2. The basic trigonometric formulas needed are the laws of cosines and sines for plane and spherical triangles which are given in Appendix B for convenience.

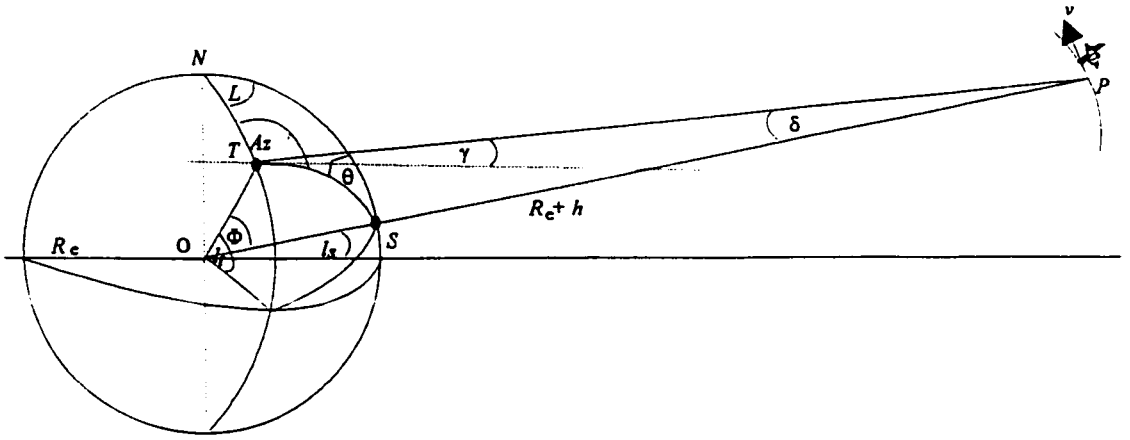


Figure 3.2: Satellite geometry

Here we assume that the earth is spherical with the same mean radius as used previously.

R_e: Mean earth radius

h: Altitude of satellite

γ: Elevation angle (angle above the horizon at which the user terminal sees the satellite)

Az: Azimuth angle of the user terminal ($\measuredangle$ *NTS* measured through east from north)

Φ: Coverage angle ($\measuredangle$ *TOS*, always non-negative)

δ: Beamwidth angle (also known as tilt or nadir angle)

s: Slant range from satellite to user terminal

l_s: Latitude of the subsatellite point

l_t: Latitude of user terminal

L: Longitude difference between subsatellite point and user terminal

P: Location of Satellite

T: Location of user terminal

S: Subsatellite point

The subsatellite point is the place where a line drawn from the center of the earth to the satellite passes through the earth's surface. From the spherical triangle $\triangle NTS$, redrawn in Figure 3.3 and the cosine law for sides (spherical triangles), the coverage angle Φ is given by [49-50]

$$\begin{aligned}\cos(\Phi) &= \cos(90-l_t)\cos(90-l_s) + \sin(90-l_t)\sin(90-l_s)\cos(L)\\ \cos(\Phi) &= \sin(l_t)\sin(l_s) + \cos(l_t)\cos(l_s)\cos(L)\end{aligned} \tag{3.4}$$

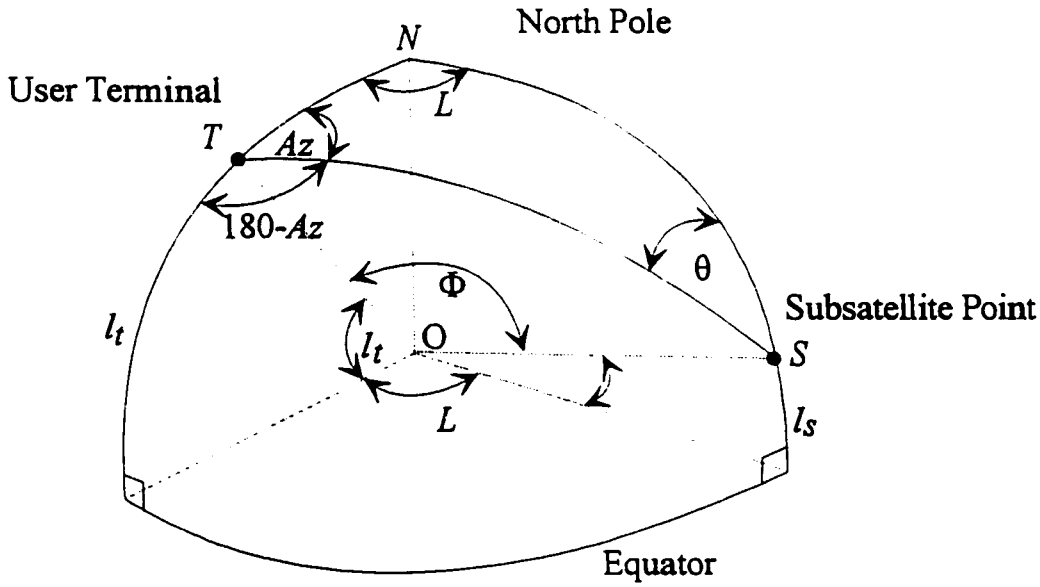


Figure 3.3: Spherical triangle $\triangle NTS$

For simulation purposes l_t and L are fixed while l_s is allowed to vary. The slant range is determined by the law of cosines applied to the plane triangle ΔTOP which is given in Figure 3.4 and using equation (3.4) we get

$$\begin{aligned} s &= \sqrt{R_e^2 + (R_e + h)^2 - 2R_e(R_e + h)\cos(\Phi)} \\ s &= \sqrt{h^2 + 2R_e(R_e + h)\,(1 - \sin(l_t)\sin(l_s) - \cos(l_t)\cos(l_s)\cos(L))} \end{aligned} \tag{3.5}$$

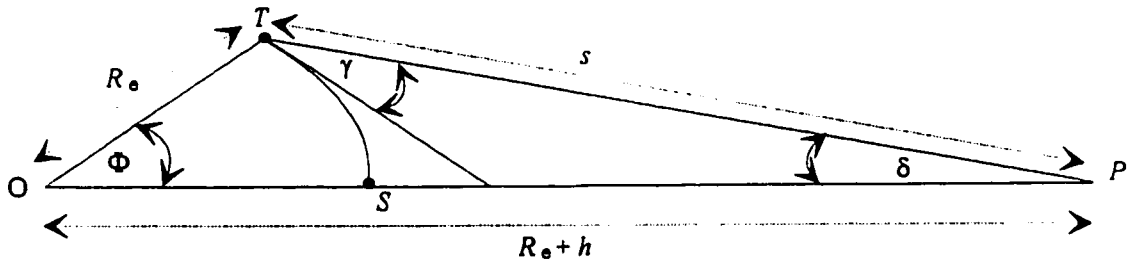


Figure 3.4: Plane triangle ΔTOP

3.1.3 Satellite Visibility

Coordinates to which a terminal must be pointed to communicate with a satellite are called the look angles. These are most commonly specified by the azimuth look angle Az, and elevation look angle γ. Azimuth is measured eastward from geographical north to the projection of the satellite path on a (locally) horizontal plane at the user terminal. From

Figure 3.3, we apply the sine law (spherical triangles) to obtain the azimuth angle Az [4, 49-51].

$$\frac{\sin(Az)}{\sin(90-l_s)} = \frac{\sin(L)}{\sin(\Phi)}$$
$$\sin(Az) = \frac{\sin(L)\cos(l_s)}{\cos(\Phi)} \tag{3.6}$$

The elevation angle is the angle above the horizon at which a user terminal sees the satellite and is given by [4, 49-51]

$$\cos(\gamma) = \left(\frac{R_e + h}{s}\right)\sin(\Phi) \tag{3.7}$$

For the satellite to be visible from a user terminal, its elevation angle γ must be greater than 8°. It is the elevation angle that determines the availability of the communication link and plays an important role in satellite to satellite handoffs. Doppler shift and Doppler rate are only calculated when the satellite is visible.

3.1.4 Expressions for Doppler Shift and Doppler Rate

Let v be the velocity vector of the satellite. The projection of this vector onto the plane TOS is $v\cos(\theta)$. Applying the cosine law for angles (spherical triangles) to triangle ΔNTS, given in Figure 3.4, we obtain

$$\cos(\theta) = -\cos(L)\cos(Az) + \sin(L)\sin(Az)\cos(90-l_t)$$
$$\cos(\theta) = -\cos(L)\cos(Az) + \sin(L)\sin(Az)\sin(l_t) \quad \textbf{(3.8)}$$

The radial velocity component v_d is given by

$$v_d = |v|\cos(\theta)\cos(\Phi + \gamma) \quad \textbf{(3.9)}$$

The Doppler shift is given by

$$f_d = \frac{|v|\cos(\theta)\cos(\Phi + \gamma) f_c}{c} \quad \textbf{(3.10)}$$

The Doppler rate is the time derivative of the Doppler shift and is given by

$$f_r = \frac{\partial f_d}{\partial t} \quad \textbf{(3.11)}$$

3.2 Simulation Results

Cross track angles $L = \{0°, 5°, 10°, 15°, 20°\}$ were chosen and $l_t = 0°$, where the latter implies the user being at the equator. $l_t = 0°$ was chosen for the following reason. For any other value of l_t, the effect is a horizontal shift of the Doppler curve on the time or l_s axis, depending which one is selected. Therefore, the Doppler shift is not affected. However, placing the user at the equator, produces Doppler curves which are centred on the l_s axis.

Note that $L = 0°$ refers to the case when the satellite passes directly overhead. The maximum cross track angle of 20° was chosen because the maximum coverage angle given by equation (3) is $\Phi_{max} = 20.10°$. Figure 3.5(a) and 3.5(b) depicts the Doppler shift and Doppler rate for the various cross track angles, plotted versus the angle of latitude of the subsatellite point l_s. A few noticeable features from the graphs are listed below:

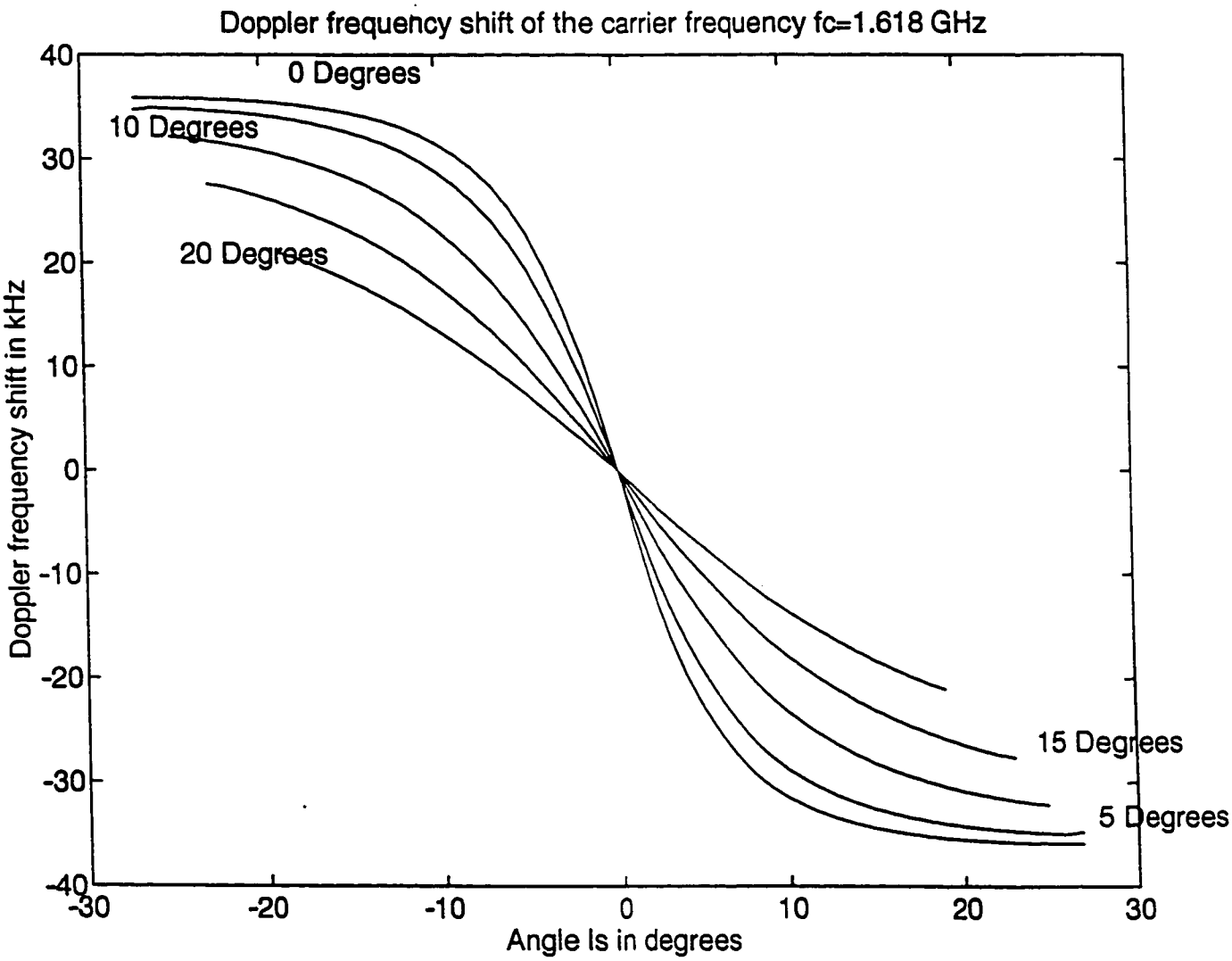


Figure 3.5(a): Doppler shift in kHz versus latitude of subsatellite point

1. The Doppler shift curves become less steep as the cross track angle increases.
2. Smaller cross track angles produce higher Doppler shift and Doppler rate.
3. The duration period for viewing a satellite become shorter as the cross track angle increases.
4. As the cross track angle increases, the Doppler shift behaves almost linearly.

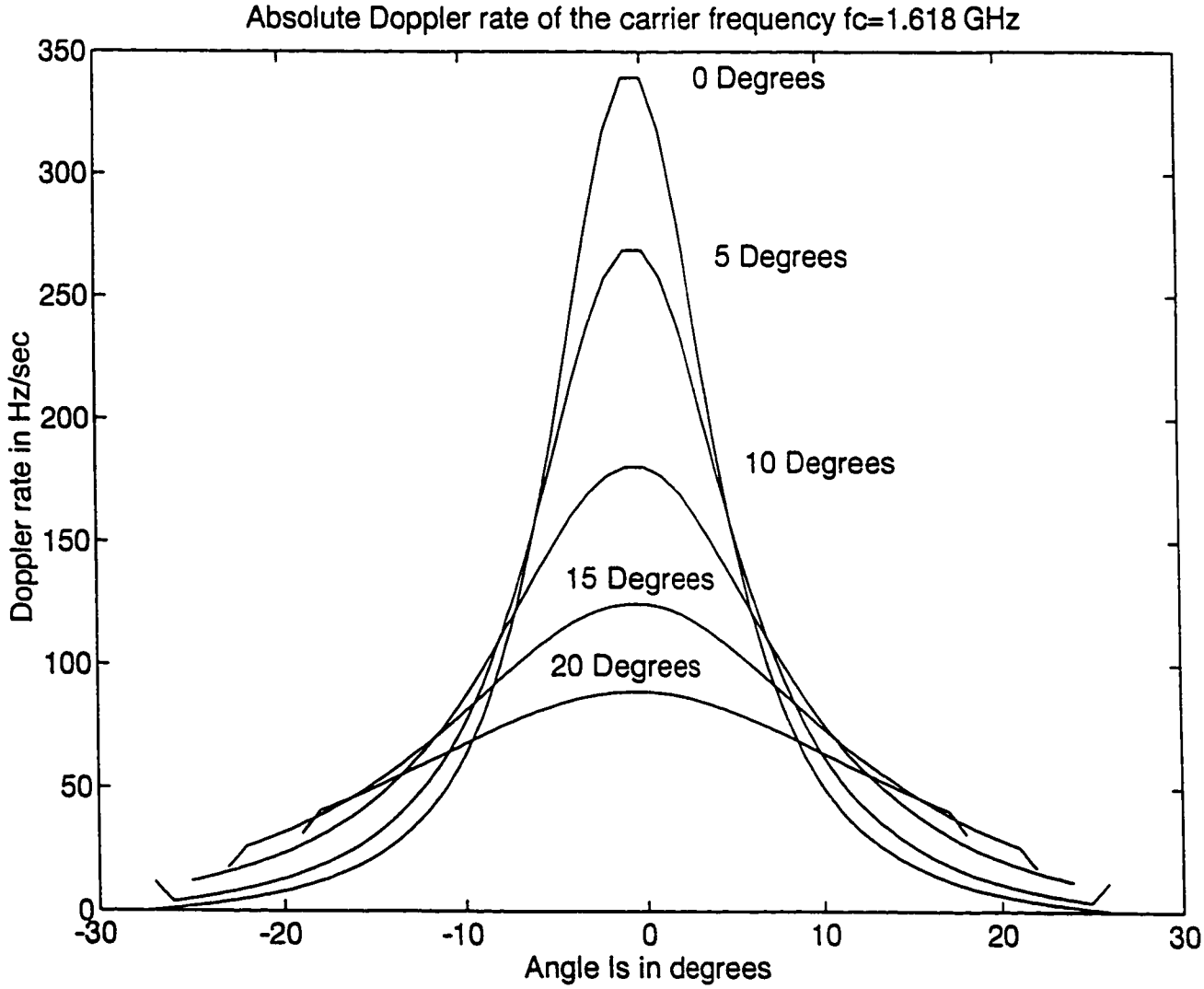


Figure 3.5(b): Doppler rate in Hz/s versus latitude of subsatellite point

Figure 3.6(a) and 3.6(b) depicts the Doppler shift and Doppler rate for the various cross track angles, but this time plotted versus time in minutes. The time axis is relative and is presented here to compare the relative duration period of viewing a satellite for the various cross track angles.

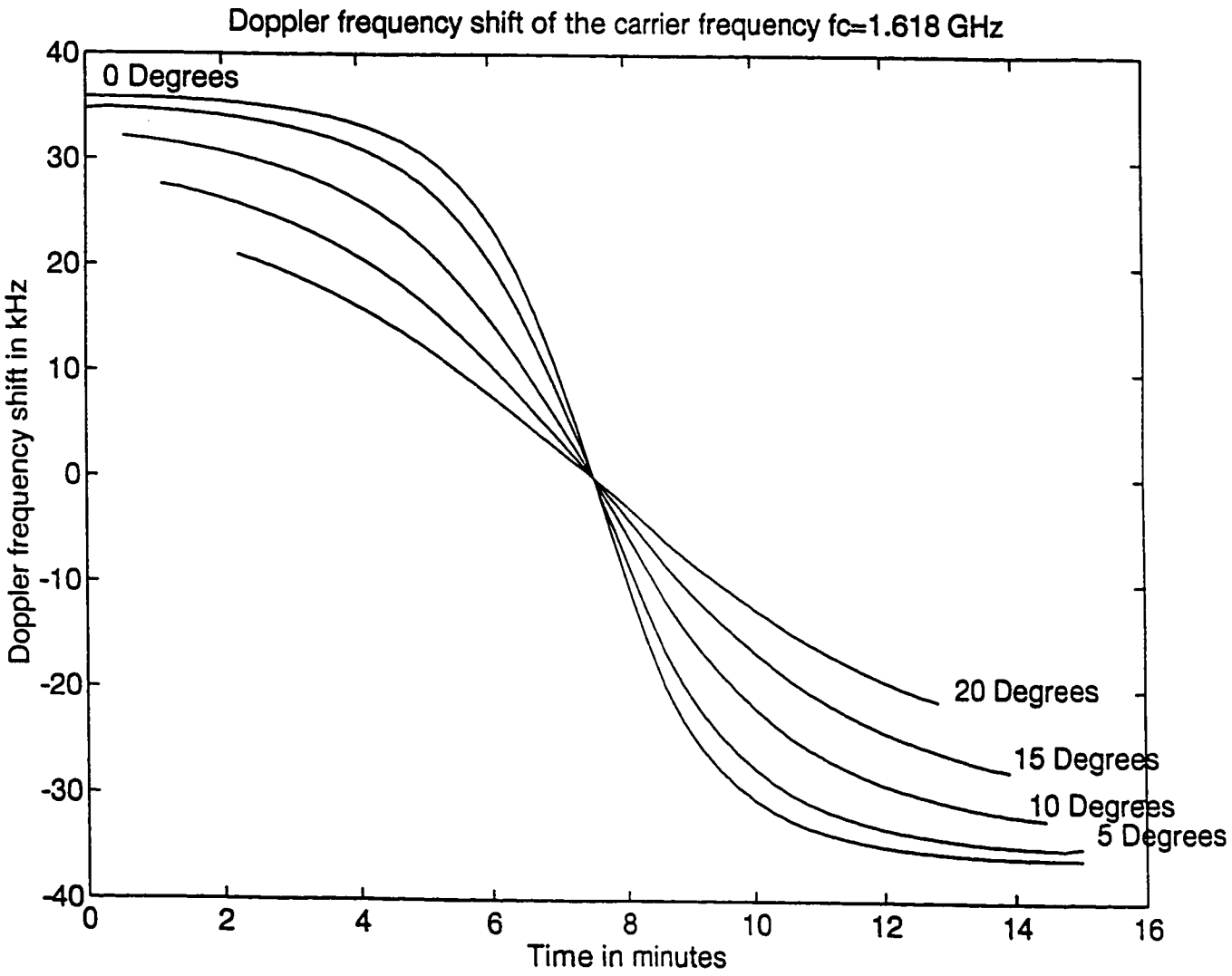


Figure 3.6(a): Doppler shift in kHz versus time in minutes

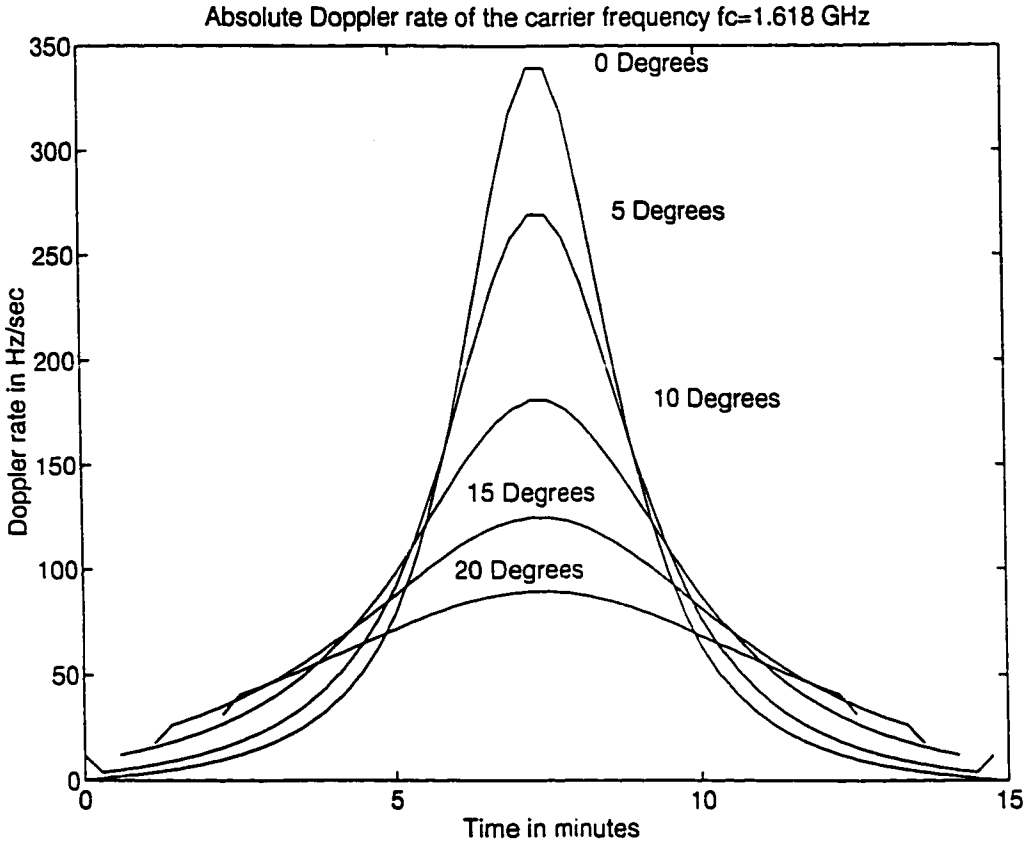


Figure 3.6(b): Doppler rate in Hz/s versus time in minutes

3.3 Introduction of the Novel Technique

Conventional carrier recovery techniques for multiple phase shift keying (MPSK) signals can cope with initial frequency offsets satisfying the fundamental limit for MPSK signal which is quoted here but its proof will given in section 4.8.2 of chapter 4.

$$|f_d T_s| \leq \frac{1}{2M} \tag{3.12}$$

Where T_s is the symbol period and $M = 4$ in the case of QPSK. Therefore, the maximum

tolerable initial frequency offset for QPSK signal is given by the fundamental limit for carrier recovery which is $|f_d T_s| \leq 1/8$. In our case, the tolerable Doppler shift on the downlink must satisfy $f_d < 25$ kHz, since the downlink coded data rate is 400 kbps.

The Doppler correction technique involves reducing the Doppler shift by applying a linearly decreasing frequency sweep to the received signal. This technique removes approximately half the Doppler shift for an overhead pass and the residual Doppler shift (compensated Doppler shift) is compared in Figure 3.7 [52]. The maximum Doppler shift has reduced from 35.5 kHz to 18.5 kHz. The residual Doppler shift is now within the tolerable limits of $f_d < 25$ kHz. Figures 3.8 show similar situations for cross track angles L = {0°, 5°, 10°, 15°, 20°}.

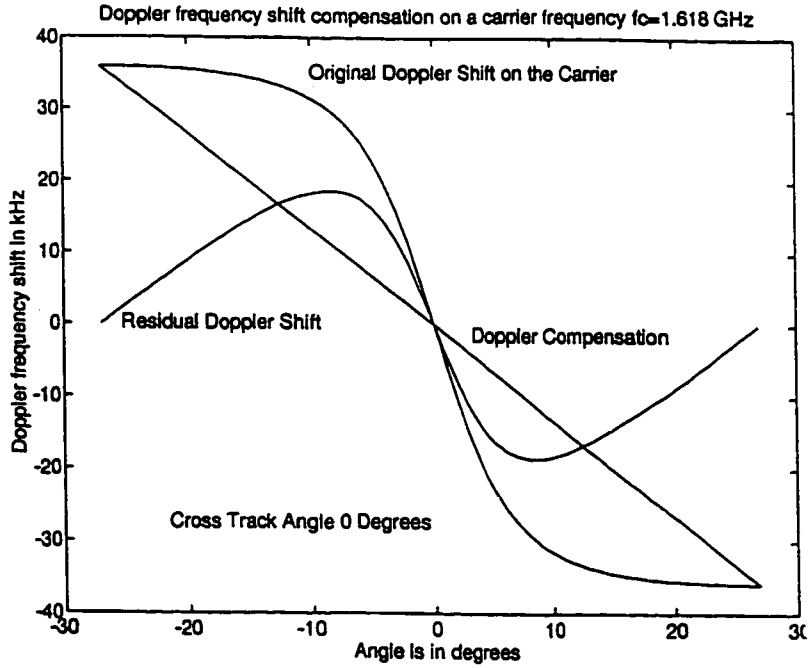


Figure 3.7: Doppler shift, linear frequency sweep and residual Doppler shift

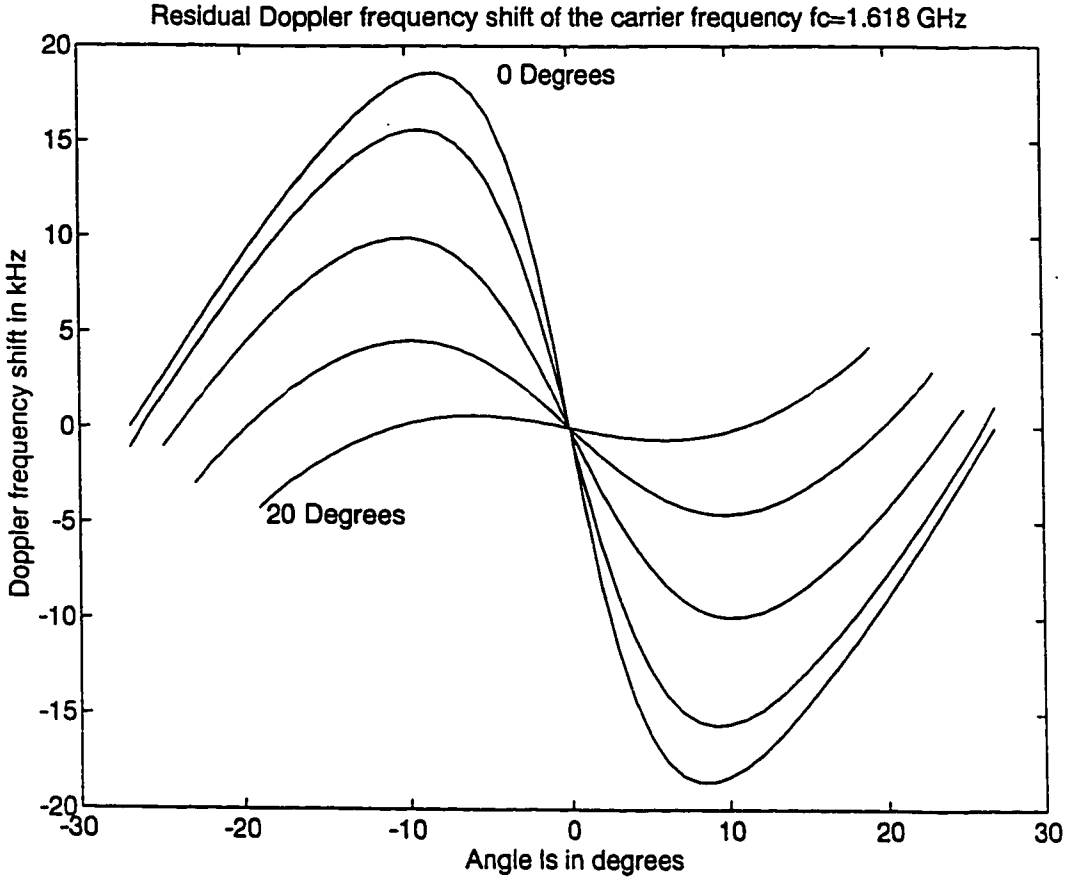


Figure 3.8: Residual Doppler shift

Satellites with largest range orbits, that is, $L = 20°$, have minimum residual Doppler shift. The graphs show that there is a progressive improvement in reducing the Doppler shift by using this novel Doppler shift compensation technique. For example, when the cross track angle $L = 20°$, the Doppler shift has reduced from 21.0 kHz to 4.25 kHz.

3.3.1 Residual Doppler Rate

Here we attempt to find the residual Doppler rate of the carrier for comparison with the original Doppler rate and the result is depicted in Figure 3.9. The maximum magnitude of Doppler rate has been reduced from 339.4 Hz/s to 259 Hz/s.

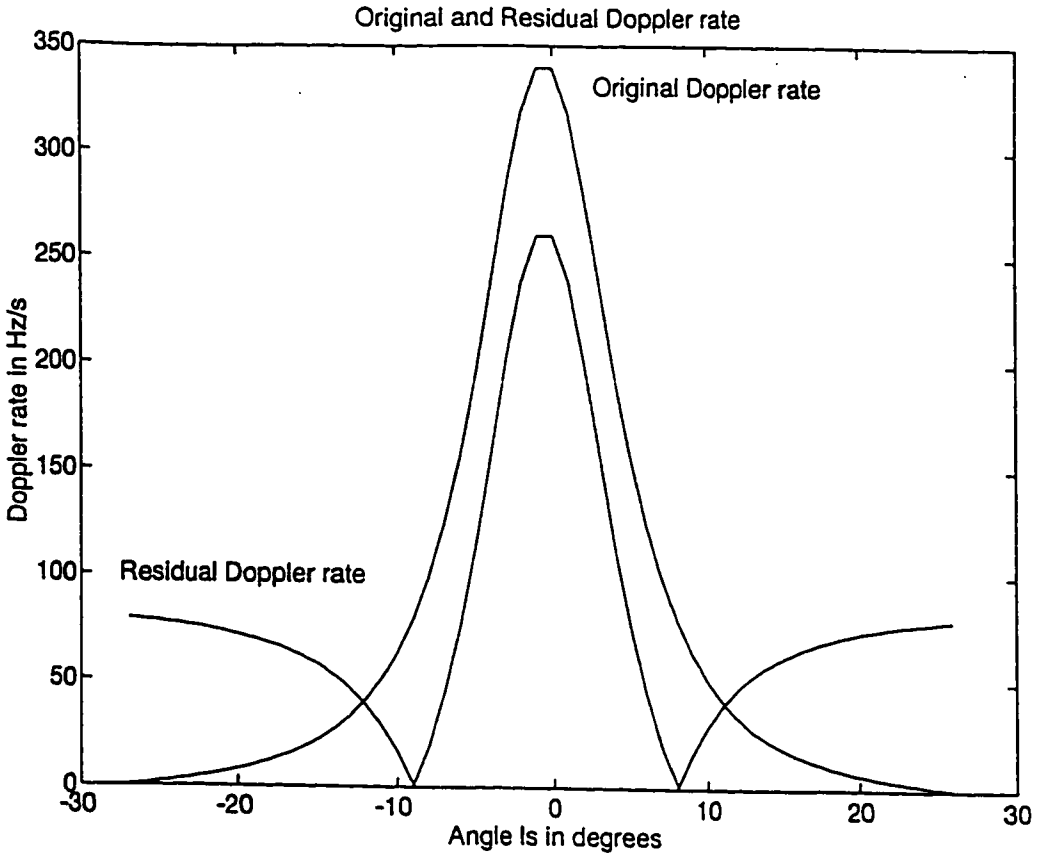


Figure 3.9: Residual Doppler rate

3.4 Modification of the Novel Technique

Doppler compensation using a linearly decreasing frequency sweep does not perform well for the overhead pass case because the Doppler shift curve is more non-linear. Therefore, it is intuitive to use a non-linear decreasing frequency sweep for small cross track angles and a linear decreasing frequency sweep for large cross track angles. For example, the following cube-root polynomial can be used as the non-linear model, where K is a constant.

$$f_{comp} = -K\, t^{\frac{1}{3}} \tag{3.13}$$

Figure 3.10 shows the residual Doppler shift before and after compensation for the case when $L = 0°$.

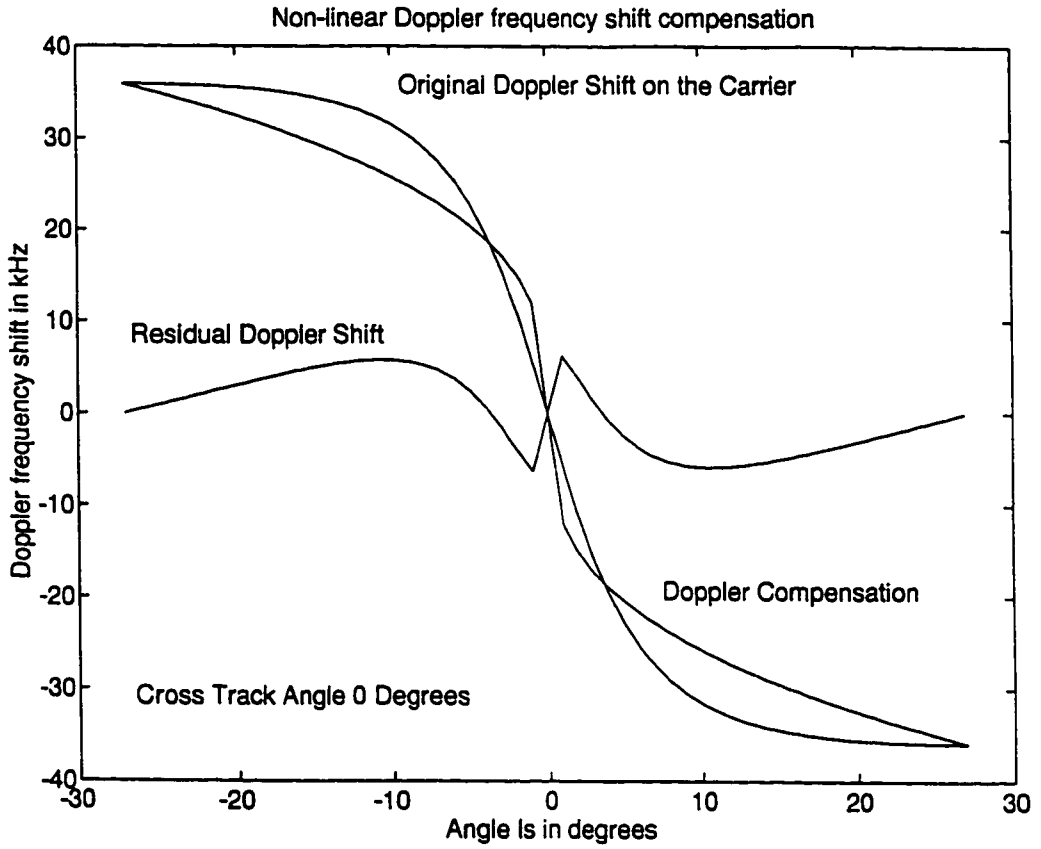


Figure 3.10: Doppler shift, non-linear frequency sweep and residual Doppler shift

Simulation has shown that the non-linear example was able to reduce the Doppler shift from 35.5 kHz to 6.29 kHz in the overhead case. This is an improvement in the overhead case but not so for large cross track angles.

3.5 Conclusions

This chapter describes a novel technique of Doppler compensation of a QPSK signal that has been transmitted over a satellite link, especially on both the uplink (user terminal to satellite) and downlink (satellite to user terminal). Simulations have shown that the extreme Doppler shifts encountered in LEO satellite systems can be reduced to tolerable levels for low data rate applications. In situations of a direct satellite overhead pass, the modified novel technique provides further improvement in reducing the Doppler shift.

If however, the mobile user terminal has prior knowledge of both the nominal carrier frequency, f_c and the motion of the satellite providing the service (case when such information is transmitted by the satellite to the user terminal over the paging channels), the mobile terminal can estimate the appropriate Doppler curve instead of choosing the direct overhead Doppler curve [53]. Choosing the right Doppler curve and applying the novel Doppler compensation presented in this chapter may provide a simple and elegant solution for downlink Doppler compensation in low cost mobile user terminals.

An investigation of Doppler shift estimation using the location of both the satellite and user terminal is provided in Appendix C. In addition, the effects of uncertainties in satellite and user terminal location on this method of Doppler shift estimation is illustrated by sample calculations.

CHAPTER 4

QPSK Receiver Synchronization for LEO Mobile Users

In this chapter, we examine methods of carrier synchronization for QPSK signal transmissions with large time varying Doppler shifts; that is, a Doppler frequency shift in the range -35.5 kHz $\leq f_d \leq$ 35.5 kHz of and Doppler rate of $0 \leq f_r \leq$ -350 Hz/s. First we introduce the necessity of carrier synchronization in coherent demodulation of QPSK signals. Basic modulation, demodulation and carrier recovery techniques for QPSK signal, in the absence of Doppler shift are briefly summarized. Then, carrier synchronizers which have been developed to combat constant Doppler shift are presented followed by the MPSK

acquisition range fundamental limit. The chapter is concluded by listing the shortcomings of current methods of mitigating dynamic Doppler shift and selecting the strategy to be used to design a Doppler resistant carrier synchronizer.

4.1 Doppler Shift in LEO Satellite Systems

At the present time, there is considerable interest in the potential of LEO satellite systems to provide a whole range of data communication services, including voice communications primarily to mobile users equipped with hand-held terminals. The frequency band allocated for mobile satellite service is in the L-band range of 1.6 - 2.5 GHz and some of these systems have been described in the previous chapter.

4.1.1 Doppler Shift and Doppler Rate

The relative velocity between the satellites and the mobile users introduces a large and time-varying Doppler shift into the transmitted signal. This is a time modulation of the signal which presents itself as a change in frequency with time into the signal carrier [54].

In section 2.3.3 of chapter 2, we investigated the Doppler shift and Doppler rate of a typical LEO satellite passing a user overhead. The Doppler rate curve in Figure 2.12 shows that the Doppler shift can change rapidly. Therefore, the phase of the signal from the satellite will have a high order function of time [55-56]. This rapid change of frequency implies that

we have to consider not only the static but dynamic characteristic of the Doppler shift in LEO satellite systems [57]. Earlier, we found that the maximum Doppler shift was around 35.5 kHz which is several times as wide as the bandwidth of a low-rate signal [58]. Typical frequency offsets in LEO systems range from 30 to 60 kHz depending on the frequency and satellite altitude [59].

4.1.2 Carrier and Clock Recovery

QPSK is a spectrally efficient format for transmitting digital data in satellite communication system. Since it is a double side band suppressed carrier (DSB-SC) modulation there is no energy transmitted at the carrier frequency. All the transmitted power goes into the modulation eliminating the discrete carrier component and creating a spectrum symmetric with respect to the carrier frequency. Since a discrete carrier component is absent, carrier regeneration is required (sometimes known as carrier recovery or phase referencing). Phase referencing is the operation of extracting a phase coherent reference carrier from an observed noisy received carrier.

Carrier recovery for coherent demodulation can be classified according to either burst or continuous transmissions. Where transmission is continuous, the demodulation signal acquisition time need not be rapid. However, in TDMA systems for mobile voice communications, short-duration bursts are commonly used. Each burst has its own independent carrier phase and consequently, synchronization is required on a burst-by-burst

basis [45].

Correct carrier recovery needs to be achieved as soon as the satellite rises above the horizon. This is essential for an effective personal communication system and should take place as rapidly as possible, since a fast synchronization corresponds to increased capacity (more users accessing the system in a period of time). Current standard average is less than one second and a maximum limit of 5 seconds [60]. However, under this condition, the received signal has the highest Doppler shift and the lowest SNR, due to the fact the satellite is at a maximum range from the user [61-62]. Furthermore, when the satellite is at the horizon, the most severe multipath environment arises between the satellite and the user terminal [63]. It is safe to say that the Doppler shift maintains a fairly constant value during that short instant of time. Carrier phase synchronization must be maintained as the satellite passes overhead and as it approaches the horizon again. Conversely, when the satellite is at zenith the SNR is the highest and Doppler shift is zero, but changing at the maximum rate.

The problem of Doppler shift during acquisition has received considerable attention [45, 54-58, 60-64] and this chapter examines some of the techniques used to mitigate the harmful effects of Doppler shift and highlights some of the problems from the viewpoint of carrier recovery at the receiver as encountered in the case of the Iridium satellite system which implements QPSK signalling. Apart from Doppler shift, local oscillators in the transmitter/receiver also cause frequency offset (oscillator instabilities). There has been considerable effort to develop a carrier recovery technique as part of a coherent demodulation scheme which is not only applicable for small initial frequency offsets, but for LEO satellite

communication systems.

Basically, one of the key requirements of a successful LEO satellite system is the development of a Doppler shift resistant demodulator. We now proceed to describe conventional carrier synchronizers for zero Doppler shift then investigate those which are specifically designed to operate in the presence of Doppler shift.

4.2 Quaternary Phase Shift Keying

The most popular forms of modulation used in satellite communications are MPSK schemes, especially QPSK. The QPSK modulation scheme is characterized by the fact that the information carried by the transmitted signal is contained in the phase of the constant envelope carrier and takes on one of the four possible values {45°, 135°, 225°, or 315°}. The general waveform for QPSK is given by [46, 48, 50, 65-69]

$$s_i(t) = A\cos(\omega_c t + \theta_i) \qquad 0 \le t \le T_s \tag{4.1}$$

where A is the signal amplitude, ω_c is the carrier angular frequency, T_s is the symbol duration and θ_i is the modulating phase, given by

$$\theta_i = (2i - 1)\frac{\pi}{4} \qquad \textit{for} \quad i = 1, 2, 3, 4 \tag{4.2}$$

with equal probability. Note that the QPSK carrier has a power level of $P_c = A^2/2$ [48].

A QPSK modulator takes two input data bits at a time and produces one of the four possible phases of the carrier. Therefore, a signalling rate reduction of one-half is achieved. The mapping of the 2-bit (dibit) symbols {00,01,10,11} to the carrier phase can be accomplished in a number of ways. The most preferred conversion of binary symbols to phase angles is usually done using Gray coding. The essential idea of Gray coding is to permit only one bit change in the assignment of binary symbols to adjacent phase angles.

In practice, phase angles {45°, 135°, 225°, or 315°} represent symbols {00, 01, 11, 10} respectively, and the corresponding transmitted signals are

$$\begin{aligned} s_1(t) &= s_{10}(t) = A\cos(\omega_c t + \frac{\pi}{4}) \\ s_2(t) &= s_{00}(t) = A\cos(\omega_c t + 3\frac{\pi}{4}) \\ s_3(t) &= s_{01}(t) = A\cos(\omega_c t + 5\frac{\pi}{4}) \\ s_4(t) &= s_{11}(t) = A\cos(\omega_c t + 7\frac{\pi}{4}) \end{aligned} \tag{4.3}$$

or equivalently

$$\begin{aligned} s_1(t) &= s_{10}(t) = \frac{A}{\sqrt{2}}\cos(\omega_c t) - \frac{A}{\sqrt{2}}\sin(\omega_c t) \\ s_2(t) &= s_{00}(t) = -\frac{A}{\sqrt{2}}\cos(\omega_c t) - \frac{A}{\sqrt{2}}\sin(\omega_c t) \\ s_3(t) &= s_{01}(t) = -\frac{A}{\sqrt{2}}\cos(\omega_c t) + \frac{A}{\sqrt{2}}\sin(\omega_c t) \\ s_4(t) &= s_{11}(t) = \frac{A}{\sqrt{2}}\cos(\omega_c t) + \frac{A}{\sqrt{2}}\sin(\omega_c t) \end{aligned} \tag{4.4}$$

The geometrical representation of the Gray coded QPSK signal is shown in Figure 4.1 [46-47]. Note that since all the signal states lie on the circle, QPSK is characterized as a constant amplitude modulation.

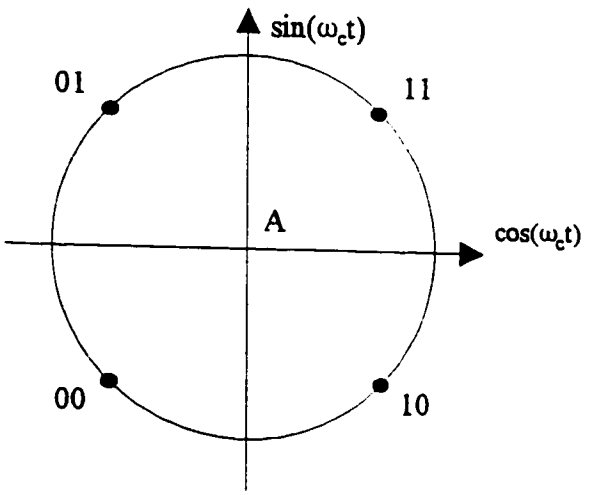


Figure 4.1: QPSK phase states using Gray coding

From (4.3), it is seen that the QPSK signal can be viewed as a linear combination of two binary phase shift keying (BPSK) signals in phase quadrature [66].

4.3 QPSK Modulator

Consider the generation of QPSK with a typical modulator as shown in Figure 4.2 [46, 50, 65-66, 69].

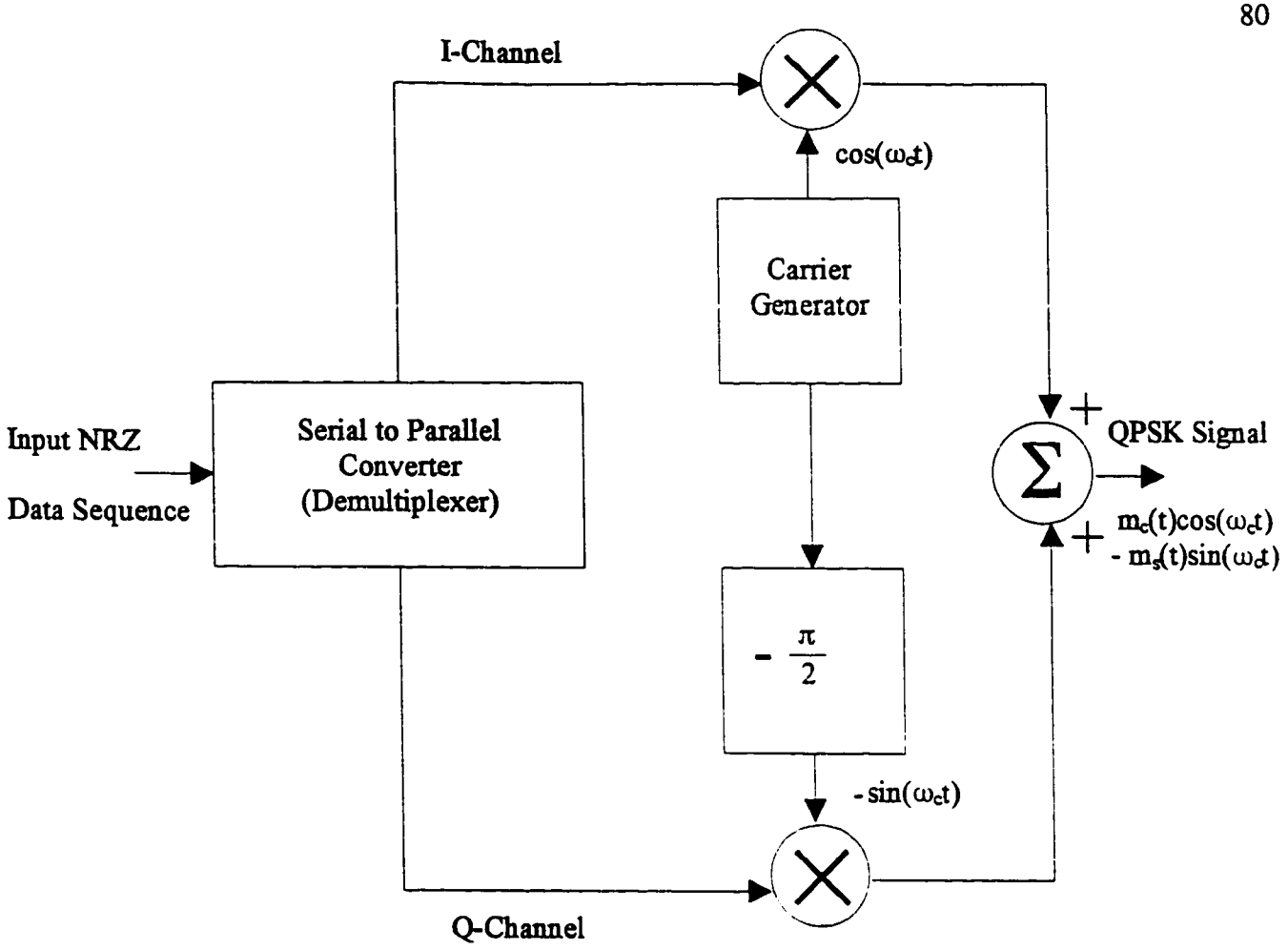


Figure 4.2: Typical QPSK modulator

We may rewrite (4.1) in the equivalent form

$$s_i(t) = A\cos(\theta_i)\cos(\omega_c t) - A\sin(\theta_i)\sin(\omega_c t) \qquad 0 \le t \le T_s$$
$$= m_c(t)\cos(\omega_c t) - m_s(t)\sin(\omega_c t) \qquad (4.5)$$

Where $m_c(t)$ and $m_s(t)$ are two separate non return to zero (NRZ) bit streams of I-channel (in-phase channel) and Q-channel (quadrature channel) respectively. These bit streams can also be the output of two separate digital sources or can be the alternative bits from a

common source. We consider the latter case. In QPSK modulation, the input NRZ enters the serial-to-parallel converter (demultiplexer) to generate $m_c(t)$ and $m_s(t)$, at one-half the original bit rate, as shown in Figure 4.3 [46].

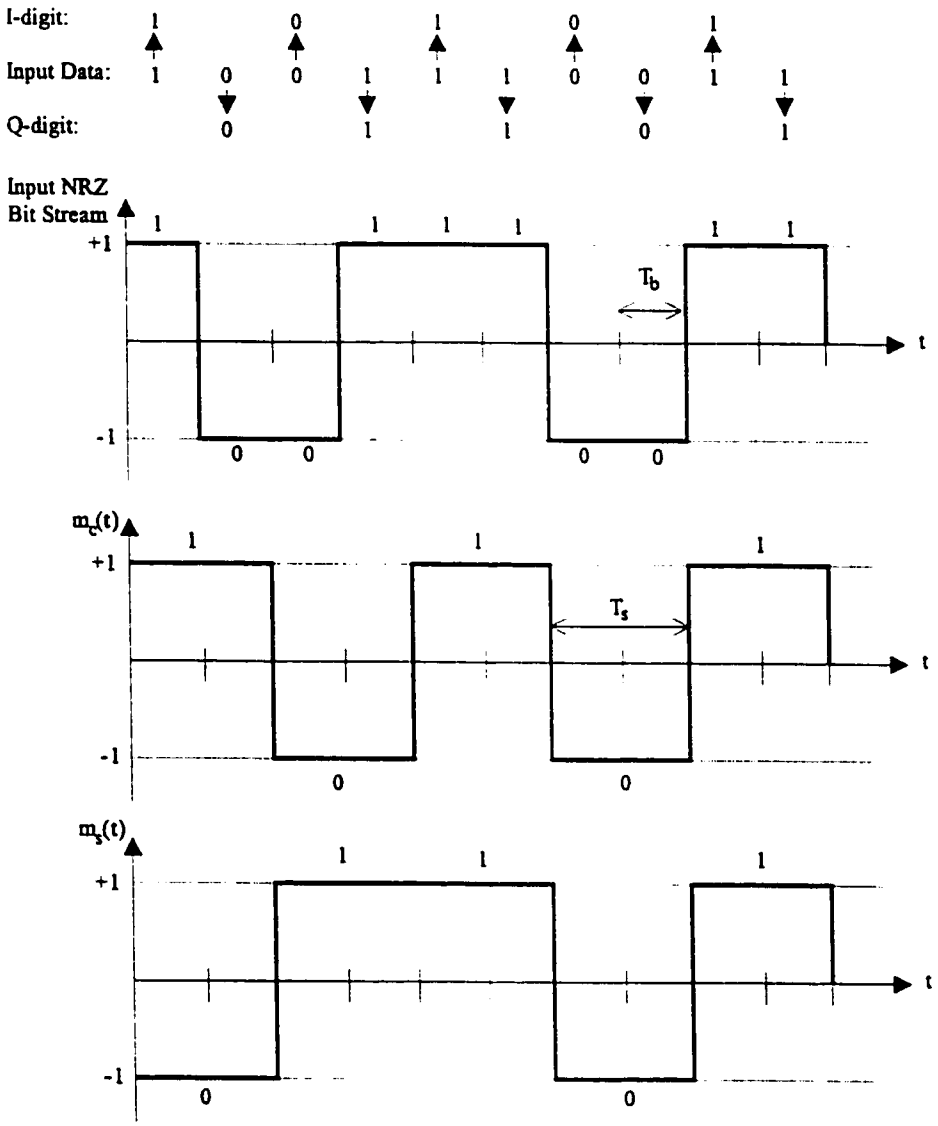


Figure 4.3: Serial to parallel conversion of input NRZ bit stream

In other words, each I-digit or Q-digit has a duration of $T_s = 2T_b$, where T_b is the bit duration of the input NRZ bit stream. A symbol is formed by combining an I-digit and a Q-digit, for example the symbol 10 is formed with the I-digit 1 and Q-digit 0.

The I-digit stream is used to modulate the carrier $\cos(\omega_c t)$ and the Q-digit stream is used to modulate the quadrature carrier $\sin(\omega_c t)$. The two resulting modulated BPSK signals, with quadrature carriers, are added to form the required QPSK signal.

Equation (4.1) can equivalently be expressed as

$$\begin{aligned} s_i(t) &= Re\left[A \exp(j\theta_i) \exp(j\omega_c t)\right] \qquad 0 \le t \le T_s \\ &= Re\left[\tilde{m}_i(t) \exp(j\omega_c t)\right] \end{aligned} \tag{4.6}$$

where the complex envelope $\tilde{m}_i(t)$ is given as follows

$$\begin{aligned} \tilde{m}_i(t) &= A \exp(j\theta_i) \qquad 0 \le t \le T_s \\ &= m_c(t) + jm_s(t) \end{aligned} \tag{4.7}$$

4.4 Coherent QPSK Demodulator

In order to achieve the optimum error performance with any digital modulation format, coherent detection must be used at the receiver. Coherent detection requires both the receiver and transmitter to be synchronous which implies knowledge of both the frequency

of the carrier and its phase. The process of estimating the carrier phase and frequency is referred to as carrier recovery or carrier synchronization, and will be addressed in sections 4.5 through 4.8.

The decoding operation is a process required to reconstruct the data bit sequence encoded onto the carrier. Figure 4.4 depicts a generalized coherent quadrature demodulator consisting of a coherent demodulation section, where the QPSK incoming signal is split into two paths namely: the I-channel and Q-channel, which are each multiplied (demodulated) by locally regenerated quadrature carrier references. The carrier reference is provided by the carrier recovery circuit [45-46, 65-67, 69].

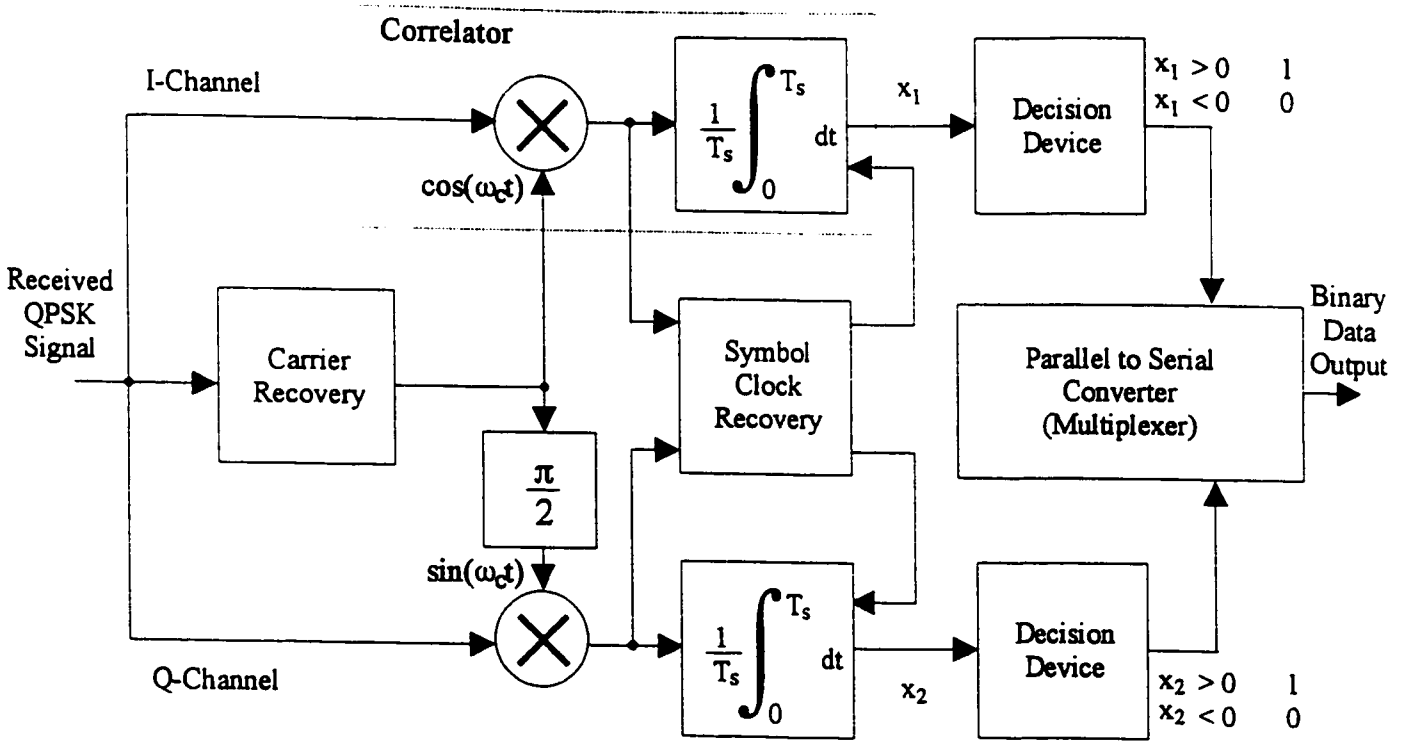


Figure 4.4: Coherent QPSK demodulator

This effectively corresponds to correlation of the received QPSK signal with the reference carriers over each symbol period. Timing for the integrate-and-dump and the data recombining is provided by the symbol clock recovery circuit. The outputs of the integrate-and-dump (correlators) x_1 and x_2 are each compared with a threshold of zero. If $x_1 > 0$, a decision is made in favour of symbol 1 for the in-phase channel output, on the other hand, if $x_1 < 0$, a decision is made in favour of symbol 0. The same applies for the quadrature channel. The original data sequence is obtained by combining the two sequences from in-phase and quadrature outputs by using a parallel-to-serial converter (multiplexer) [65, 68].

For the coherent demodulator shown in Figure 4.4 to operate successfully, we must implement an effective and accurate technique for generating a pair of coherent reference carriers. For this purpose, various types of carrier recovery techniques (synchronizing loops) include: Mth power, quadriphase Costas loop, remodulator, and decision directed loop. These will addressed in section 4.6.

4.5 Synchronization

Synchronization poses a fundamental problem in digital communications systems. In a digital communication system, there exists a hierarchy of synchronization problems to be considered. For example if a carrier type system is involved, the receiver must synchronize its reference carrier with that of the received signal in order to perform coherent demodulation [68]. Therefore, the first step of synchronization is carrier synchronization.

The next problem in the hierarchy is that of synchronizing the receiver's symbol clock with the baseband data-symbol sequence, is known as bit synchronization. Beyond the bit synchronization, other types of problems occurs. For example, depending on the type of communication system, one may encounter sync problems such as word, frame and packet synchronization. Synchronization of these type will not be investigated here [66, 70-71].

In systems where a discrete component of the carrier is present in the received signal (unmodulated sinusoidal component), the standard approach to achieve carrier synchronization is to make use of a phase lock loop (PLL). A PLL is designed with a narrow loop bandwidth such that it will lock onto and track the carrier and not be disturbed by the sideband components of the modulated signal. For systems in which the carrier is totally suppressed as in DSB-SC (eg. PSK), it must be recreated from the sideband data. A Costas loop or Mth order loop is capable of achieving such a feat in the case of PSK system [72-73].

In the case of double side band amplitude modulation (DSB-AM) except in the case of quadrature amplitude modulation (QAM), a residual carrier component is available for synchronization purposes. Isolation of the carrier from the received signal implies the use of narrow-band filtering. As a matter of fact, the actual technique commonly used to achieve carrier synchronization is the addition of a pilot carrier to the DSB-SC modulation. This technique allows the carrier level to be adjusted independently of the data modulation. In power-limited situations such as encountered in satellite communications, the carrier power is a major part of the overall transmitted power. For a given total RF power, the performance of the system is reduced but carrier synchronization is simplified. This technique is

sometimes used for low data rate applications. An automatic frequency control (AFC) circuit compares the pilot frequency with an accurate local reference and uses the difference to centre the received signal as shown in Figure 4.5.

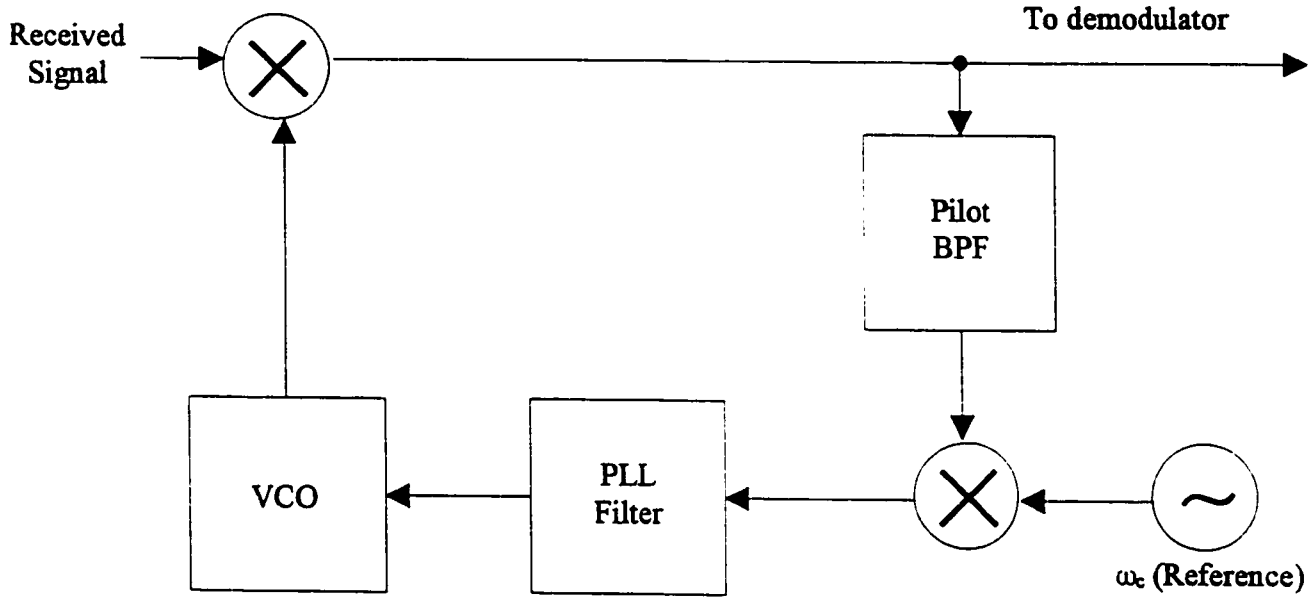


Figure 4.5: Pilot tone recovery for DSB-SC systems

4.5.1 Data-Aided Carrier Recovery Techniques (Pilot Tone and Preamble)

In TDMA satellite communication systems, carrier and clock synchronization is achieved by prefixing a preamble to each data packet. The preamble might contain an unmodulated component of the sinusoidal RF carrier (pilot tone also known as beacon) or

known data bits that accentuate carrier and clock line spectra which would assist synchronization [45, 73-74]. These systems are referred to as data-aided (DA) system and the carrier acquisition time for such systems is far in excess of the minimum required for optimal linear phase estimation. For typical voice communications, it would be beneficial to avoid the use of an acquisition preamble because of the overhead involved. In other words, no special timing information is multiplexed onto the data signal, since it would use up a portion of the available channel capacity [58, 75-79].

Our investigation for the moment will only address carrier recovery. Therefore, ideal data-timing is assumed. After carrier recovery is achieved, the phase of the modulated carrier can be demodulated to baseband, but the data bits (symbols) cannot be decoded until bit timing (bit (symbol) recovery) is achieved. Bit timing is the operation of extracting a time coherent bit rate clock from an observed noisy data-modulated waveform.

4.6 Carrier Synchronizers (Zero Doppler Shift)

There are four main types of carrier synchronizers, Mth power loop, Costas loop, remodulator, and decision directed loop. They differ in the position of the nonlinearity, which is entirely separated from the PLL in the Mth power loop, but included in the phase detector for the Costas loop, remodulator and the decision directed loop [66]. Only the first three are simple and will be addressed here, however, the last one is more complicated and involves estimation of the data symbols.

All QPSK carrier recovery techniques eliminate the modulation on the received signal to create a carrier component having a phase variation proportional to that of the received carrier. Subsequent tracking of this residual carrier component then generates the desired carrier reference.

4.6.1 Mth Power Loop

Squaring of a QPSK signal with equiprobable phases yields a BPSK signal [47]. To recover a discrete carrier frequency component, therefore, it is necessary to take the fourth power of the received QPSK. The QPSK signal $s(t)$ is band-pass filtered to minimize the effects of noise. Then it is passed into a fourth-power device, and a conventional PLL can be used to track the fourth harmonic [65-66, 72]. The voltage controlled oscillator (VCO) output is divided-by-4 to provide the desired carrier. The divide-by-4 operation creates four phase ambiguities in the interval $(0, 2\pi)$. In other words a phase ambiguity of an integer multiple $\pi/2$. Figure 4.6(a,b) depicts the block diagrams of a fourth power loop for QPSK carrier recovery.

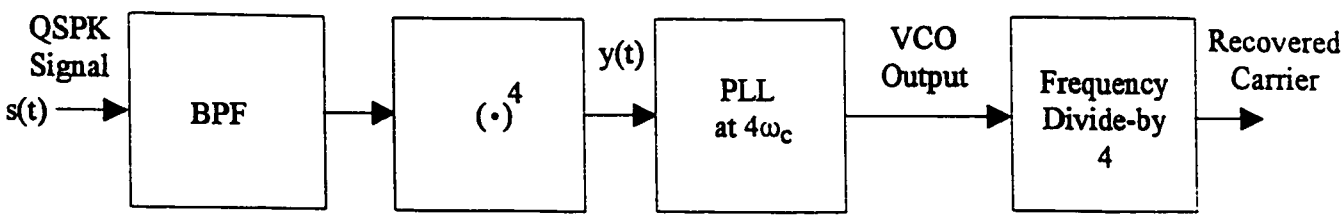
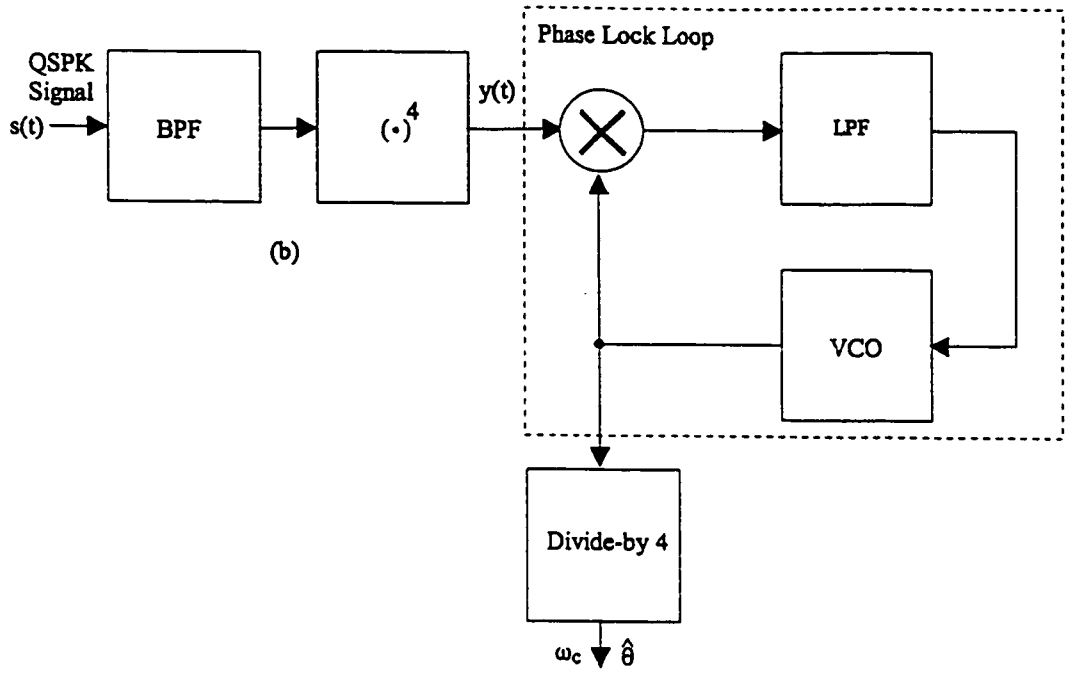


Figure 4.6(a): Block diagram of fourth power loop for QPSK carrier recovery



(b)

Figure 4.6(b): Fourth power loop for QPSK carrier recovery

To simplify, let us assume the received QPSK signal is of the form (noiseless case)

$$s(t) = Re\left[m(t)\ \exp(j\omega_c t + j\theta_0)\right] \tag{4.8}$$

where θ_o is the arbitrary carrier phase introduced by the channel and assumed to be independent of time, $m(t)$ is the complex data signal, given by

$$m(t) = m_c(t) + jm_s(t) \tag{4.9}$$

and $m_c(t)$ and $m_s(t)$ take values ±1 therefore

$$\begin{aligned} m(t) &= \pm 1 \pm j \\ &= \sqrt{2}\ \exp\left(j\frac{\pi}{4} + j(\frac{k\pi}{2})\right) \qquad for\ \ k = 0,\ 1,\ 2,\ 3 \end{aligned} \tag{4.10}$$

The output of the fourth power device is given by [46]

$$\begin{aligned} y(t) &= s^4(t) \\ &= \frac{3}{8}\,|m(t)|^4 + \frac{1}{2}\,Re\left[|m(t)|^2\, m^2(t)\, \exp(j2\omega_c t + j2\theta_0)\right] \\ &\quad + \frac{1}{8}\,Re\left[m^4(t)\, \exp(j4\omega_c t + j4\theta_0)\right] \end{aligned} \tag{4.11}$$

From (10), $|m(t)|^4 = 4$, $|m(t)|^2 = 2$ and $m^2(t) = \pm j2$

Therefore,

$$|m(t)|^2\, m^2(t) = 4\, \exp\left(\pm j\frac{\pi}{2}\right) \tag{4.12}$$

The first term of (4.11) is a DC component which is filtered out by the narrow band PLL. The second term is still a modulated signal. Since $m^4(t) = -4$, the third term is essentially a pure sinusoid signal at $4\omega_c$, phase $4\theta_o$ with no fluctuations in amplitude. It is this term which is tracked by the PLL [47].

4.6.2 Quadriphase Costas Loop

The system involves two parallel tracking loops, namely the in-phase and quadrature loops, operating simultaneously from the same VCO. There are two quadriphase Costas

loops that can be used for carrier recovery. Figure 4.7 depicts the quadriphase Costas crossover loop and Figure 4.8 shows the modified QPSK loop [48, 66, 72]. For small phase errors and high SNR, the outputs of the limiters corresponds to the cosine term of the arm filter outputs. The hard limiters can be denoted by the Signum function ($SGN[\ \cdot\ ]$).

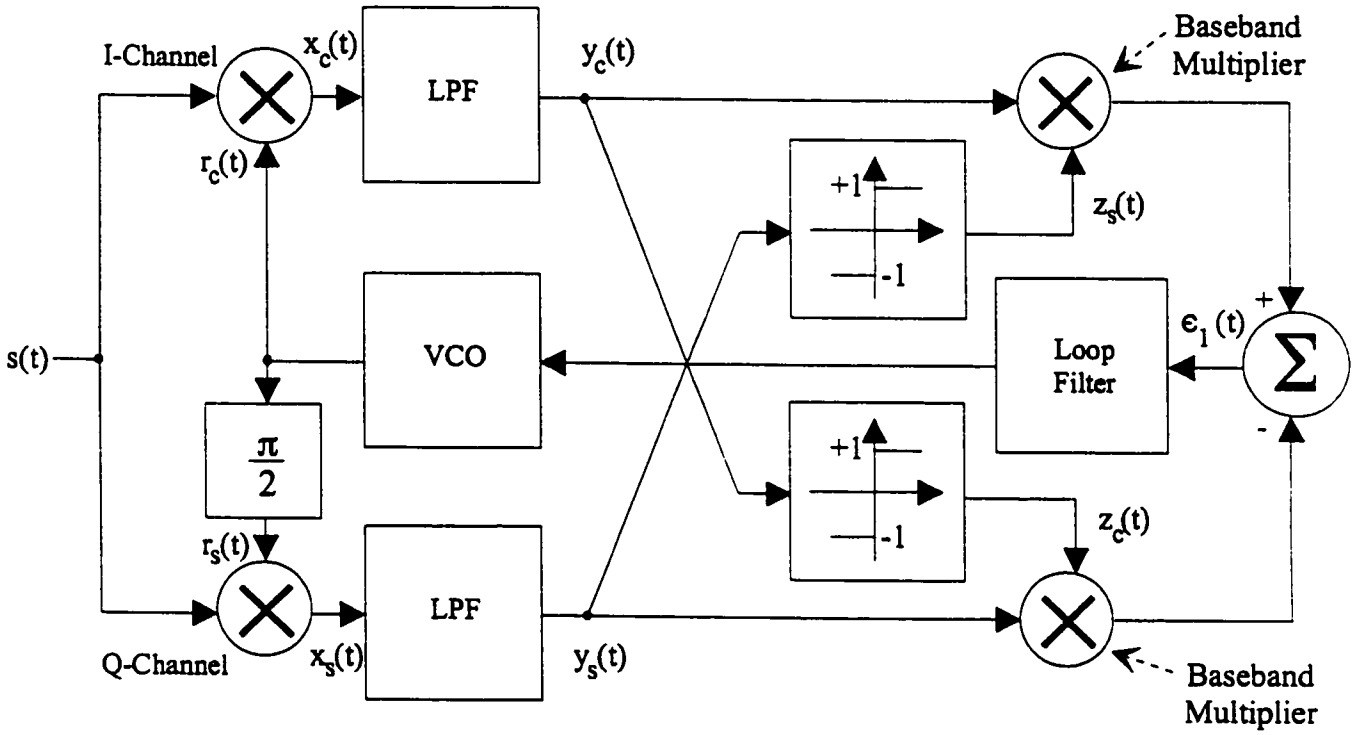


Figure 4.7: Block diagram of quadriphase Costas crossover loop

The received QPSK for the noiseless case is given by

$$\begin{aligned} s(t) &= Re\left[(m_c(t) + jm_s(t))\ \exp(j\omega_c t + j\theta_0)\right] \\ &= m_c(t)\ \cos(\omega_c t + \theta_0) - m_s(t)\ \sin(\omega_c t + \theta_0) \end{aligned} \tag{4.13}$$

The VCO produces the local carrier references and are assumed to be of the form

$$r_c(t) = \cos(\omega_c t + \theta_1) \qquad \textit{and}$$
$$r_s(t) = \sin(\omega_c t + \theta_1) \tag{4.14}$$

The output of the modulators are given by

$$x_c(t) = \frac{1}{2}\left[m_c(t)\cos(\theta_\epsilon) - m_s(t)\sin(\theta_\epsilon) + m_c(t)\cos(2\omega_c t + \theta_0 + \theta_1) - m_s(t)\sin(2\omega_c t + \theta_0 + \theta_1)\right]$$
$$x_s(t) = \frac{1}{2}\left[-m_c(t)\sin(\theta_\epsilon) - m_s(t)\cos(\theta_\epsilon) + m_c(t)\sin(2\omega_c t + \theta_0 + \theta_1) + m_s(t)\cos(2\omega_c t + \theta_0 + \theta_1)\right] \tag{4.15}$$

and the phase error is given by

$$\theta_\epsilon = \theta_0 - \theta_1 \tag{4.16}$$

The outputs of the modulator $x_c(t)$ and $x_s(t)$ are low pass filtered producing baseband signals

$y_c(t)$ and $y_s(t)$ respectively.

$$y_c(t) = LPF\left[x_c(t)\right] = \frac{1}{2}\left[m_c(t)\cos(\theta_\epsilon) - m_s(t)\sin(\theta_\epsilon)\right] \tag{4.17}$$

Similarly

$$y_s(t) = LPF\left[x_s(t)\right] = -\frac{1}{2}\left[m_c(t)\sin(\theta_\epsilon) + m_s(t)\cos(\theta_\epsilon)\right] \tag{4.18}$$

For high SNR, the limiter outputs are $z_c(t)$ and $z_s(t)$ respectively

$$\begin{aligned} z_c(t) &= SGN[y_c(t)] = \frac{1}{2}m_c(t) \\ z_s(t) &= SGN[y_s(t)] = -\frac{1}{2}m_s(t) \end{aligned} \tag{4.19}$$

The subtraction removes the DC bias effects and the resulting phase discriminator output is an error signal to phase control the loop VCO, thereby closing the quadriphase crossover Costas loop. The tracking error signal is given by

$$\begin{aligned} \epsilon_1(t) &= y_c(t)\ SGN[y_s(t)] - y_s(t)\ SGN[y_c(t)] \\ &= y_c(t)\ z_s(t) - y_s(t)\ z_c(t) \\ &= \frac{1}{4}\left[m_c^2(t)\ \sin(\theta_\epsilon) + m_s^2(t)\ \sin(\theta_\epsilon)\right] \\ &= \frac{1}{4}\ \sin(\theta_\epsilon)\left[m_c^2(t) + m_s^2(t)\right] \end{aligned} \tag{4.20}$$

Since

$$m_c^2(t) = m_s^2(t) = 1 \tag{4.21}$$

We obtain the final expression

$$\epsilon_1 = \frac{1}{2}\ \sin(\theta_\epsilon) \tag{4.22}$$

Figure 4.8 depicts a modified QPSK Costas loop [45, 73, 80-81].

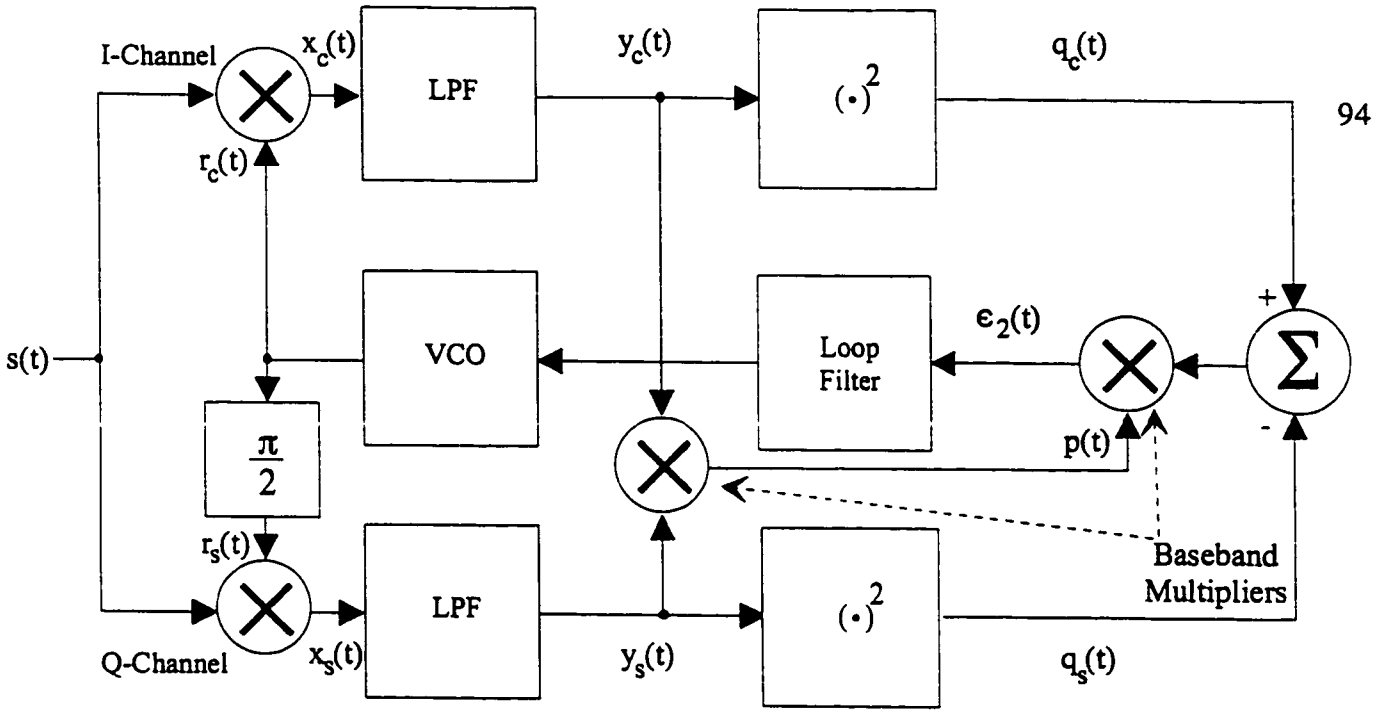


Figure 4.8: Block diagram of modified QPSK Costas loop

The signal component at the output of the squarer from the in-phase channel of the modified Costas loop is given by

$$q_c(t) = y_c^2(t) = \frac{1}{4}\left[m_c^2(t)\cos^2(\theta_\epsilon) + m_s^2(t)\sin^2(\theta_\epsilon) - 2m_c(t)\,m_s(t)\cos(\theta_\epsilon)\sin(\theta_\epsilon)\right] \quad (4.23)$$

and using (4.21) we obtain

$$q_c(t) = \frac{1}{4}\left[1 - m_c(t)\,m_s(t)\sin(2\theta_\epsilon)\right] \quad (4.24)$$

Similarly, the signal component from the quadrature channel is given by

$$q_s(t) = y_s^2(t) = \frac{1}{4}\left[m_c^2(t)\,\sin^2(\theta_\epsilon) + m_s^2(t)\,\cos^2(\theta_\epsilon) + 2m_c(t)\,m_s(t)\,\sin(\theta_\epsilon)\,\cos(\theta_\epsilon)\right] \tag{4.25}$$

Using (4.25) we simplify the expression to

$$q_s(t) = \frac{1}{4}\left[1 + m_c(t)\,m_s(t)\,\sin(2\theta_\epsilon)\right] \tag{4.26}$$

The inputs to the very last baseband multiplier is given by

$$q_c(t) - q_s(t) = -\frac{1}{2}\,m_c(t)\,m_s(t)\,\sin(2\theta_\epsilon) \tag{4.27}$$

and

$$\begin{aligned} p(t) &= y_c(t)\,y_s(t) \\ &= -\frac{1}{4}\left[m_c^2(t)\,\sin(\theta_\epsilon)\,\cos(\theta_\epsilon) + m_c(t)\,m_s(t)\,\cos^2(\theta_\epsilon) - m_c(t)\,m_s(t)\,\sin^2(\theta_\epsilon) - m_s^2(t)\,\sin(\theta_\epsilon)\,\cos(\theta_\epsilon)\right] \\ &= -\frac{1}{4}\left[m_c(t)\,m_s(t)\,\cos^2(\theta_\epsilon) - m_c(t)\,m_s(t)\,\sin^2(\theta_\epsilon)\right] \\ &= -\frac{1}{4}\,m_c(t)\,m_s(t)\,\cos(2\theta_\epsilon) \end{aligned} \tag{4.28}$$

which is the also the output of the first baseband multiplier. The tracking error signal fed into the VCO is given by

$$\begin{aligned}\epsilon_2(t) &= \left[y_c^2(t) - y_s^2(t)\right] y_c(t)\, y_s(t) \\ &= \left[q_c(t) - q_s(t)\right] p(t) \\ &= \frac{1}{8} m_c^2(t)\, m_s^2(t) \sin(2\theta_\epsilon) \cos(2\theta_\epsilon) \\ &= \frac{1}{16} \sin(4\theta_\epsilon)\end{aligned} \tag{4.29}$$

From the use of the tracking error signal format $\epsilon_1 = y_c(t)\, SGM[y_s(t)] - y_s(t)\, SGM[y_c(t)]$ from (4.20) and $\epsilon_2 = \left[y_c^2(t) - y_s^2(t)\right] y_c(t)\, y_s(t)$ from (4.29), the former is most commonly used. The quadriphase Costas loop is often preferred over the Mth power loop because its circuits are less sensitive to Doppler shift and are generally capable of wider bandwidth operation.

4.6.3 Remodulator

The outputs of the hard limiters of the QPSK remodulator shown in Figure 4.9 is identical to those from the quadriphase Costas crossover loop [73, 82]. This carrier recovery technique involves demodulating the received signal, and low pass filtering to remove unwanted components of the mixing operation. As a result $y_c(t)$ and $y_s(t)$ are recovered and used to remodulate the received signal so as to remove modulation. The baseband waveforms $y_c(t)$ and $y_s(t)$ are hard limited to produce $z_c(t)$ and $z_s(t)$ respectively.

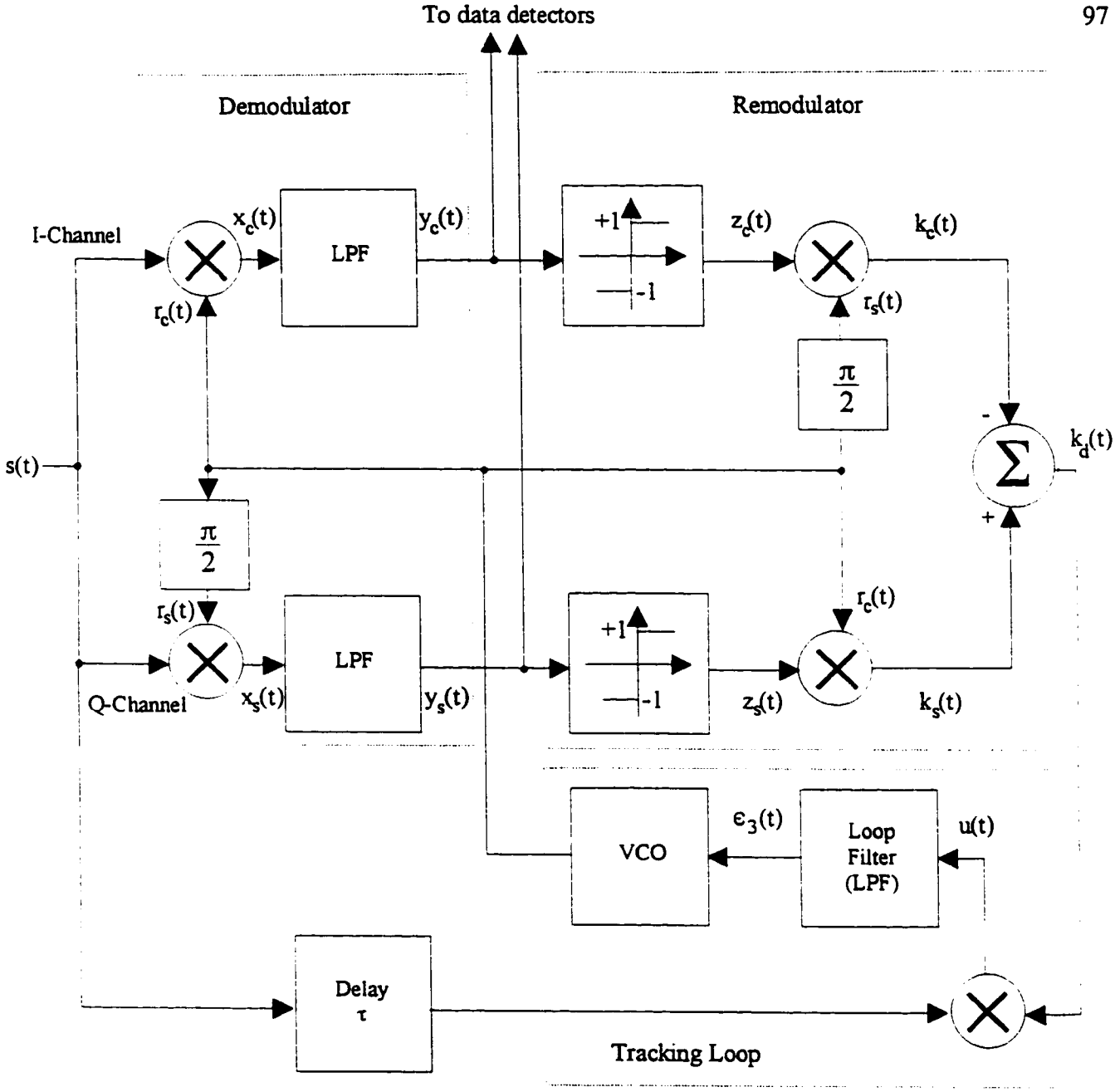


Figure 4.9: Block diagram of QPSK remodulator

The reconstructed signal $k_d(t)$ is the reference signal for the tracking loop phase detector.

The delay τ in Figure 4.9 is the time required for $s(t)$ to be processed by the

demodulator/remodulator. In the analysis, we assume that the circuit delay is zero.

The output of the remodulator is given by

$$\begin{aligned} k_c(t) &= z_c(t)\, r_s(t) \\ &= \frac{1}{2}\, m_c(t)\, \sin(\omega_c t + \theta_1) \end{aligned} \tag{4.30}$$

and

$$\begin{aligned} k_s(t) &= z_s(t)\, r_c(t) \\ &= -\frac{1}{2}\, m_s(t)\, \cos(\omega_c t + \theta_1) \end{aligned} \tag{4.31}$$

The output of the summing device produces the reference signal $k_d(t)$ which can be compared to (4.13), is a QPSK signal with carrier angular frequency ω_c and is given by

$$\begin{aligned} k_d(t) &= k_s(t) - k_c(t) \\ &= -\frac{1}{2}\left[m_c(t)\, \sin(\omega_c t + \theta_1) + m_s(t)\, \cos(\omega_c t + \theta_1)\right] \\ &= Re\left[(m_c(t) + j m_s(t))\, \exp(j\omega_c t + j\theta_1 + j\frac{\pi}{2})\right] \end{aligned} \tag{4.32}$$

Using (4.13), (4.21) and (4.32), the output $u(t)$, of the phase detector (multiplier) is simplified to

$$\begin{aligned} u(t) &= k_d(t)\, s(t) \\ &= \frac{1}{2}\left[\sin(\theta_\epsilon) - m_c(t)\, m_s(t)\, \cos(2\omega_c t + \theta_0 + \theta_1)\right] \end{aligned} \tag{4.33}$$

And when the harmonic component at angular frequency $2\omega_c$ is eliminated by the loop filter, the tracking error signal fed into the VCO circuitry is given by

$$\begin{aligned} \epsilon_3(t) &= LPF[u(t)] \\ &= \frac{1}{2}\sin(\theta_\epsilon) \end{aligned} \tag{4.34}$$

As we can see, the remodulator generates a tracking error signal similar to the quadriphase Costas loop, shown earlier. Furthermore, this method avoids the use of baseband multipliers as required in the quadriphase Costas carrier recovery loop. The remodulator technique exhibits a somewhat faster acquisition time compared to the conventional quadriphase Costas loop.

4.7 Coherent v/s Noncoherent QPSK

The coherent demodulation of PSK signals requires the local generation of a reference carrier. This may be undesirable either because of the complexity of the circuitry required for carrier recovery or in applications devoid of sufficient time for carrier acquisition. A simple approach called Differentially Phase Shift Keying (DPSK) cleverly avoids the need for a reference carrier consists of performing the demodulation by looking at the phases of the received signal in two successive intervals and estimating their difference. If the information digits have been differentially encoded at the transmitter, then

the observed phase difference at the receiver allows the recovery of the information and the removal of the phase ambiguity. DPSK can be viewed as a noncoherent version of PSK [65-66, 68-69].

4.8 Carrier Synchronizers for Constant Doppler Shift

Various synchronization schemes have been adopted to combat constant Doppler frequency shift and these techniques have been summarised in the Table 4.1 below. The table shows which methods employ the ML and phase detector (PD), and frequency detector (FD) combination, by using a check mark ✔ to indicate yes and a cross ✗ for no. In addition, the acquisition range of the scheme is also tabulated. Details about their operation can be found in the appropriate references [55, 57-58, 77-78, 83-92].

4.8.1 Frequency Detector and Phase Detector Combination

In the presence of a large initial Doppler shift, initial acquisition of a PLL, used for carrier extraction is a significant problem. A very narrow loop bandwidth is required to control the phase jitter and in the process restricts acquisition (pull-in) range. This also accounts for long acquisition times. Bandwidth efficient modulations techniques like QPSK and QAM are very sensitive to phase jitter and require very narrow loop bandwidths, in their carrier recovery [92].

Index	Reference	Schemes	Method of Operation	PD + FD	ML	$\lvert\Delta f\rvert$ [Hz]
1	[57]	Block Demodulator Schemes (constant f_d)	Conventional PLL method	✗	✗	$\leq 1/(8T_s)$
2	[55]	Block Demodulator Schemes (constant f_d)	Carrier regeneration in a block demodulator (Phase estimation)	✗	✗	$\leq 1/(8T_s)$
3	[58]		Block demodulator for LEO satellite systems (Phase and frequency estimation)	✗	✗	$\leq 1/(8T_s)$
4	[78, 83]	Frequency and Phase Estimation Schemes (constant f_d)	Sliding window frequency estimation	✗	✗	$\leq 1/(8T_s)$
5	[84]	Frequency Acquisition Aiding Schemes (constant f_d)	Modified Costas loop	✗	✗	$\leq 1/(8T_s)$
6	[85]		Simplified quadricorrelator (frequency difference detector) aid for phase-lock coherent reception	✓	✗	$\leq 1/(8T_s)$
7	[86]		Modified PLL	✓	✗	$\leq 1/(8T_s)$
8	[87]		Linearized loop models in a modified PLL	✓	✗	$\leq 1/(8T_s)$
9	[88-89]		Simplified Maximum Likelihood frequency detector	✓	✓	$\leq 1/T_s$
10	[90]		Maximum Likelihood frequency detector	✓	✓	$\leq (1+\alpha)/T_s$
11	[91]		Balanced quadricorrelator using time delay	✓	✗	$\leq 1/(8T_s)$
12	[92]		Balanced quadricorrelator utilizing filters	✓	✗	$\leq 1/(2T_s)$
13	[77]	Data-Aided Schemes (constant f_d)	Data-aided frequency estimation (Exploits the correlation of samples, 128 symbols). Result is for rolloff factor $\alpha = 0.5$	✗	✓	$\leq 1/(5T_s)$

PD Phase Detector — f_d Doppler frequency shift — PLL Phase Lock Loop — α Rolloff Factor
FD Frequency Detector — $\lvert\Delta f\rvert$ Acquisition range — LEO Low Earth Orbit — ML Maximum Likelihood

Table 4.1: Frequency aiding acquisition schemes

The two requirements, large acquisition range due to Doppler frequency shift and small phase jitter are not compatible in conventional PLL design. A common approach to overcome this difficulty, is to design a narrow-band PLL satisfying the phase jitter requirement, and to extend the acquisition range using an auxiliary frequency detector as an acquisition aiding technique. Hence, in the presence of large initial frequency offset, acquisition can be effected with the addition of a frequency detector to the conventional PLL phase detector. This method leads to short acquisition times when a large initial frequency

offset is present. The conventional FD is only useful for frequency acquisition and tracking and cannot be used for phase synchronization. The FD is used in parallel with a PD, and their respective outputs summed after separate filtering, and sent to the VCO as depicted in Figure 4.10 below.

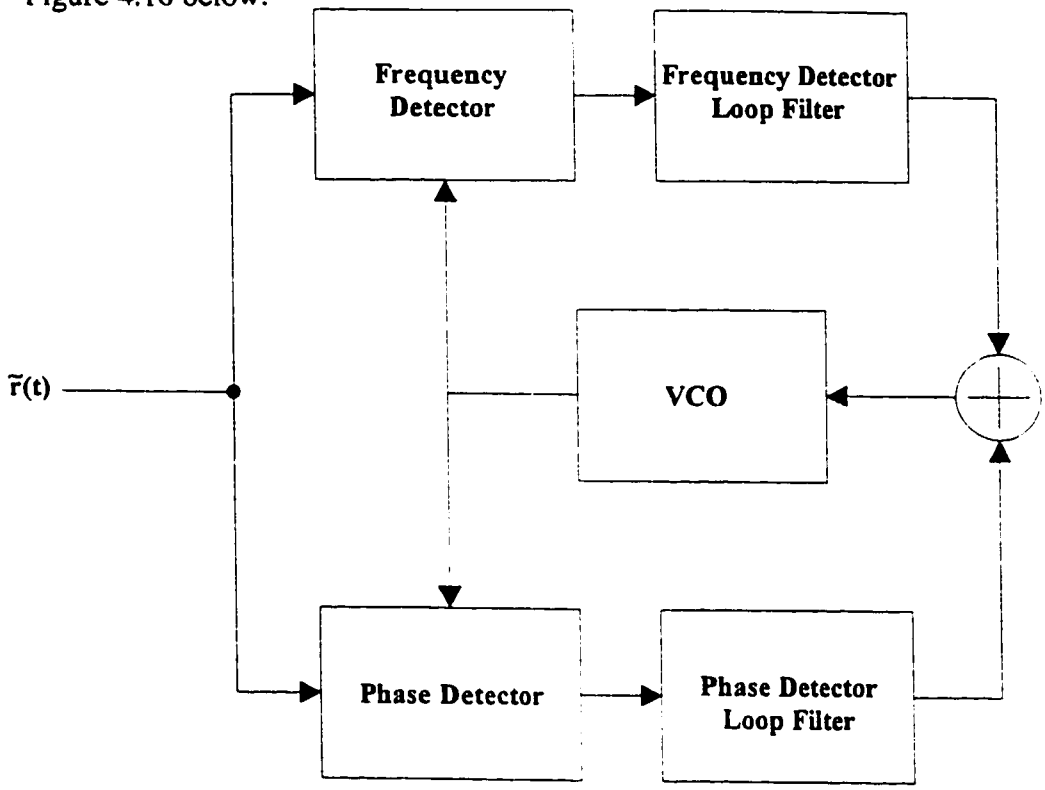


Figure 4.10: Carrier synchronizer employing PD in parallel with a FD

During the acquisition mode, the loop acquisition behaviour is governed by the FD and by the PD during steady state (lock or tracking mode). With large initial frequency offset, the FD generates a voltage (error signal) proportional to the frequency difference between the input and the VCO, and driving that difference to zero. At the same time, the output of the PD is essentially zero, thus the overall circuit behaves as a frequency-tracking loop. However, when the contribution from the FD is negligible (small), the PD takes

control over the VCO, completing the acquisition and entering the tracking mode. Ideally, when the PLL is in-lock, the output of the FD will have at least a zero mean [93].

4.8.2 QPSK Acquisition Range Fundamental Limit

Consider a received QPSK signal given by

$$r_i(t) = A\cos(\omega_c t + 2\pi\,\Delta f\, t + \theta_i) \qquad 0 \le t \le T_s \tag{4.35}$$

where noise has been ignored and Δf is the constant Doppler frequency shift experienced by the signal. The decision boundaries for demodulating the symbols $\{\theta_i;\ i = 1, 2, 3, 4\}$ are the appropriate quadrants formed by the $\sin(\omega_c t)$ and $\cos(\omega_c t)$ axes of the signal space diagram of Fig 4.1. At the end of a symbol period T_s, a symbol may be erroneously decoded if the total phase of (4.35) were to fall within either of the adjacent quadrants. To avoid this event from occurring, the maximum tolerable phase error introduced in a QPSK signal due to an initial constant Doppler frequency shift is given by

$$|2\pi\,\Delta f\, T_s| \le \frac{\pi}{4} \tag{4.36}$$

which can be re-arranged to establish the expression for the acquisition range fundamental limit of a QPSK signal [88, 91, 94].

$$|\Delta f| \leq \frac{1}{8T_s} \tag{4.37}$$

Similarly, the acquisition range fundamental limit of MPSK is given by [91]

$$|\Delta f| \leq \frac{1}{2\,M\,T_s} \tag{4.38}$$

4.8.3 LEO Doppler Frequency Shift Problem Summary

The Doppler problem experienced by LEO satellite can be described as follows with some help from Figure 4.11. In regions A and C, the Doppler shift is fairly constant and the Doppler rate is zero. However, in region B the Doppler shift is not constant and the Doppler rate is not equal to zero. In this range, where most communication events arises, Doppler is dynamic and can be considered to be a linear function of time for short period durations. In addition, the maximum Doppler shift is around f_d = 40 kHz. Applying (4.37) for low data rate applications, that is for relatively large T_s, the acquisition range $|\Delta f|$ will be of small value. Hence, such systems fail to acquire synchronization using conventional techniques, since the Doppler shift is greater than the QPSK acquisition range.

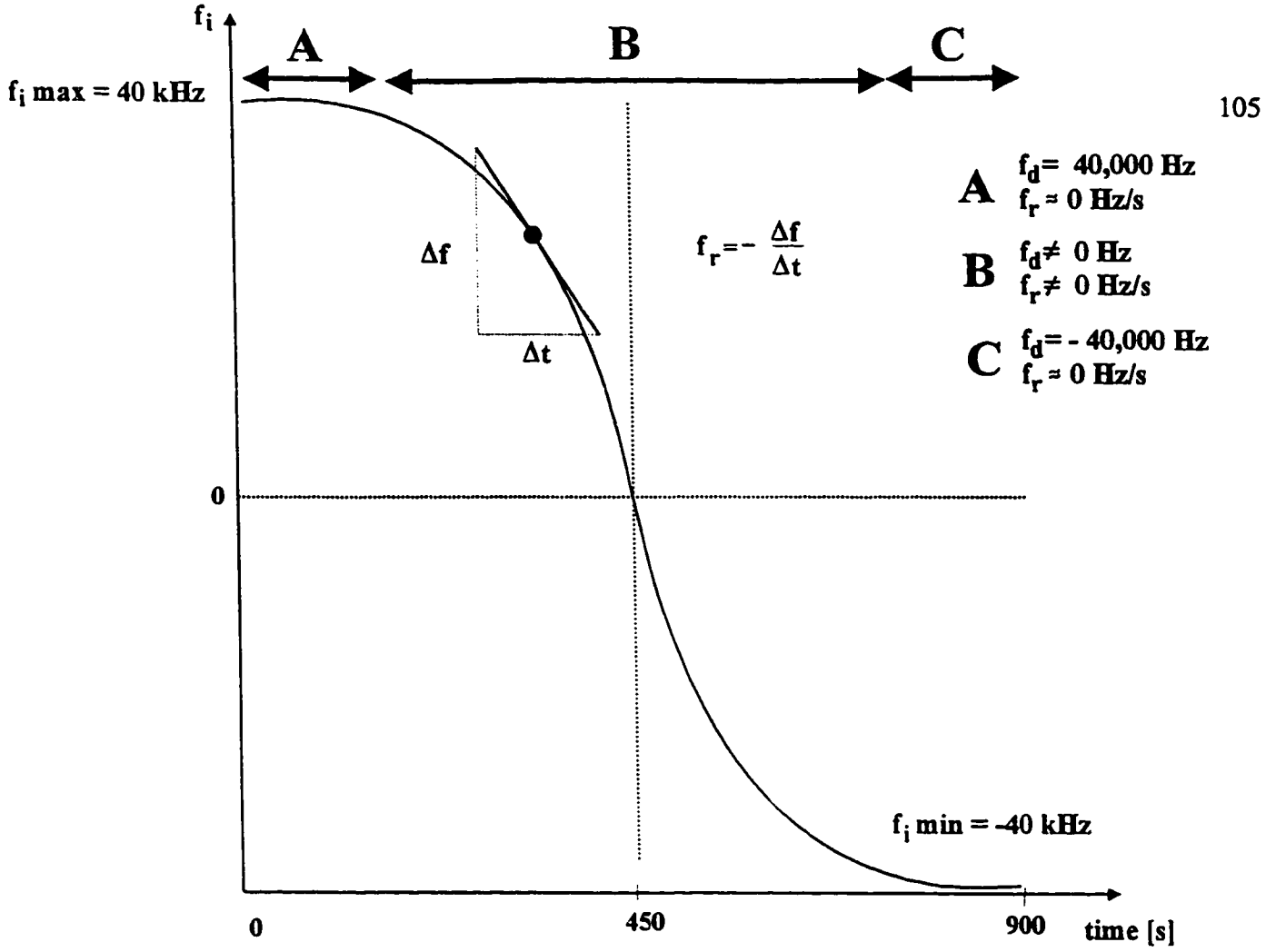


Figure 4.11: Typical LEO Doppler curve

At the present time, the maximum duration allowed to establish synchronization in personal communication systems is 5 seconds and this is due to the current synchronization methods utilized. However, reducing this period to the order of milliseconds would greatly increase the efficiency of the network. In order words, the system will have the ability to provide the service faster thus, increasing its capacity. A typical Iridium handheld receiver downlink budget calculation performed in section 2.3.6 of chapter 2 has shown that the system has a downlink margin of 2.86 dB. Furthermore, when an Iridium satellite is at minimum grazing angle, its slant range is 2480 km and this corresponds to an extra 3 dB link

margin (20log(2480E3/780E3) = 3 dB) compared to an overhead pass satellite communication link. Both cases accounts for low SNR environment. Therefore, synchronization must be effected in low SNR (or E_b/N_o) environment.

4.8.4 Shortcomings of Previously Proposed Methods

The Doppler shift encountered in typical LEO satellite systems for mobile communications (low data rate applications) is a strong function of time. Most of the solutions investigated assume that the Doppler frequency shift, f_d is a fixed value, small compared to the symbol rate, and the Doppler rate, f_r to be equal to zero. Furthermore, when the satellite is in the zenith region, Doppler rate f_r is significant.

In some of the methods a preamble of known symbols (training sequence) is appended to the beginning of each burst for carrier and clock recovery. Often, these same symbols are also used for channel equalization. The main purpose of using a preamble is to allow modulation stripping from the received signal. In some situations, modulation is stripped using an approximation or they are modelled as a random data sequence. Often the data is modelled as equiprobable symbols. The maximum length of symbols used as a preamble is 128 [77]. The use of preambles reduces the time available for passing data and is therefore not suitable for low data rate applications. Although the use of a preamble affects the efficiency of the transmission, and the effect being proportional to the length of

the preamble, the estimation process (short acquisition times) is easier to perform and works for varying SNR conditions [75]. Conversely, a fully modulated carrier (no preamble) provides a high degree of channel efficiency, but the estimation process is usually tedious, and performance degrades with low SNR. The use of a preamble in a DA algorithm which guarantees a fast and reliable carrier synchronization can only be justified if it is possible to reduce its length. Synchronization may sometimes be lost during fading or blockage. When the signal returns, there should be a fast reacquisition. This acquisition must be effected with random data, as it will be performed during data transmission. Thus the reason why for fast synchronization preambles are not required.

Most carrier synchronizers presented in Table 4.1 can achieve a tracking range of $|\Delta f| \leq 1/(8T_s)$. These methods are ideal for medium and large capacity digital radio systems but inadequate for low data rate used in mobile communications, where T_s is relatively large.

When ML techniques are employed to enable receivers to mitigate the effects of Doppler frequency shift, an observation period (estimation period) T_o of the received signal is required. Processing time has to be taken into account as well. Allowing T_o to be large yields an inefficient system because of the long delays in acquiring synchronization. In the case of low data rate receivers, as in the case of mobile satellite communications, T_o must be as small as possible. Therefore, T_o is a design parameter whose value must be set

according to acquisition time (pull-in time) requirements of the receiver.

4.9 Why Use Maximum Likelihood Approach?

QPSK represents an efficient scheme for satellite communications. In such TDMA systems, rapid carrier and clock synchronization is a must. In the presence of Doppler frequency shift, Doppler rate and burst mode operation, methods of synchronization outlined in Table 4.1 cannot provide fast acquisition with high probability. In robust coherent detection receivers, the carrier phase synchronizers, often implement ML estimation algorithms. Table 4.1 clearly shows that from the 13 schemes, those implementing ML offer superior performance. Furthermore, in general methods employing ML estimation provide design guidelines for practical estimators (receiver design). Table 4.2 depicts the carrier synchronizers that would mitigate the effects of constant Doppler shift beyond the fundamental limit of QPSK acquisition range.

Reference	Method of Operation	PD + FD	ML	$\lvert\Delta f\rvert$ [Hz]
[88-89]	Simplified Maximum Likelihood frequency detector	✓	✓	$\leq 1/T_s$
[90]	Maximum Likelihood frequency detector	✓	✓	$\leq (1+\alpha)/T_s$
[92]	Balanced quadricorrelator utilizing filters	✓	✗	$\leq 1/(2T_s)$
[77]	Data-aided frequency estimation (Exploits the correlation of samples, 128 symbols). Result is for rolloff factor $\alpha = 0.5$	✗	✓	$\leq 1/(5T_s)$

Table 4.2: Constant Doppler compensators with acquisition range $|\Delta f| > 1/(8T_s)$

Three out of 4 of these schemes employ ML approach [77, 88-90] and similarly for using an FD and PD combination [88-90, 92]. It is safe to say that the best method utilizes ML.

Low data rate QPSK signal combined with coding, where power is a scarce resource, and at the same time, low SNR environments of today's world, makes the task of a coherent receiver more challenging. ML techniques employing the use of preambles (Data-aided technique) are normally employed in medium and high data rate communication systems to attain reasonable performance with short preambles. Synchronization is still difficult, especially at low SNR values and in the case of LEO satellite systems where high Doppler shift and Doppler rate are present. In the case when synchronization needs to be effected during random data transmission, as in the situation of a rapid frequency re-acquisition, ML schemes are more likely to be successful since they assume the received data symbols as a random process. From this point of view, the ML scheme is actually a non data-aided approach compared to the complete opposite as in [88-90].

It must be noted that often ML solutions are non-implementable and therefore, physically unrealizable, but making an approximation, usually by specializing to limits of high and low SNR, for the purpose of implementation leads to near optimum (sub-optimum) realizable receiver structures.

CHAPTER 5

ML Approach to Doppler Compensation

Chapter 4 provides a rational for selecting ML as the choice of solving the Doppler shift encountered in low data rate LEO communication satellite systems. In this chapter, we first introduce Maximum Likelihood estimation of multiple parameters. Then we develop a model for the phase error introduced by Doppler frequency shift and Doppler rate and use ML to estimate the two unknowns in the phase error model. In section 5.2 we present the two implementable ML estimation solutions: grid search and ML QPSK Doppler compensator. Then in Section 5.3 we discuss how to implement the grid search and introduce an improved grid search for better results. In section 5.4 we explore all parts of the structure and implementation of the ML QPSK Doppler compensator while in section 5.5 we talk about its digital implementation.

5.1 Parameter Estimation Technique Via Maximum Likelihood

This section is devoted to introducing Maximum Likelihood estimation of multiple parameters. The likelihood function derived here will be used in the forthcoming sections of this chapter. We formulate the ML problem of estimating signal parameters in the presence of additive noise by letting the received signal $r(t)$ be represented as follows

$$r(t) = s(t,\underline{\alpha}) + n(t) \tag{5.1}$$

where $s(t,\underline{\alpha})$ is the trial signal component and $n(t)$ denotes the additive white Gaussian noise (AWGN) with zero mean and two-sided power spectral density $N_o/2$ W/Hz . The signal parameters to be estimated are contained in the vector $\underline{\alpha}$ and are unknown but deterministic (non-random). When $\underline{\alpha}$ contains more than one element, the estimation is referred to as composite hypothesis parameter estimation.

In ML estimation of signal parameters, we require that the receiver extract the estimate by observing the received signal over a time interval T_o which is called the observation interval. Usually, a total number of N symbols of the received signal is taken, one every T_s seconds, where T_s is the symbol period and $T_o = NT_s$. Estimates obtained from a single observation interval are sometimes called one shot estimates. In practice, however, the estimation is performed on a continuous basis by using tracking loops, which

can either be analog or digital that continuously update the estimates. On the other hand, one shot estimates yield insight for tracking loop implementation. In addition, they prove useful in the analysis of the performance of ML estimation and their performance can be related to that obtained with a tracking loop. Both techniques will be investigated and feasibility of implementation will be addressed.

In the limit as the observation interval T_o approaches infinity, that is for a long observation period, the conditional probability density function $p(r|\underline{\alpha})$ is proportional to the likelihood function $\Lambda(r|\underline{\alpha})$, defined as [95-97]

$$\Lambda(r|\underline{\alpha}) = \exp\left(-\frac{1}{N_o}\int_{T_o}[r(t) - s(t,\underline{\alpha})]^2 dt \right) \tag{5.2}$$

Expanding the above expression we obtain [98-99]

$$\Lambda(r|\underline{\alpha}) = \exp\left(-\frac{1}{N_o}\int_{T_o} r^2(t)dt + \frac{2}{N_o}\int_{T_o} r(t)\, s(t,\underline{\alpha})dt - \frac{1}{N_o}\int_{T_o} s^2(t,\underline{\alpha})dt \right) \tag{5.3}$$

The first term of the exponential factor is a constant, does not involve the signal parameters $\underline{\alpha}$ and equal to the incoming signal energy to noise ratio. The third term, which contains the integral $s^2(t,\underline{\alpha})$, is a constant equal to the trial signal energy to noise ratio over the observation interval T_o for any value of $\underline{\alpha}$ since QPSK signals have a constant envelope.

For signals with non-constant envelope, such as QAM signals, the third term is still a constant over a per-symbol basis. Only the second term which involves the cross correlation of the received signal $r(t)$ and the trial signal $s(t,\underline{\alpha})$ depends on the choice of $\underline{\alpha}$. Therefore, the likelihood function $\Lambda(r|\underline{\alpha})$ may be simply expressed as

$$\Lambda(r|\underline{\alpha}) = \exp\left(\frac{2}{N_o} Re\left[\int_{T_o} r(t)\ s(t,\underline{\alpha})dt \right] \right) \tag{5.4}$$

or equivalently [88, 95-96]

$$\Lambda(r|\underline{\alpha}) = \exp\left(\frac{2}{N_o} Re\left[\int_{T_o} r(t)\ s^*(t,\underline{\alpha})dt \right] \right) \tag{5.5}$$

ML theory states that $\underline{\alpha}$ is varied so as to maximize (5.5). From this point onwards, instead of formulating the ML estimation problem by using (5.2), we will refer to the simplified result given by (5.5) to avoid unnecessary large algebraic manipulations.

Note that we can divide a likelihood function by any expression that does not depend on $\underline{\alpha}$ and still have a likelihood function. Therefore, the same result given by (5.5) can also be obtained if we divide the likelihood function (5.2), by $p(r)$

$$p(r) = \exp\left(-\frac{1}{N_o} \int_{T_o} r(t)^2 dt \right) \tag{5.6}$$

where the terms remaining after performing the division are the second and third terms given in (5.3), since (5.2) and (5.6) are in exponential form [98]. This is equivalent to normalizing (5.2) with respect to the incoming SNR.

5.2 Defining the Doppler Shift Problem

The prime motivation of the work in this section is to present insight on the effects of Doppler in a QPSK signal on the downlink path. Our aim is to look into the possibility of a ML solution for the Doppler shift problem. If however a closed form solution is feasible, its performance compared to other existing methods will be investigated. Furthermore, simplification of implementation of the method is also of great concern, since we are considering its usage in low cost mobile terminals.

5.2.1 Doppler Shifted QPSK Signal Model

Let $f_i(t)$ be the dynamic Doppler frequency shift and θ_o is an arbitrary phase shift introduced by the channel, and in the absence of any side information, it is assumed to be an unknown constant, independent of time over a specified short period interval, but generally a random variable which is assumed to be uniformly distributed in the interval $(0, 2\pi)$. The total phase shift, $\mu(t)$ of the carrier, f_c caused by Doppler shift and channel transmission can

be written as

$$\mu(t) = 2\pi \int_0^t f_i(t)\, dt + \theta_o = \int_0^t \Delta(t)\, \omega_c\, dt + \theta_o = \Psi(t) + \theta_o \tag{5.7}$$

where $\Psi(t)$ is the phase error introduced solely due to Doppler shift and Doppler rate and the normalized Doppler shift is given by

$$\Delta(t) = \frac{f_i(t)}{f_c} \tag{5.8}$$

The received bandpass signal is of the form

$$r(t) = s(t) + n(t) \tag{5.9}$$

where the noise $n(t)$ is assumed to be additive white Gaussian noise with zero mean and variance $\sigma_n^2 = (\frac{1}{2}N_o)/E_s$, that is, two-sided power spectral density of $(\frac{1}{2}N_o)/E_s$ W/Hz, where E_s is the symbol energy. The AWGN $n(t)$ is represented in its canonical form by

$$\begin{aligned} n(t) &= n_c(t)\cos(\omega_c t) - n_s(t)\sin(\omega_c t) \\ &= Re\left[\tilde{n}(t)\exp(j\omega_c t)\right] \end{aligned} \tag{5.10}$$

where the complex envelope of the noise is given by

$$\tilde{n}(t) = n_c(t) + j\, n_s(t) \tag{5.11}$$

where $n_c(t)$ is the in-phase component and $n_s(t)$ is the quadrature component of the $n(t)$. Both $n_c(t)$ and $n_s(t)$ are zero mean independent Gaussian processes with variance $\sigma_{n_c}^2 = \sigma_{n_s}^2 = (\frac{1}{2}N_o)/E_s$ [100-103].

The noisy Doppler shifted QPSK signal is given in the following format

$$s(t) = \sum_{k=-\infty}^{\infty} h(t-kT_s)\cos\left(\omega_c t + \Psi(t) + \theta_o + \phi_k\right) \qquad for \quad k = 0, \pm1, \pm2, \pm3,\ldots \tag{5.12}$$

and the information bearing function is

$$\phi_k = (2n-1)\frac{\pi}{4} \qquad for \quad n = 1, 2, 3, 4 \tag{5.13}$$

The discrete random variable ϕ_k is the phase shift of the k^{th} data symbol and can assume only one of four values. If the four symbols are statistically independent and equiprobable, the probability density function of the discrete random variable ϕ_k is expressed as the following and shown in Figure 5.1.

$$p(\phi_k) = \sum_{n=1}^{4} P\left(\phi_k-(2n-1)\frac{\pi}{4}\right)\ \delta\left(\phi_k-(2n-1)\frac{\pi}{4}\right) \tag{5.14}$$

where $\delta(x)$ denotes a unit impulse at $x = 0$ and $P(\phi_k)$ denotes the a priori probability of the

$$P\left(\phi_k - (2n-1)\frac{\pi}{4}\right) = \frac{1}{4} \qquad \textit{for} \quad n = 1, 2, 3, 4 \tag{5.15}$$

indicated discrete random variable ϕ_k and given by

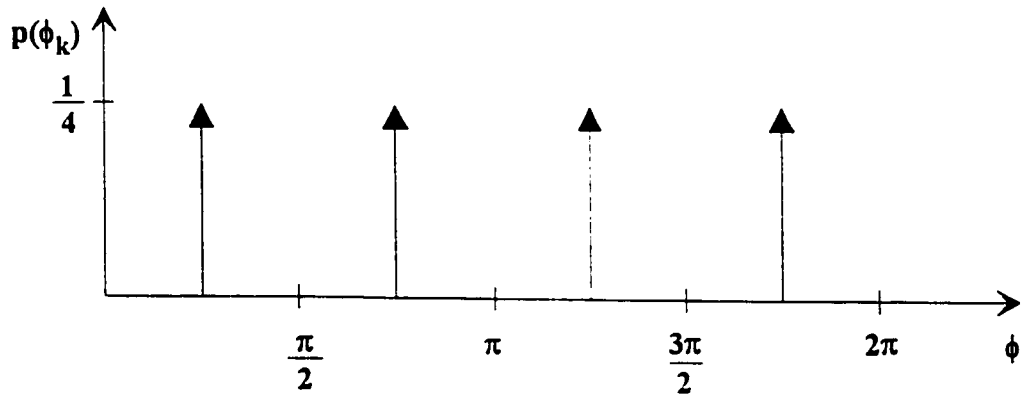


Figure 5.1: Probability density function of the discrete random variable ϕ_k

Furthermore, T_s is the symbol period and $h(t)$ represents a real-valued signal pulse shape. For simulation purposes we shall investigate cases when $h(t)$ is a rectangular unit pulse of duration T_s and raised cosine with rolloff factor $\alpha = 0.4$.

An alternate representation of $s(t)$ is given by

$$s(t) = \sum_{k=-\infty}^{\infty} \left\{ a_k\, h(t-kT_s)\, \cos\left(\omega_c t + \Psi(t) + \theta_o\right) - b_k\, h(t-kT_s)\, \sin\left(\omega_c t + \Psi(t) + \theta_o\right) \right\} \tag{5.16}$$

where a_k and b_k are statistically independent and equally likely data symbols and have values $\pm 1/\sqrt{2}$ and $|a_k| = |b_k| = 1/\sqrt{2}$.

5.2.2 Maximum Likelihood Strategy

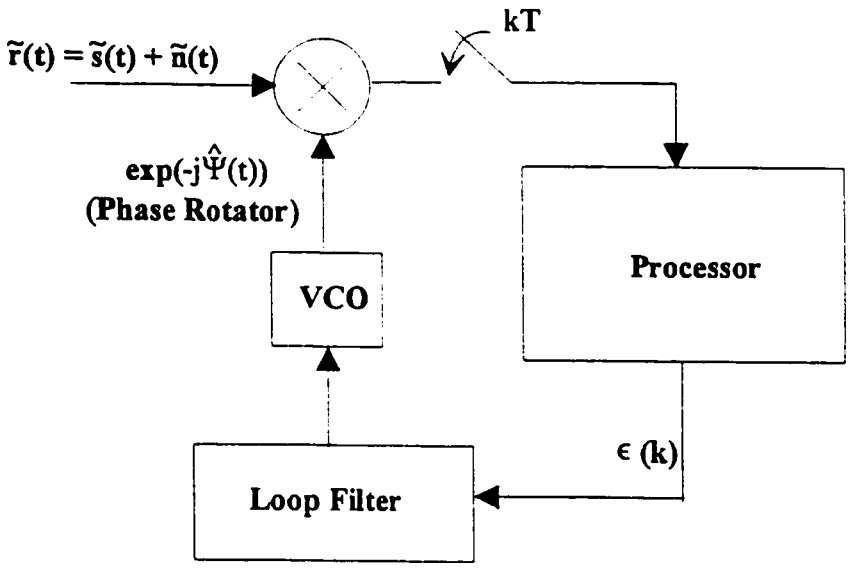


Figure 5.2: Block diagram of receiver structure implementing ML

The proposed ML carrier frequency compensation is implemented by first deriving an estimate $\hat{\Psi}(t)$ from its nominal value then compensating for Doppler shift by using this estimate to remove the phase introduced into the transmitted signal due to Doppler shift and Doppler rate. An analog block diagram to illustrate how the result from the ML strategy can

be used to compensate for the initial Doppler shift and Doppler rate is given in Figure 5.2. Since

$$r(t) = Re\left[\tilde{r}(t)\ \exp(j\omega_c\ t)\right] \tag{5.17}$$

and

$$s(t) = Re\left[\tilde{s}(t)\ \exp(j\omega_c\ t)\right] \tag{5.18}$$

then the input to the receiver structure is the complex envelope of $r(t)$, that is the lowpass equivalent of $r(t)$, and is given by

$$\tilde{r}(t) = \tilde{s}(t) + \tilde{n}(t) \tag{5.19}$$

where the signal $\tilde{s}(t)$ is given by

$$\tilde{s}(t) = \sum_{k=-\infty}^{\infty} h(t-kT_s)\ \exp\left(j[\Psi(t) + \theta_o + \phi_k]\right) \qquad for\ \ k = 0, \pm1, \pm2, \pm3,... \tag{5.20}$$

and similarly

$$\tilde{s}^*(t) = \sum_{k=-\infty}^{\infty} h(t-kT_s)\ \exp\left(-j[\Psi(t) + \theta_o + \phi_k]\right) \qquad for\ \ k = 0, \pm1, \pm2, \pm3,... \tag{5.21}$$

To simplify the presentation of the ML technique, the following assumptions are made;

1. Clock recovery has been performed prior to carrier synchronization (symbol timing is known).
2. Discrete time samples of the received signal are taken from the output of a pulsed shaped matched filter every $t = kT_s$ seconds.
3. The pulse shape satisfies the Nyquist criterion for zero interference.

Furthermore, we mentioned earlier that; the carrier $f_c = \omega_c/(2\pi)$ is a deterministic (non-random) variable; the symbol phase ϕ_k is independent and equally likely and the arbitrary phase shift θ_o is a uniformly distributed random variable. However, the low pass equivalent noise, $\tilde{n}(t)$ is AWGN with zero mean and variance $\sigma_{\tilde{n}}^2 = N_o/E_s$, where the latter is twice the variance of $n(t)$. Furthermore, $1/\sigma_{\tilde{n}}^2 = E_s/N_o$ is also the symbol SNR [77, 102]. The variable $\Psi(t)$ is deterministic (non-random), unknown, varies with time but constant during every symbol period and depends on the two parameters f_d and f_r which are estimated via the ML process.

5.2.3 Maximum Likelihood Formulation

Our main goal is to use the ML technique to estimate $\Psi(t)$ based on the observation of a finite interval T_o record length of the received signal. We would like to obtain joint ML estimates of f_d and f_r, namely; $\hat{f}_d$ and $\hat{f}_r$ which would allow us to compute the estimate $\hat{\Psi}(t)$. Referring to Figure 5.3, and focussing on the region labelled B, we suggest the relationship that the instantaneous frequency, $f_i(t)$ (deviation from its nominal value of f_c) is dependent on the Doppler rate which is linear function of time over short time durations and can be expressed as follows

$$f_i(t) = f_d + f_r\, t \tag{5.22}$$

That is, the received frequency contains a Doppler shift component as well as a Doppler rate component, where the latter accounts for the Doppler being dynamic. Referring to the S-shape Doppler curve for typical LEO satellite systems, the flat regions (satellite at horizon, regions A and C) present a relatively constant Doppler frequency shift and negligible Doppler rate. But as the satellite approaches zenith (centre of region B), with respect to the user terminal, the Doppler shift is modelled as a linear function of time.

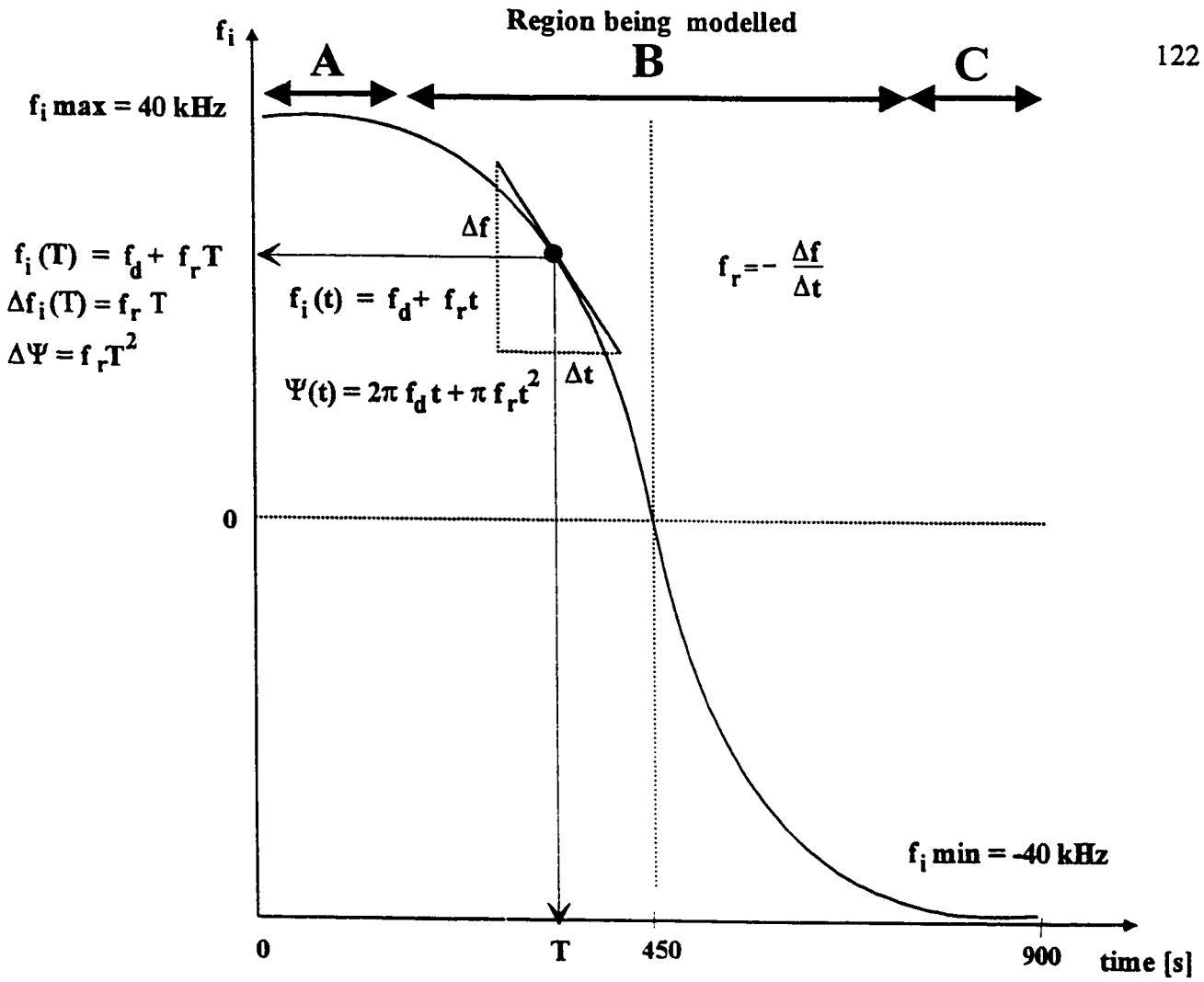


Figure 5.3: Doppler shift and Doppler rate relationship

The relationship between f_d, f_r and the phase, $\Psi(t)$ due to Doppler can now be obtained by using (5.22) and is given as follows

$$\begin{aligned}\Psi(t) &= 2\pi \int_0^t f_i(t)\, dt \\ &= 2\pi \int_0^t \{f_d + f_r\, t\}\, dt \\ &= 2\pi f_d\, t + \pi f_r\, t^2\end{aligned} \tag{5.23}$$

At each time $t = kT_s$, we would like to obtain a ML estimate of the phase error introduced by Doppler shift and Doppler rate, based on the received signal over the immediate past N symbol interval. Now we turn our attention to the development of the likelihood function. First, we rewrite (5.5) in terms of the variance of the noise process, low pass equivalent in our case, thus we have

$$\Lambda(r|\underline{\alpha}) = \exp\left(\frac{1}{\sigma^2} Re\left[\int_{T_o} r(t)\ s^*(t,\underline{\alpha}) dt \right] \right) \tag{5.24}$$

Assuming that we have agreed on the above points, we formulate the ML problem by first expressing the simplified joint conditional likelihood function (5.24) which was developed in section 5.1 of $\tilde{r}(t)$ given, ϕ_k, θ_o, $\sigma_{\tilde{n}}$ and $\Psi(t)$ and using (5.19), (5.21) and (5.24) we obtain

$$\Lambda(\tilde{r}(T_o)|\Psi(t),\ \theta_o,\ \phi_k) = \prod_{k=0}^{N-1} \exp\left(\frac{E_s}{N_o} Re\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[\Psi(t) + \theta_o + \phi_k]\right) dt \right] \right) \tag{5.25}$$

Making use of a derivative of Euler's formula [104]

$$\exp(-j\ \phi_k) = \cos(\phi_k) - j\ \sin(\phi_k) \tag{5.26}$$

it is possible to expand (5.25) to obtain

$$\Lambda(\tilde{r}(T_o)|\Psi(t),\ \theta_o,\ \phi_k) = \prod_{k=0}^{N-1} \cos(\phi_k)\ \exp\left(\frac{E_s}{N_o}\ Re\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[\Psi(t) + \theta_o]\right)dt \right] \right)$$
$$+ \sin(\phi_k)\ \exp\left(\frac{E_s}{N_o}\ Im\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[\Psi(t) + \theta_o]\right)dt \right] \right) \quad \textbf{(5.27)}$$

Data dependence of the likelihood function is achieved by averaging the likelihood function over all possible values of ϕ_k [105].

$$\Lambda(\tilde{r}(T_o)|\Psi(t),\ \theta_o) = \int_{-\infty}^{\infty} \Lambda(\tilde{r}(T_o)|\Psi(t),\ \theta_o,\ \phi_k)\ p(\phi_k)\ d\phi_k \quad \textbf{(5.28)}$$

Using (5.13-5.16) we substitute for $p(\phi_k)$ into (5.28) and the result is given by

$$\Lambda(\tilde{r}(T_o)|\Psi(t),\ \theta_o) = \prod_{k=0}^{N-1} \sum_{n=1}^{4} \frac{1}{4}\ \exp\left(\frac{E_s}{N_o}\ Re\left[U(kT_s)\right] \cos\left((2n-1)\frac{\pi}{4} \right) \right.$$
$$\left. + \frac{E_s}{N_o}\ Im\left[U(kT_s)\right] \sin\left((2n-1)\frac{\pi}{4} \right) \right) \quad \textbf{(5.29)}$$

where $U(kT_s)$ is given by

$$
\begin{aligned}
U(kT_s) &= \int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\; h(t-kT_s)\; \exp\Big(-j[\Psi(t) + \theta_o]\Big)dt \\
&= \int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\; h(t-kT_s)\; \exp\Big(-j[2\pi f_d\, t + \pi f_r\, t^2 + \theta_o]\Big)dt
\end{aligned}
\tag{5.30}
$$

We first perform the summation and the resultant is given by

$$
\begin{aligned}
\Lambda(\tilde{r}(T_o)|\Psi(t),\, \theta_o) = \prod_{k=0}^{N-1} & \frac{1}{4} \exp\left(\frac{E_s}{N_o} Re\big[U(kT_s)\big] \cos\left(\frac{\pi}{4}\right) + \frac{E_s}{N_o} Im\big[U(kT_s)\big] \sin\left(\frac{\pi}{4}\right) \right) \\
& + \frac{1}{4} \exp\left(\frac{E_s}{N_o} Re\big[U(kT_s)\big] \cos\left(\frac{3\pi}{4}\right) + \frac{E_s}{N_o} Im\big[U(kT_s)\big] \sin\left(\frac{3\pi}{4}\right) \right) \\
& + \frac{1}{4} \exp\left(\frac{E_s}{N_o} Re\big[U(kT_s)\big] \cos\left(\frac{5\pi}{4}\right) + \frac{E_s}{N_o} Im\big[U(kT_s)\big] \sin\left(\frac{5\pi}{4}\right) \right) \\
& + \frac{1}{4} \exp\left(\frac{E_s}{N_o} Re\big[U(kT_s)\big] \cos\left(\frac{7\pi}{4}\right) + \frac{E_s}{N_o} Im\big[U(kT_s)\big] \sin\left(\frac{7\pi}{4}\right) \right)
\end{aligned}
\tag{5.31}
$$

which simplifies further to the following

$$
\begin{aligned}
\Lambda(\tilde{r}(T_o)|\Psi(t),\, \theta_o) = \prod_{k=0}^{N-1} & \frac{1}{2} \cosh\left(\frac{E_s}{\sqrt{2}N_o} Re\big[U(kT_s)\big] + \frac{E_s}{\sqrt{2}N_o} Im\big[U(kT_s)\big] \right) \\
& + \frac{1}{2} \cosh\left(\frac{E_s}{\sqrt{2}N_o} Re\big[U(kT_s)\big] - \frac{E_s}{\sqrt{2}N_o} Im\big[U(kT_s)\big] \right)
\end{aligned}
\tag{5.32}
$$

To further simplify (5.32) we make use of the hyperbolic identity

$$2\cosh(A)\cosh(B) = \cosh(A+B) + \cosh(A-B) \tag{5.33}$$

and we obtain the data averaged likelihood function given by

$$\Lambda(\bar{r}(T_o)|\Psi(t),\theta_o) = \prod_{k=0}^{N-1} \cosh\left(\frac{E_s}{\sqrt{2}N_o} Re[U(kT_s)]\right) \cosh\left(\frac{E_s}{\sqrt{2}N_o} Im[U(kT_s)]\right) \tag{5.34}$$

and taking the natural log of (5.34) we get the corresponding log-likelihood function which is given by

$$L(\bar{r}(T_o)|\Psi(t),\theta_o) = \sum_{k=0}^{N-1} \ln\cosh\left(\frac{E_s}{\sqrt{2}N_o} Re[U(kT_s)]\right) + \ln\cosh\left(\frac{E_s}{\sqrt{2}N_o} Im[U(kT_s)]\right) \tag{5.35}$$

Results from [95, 102, 106] has similar presentation as the one presented here in (5.35). Methods [102] and [106] use ML estimates of phase θ_o and assumes independent and equally likely data symbols. Conversely, [95] method uses ML estimates of phase θ_o and time delay τ, and assumes the data symbols to be independent and Gaussian.

Unfortunately (5.35) is highly non-linear and an exact optimal solution is difficult to obtain, hence the ML estimation of $\hat{f}_d$ and $\hat{f}_r$ is not easily implemented.

5.2.4 Approximations and Implementable Solutions

In typical LEO satellite systems, power is a scarce resource. This low SNR environment makes the synchronization process in receivers a more challenging task [100]. Furthermore, using (5.35) to obtain the ML estimates, requires the knowledge about the operating E_s/N_o to evaluate ln cosh(x) given by (5.35) [101, 107]. Therefore, the estimator is not practical. But looking at the asymptotic behaviour of ln cosh(x) function, in terms of high and low SNR, which are commonly used in communications, it provides us with readily implementable results which are practicable.

The first order approximation for cosh(x) is given by [72, 95, 97, 108-109]

$$\begin{aligned} \cosh(x) &\approx \exp\left(\frac{x^2}{2}\right) && |x| \ll 1 \\ &\approx \frac{1}{2}\exp(|x|) && |x| \gg 1 \end{aligned} \tag{5.36}$$

Hence, by taking the natural log yields [99, 109]

$$\begin{aligned} \ln\cosh(x) &\approx \frac{x^2}{2} && |x| \ll 1 \\ &\approx |x| && |x| \gg 1 \end{aligned} \tag{5.37}$$

Using the quadratic approximation from (5.37) for $|x| << 1$, which applies for situations of

low SNR (the case when the satellite is at the horizon) and the log-likelihood function, we obtain an approximate log-likelihood function given by

$$L(\bar{r}(T_o)|\Psi(t)) = \sum_{k=0}^{N-1} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s)\right]^2 + \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 Im\left[U(kT_s)\right]^2 \qquad \textbf{(5.38)}$$

Equivalently, we can also say

$$\begin{aligned} L(\bar{r}(T_o)|\Psi(t)) &= \sum_{k=0}^{N-1} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left|U(kT_s)\right|^2 \\ &= \sum_{k=0}^{N-1} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 U(kT_s)\ U(kT_s)^* \end{aligned} \qquad \textbf{(5.39)}$$

Investigating (5.39), we notice that it is independent of θ_o and this is due mainly to (5.35) having been transformed into quadratic expressions (5.38) and (5.39). This can be shown by using the identity

$$\left|U(kT_s)\right|^2 = U(kT_s)\ U(kT_s)^* \qquad \textbf{(5.40)}$$

and by using (5.30) and the above identity on (5.39) and taking the time invariant random variable θ_o out of the integration operation, we can show that (5.39) is independent of θ_o.

$$L(\tilde{r}(T_o)|\Psi(t)) = \sum_{k=0}^{N-1} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \exp(-j\theta_o)\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\, h(t-kT_s)\, \exp(-j\Psi(t))dt\right] \times \exp(j\theta_o)\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\, h(t-kT_s)\, \exp(-j\Psi(t))dt\right]^* \quad \textbf{(5.41)}$$

Conversely, $\Psi(t)$ cannot be removed from the integration as was θ_o, since the former is time dependent. For joint ML estimation, the trial parameter $\Psi(t)$ has to be chosen in such a way that the likelihood function (5.35) attains its maximum. Hence we maximize the log likelihood function (5.39) (which is mathematically equivalent to maximizing the likelihood function and for simplicity), to obtain $\hat{\Psi}(t)$. One technique to determine $\hat{\Psi}(t)$, is to compute (5.38) for various initial guesses of $\Psi(t)$ by using various combinations of f_d and f_r. Then select the one combination of $\hat{f}_d$ and $\hat{f}_r$ that maximizes (5.39).

Strictly speaking, $\hat{\Psi}(t)$ is not a ML estimate but $\hat{f}_d$ and $\hat{f}_r$ are joint ML estimates. The method employed to obtain ML estimates in this manner is referred to as a grid search technique. This is the first implementable solution and it is often computationally intensive but the only option for large initial frequency offsets. Furthermore, this technique does not allow tracking of changes in f_d and f_r during the observation period T_o. Therefore, to obtain the necessary condition for the above technique to work, we must assume that f_d and f_r have

constant values throughout the observation period T_o. Details on how to perform the grid search will be described in section 5.3. One characteristic of such an algorithm consists of the fact that the ML estimates $\hat{f}_d$ and $\hat{f}_r$ are explicitly computed, and depends only on direct processing of the received signal $\tilde{r}(t)$ prior to any Doppler compensation. Hence, this method does not require a negative feedback loop and VCO for coherent detection as in PLL [106].

We now look into the possibility of a technique that would allow us to track changes in f_d and f_r. The necessary condition for this is we must assume that f_d and f_r have constant values during each observable symbol period. We differentiate (5.35) with respect to f_d and f_r and we obtain

$$\frac{\partial L(\tilde{r}(T_o)|f_d, f_r, \theta_o)}{\partial f_d} = \sum_{k=0}^{N-1} \frac{E_s}{\sqrt{2}N_o} Re\left[\frac{\partial U(kT_s)}{\partial f_d}\right] \tanh\left(\frac{E_s}{\sqrt{2}N_o} Re[U(kT_s)]\right) + \sum_{k=0}^{N-1} \frac{E_s}{\sqrt{2}N_o} Im\left[\frac{\partial U(kT_s)}{\partial f_d}\right] \tanh\left(\frac{E_s}{\sqrt{2}N_o} Im[U(kT_s)]\right) \tag{5.42}$$

and similarly

$$\frac{\partial L(\tilde{r}(T_o)|f_d, f_r, \theta_o)}{\partial f_r} = \sum_{k=0}^{N-1} \frac{E_s}{\sqrt{2}N_o} Re\left[\frac{\partial U(kT_s)}{\partial f_r}\right] \tanh\left(\frac{E_s}{\sqrt{2}N_o} Re[U(kT_s)]\right) + \sum_{k=0}^{N-1} \frac{E_s}{\sqrt{2}N_o} Im\left[\frac{\partial U(kT_s)}{\partial f_r}\right] \tanh\left(\frac{E_s}{\sqrt{2}N_o} Im[U(kT_s)]\right) \tag{5.43}$$

To obtain the ML estimates, we set (5.42) and (5.43) to zero. However, both expressions are non-linear and, hence, an exact optimal solution is still not feasible. The partial derivative components are given by [110]

$$\frac{\partial U(kT_s)}{\partial f_d} = -j2\pi \int_{kT_s}^{(k+1)T_s} t\, \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d\, t + \pi f_r\, t^2 + \theta_o]\right) dt \qquad \textbf{(5.44)}$$

and

$$\frac{\partial U(kT_s)}{\partial f_r} = -j\pi \int_{kT_s}^{(k+1)T_s} t^2\, \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d\, t + \pi f_r\, t^2 + \theta_o]\right) dt \qquad \textbf{(5.45)}$$

The expression given by (5.42) and (5.43) cannot be simplified exactly in any obvious way. Instead of using the tanh(x) non-linearity, we can make use of the following approximations for high and low SNR situations which would imply practical implementations [72, 106, 108-109]

$$\begin{aligned} \tanh(x) &\approx x && |x| \ll 1 \\ &\approx sgn\{x\} && |x| \gg 1 \end{aligned} \qquad \textbf{(5.46)}$$

The second implementable solution is obtained by using the approximation (5.46) on (5.42) and (5.43) and employing the result for a feedback loop implementation, we set $N = 1$ and get an expression which is equivalent to a frequency-error detector producing an error $\underline{\epsilon}(k)$ which is given in a vector format. The frequency-error goes to zero when $\Psi(t) = \hat{\Psi}(t)$ and

is given as follows

$$\underline{\epsilon}(k) = \begin{bmatrix} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_d}\right] Re[U(kT_s)] + \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Im\left[\frac{\partial U(kT_s)}{\partial f_d}\right] Im[U(kT_s)] \\ \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_r}\right] Re[U(kT_s)] + \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Im\left[\frac{\partial U(kT_s)}{\partial f_r}\right] Im[U(kT_s)] \end{bmatrix} \quad \textbf{(5.47)}$$

and using the identity below

$$\begin{aligned} Re[A \times B^*] &= Re\left[Re[A]\ Re[B^*] - Im[A]\ Im[B^*] + j\ Re[A]\ Im[B^*] + j\ Im[A]\ Re[B]\right] \\ Re[A \times B^*] &= Re\left[Re[A]\ Re[B^*] + Im[A]\ Im[B] - j\ Re[A]\ Im[B] + j\ Im[A]\ Re[B]\right] \\ Re[A \times B^*] &= Re[A]\ Re[B] + Im[A]\ Im[B] \end{aligned} \quad \textbf{(5.48)}$$

and (5.47) can be further simplified to the following

$$\underline{\epsilon}(k) = \begin{bmatrix} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s)\ \frac{\partial U(kT_s)^*}{\partial f_d}\right] \\ \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s)\ \frac{\partial U(kT_s)^*}{\partial f_r}\right] \end{bmatrix} \quad \textbf{(5.49)}$$

and has the same form as produced by [88-90] if we set $f_r = 0$ and except for a scaling factor. Furthermore, their methods involve ML estimates of the phase θ_o. The summation over N is performed by the feedback loop which is equivalent to the expectation (time average) $E[\underline{\epsilon}(k)]$.

It is interesting to note that if we differentiate (5.39) we obtain the same result produced in (5.49) for the case when $N = 1$.

$$\underline{\epsilon}(k) = \begin{bmatrix} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left[\frac{\partial U(kT_s)}{\partial f_d}\right] \left[U(kT_s)\right]^* + \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left[U(kT_s)\right] \left[\frac{\partial U(kT_s)}{\partial f_d}\right]^* \\ \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left[\frac{\partial U(kT_s)}{\partial f_r}\right] \left[U(kT_s)\right]^* + \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left[U(kT_s)\right] \left[\frac{\partial U(kT_s)}{\partial f_r}\right]^* \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s) \frac{\partial U(kT_s)^*}{\partial f_d}\right] \\ \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s) \frac{\partial U(kT_s)^*}{\partial f_r}\right] \end{bmatrix} \qquad \textbf{(5.50)}$$

T_o is also a design parameter and allowing a large value of T_o yields an inefficient system due to lengthy delays in acquiring synchronization. For LEO personal communications systems, time requirements T_o max < 5 seconds and this should include processing overhead.

Now we turn our attention to the comparison of the result given by (5.49) with tested results by [88-90]. All three performed the ML estimation of the constant Doppler shift f_d, and the phase θ_o. Their ML based frequency detector is implemented by the block diagram depicted in Figure 5.4.

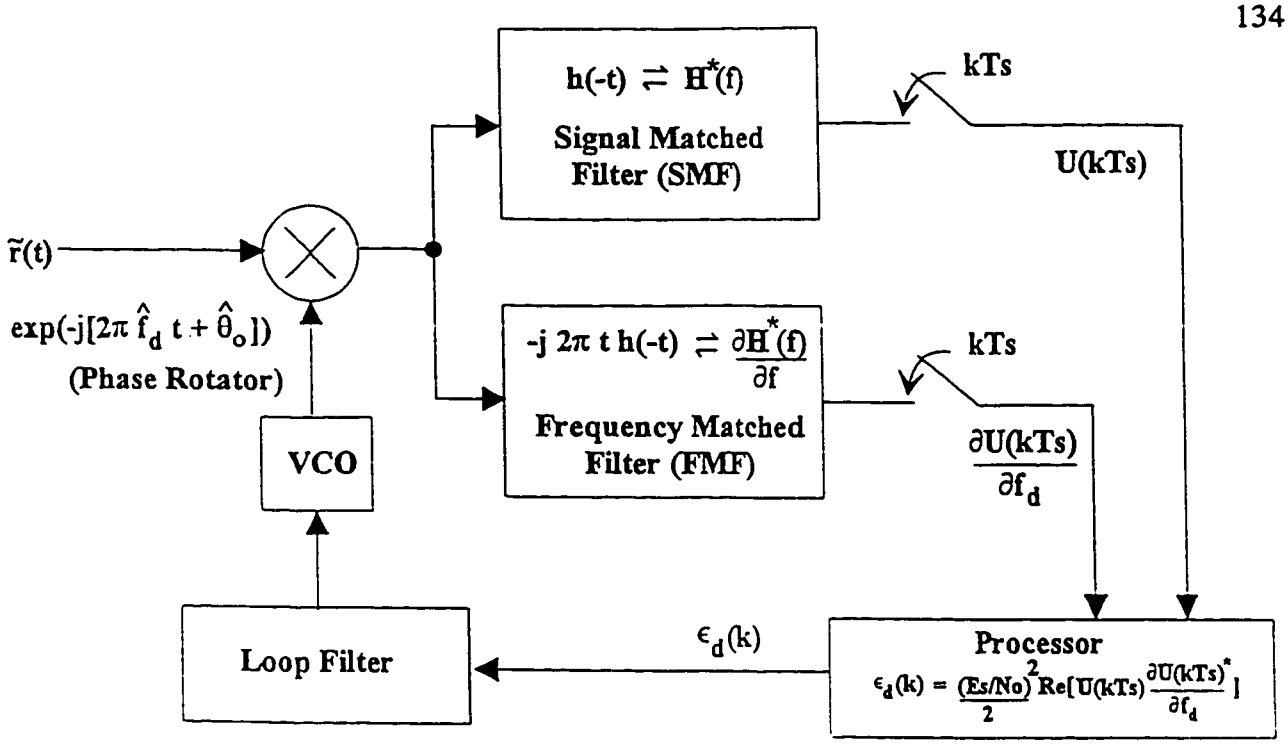


Figure 5.4: Block diagram of ML based frequency detector

5.3 Estimation of $\Psi(t)$ Using Grid Search Technique (Initial Estimator)

The trial parameter $\Psi(t)$ depends on two variables, namely; f_d and f_r. Therefore, locating the maximum of (5.39) involves a two-dimensional search, known as the grid search. The basic two-dimensional grid search is explained in section 5.3.1, while an improved version for a better solution estimate is discussed in section 5.3.2.

5.3.1 Grid Search

Here we attempt to find the initial estimating parameters $\hat{f}_d$ and $\hat{f}_r$ which would be used by the ML QPSK Doppler tracking algorithm. We proceed as follows;

1. Select starting trial values for both $\hat{f}_d$ and $\hat{f}_r$ namely $\bar{f}_d$ and $\bar{f}_d$.

2. Compute the log-likelihood function, which is repeated below with all the variables substituted in, for the above initial value of parameters $\bar{f}_d$ and $\bar{f}_r$.

$$L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r\ ) = \sum_{k=0}^{N-1} \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \left[\int_{kT_s}^{(k+1)T_s} \bar{r}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ \bar{f}_d\ t\ +\ \pi\ \bar{f}_r\ t^2]\right) dt\right] \times \left[\int_{kT_s}^{(k+1)T_s} \bar{r}(t)^*\ h(t-kT_s)\ \exp\left(+j[2\pi\ \bar{f}_d\ t\ +\ \pi\ \bar{f}_r\ t^2]\right) dt\right] \quad \textbf{(5.51)}$$

3. Go to step 1 and increase $\bar{f}_d$ by $\Delta\bar{f}_d$ while leaving $\bar{f}_r$ a constant. That is

$$\bar{f}_d(i)\ =\ \bar{f}_d(i-1)\ +\ \Delta\bar{f}_d \quad \textbf{(5.52)}$$

Save all values of $L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r\ )$ for later usage.

4. Repeat steps 2 and 3 until the range of interest for the parameter $\hat{f}_d$ has been

searched.

5. Go to step 1 and increase $\bar{f}_r$ by $\Delta\bar{f}_r$ then perform steps 2, 3 and 4.

$$\bar{f}_r(i) = \bar{f}_r(i-1) + \Delta\bar{f}_r \tag{5.53}$$

Repeat step 5 until the range of interest for the parameter $\hat{f}_r$ has been searched. Here we are trying to obtain $L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r\ )$ for all combinations of $\bar{f}_d$ and $\bar{f}_r$, which is of our interest.

6. Choose the largest stored value of $L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r\ )$. The trial parameter combination corresponding to this, are the desired parameter estimates we are seeking.

The reason for this is as follows. We would expect that the log-likelihood function $L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r\ )$ to have a local maximum at $\hat{f}_d = \bar{f}_d(n)$ and $\hat{f}_r = \bar{f}_r(m)$, and fall off in all directions from the sides of the maximum as shown in Figure 5.5 below.

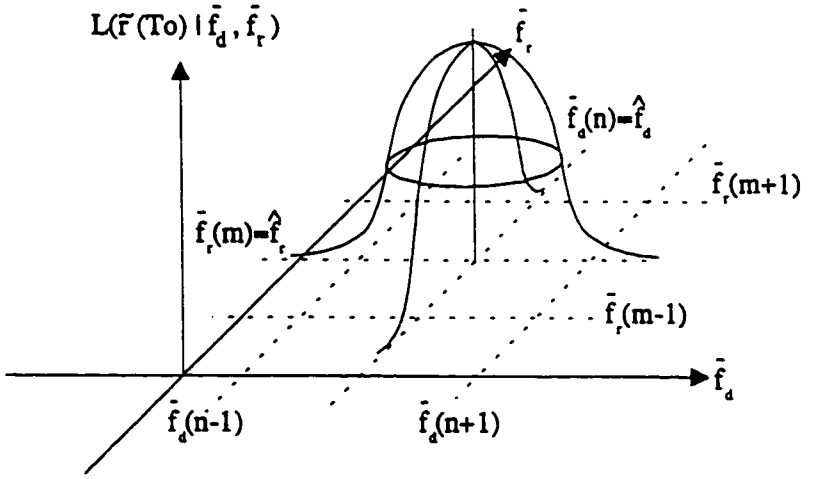


Figure 5.5: Typical joint 2-dimensional log-likelihood function $L(\bar{r}(T_o)|\ \bar{f}_d, \bar{f}_r)$

5.3.2 Improved Grid Search

This technique is based on decreasing the step size of the parameters we are searching for. Performing the grid search outlined in section 5.3 with a finer step size for the range of interest for the parameters, would involve excessive unnecessary computations. We could instead perform a finer grid search only in the region where the local maxima was estimated to be.

We now define new ranges of interest for the parameters $\hat{f}_d$ and $\hat{f}_r$ given by

$$\Delta\bar{f}_d = \left\lfloor \frac{\bar{f}_d(n+1) - \bar{f}_d(n-1)}{K} \right\rfloor \tag{5.54}$$

and

$$\Delta\bar{f}_r = \left\lfloor \frac{\bar{f}_r(m+1) - \bar{f}_r(m-1)}{K} \right\rfloor \tag{5.55}$$

where K denotes how fine the step size of the parameter in the grid search algorithm. Furthermore, $K \geq 2$, and when the equality condition holds, that is, $K = 2$, no improvement in the search is achieved.

We now perform a grid search similar to the one outlined in section 4.10 but starting with $\bar{f}_d(1) = \bar{f}_d(n-1)$ and $\bar{f}_r(1) = \bar{f}_r(m-1)$. The step sizes for the search are given by (5.54)

and (5.55).

5.3.3 Interpretation of the Grid Search Technique

Expressions (5.42) and (5.43) are really components of the gradient vector of the Likelihood function (5.35) with respect to the two variables we want to estimate $\hat{f}_d$ and $\hat{f}_r$, respectively. This suggests the implementation of the steepest descent method as a two-dimensional minimization of $-L(\bar{r}(T_o)|\ \tilde{f}_d, \tilde{f}_r\ )$, which is equivalent to the maximization of $L(\bar{r}(T_o)|\ f_d, f_r\ )$ [68, 111-112].

5.4 Proposed ML QPSK Doppler Compensator Detector (Fine Estimator)

We now investigate the possibility of obtaining a ML QPSK Doppler compensator from the results from section 5.2.

5.4.1 Interpretation of $U(kT_s)$

Careful examination of the expression given by (5.30), rewritten here without the

variable θ_o, since we have shown in (5.41) that it can be ignored.

$$U(kT_s) = \int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d\, t + \pi f_r\, t^2]\right) dt \qquad \textbf{(5.56)}$$

The above expression is the received signal $\tilde{r}(t)$ which has been phase rotated by $\exp(-j[2\pi f_d\, t + \pi f_r\, t^2])$ convolved with the impulse response of the signal matched filter (SMF) namely, $h(-t)$. Hence, $U(kT_s)$ can be obtained by taking samples from the SMF output while using the phase rotated received signal as the input. The process is shown in Figure 5.6.

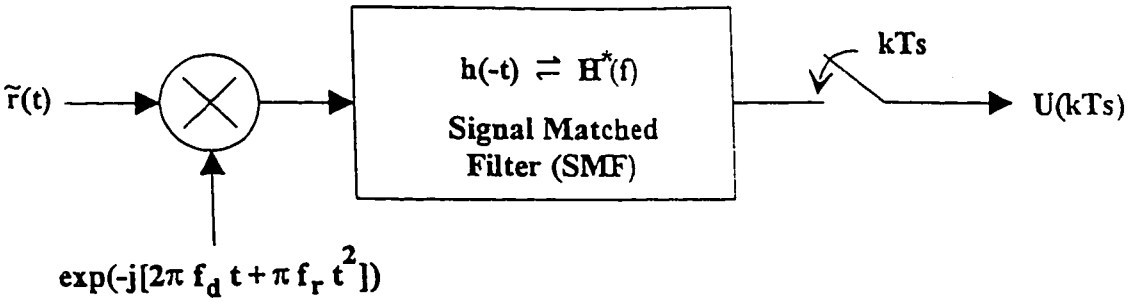


Figure 5.6: Signal matched filter operation on phase rotated $\tilde{r}(t)$

Looking back at (5.39) we can say that it is equal to the non-coherent sum of samples from the output of a SMF envelope demodulator [66].

5.4.2 Interpretation of $\partial U(kT_s)/\partial f_d$

Using (5.44) without the variable θ_o, we obtain

$$\frac{\partial U(kT_s)}{\partial f_d} = -j2\pi \int_{kT_s}^{(k+1)T_s} t\, \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d\, t + \pi f_r\, t^2]\right)dt \qquad \textbf{(5.57)}$$

Since the matched filter has an impulse response $h(-t)$, then its corresponding frequency response if given by

$$H^*(f) = \int_{-\infty}^{\infty} h(-t)\, \exp(-j2\pi f\, t)dt \qquad \textbf{(5.58)}$$

and if we differentiate $H^*(f)$ with respect to the frequency f we get

$$\frac{\partial H^*(f)}{\partial f} = -j2\pi \int_{-\infty}^{\infty} t\, h(-t)\, \exp(-j2\pi f\, t)dt = -j2\pi t\, h(-t) \qquad \textbf{(5.59)}$$

which is referred to as the frequency matched filter (FMF). Therefore, in the frequency domain, the FMF has a frequency response which is the derivative of the frequency response of SMF. Conversely, in time domain it can be stated that the impulse response of the FMF is the same as that of the SMF multiplied by $-j2\pi t$. Hence, (5.57) can be viewed as the convolution of the received signal which has been phase rotated and the impulse response

of the FMF. We conclude that $\partial U(kT_s)/\partial f_d$ can be obtained by taking samples from the FMF output while using the phase rotated received signal as the input. The process is illustrated in Figure 5.7.

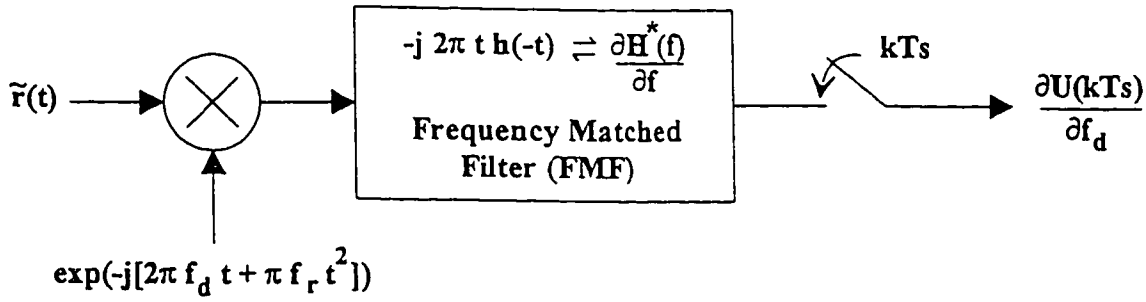


Figure 5.7: Frequency matched filter operation on phase rotated $\tilde{r}(t)$

5.4.3 Interpretation of $\partial U(kT_s)/\partial f_r$

Using the same idea presented in the previous section applied to (5.45) we have

$$\frac{\partial U(kT_s)}{\partial f_r} = -j\pi \int_{kT_s}^{(k+1)T_s} t^2\, \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d\, t + \pi f_r\, t^2]\right) dt \qquad \textbf{(5.60)}$$

The second derivative of $H^*(f)$ with respect to the frequency f is given by

$$\frac{\partial^2 H^*(f)}{\partial f^2} = -(2\pi)^2 \int_{-\infty}^{\infty} t^2\, h(-t)\, \exp(-j2\pi f\, t)dt = (-j2\pi t)^2\, h(-t) \qquad \textbf{(5.61)}$$

where the left hand side of (5.61) can be considered as the frequency rate matched filter (FRMF). Rearranging (5.61) to suit our purposes, we have

$$\left(\frac{1}{-j4\pi}\right)\frac{\partial^2 H^*(f)}{\partial f^2} = -j\pi t^2\, h(-t) \qquad \textbf{(5.62)}$$

Comparing (5.60) and (5.62), we can clearly see that apart from a reduction factor of $-j4\pi$, $\partial U(kT_s)/\partial f_r$ can be obtained by taking samples from the output of a FRMF while using the phase rotated received signal as the input. The process is shown in Figure 5.8 below.

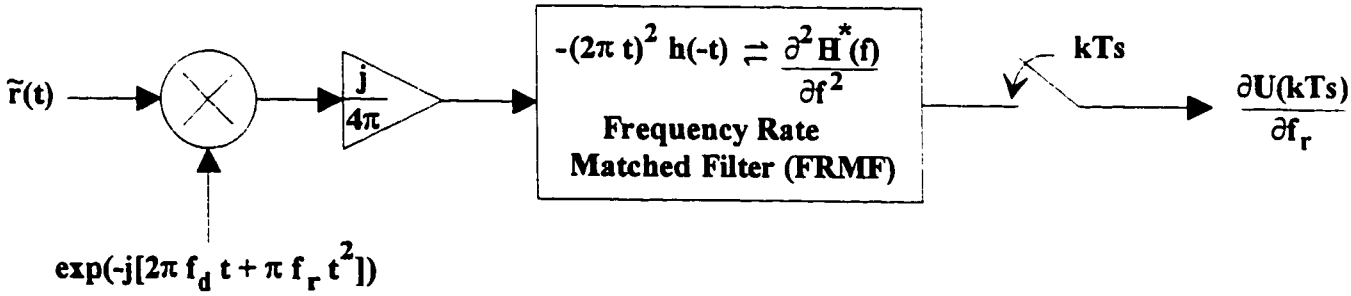


Figure 5.8: Frequency rate matched filter operation on phase rotated $\tilde{r}(t)$

5.4.4 ML QPSK Doppler Compensator

Let the components of the error $\underline{\epsilon}(k)$ given by (5.49) be $\underline{\epsilon}(k) = [\epsilon_d(k)\ \epsilon_r(k)]'$, where $\epsilon_d(k)$ is the frequency error feeding the frequency error control loop filter and $\epsilon_r(k)$ the frequency rate error feeding the frequency rate error control loop filter.

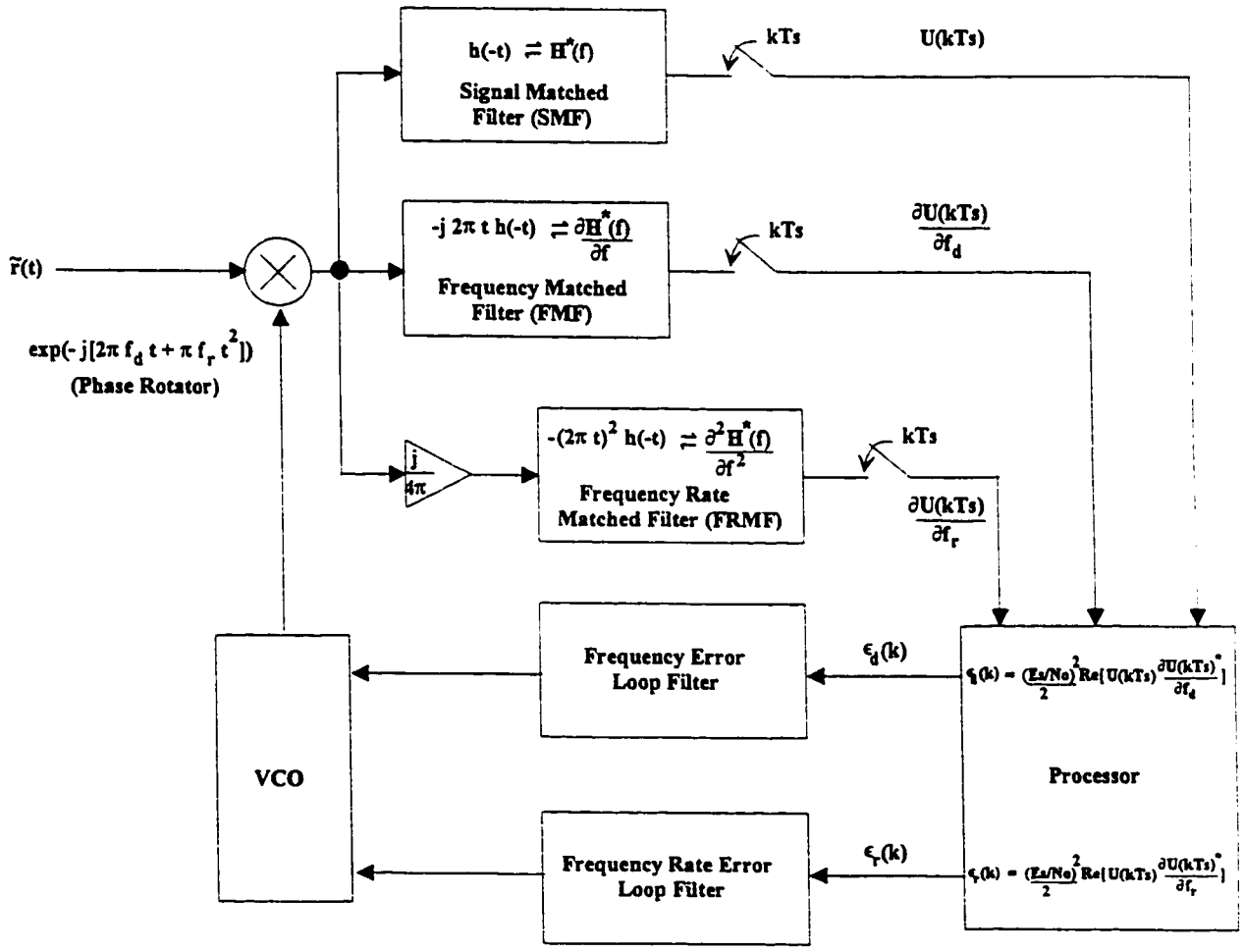


Figure 5.9: Block diagram of ML QPSK Doppler compensator

Applying the three filters and the feedback loop error components we have just

discussed we obtain the block diagram representation of the ML estimation process of the proposed method of compensating for dynamic Doppler shift and it is depicted in Figure 5.9.

5.4.5 Comparison with Previous Work on Doppler Compensation

Others have assumed the Doppler shift as a constant, hence not varying with time, which is contrary to what is presented here. To show how the proposed method of Doppler compensation incorporates theirs, that is solving the case of constant Doppler shift, we set f_r to zero and the only error component of interest from $\underline{\epsilon}(k)$ would be $\epsilon_d(k)$. The block diagram of the ML Doppler compensator presented in Figure 5.9 is simplified to the one presented earlier in Figure 5.4. It is intuitively satisfying to say that it is from the approximation (5.37) that has produced an implementable ML solution, which is realized by frequency detector schemes presented by block diagrams in Figures 5.4 and 5.9. Furthermore, all frequency detection schemes must be phase independent, and this is made possible due to the nature of (5.49). However, using (5.35) or (5.42) and (5.43), the frequency detection scheme is phase dependent.

5.5 Digital Implementation

To check the theoretical expressions developed in the previous sections, computer

programs are being developed that operates according to the scheme outlined in Figure 5.10.

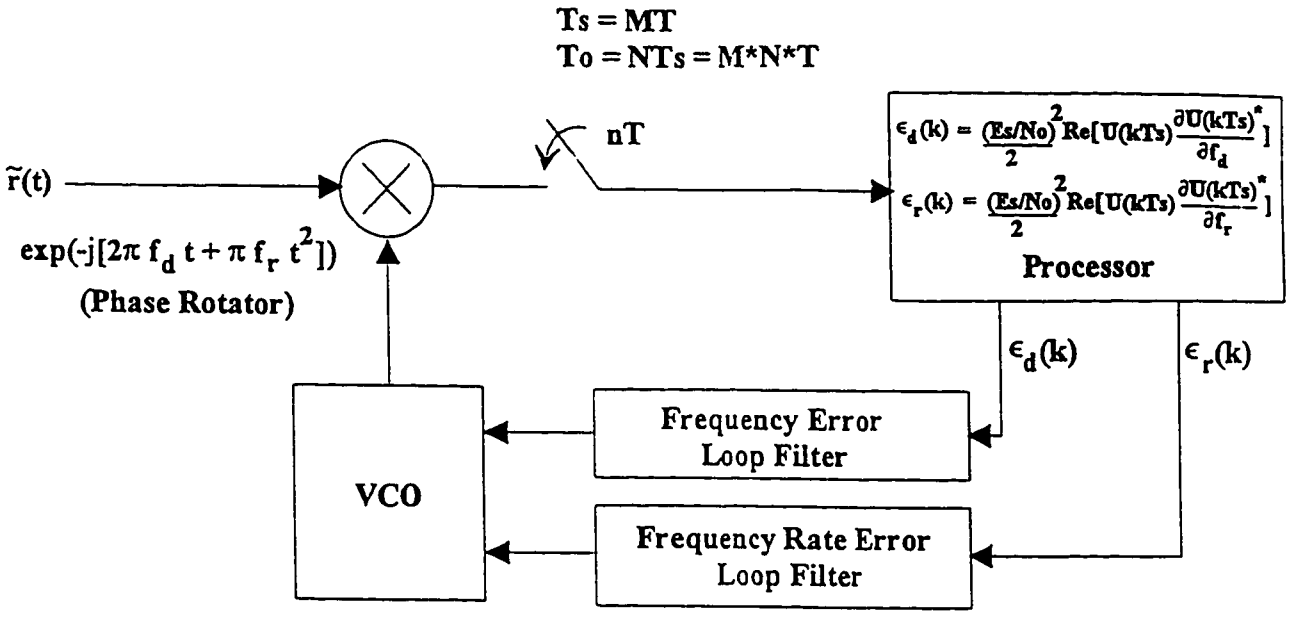


Figure 5.10: Digital implementation of ML QPSK Doppler compensator

The received signal $\tilde{r}(t)$ is sampled at a rate $1/T$, where we assume that $T_s = MT$ for some integer $M \geq 1$. Here M is a design parameter also known as the over-sample factor, that is, the sample rate of M samples per symbol, for computer representation. The block diagram labelled as processor, produces output $\epsilon_d(k)$ and $\epsilon_r(k)$ every T_s seconds. The total number of samples taken during the observation interval T_o is $N \times M$, since N symbols are observed, each represented by M samples. The compensation technique implemented here is also referred to as phase rotator technique.

5.5.1 Proposed Compensation Process

First we perform a grid search to obtain a coarse estimate of $\hat{f}_d$ and $\hat{f}_r$ as outline in detail in section 5.3. The grid search algorithm can be triggered by the detection of the start of burst (SB) of the data burst [75]. Then we use these coarse estimates as starting values in the ML QPSK Doppler compensator, which utilizes a feedback loop.

The proposed ML estimation technique being implemented by the block diagram in Figure 5.9 and represented digitally in Figure 5.10, can be used to synchronize the local oscillator with the carrier frequency of the incoming signal. This method cannot be used for phase synchronization, hence in order to achieve phase synchronization operation, a phase detector and a loop filter are required in parallel as shown earlier in Figure 4.10 [88].

5.6 Conclusions

From this chapter, two implementable solutions have been identified based on the Maximum Likelihood estimation of multiple parameters: grid search and the ML QPSK Doppler compensator. Both solutions have come about by using approximations of low SNR conditions. In addition, since they are both ML estimators, they are therefore efficient estimators. Hence they are unbiased estimates and satisfy the Cramér-Rao bound with equality for all SNRs [98]. It should be noted that the grid search and ML QPSK Doppler

compensator are mathematically equivalent to each other since they are based on maximizing the same objective function, which is the log-likelihood given by (5.39). Accuracy of the grid search is based on how fine the grid is made. Hence, there is a trade-off between estimation accuracy and the number of calculations. Furthermore, the grid search is not capable of tracking f_d and f_r. Conversely, the ML QPSK Doppler compensator is capable of tracking, utilizing two feedback loops and employs a signal matched filter, frequency matched filter and frequency rate matched filter. Between the two, the grid search is not suitable for real-time applications, due to the amount of computation involved.

The matched filtering and digital implementation of the ML QPSK Doppler compensator were introduced and also showing the feedback tracking loops. Interpretations of all three matched filters were given and a proposed strategy to compensate for Doppler has been suggested.

CHAPTER 6

Applications, Results and Analysis

In the previous chapter, the theory for the ML solutions for compensating rapidly changing Doppler shift have been presented. In Chapter 6, computer simulations will be carried out to study the performance of both the grid search and the ML QPSK Doppler compensator. Since the latter employs a tracking technique, three major modes of operation will be compared. Simulations are performed to evaluate the performance with respect to lock-in time for the adaptive optimized steepest descent method. This is then compared with the currently proposed carrier synchronizer.

6.1 S-Curve (Discriminator Characteristic)

The s-curve for a frequency detector is most oftenly used to show the important characteristics, especially its acquisition properties. As a matter of fact, the s-curve can be defined as the expectation or average of the frequency and frequency rate error applied to the frequency and frequency rate error control loop filters of the ML Doppler compensator versus the frequency offset [113-114].

6.1.1 S-Curve of the Proposed ML Doppler Compensator (Theory)

We now proceed to derive the expression of the s-curve. The output (tracking error) at the n^{th} symbol interval of the frequency error detector is $\epsilon_d(n)$ whereas the output of the frequency rate detector is $\epsilon_r(n)$. By definition, the s-curve is the mean value of $\underline{\epsilon}(n)$. Hence, in our case we would have two s-curves.

Rewriting the complex envelope of (5.16) in terms of the actual Doppler shift f_D and Doppler rate f_R, we have

$$\tilde{s}(t) = \sum_{k=-\infty}^{\infty} \left\{ c_k \, h(t-kT_s) \, \exp\left(j[2\pi f_D \, t + \pi f_R \, t^2 + \theta_o]\right)\right\} \qquad \textbf{(6.1)}$$

where $c_k = a_k + j b_k$, is the complex data symbol of the QPSK signal. For the time being,

let us consider the noiseless case and the s-curve after the n^{th} symbol using $E[\epsilon_d(n)]$.

Substituting (5.56-5.57) into (5.49) yields

$$\epsilon_d(n) = \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi f_d(k)\, t + \pi f_r(k)\, t^2]\right)dt \right.$$
$$\left. \times (-2\pi) \int_{kT_s}^{(k+1)T_s} t\, \tilde{r}^*(t)\, h(t-kT_s)\, \exp\left(j[2\pi f_d(k)\, t + \pi f_r(k)\, t^2]\right)dt \right] \qquad (6.2)$$

where $f_d(k)$ and $f_r(k)$ are the trial Doppler shift and Doppler rate estimates at the k^{th} symbol interval. Substituting (6.1) as $\tilde{r}(t)$ in (6.2), we obtain

$$\epsilon_d(n) = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} \sum_{m=-\infty}^{\infty} c_m\, h(t-mT_s)\, \exp\left(j[2\pi f_D\, t + \pi f_R\, t^2 + \theta_o]\right)\right.$$
$$\times h(t-kT_s)\, \exp\left(-j[2\pi f_d(k)\, t + \pi f_r(k)\, t^2]\right)dt \times \int_{kT_s}^{(k+1)T_s} \sum_{i=-\infty}^{\infty} t\, c_i^*\, h(t-iT_s) \qquad (6.3)$$
$$\left. \times \exp\left(-j[2\pi f_D\, t + \pi f_R\, t^2 + \theta_o]\right) h(t-kT_s)\, \exp\left(j[2\pi f_d(k)\, t + \pi f_r(k)\, t^2]\right)dt \right]$$

It is obvious that the initial phase θ_o is removed from (6.3) due to the multiplication of the exponents, and taking the integral operation inside the summation (i.e. reversing the order) we obtain

$$\epsilon_d(n) = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\sum_{m=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} c_m c_i^* \int_{kT_s}^{(k+1)T_s} h(t-mT_s)\, h(t-kT_s)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} t\, h(t-iT_s)\, h(t-kT_s)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right] \qquad \textbf{(6.4)}$$

where the change in Doppler shift at the k^{th} symbol interval is given by

$$\Delta f_d = f_D - f_d(k) \qquad \textbf{(6.5)}$$

and the change in Doppler rate at the k^{th} symbol interval is given by

$$\Delta f_r = f_R - f_r(k) \qquad \textbf{(6.6)}$$

Careful inspection of (6.4) shows that it is dependent on the random transmitted complex data symbols c_m and c_i. However, the data symbols are independent, therefore

$$\begin{aligned} E(c_m\, c_i^*) &= 1 \qquad && m = i \\ &= 0 && m \neq i \end{aligned} \qquad \textbf{(6.7)}$$

since

$$\begin{aligned} &E[a_m\, a_i] = E[b_m\, b_i] = 0 && m \neq i \\ &E[a_m^2] = E[b_m^2] = \frac{1}{2} \\ &E[a_m\, b_i] = 0 \end{aligned} \qquad \textbf{(6.8)}$$

Taking the expectation $E[\epsilon_d(n)]$ we obtain the general expression for the s-curve of the proposed ML QPSK compensator which is given by

$$E[\epsilon_d(n)] = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\sum_{i=-\infty}^{\infty} \int_{kT_s}^{(k+1)T_s} h(t-iT_s)\, h(t-kT_s)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} t\, h(t-iT_s)\, h(t-kT_s)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right] \tag{6.9}$$

For a time-limited pulse, such as the rectangular pulse $h(t)$, the product in the integral has the following property

$$h(t-iT_s)\, h(t-kT_s) = 0 \qquad i \neq k \tag{6.10}$$

Hence (6.9) simplifies further to

$$E[\epsilon_d(n)] = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} h^2(t-kT_s)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} t\, h^2(t-kT_s)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right] \tag{6.11}$$

For the case when $n = 1$, that is for the first symbol ($k = 0$), we have

$$E[\epsilon_d] = -\pi \left(\frac{E_s}{N_o}\right)^2 Im\left[\int_0^{T_s} h^2(t)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right.$$
$$\left. \times \int_0^{T_s} t\, h^2(t)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \right] \tag{6.12}$$

It can also be shown that substituting (5.56, 5.60) and (6.1) as $\tilde{r}(t)$ into (5.49) we obtain

$$E[\epsilon_r(n)] = -\frac{\pi}{2}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} h^2(t-kT_s)\exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \times \int_{kT_s}^{(k+1)T_s} t^2\, h^2(t-kT_s)\exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt\right] \quad (6.13)$$

Furthermore, for the case when $n = 1$, we have

$$E[\epsilon_r] = -\frac{\pi}{2}\left(\frac{E_s}{N_o}\right)^2 Im\left[\int_0^{T_s} h^2(t)\exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \times \int_0^{T_s} t^2\, h^2(t)\exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt\right] \quad (6.14)$$

Examining both (6.12) and (6.14), we can say that the s-curve is dependent on the symbol SNR and the signal pulse shape $h(t)$ [115]. However, the PSK modulation has the effect of only scaling the amplitude due to $|c_i\, c_i^*|$, and has no effect on the shape of the s-curve.

6.1.2 S-Curve in the Presence of AWGN (Simulation)

Let the variance of the low pass equivalent noise $\tilde{n}(t)$ which is sampled every T_s be given by $\sigma_{\tilde{n}}^2 = N_o/E_s = 1/SNR$. However, if we over-sample the noise to produce M samples per symbol (a sample every T seconds), as discussed in section 5.5, the variance of

each sample of this new noise process is given by [116]

$$\hat{\sigma}_{\tilde{n}}^2 = \sigma_{\tilde{n}}^2 \left(\frac{T_s}{T} \right) = \sigma_{\tilde{n}}^2 M \tag{6.15}$$

The above identity can be verified by the formula used to numerically calculate the variance of a random process and is given by

$$\sigma_{\tilde{n}}^2 = \frac{1}{M} \sum_{j=1}^{M} \hat{\sigma}_{\tilde{n}_j}^2 \tag{6.16}$$

In the presence of AWGN, $\tilde{r}(t)$ is given by

$$\tilde{r}(t) = \sum_{k=0}^{N-1} \left\{ c_k \, h(t-kT_s) \, \exp\left(j[2\pi f_D \, t + \pi f_R \, t^2 + \theta_o] \right) \right\} + \tilde{n}(t) \tag{6.17}$$

Substituting (6.17) into (6.2) we obtain the expression for the frequency error detector and is given by

$$\begin{aligned} \epsilon_d(n) = -\pi \left(\frac{E_s}{N_o} \right)^2 \sum_{k=0}^{n-1} Im \Bigg[\int_{kT_s}^{(k+1)T_s} \left(\sum_{m=0}^{\infty} c_m \, h(t-mT_s) \, \exp\left(j[2\pi \, f_D \, t + \pi \, f_R \, t^2 + \theta_o] \right) + \tilde{n}(t) \right) \\ \times h(t-kT_s) \, \exp\left(-j[2\pi \, f_d(k) \, t + \pi \, f_r(k) \, t^2] \right) dt \times \int_{kT_s}^{(k+1)T_s} t \left(\sum_{i=0}^{\infty} c_i^* \, h(t-iT_s) \right. \\ \left. \times \exp\left(-j[2\pi \, f_D \, t + \pi \, f_R \, t^2 + \theta_o] \right) + \tilde{n}^*(t) \right) h(t-kT_s) \, \exp\left(j[2\pi \, f_d(k) \, t + \pi \, f_r(k) \, t^2] \right) dt \Bigg] \end{aligned} \tag{6.18}$$

The resultant is composed of three parts, $S \times S_1$ (signal × signal), $N \times N_1$ (noise × noise) and $S \times N_1$ (signal × noise) and is given by

$$\epsilon_d(n) = S \times S_1 + N \times N_1 + S \times N_1 \tag{6.19}$$

Each term will be addressed in the next three subsections.

Signal × Signal Contribution

Taking the signal × signal component from (6.18) and taking the integral operation inside the summation, yields

$$S{\times}S_1 = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\sum_{m=0}^{\infty}\sum_{i=0}^{\infty} c_m c_i^* \int_{kT_s}^{(k+1)T_s} h(t-mT_s)\, h(t-kT_s)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \times \int_{kT_s}^{(k+1)T_s} t\, h(t-iT_s)\, h(t-kT_s)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt\right] \tag{6.20}$$

Now, taking the expectation of $S{\times}S_1$ results in the following

$$E[S{\times}S_1] = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} h^2(t-kT_s)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \times \int_{kT_s}^{(k+1)T_s} t\, h^2(t-kT_s)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt\right] \tag{6.21}$$

Since the s-curve involves only the first symbol, we substitute $n = 1$ into (6.21) and the result is identical to the same expression obtained in section 6.1.1, specifically (6.12).

$$E[S{\times}S_1] = -\pi \left(\frac{E_s}{N_o}\right)^2 Im\left[\int_0^{T_s} h^2(t)\, \exp\left(j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt \times \int_0^{T_s} t\, h^2(t)\, \exp\left(-j[2\pi\Delta f_d\, t + \pi\Delta f_r\, t^2]\right)dt\right] \tag{6.22}$$

Noise × Noise Contribution

Similarly, it can be shown that the noise × noise contribution is given by

$$N\times N_1 = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im\left[\int_{kT_s}^{(k+1)T_s} \tilde{n}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi f_d(k)\ t + \pi f_r(k)\ t^2]\right)dt \times \int_{kT_s}^{(k+1)T_s} t\ \tilde{n}^*(t)\ h(t-kT_s)\ \exp\left(j[2\pi f_d(k)\ t + \pi f_r(k)\ t^2]\right)dt\right] \quad \textbf{(6.23)}$$

We have modelled the $\tilde{n}(t)$ as a zero mean AWGN. Since $\tilde{r}(t)$ is a real valued function we can rewrite (6.23) as

$$N\times N_1 = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} \tilde{n}_1(k)\ \tilde{n}_2(k) \quad \textbf{(6.24)}$$

where $\tilde{n}_1(k)$ and $\tilde{n}_2(k)$ can be considered as filtered versions of $\tilde{n}(t)$ and $\tilde{n}^*(t)$ respectively. The latter is the phase rotated noise passing through a frequency matched filter. Hence both $\tilde{n}_1(k)$ and $\tilde{n}_2(k)$ are independent of each other and still have a zero mean value. This implies that the expectation of $N\times N_1$ is in fact equal to zero, hence

$$E[N\times N_1] \equiv 0 \quad \textbf{(6.25)}$$

Signal x Noise Contribution

Similarly $S\times N_1$ can be expressed as

$$
\begin{aligned}
S\times N_1 = -\pi \left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{n-1} Im \Bigg[& \int_{kT_s}^{(k+1)T_s} \left(\sum_{m=0}^{\infty} c_m \, h(t-mT_s) \, h(t-kT_s) \exp\left(j[2\pi\Delta f_d \, t + \pi\Delta f_r \, t^2 + \theta_o]\right)\right) dt \\
& \times \int_{kT_s}^{(k+1)T_s} \left(t \, \tilde{n}(t)^* \, h(t-kT_s) \exp\left(j[2\pi \, f_d(k) \, t + \pi \, f_r(k) \, t^2]\right)\right) dt \\
& + \int_{kT_s}^{(k+1)T_s} \left(\tilde{n}(t) \, h(t-kT_s) \exp\left(-j[2\pi \, f_d(k) \, t + \pi \, f_r(k) \, t^2]\right)\right) dt \\
& \times \int_{kT_s}^{(k+1)T_s} \left(t \sum_{i=0}^{\infty} c_i^* \, h(t-iT_s) \, h(t-kT_s) \exp\left(-j[2\pi\Delta f_d \, t + \pi\Delta f_r \, t^2 + \theta_o]\right)\right) dt \Bigg]
\end{aligned}
\tag{6.26}
$$

Taking the expectation of (6.26) results to zero since the expectation of the filtered noise process equates to zero.

$$E[S\times N_1] \equiv 0 \tag{6.27}$$

Therefore, the only contribution is from the signal noise interaction as shown earlier in (6.12). Moreover, in the presence of AWGN, the expectation of the frequency rate error can be shown to reduce to (6.14), thus not being affected by the presence of Gaussian noise [90, 92].

6.1.3 S-Curve Theory and Simulation Analysis

S-Curves $E[\epsilon_d]$ have been plotted for various values of symbol SNR to compare the analysis of simulation and theory. A 400 kbps QPSK signal with Doppler frequency shift f_D = 40 kHz, and Doppler rate of f_R = -350 Hz/s, having a selected initial trial value of f_r set to zero, hence Δf_r = -350 Hz/s. Furthermore, the over-sampling factor M = 50, and the range of interest for f_d was, $-2R_s \leq f_d \leq 2R_s$ where the symbol rate R_s = 200 k symbol/s. Figure 6.1(a) depicts the theoretical s-curve while Figure 6.1(b) depicts the results from simulation.

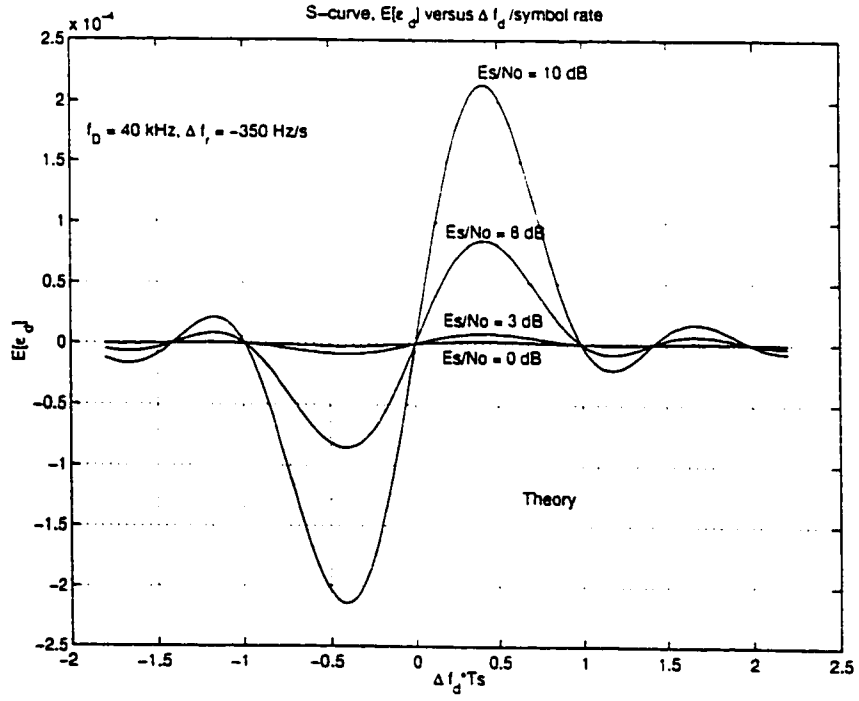


Figure 6.1(a): S-curve (theory) f_D = 40 kHz and Δf_r = -350 Hz/s

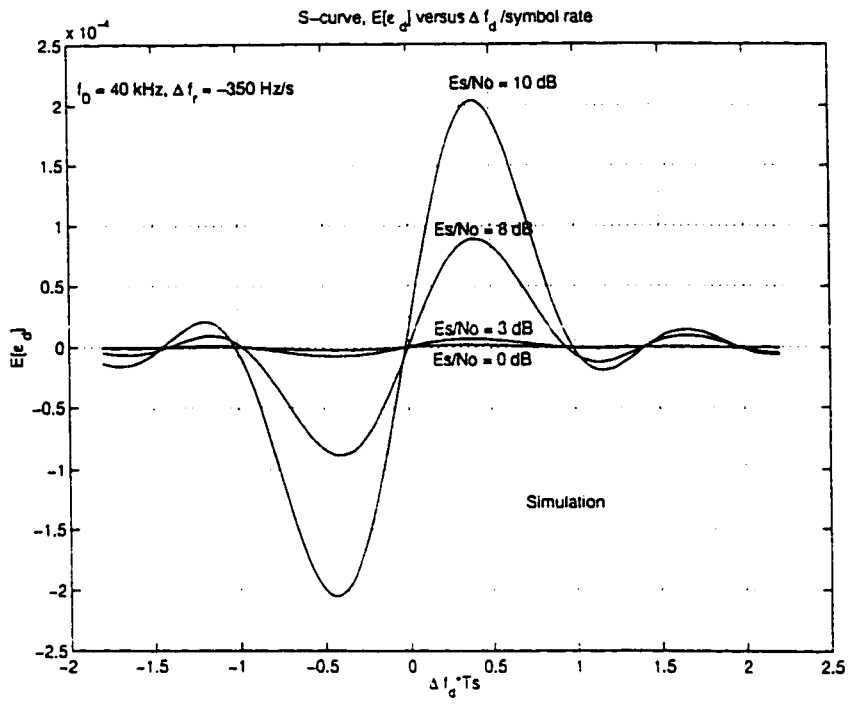


Figure 6.1(b): S-curve (simulation) f_D = 40 kHz and Δf_r = -350 Hz/s

The theoretical s-curve crosses, at the origin for all four levels of symbol SNR and shows that the ML QPSK Doppler compensator has a useful range of $-R_s \leq \Delta f_d \leq R_s$. However, for practical applications, only the linear region of the discriminator characteristic curve is utilized, hence the practical useful range of the ML QPSK Doppler compensator is $-0.4R_s \leq \Delta f_d \leq 0.4R_s$.

The s-curve simulated does not cross the origin perfectly and this is accounted for by the effects of noise, since it is the only factor not included in the s-curve generated from theory. Figure 6.2(a) and 6.2(b) are similar to Figures 6.1(a) and 6.1(b) except that in the former case, the Doppler rate f_R = 0 Hz/s.

The simulation results at low symbol SNR, demonstrate the effects of noise. Note that symbol SNR comes into effect in the theory since it is presented as a scaling factor in (6.12). Signals with higher SNR produced s-curves with larger amplitudes. This makes sense, since we would expect the detector to produce a larger initial frequency correction to reduce the Doppler frequency error. This can also be considered as the detector taking larger steps in the direction of the steepest descent to correct the frequency error more rapidly. However, at higher symbol SNR, simulation results are very close to the theoretical results. This can be explained by the fact that in this situation, the effects of noise is insignificant with respect to the signal power, hence noise can be ignored for higher signal power.

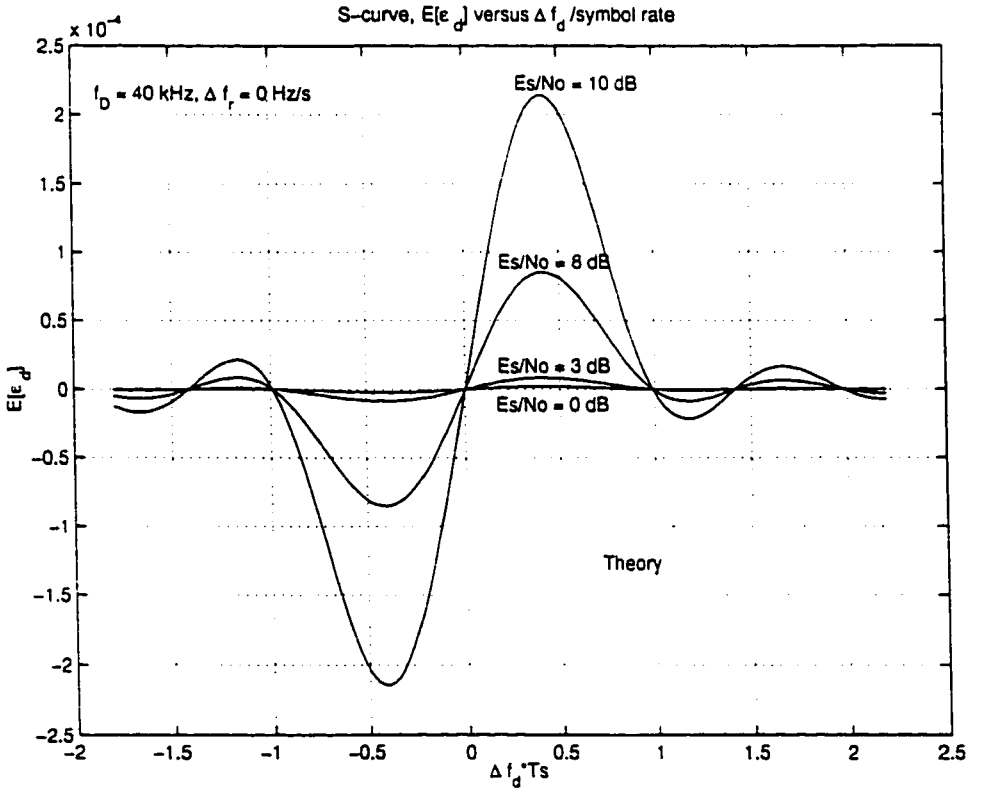


Figure 6.2(a): S-curve (theory) f_D = 40 kHz and Δf_r = 0 Hz/s

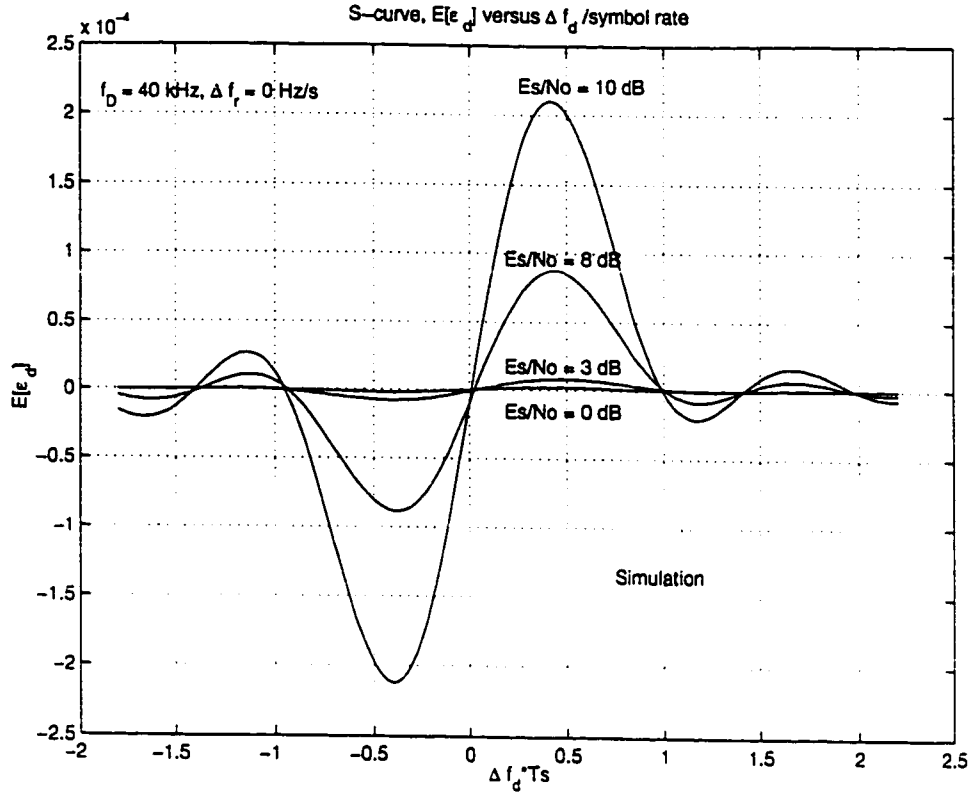


Figure 6.2(b): S-curve (simulation) f_D = 40 kHz and Δf_r = 0 Hz/s

Once again, the theoretical s-curve crosses at the origin for all values of E_s/N_o while the simulation s-curve does not. In Figure 6.2(b), it is clear that the frequency feedback errors become smaller with increased symbol SNR.

6.1.4 Discussion on the Results of S-Curve Simulations

S-curves $E[\epsilon_d]$ versus frequency error Δf_d and $E[\epsilon_r]$ versus frequency rate error Δf_r, which are the average error used in the feedback control loops, are shown in Figure 5.10.

The s-curve demonstrates how the negative feedback loop actually operates. The loop has the correct negative feedback sense, since $E[\epsilon_d]$ is positive for values $\Delta f_d > 0$, and therefore, it attempts to increase f_d immediately. Conversely, $E[\epsilon_r]$ is negative for values $\Delta f_d < 0$, and it attempts to decrease f_d likewise.

The detector loop maintains the correct operation mode for frequency errors Δf_d as large as $\pm\ 1/T_s$. This can also be considered as the useful acquisition range of the detector. At equilibrium, that is when $\Delta f_d = 0$ Hz and $\Delta f_r = 0$ Hz/s, $E[\epsilon_d] = 0$ as expected and this is due to the s-curve crossing the origin. Beyond $|\Delta f_d| \geq 1/T_s$, the detector reverses polarity and therefore does not maintain correct mode of operation. Hence, if any frequency error or frequency rate error exists beyond this condition, the loop will pull out of lock instead of pulling in.

From Figure 6.1 and Figure 6.2, more zero crossings with a positive slope are observed at larger frequency errors, $\Delta f_d = \pm 1.4 R_s$. These points are also stable equilibriums but at the wrong frequencies (false lock).

Increasing the symbol SNR has the effect of increasing the amplitude of $E[\epsilon_d]$ and increasing the slope at the equilibrium condition. A steeper slope basically implies that the detector will perform a faster acquisition. This can be visualized as larger steps taken towards a better estimate of either $\hat{f}_d$ or $\hat{f}_r$, in the direction of the steepest descent.

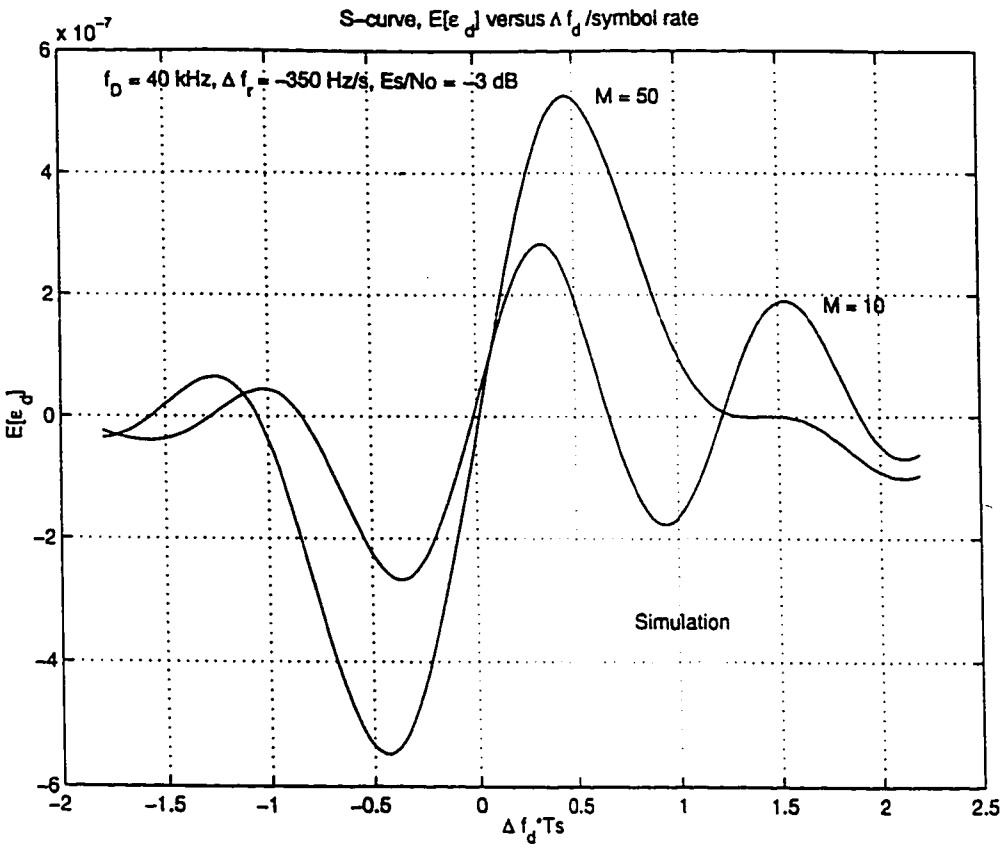


Figure 6.3: S-curve comparison for different over-sampling factors

The symbol SNR is given by E_s/N_o, which is also equal to the reciprocal of the variance $E_s/N_o = 1/\sigma_{\tilde{n}}^2$, of the low pass equivalent Gaussian noise $\tilde{n}(t)$. Using only a few samples per symbol to represent the noise process may not be adequate. Figure 6.3 shows the effects where the over-sample factor is changed. For larger M, typically $M = 50$ the noise is better represented, hence producing a better s-curve. However, when a smaller value is chosen, say $M = 10$, the noise process is not adequately represented, resulting in a distorted s-curve.

6.1.5 Closed Form S-Curve Representation for the Case $\Delta f_r = 0$ Hz/s

Setting $\Delta f_r = 0$ Hz/s in (6.12) and using a rectangular signal pulse shape we obtain

$$E[\epsilon_d] = -\pi \left(\frac{E_s}{N_o}\right)^2 Im\left[\int_0^{T_s} \exp(j2\pi\Delta f_d\, t)dt \times \int_0^{T_s} t\, \exp(-j2\pi\Delta f_d\, t)dt\right] \tag{6.28}$$

Evaluating the first integral we have

$$\int_0^{T_s} \exp(j2\pi\Delta f_d\, t)dt = \frac{\exp(j\pi\Delta f_d\, T_s)\sin(\pi\Delta f_d\, T_s)}{\pi\Delta f_d} \tag{6.29}$$

Evaluating the second integral with some manipulation results in

$$\int_0^{T_s} t\, \exp(-j2\pi\Delta f_d\, t)dt = j\left[\frac{\exp(-j\pi\Delta f_d\, T_s)\sin(\pi\Delta f_d\, T_s)}{2\pi^2\Delta f_d^2} - \frac{T_s \exp(-j2\pi\Delta f_d\, T_s)}{2\pi\Delta f_d}\right] \tag{6.30}$$

Substituting (6.29) and (6.30) into (6.28) we obtain the closed form expression given by

$$\begin{aligned} E[\epsilon_d] &= -\pi \left(\frac{E_s}{N_o}\right)^2 Im\left[\left(\frac{\exp(j\pi\Delta f_d\, T_s)\sin(\pi\Delta f_d\, T_s)}{\pi\Delta f_d}\right)\right. \\ &\quad \left.\times j\left(\frac{\exp(-j\pi\Delta f_d\, T_s)\sin(\pi\Delta f_d\, T_s)}{2\pi^2\Delta f_d^2} - \frac{T_s \exp(-j2\pi\Delta f_d\, T_s)}{2\pi\Delta f_d}\right)\right] \\ &= -\left(\frac{E_s}{N_o}\right)^2 Re\left[\left(\frac{\sin(\pi\Delta f_d\, T_s)}{\pi\Delta f_d}\right)\times\left(\frac{\sin(\pi\Delta f_d\, T_s)}{2\Delta f_d\,(\pi\Delta f_d)} - \frac{T_s \exp(-j\pi\Delta f_d\, T_s)}{2\Delta f_d}\right)\right] \end{aligned} \tag{6.31}$$

From (5.26), we have

$$
\begin{aligned}
E[\epsilon_d] &= \left(\frac{E_s}{N_o}\right)^2 Re\left[\left(\frac{T_s \sin(\pi\Delta f_d T_s)}{\pi\Delta f_d T_s}\right)\times\left(\frac{T_s\left[\cos(\pi\Delta f_d T_s) - j\sin(\pi\Delta f_d T_s)\right]}{2\Delta f_d} - \frac{T_s \sin(\pi\Delta f_d T_s)}{2\Delta f_d (\pi\Delta f_d T_s)}\right)\right] \\
&= \left(\frac{E_s}{N_o}\right)^2 \left[\left(\frac{T_s \sin(\pi\Delta f_d T_s)}{\pi\Delta f_d T_s}\right)\times\left(\frac{T_s \cos(\pi\Delta f_d T_s)}{2\Delta f_d} - \frac{T_s \sin(\pi\Delta f_d T_s)}{2\Delta f_d (\pi\Delta f_d T_s)}\right)\right] \\
&= \left(\frac{E_s}{N_o}\right)^2 \frac{T_s \sin(\pi\Delta f_d T_s)}{2\Delta f_d (\pi\Delta f_d T_s)}\left[T_s \cos(\pi\Delta f_d T_s) - \frac{T_s \sin(\pi\Delta f_d T_s)}{(\pi\Delta f_d T_s)}\right] \\
&= \left(\frac{E_s}{N_o}\right)^2 \frac{T_s^2\, sinc(\Delta f_d T_s)}{2\Delta f_d}\left[\cos(\pi\Delta f_d T_s) - sinc(\Delta f_d T_s)\right]
\end{aligned}
\tag{6.32}
$$

6.1.6 Error Analysis at Equilibrium

At equilibrium, that is, when the frequency error $\Delta f_d = 0$ Hz and frequency rate error $\Delta f_r = 0$ Hz/s, using (6.9) with these conditions for $n = 1$, the s-curve (discriminator characteristic) has the following format

$$
E[\epsilon_d] = -\pi\left(\frac{E_s}{N_o}\right)^2 Im\left[\sum_{i=-\infty}^{\infty}\int_0^{T_s} h(t-iT_s)\, h(t)\, dt \times \int_0^{T_s} t\, h(t-iT_s)\, h(t)\, dt\right] \tag{6.33}
$$

In order for the frequency feedback error to equal to zero, one of the integrals must be zero. We proceed to work with the first integral which is given by X_i

$$X_i = \int_0^{T_s} h(t-iT_s)\ h(t)\ dt \tag{6.34}$$

The Fourier transform of $h(t - iT_s)$ is given by

$$h(t-iT_s) = \int_{-\infty}^{\infty} H(f)\ \exp\left(j2\pi f(t-iT_s)\right)\ df \tag{6.35}$$

Now substitute (6.35) into (6.34) which is the first integral of (6.33) with the resultant being

$$\begin{aligned} X_i &= \int_0^{T_s}\int_{-\infty}^{\infty} h(t)\ H(f)\ \exp\left(-j2\pi fiT_s\right)\ \exp\left(j2\pi fi\right)\ df\ dt \\ &= \int_{-\infty}^{\infty} H(-f)\ H(f)\ \exp\left(-j2\pi fiT_s\right)\ df \\ &= \int_{-\infty}^{\infty} H^*(f)\ H(f)\ \exp\left(-j2\pi fiT_s\right)\ df \\ &= \int_{-\infty}^{\infty} P(f)\ \exp\left(-j2\pi fiT_s\right)\ df \\ &= p(-iT_s) \end{aligned} \tag{6.36}$$

Since we have chosen the impulse response $h(t)$ as a real function, $H(-f) = H^*(f)$. Furthermore, if $p(iT_s)$ satisfies the Nyquist criterion for zero inter symbol interference (ISI) then

$$p(-iT_s) = \begin{cases} 0 & i \neq 0 \\ 1 & i = 0 \end{cases} \tag{6.37}$$

As a matter of fact, if $p(iT_s)$ does not satisfy the Nyquist criterion, a self inflicting error is introduced in each feedback loop when $\Delta f_d = 0$ Hz and $\Delta f_r = 0$ Hz/s.

6.1.7 Grid Search Algorithm (Theory)

Using the same principle as the formulation of s-curve expressions and (5.51) we obtain the expected value for the likelihood function given by

$$E[L(\bar{f}_d, \bar{f}_r)] = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left[\int_{kT_s}^{(k+1)T_s} h^2(t-kT_s) \exp\left(j[2\pi(f_D - \bar{f}_d(k))\ t + \pi(f_R - \bar{f}_r(k))\ t^2]\right) dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} h^2(t-kT_s) \exp\left(-j[2\pi(f_D - \bar{f}_d(k))\ t + \pi(f_R - \bar{f}_r(k))\ t^2]\right) dt \right] \tag{6.38}$$

Substituting $t = (mT + kT_s)$ in (6.38) results in the expression for the likelihood function after observing N symbols given by

$$E[L(\bar{f}_d, \bar{f}_r)] = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left[I(k) \times I^*(k)\right] \tag{6.39}$$

where the integral $I(k)$ at the k^{th} symbol is calculated by using M samples within that

symbol period and calculated numerically as follows

$$I(k) = T \sum_{m=0}^{M-1} h^2(mT) \exp\left(j[2\pi(f_D - \bar{f}_d(k))\ (mT+kT_s) + \pi(f_R - \bar{f}_r(k))\ (mT+kT_s)^2]\right) \quad \textbf{(6.40)}$$

Usually, when evaluating integrals over short durations periods such as T_s seconds, one normalizes the integral by multiplying it with $1/T_s$, that is

$$A = \frac{1}{T_s} \int_{kT_s}^{(k+1)T_s} (\ \bullet\ ) dt \quad \textbf{(6.41)}$$

Performing this operation on (6.40) gives

$$\begin{aligned} I(k) &= \left(\frac{T}{T_s}\right) \sum_{m=0}^{M-1} h^2(mT) \exp\left(j[2\pi(f_D - \bar{f}_d(k))\ (mT+kT_s) + \pi(f_R - \bar{f}_r(k))\ (mT+kT_s)^2]\right) \\ &= \left(\frac{1}{M}\right) \sum_{m=0}^{M-1} h^2(mT) \exp\left(j[2\pi\Delta\bar{f}_d\ (mT+kT_s) + \pi\Delta\bar{f}_r\ (mT+kT_s)^2]\right) \end{aligned} \quad \textbf{(6.42)}$$

where

$$\Delta\bar{f}_d = f_D - \bar{f}_d(k) \quad \textbf{(6.43)}$$

and

$$\Delta\bar{f}_r = f_R - \bar{f}_r(k) \quad \textbf{(6.44)}$$

6.1.8 Grid Search Algorithm in the Presence of AWGN (Simulation)

By substituting $\tilde{r}(t)$, which is the actual received signal, composed of the Doppler shifted QPSK signal plus AWGN, given by (6.17), into (5.51), we obtain the lengthy expression representing the log-likelihood function given by

$$L(\tilde{f}_d, \tilde{f}_r) = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left[\int_{kT_s}^{(k+1)T_s} \left(\sum_{m=0}^{N-1} c_m \, h(t-mT_s) \exp\left(j[2\pi f_D \, t + \pi f_R \, t^2 + \theta_o]\right) + \tilde{n}(t) \right) \right.$$
$$\times h(t-kT_s) \exp\left(-j[2\pi \tilde{f}_d(k) \, t + \pi \tilde{f}_r(k) \, t^2]\right) dt \times \int_{kT_s}^{(k+1)T_s} \left(\sum_{i=0}^{N-1} c_i^* \, h(t-iT_s) \right. \qquad \textbf{(6.45)}$$
$$\left. \left. \times \exp\left(-j[2\pi f_D \, t + \pi f_R \, t^2 + \theta_o]\right) + \tilde{n}^*(t) \right) h(t-kT_s) \exp\left(j[2\pi \tilde{f}_d(k) \, t + \pi \tilde{f}_r(k) \, t^2]\right) dt \right]$$

where N is the number of symbols observed. The resultant algebraic expansion of the log-likelihood function is now composed of three parts, as explained earlier in section 6.1.2, namely; $S{\times}S_2$ (signal × signal), $N{\times}N_2$ (noise × noise) and $S{\times}N_2$ (signal × noise) and is given by

$$L(\tilde{f}_d, \tilde{f}_r) = S{\times}S_2 + N{\times}N_2 + S{\times}N_2 \qquad \textbf{(6.46)}$$

Signal × Signal Contribution

Reversing the order of the integral and summation operations for the signal × signal contribution, we obtain

$$S\times S_2 = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left| \sum_{m=0}^{N-1}\sum_{i=0}^{N-1} c_m c_i^* \int_{kT_s}^{(k+1)T_s} h(t-mT_s)\, h(t-kT_s)\, \exp\left(j[2\pi\Delta\bar{f}_d\, t + \pi\Delta\bar{f}_r\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} h(t-iT_s)\, h(t-kT_s)\, \exp\left(-j[2\pi\Delta\bar{f}_d\, t + \pi\Delta\bar{f}_r\, t^2]\right)dt \right| \qquad (6.47)$$

Taking the expectation of $S\times S_2$ results in the following

$$E[S\times S_2] = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left[\int_{kT_s}^{(k+1)T_s} h^2(t-kT_s)\, \exp\left(j[2\pi\Delta\bar{f}_d\, t + \pi\Delta\bar{f}_r\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} h^2(t-kT_s)\, \exp\left(-j[2\pi\Delta\bar{f}_d\, t + \pi\Delta\bar{f}_r\, t^2]\right)dt \right] \qquad (6.48)$$

which is the same expression obtained in section 6.1.7, specifically (6.38), written in a slightly different format.

Noise x Noise Contribution

Similarly, the noise x noise contribution is given by

$$N\times N_2 = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \left[\int_{kT_s}^{(k+1)T_s} \bar{n}(t)\, h(t-kT_s)\, \exp\left(-j[2\pi\bar{f}_d(k)\, t + \pi\bar{f}_r(k)\, t^2]\right)dt \right.$$
$$\left. \times \int_{kT_s}^{(k+1)T_s} \bar{n}^*(t)\, h(t-kT_s)\, \exp\left(j[2\pi\bar{f}_d(k)\, t + \pi\bar{f}_r(k)\, t^2]\right)dt \right] \qquad (6.49)$$

We have modelled the $\tilde{n}(t)$ as a zero mean AWGN. Since $\tilde{r}(t)$ is a real valued function we can rewrite (6.49) as

$$N\times N_2 = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \tilde{n}_a(k)\ \tilde{n}_b(k) \tag{6.50}$$

where $\tilde{n}_a(k)$ and $\tilde{n}_b(k)$ can be considered as filtered versions of $\tilde{n}(t)$ and $\tilde{n}^*(t)$ respectively, thus independent of each other and still have a zero mean value. This implies that the expectation of $N\times N_2$ is in fact equal to zero, hence

$$E[N\times N_2] \equiv 0 \tag{6.51}$$

Signal × Noise Contribution

Similarly $S\times N_2$ can be expressed as

$$\begin{aligned} S\times N_2 = \frac{1}{4}\left(\frac{E_s}{N_o}\right)^2 \sum_{k=0}^{N-1} \Bigg[& \int_{kT_s}^{(k+1)T_s} \left(\sum_{m=0}^{N-1} c_m\ h(t-mT_s)\ h(t-kT_s)\ \exp\left(j[2\pi\Delta\bar{f}_d\ t+\pi\Delta\bar{f}_r\ t^2+\theta_o]\right)\right) a \\ & \times \int_{kT_s}^{(k+1)T_s} \left(\tilde{n}(t)^*\ h(t-kT_s)\ \exp\left(j[2\pi\ \bar{f}_d(k)\ t+\pi\ \bar{f}_r(k)\ t^2]\right)\right) dt \\ & + \int_{kT_s}^{(k+1)T_s} \left(\tilde{n}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ \bar{f}_d(k)\ t+\pi\ \bar{f}_r(k)\ t^2]\right)\right) dt \\ & \times \int_{kT_s}^{(k+1)T_s} \left(\sum_{i=0}^{N-1} c_i^*\ h(t-iT_s)\ h(t-kT_s)\ \exp\left(-j[2\pi\Delta\bar{f}_d\ t+\pi\Delta\bar{f}_r\ t^2+\theta_o]\right)\right) dt \Bigg] \end{aligned} \tag{6.52}$$

Taking the expectation of (6.52) results to zero since the expectation of the filtered noise process equates to zero.

$$E[S \times N_2] \equiv 0 \tag{6.53}$$

Hence, the only contribution is from the interaction of $S \times S_2$ as shown earlier in (6.38). The likelihood functions for various Doppler frequency shift, Doppler rates, and for N = 1, 5, 10, 25, 50, 100 are shown in Figure 6.4(a-c).

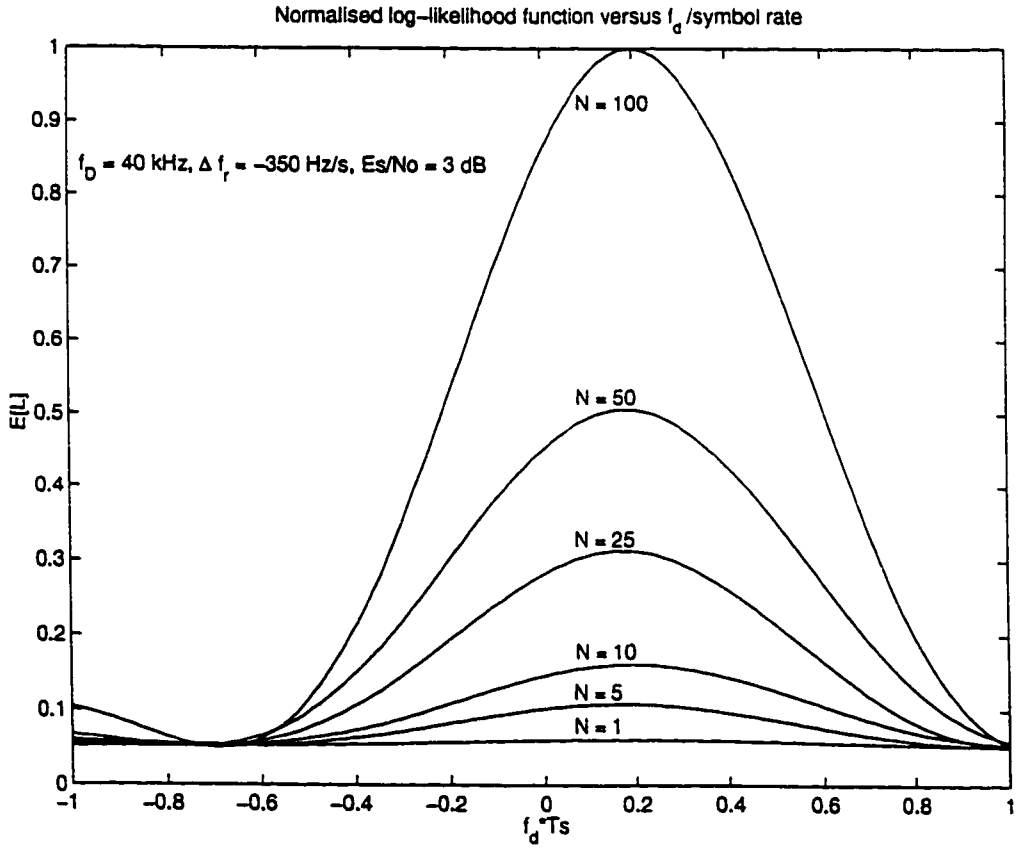


Figure 6.4(a): Likelihood function f_D = 40 kHz, Δf_r = -350 Hz/s and E_s/N_o = 3 dB

Figure 6.4(a) shows that the maximum of the likelihood function is located at $f_d T_s = 0.2$, corresponding to $f_d = 40$ kHz and is more defined for higher E_s/N_o values. A similar situation arises in Figure 6.4(b) where it is seen that the peak of the likelihood function is again located at $f_d T_s = 0.2$, that is at $f_d = 40$ kHz.

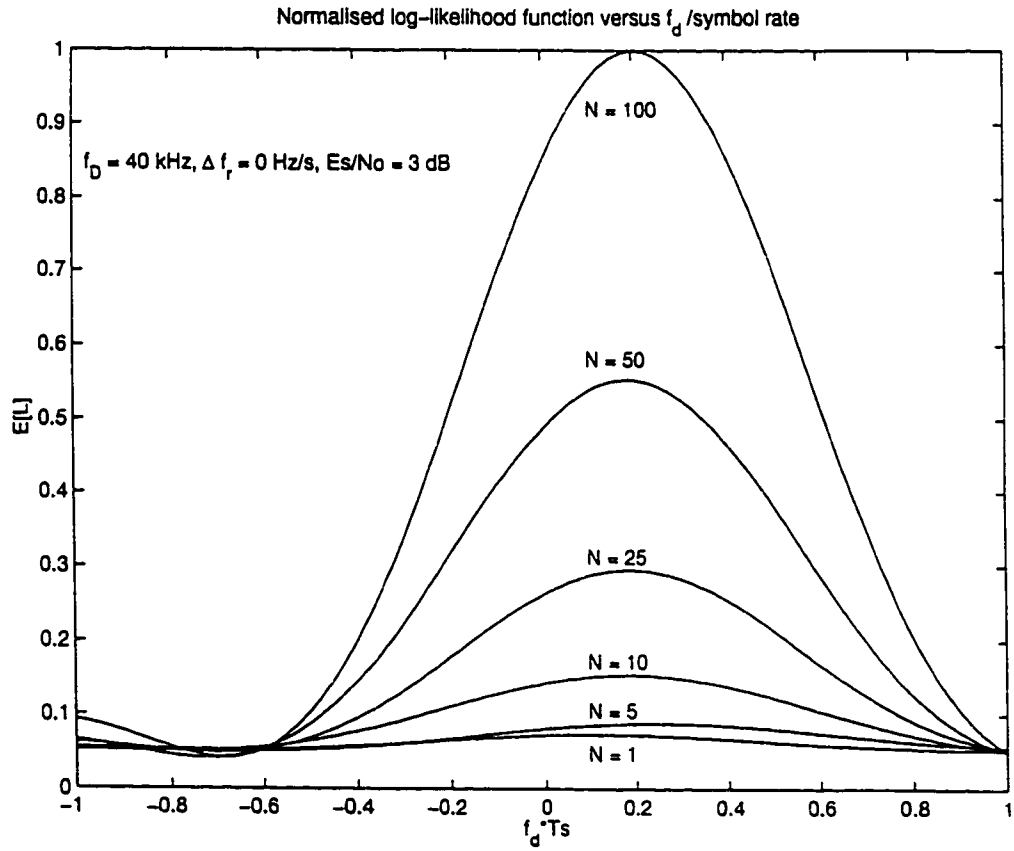


Figure 6.4(b): Likelihood function $f_D = 40$ kHz, $\Delta f_r = 0$ Hz/s and $E_s/N_o = 3$ dB

However, in Figure 6.4(c), the maximum of the likelihood function is located at $f_d T_s = 0$, corresponding to $f_d = 0$ Hz, since $f_D = 0$ Hz. We conclude that the effect of $f_R = -350$ Hz/s

is negligible; however if a very large number of symbols are observed, implying N being very large, the peak of the likelihood function would be located elsewhere.

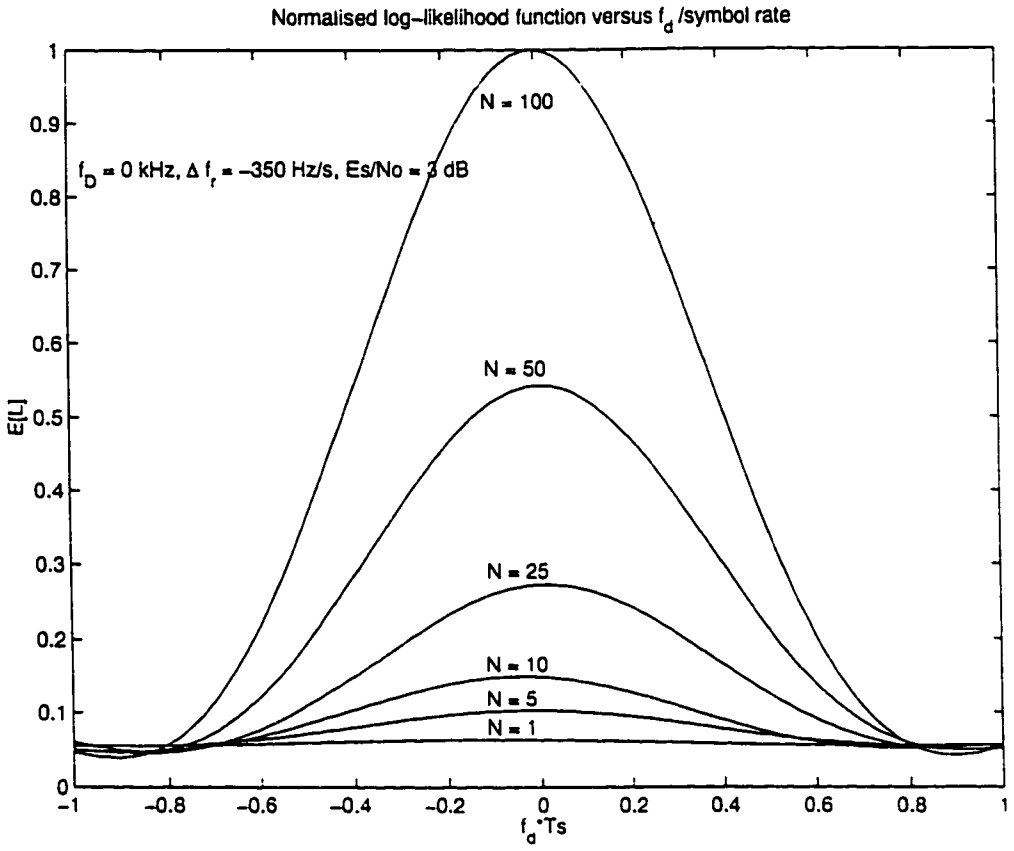


Figure 6.4(c): Likelihood function f_D = 0 Hz, Δf_r = -350 Hz/s and E_s/N_o = 3 dB

6.1.8 SNR Considerations

The likelihood function is proportional to the symbol SNR squared and has the effect of only scaling the amplitude of the likelihood function. Figure 6.5 shows how the likelihood

function differs when E_s/N_o varies, while f_D = 40 kHz, zero Doppler rate and N = 100. The peak of the likelihood function is more defined and is located more precisely at f_d = 40 kHz. Hence, at higher E_s/N_o, it is easier to compensate for Doppler shift using the grid search technique.

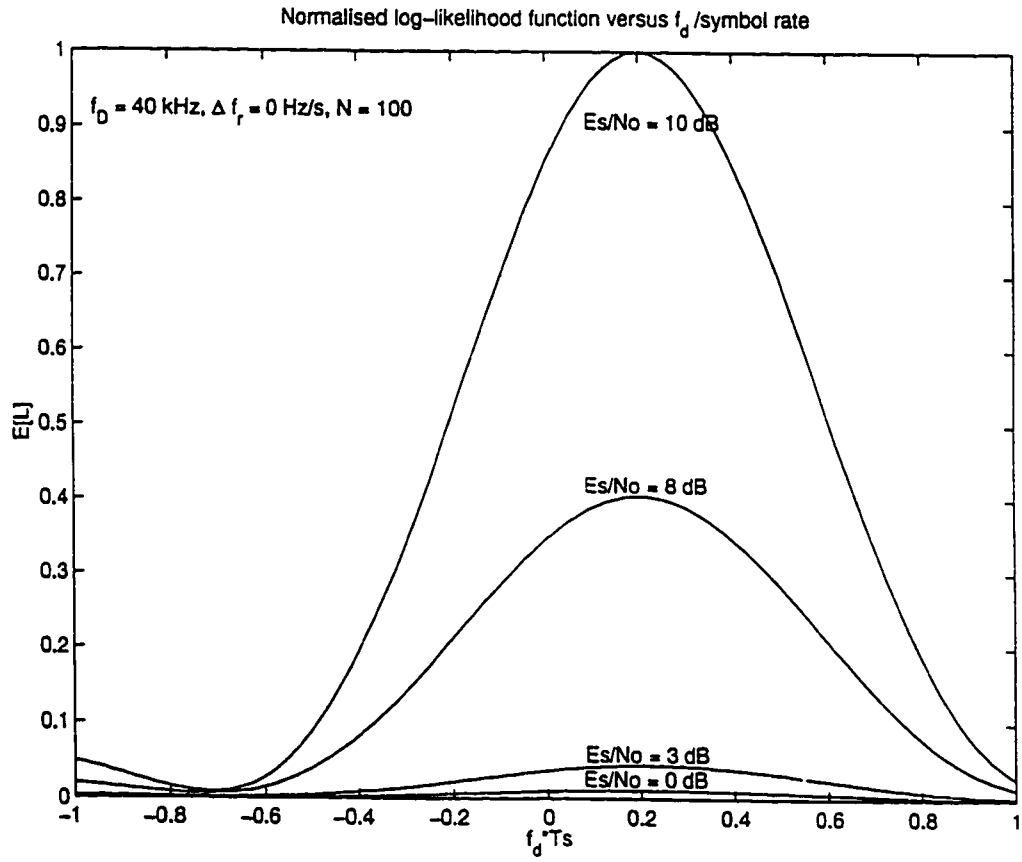


Figure 6.5: Likelihood function for various symbol SNR (E_s/N_o)

6.2 Modes of Operation

In this section three modes of operation are investigated for the ML QPSK Doppler compensator and these include: (1) conventional steepest descent, (2) optimized steepest descent, and (3) Newton Raphson method [111-112]. These various gradient methods of function minimization. The modes of operation refer to how tracking strategy of the detector and therefore, the process by which the feedback loop errors, ϵ_d and ϵ_r are used to obtain future estimates of f_d and f_r. In other words, these are three optimization techniques used to locate the maximum of the likelihood function given by (5.39). In this case we are trying to obtain the local maximum of a 3-D surface, hence we have two unknowns namely, f_d and f_r. All three methods require starting values for $f_d(0)$ and $f_r(0)$.

6.2.1 Conventional Steepest Descent Strategy

This method involves updating $f_d(n)$ and $f_r(n)$ as follows

$$\begin{bmatrix} f_d(n+1) \\ f_r(n+1) \end{bmatrix} = \begin{bmatrix} f_d(n) \\ f_r(n) \end{bmatrix} + \begin{bmatrix} \epsilon_d(n) \\ \epsilon_r(n) \end{bmatrix} \tag{6.54}$$

The gradient vector components, which are the first partial derivatives, $\epsilon_d(n)$ and $\epsilon_r(n)$ of

the objective function (5.39) are computed from

$$\underline{\epsilon}(n) = \begin{bmatrix} \epsilon_d(n) \\ \epsilon_r(n) \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s)\ \frac{\partial U(kT_s)^*}{\partial f_d}\right] \\ \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[U(kT_s)\ \frac{\partial U(kT_s)^*}{\partial f_r}\right] \end{bmatrix} \tag{6.55}$$

with

$$U(kT_s) = \int_{kT_s}^{(k+1)T_s} \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ f_d(k)\ t\ +\ \pi\ f_r(k)\ t^2]\right)dt \tag{6.56}$$

$$\frac{\partial U(kT_s)}{\partial f_d} = -j2\pi \int_{kT_s}^{(k+1)T_s} t\ \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ f_d(k)\ t\ +\ \pi\ f_r(k)\ t^2]\right)dt \tag{6.57}$$

$$\frac{\partial U(kT_s)}{\partial f_r} = -j\pi \int_{kT_s}^{(k+1)T_s} t^2\ \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ f_d(k)\ t\ +\ \pi\ f_r(k)\ t^2]\right)dt \tag{6.58}$$

where $\tilde{r}(t)$ is given by

$$\tilde{r}(t) = \sum_{k=0}^{n-1}\left\{c_k\ h(t-kT_s)\ \exp\left(j[2\pi f_D\ t\ +\ \pi f_R\ t^2\ +\ \theta_o]\right)\right\}\ +\ \tilde{n}(t) \tag{6.59}$$

This particular optimization technique uses a constant step gain of unity. The rate of obtaining an acceptable solution may depend on the step gain. Next we examine a method employing variable step gain.

6.2.2 Optimized Steepest Descent Strategy

The optimized steepest descent employs a variable step gain and is given as follows

$$\begin{bmatrix} f_d(n+1) \\ f_r(n+1) \end{bmatrix} = \begin{bmatrix} f_d(n) \\ f_r(n) \end{bmatrix} - \alpha_{opt}(n) \begin{bmatrix} \epsilon_d(n) \\ \epsilon_r(n) \end{bmatrix} \tag{6.60}$$

where the optimum step size $\alpha_{opt}(n)$ at the n^{th} symbol interval is given by

$$\alpha_{opt}(n) = \left(\frac{\underline{\epsilon}(n)^T \, \underline{\epsilon}(n)}{\underline{\epsilon}(n)^T \, \underline{H}(n) \, \underline{\epsilon}(n)} \right) \tag{6.61}$$

This involves computing the Hessian matrix $\underline{H}(n)$, which is symmetric and contains the second partial derivatives of the objective function mentioned earlier, represented here by L, corresponding to the likelihood function. The idea used here is to first calculate the optimum step gain for each symbol interval by using (6.61), then compute the estimated Doppler frequency shift f_d and Doppler rate f_r. In situations where computational power is a major factor, approximation of the Hessian matrix can be used [112]. In the simulations, the Hessian matrix is calculated after every symbol observation rather than approximated and is given by

$$H(n) = \begin{bmatrix} H_{1,1}(n) & H_{1,2}(n) \\ H_{2,1}(n) & H_{2,2}(n) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2 L(\bar{r}(T_o)|\ f_d, f_r)}{\partial f_d^2} & \dfrac{\partial^2 L(\bar{r}(T_o)|\ f_d, f_r)}{\partial f_d\ \partial f_r} \\ \dfrac{\partial^2 L(\bar{r}(T_o)|\ f_d, f_r)}{\partial f_r\ \partial f_d} & \dfrac{\partial^2 L(\bar{r}(T_o)|\ f_d, f_r)}{\partial f_r^2} \end{bmatrix} \tag{6.62}$$

The components of the Hessian matrix can be computed from the expressions given below

$$H_{1,1}(n) = \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_d}\ \frac{\partial U(kT_s)^*}{\partial f_d} + U(kT_s)\ \frac{\partial^2 U(kT_s)^*}{\partial f_d^2}\right] \tag{6.63}$$

$$H_{2,2}(n) = \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_r}\ \frac{\partial U(kT_s)^*}{\partial f_r} + U(kT_s)\ \frac{\partial^2 U(kT_s)^*}{\partial f_r^2}\right] \tag{6.64}$$

$$H_{1,2}(n) = \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_d}\ \frac{\partial U(kT_s)^*}{\partial f_r} + U(kT_s)\ \frac{\partial^2 U(kT_s)^*}{\partial f_d\ \partial f_r}\right] \tag{6.65}$$

$$H_{2,1}(n) = \sum_{k=0}^{n-1} \frac{1}{2}\left(\frac{E_s}{N_o}\right)^2 Re\left[\frac{\partial U(kT_s)}{\partial f_r}\ \frac{\partial U(kT_s)^*}{\partial f_d} + U(kT_s)\ \frac{\partial^2 U(kT_s)^*}{\partial f_r\ \partial f_d}\right] \tag{6.66}$$

However, to evaluate the above expressions we need

$$\frac{\partial^2 U(kT_s)}{\partial f_d^2} = (-j2\pi)^2 \int_{kT_s}^{(k+1)T_s} t^2 \, \tilde{r}(t) \, h(t-kT_s) \exp\left(-j[2\pi f_d(k) \, t + \pi f_r(k) \, t^2]\right) dt \qquad \textbf{(6.67)}$$

$$\frac{\partial^2 U(kT_s)}{\partial f_r^2} = (-j\pi)^2 \int_{kT_s}^{(k+1)T_s} t^4 \, \tilde{r}(t) \, h(t-kT_s) \exp\left(-j[2\pi f_d(k) \, t + \pi f_r(k) \, t^2]\right) dt \qquad \textbf{(6.68)}$$

$$\frac{\partial^2 U(kT_s)}{\partial f_d \, \partial f_r} = 2 \, (-j\pi)^2 \int_{kT_s}^{(k+1)T_s} t^3 \, \tilde{r}(t) \, h(t-kT_s) \exp\left(-j[2\pi f_d(k) \, t + \pi f_r(k) \, t^2]\right) dt \qquad \textbf{(6.69)}$$

$$\frac{\partial^2 U(kT_s)}{\partial f_r \, \partial f_d} = 2 \, (-j\pi)^2 \int_{kT_s}^{(k+1)T_s} t^3 \, \tilde{r}(t) \, h(t-kT_s) \exp\left(-j[2\pi f_d(k) \, t + \pi f_r(k) \, t^2]\right) dt \qquad \textbf{(6.70)}$$

Note that

$$\frac{\partial^2 U(kT_s)}{\partial f_d \, \partial f_r} = \frac{\partial^2 U(kT_s)}{\partial f_r \, \partial f_d} \qquad \textbf{(6.71)}$$

Computational effort can be reduced by performing the multiplication of time t, just before evaluating the integrals in (6.56-6.58) and (6.67-6.70). To illustrate how this process is

accomplished we compute A which given by

$$A = \tilde{r}(t)\ h(t-kT_s)\ \exp\left(-j[2\pi\ f_d(k)\ t + \pi\ f_r(k)\ t^2]\right) \tag{6.72}$$

Now, (6.56) can be obtained as follows

$$U(kT_s) = \int_{kT_s}^{(k+1)T_s} A\ dt \tag{6.73}$$

Similarly, from (6.57) and (6.58)

$$\begin{aligned} \frac{\partial U(kT_s)}{\partial f_d} &= -j2\pi \int_{kT_s}^{(k+1)T_s} t\ A\ dt \\ \frac{\partial U(kT_s)}{\partial f_r} &= -j\pi \int_{kT_s}^{(k+1)T_s} t^2\ A\ dt \end{aligned} \tag{6.74}$$

Using the same idea on (6.67) we obtain

$$\frac{\partial^2 U(kT_s)}{\partial f_d^2} = (-j2\pi)^2 \int_{kT_s}^{(k+1)T_s} t^2\ A\ dt \tag{6.75}$$

and from (6.68)

$$\frac{\partial^2 U(kT_s)}{\partial f_r^2} = (-j\pi)^2 \int_{kT_s}^{(k+1)T_s} t^4\ A\ dt \tag{6.76}$$

$$\frac{\partial^2 U(kT_s)}{\partial f_d \, \partial f_r} = \frac{\partial^2 U(kT_s)}{\partial f_r \, \partial f_d} = 2\,(-j\pi)^2 \int_{kT_s}^{(k+1)T_s} t^3 \, A \, dt \qquad \textbf{(6.77)}$$

and finally, from (6.69) and (6.70)

6.2.3 Newton Raphson Method

The Newton Raphson method of function minimization is given by

$$\begin{bmatrix} f_d(n+1) \\ f_r(n+1) \end{bmatrix} = \begin{bmatrix} f_d(n) \\ f_r(n) \end{bmatrix} - H^{-1}(n) \begin{bmatrix} \epsilon_d(n) \\ \epsilon_r(n) \end{bmatrix} \qquad \textbf{(6.78)}$$

Since the objective function given by (5.39) is a quadratic in nature, the Newton Raphson method takes us to a maximum after only one symbol observation. The solution will therefore be very coarse but may be useful for low cost equipment.

6.2.4 Adaptive Operation

The conventional procedure for computing $\epsilon_d(n)$ and $\epsilon_r(n)$ involves using estimates of $f_d(n)$ and $f_r(n)$ during the n^{th} symbol interval $(n-1)T_s$ to T_s, estimates of $f_d(n-1)$ and $f_r(n-1)$ during the $(n-1)^{th}$ symbol interval $(n-2)T_s$ to $(n-1)T_s$ and so on, up to $f_d(1)$ and $f_r(1)$

during the first symbol interval (0 to T_s), by using (6.55-6.59). Note that here $f_d(1)$ and $f_r(1)$ are initial estimates. The above process for f_d is illustrated with the help of Figure 6.6(a).

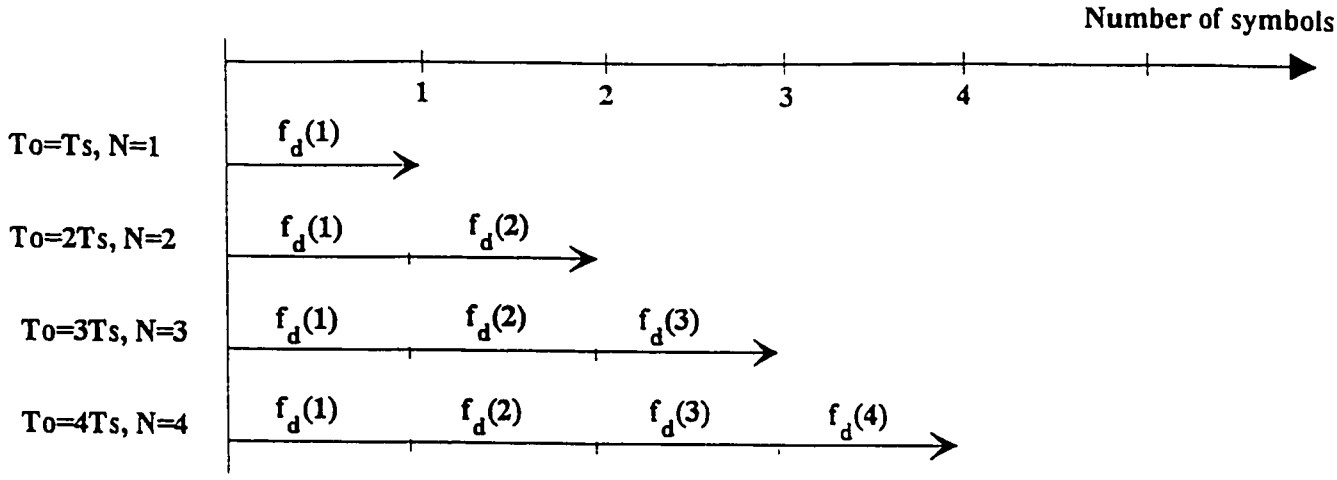


Figure 6.6(a): Conventional approach

However, an adaptive approach can also be employed whereby, at the n^{th} symbol interval, estimates $f_d(n)$ and $f_r(n)$ are used throughout the whole observation period (0 to nT_s) instead of only during the n^{th} symbol interval ($(n-1)T_s$ to nT_s). This is best shown in Figure 6.6(b). Both methods are explored in the next section.

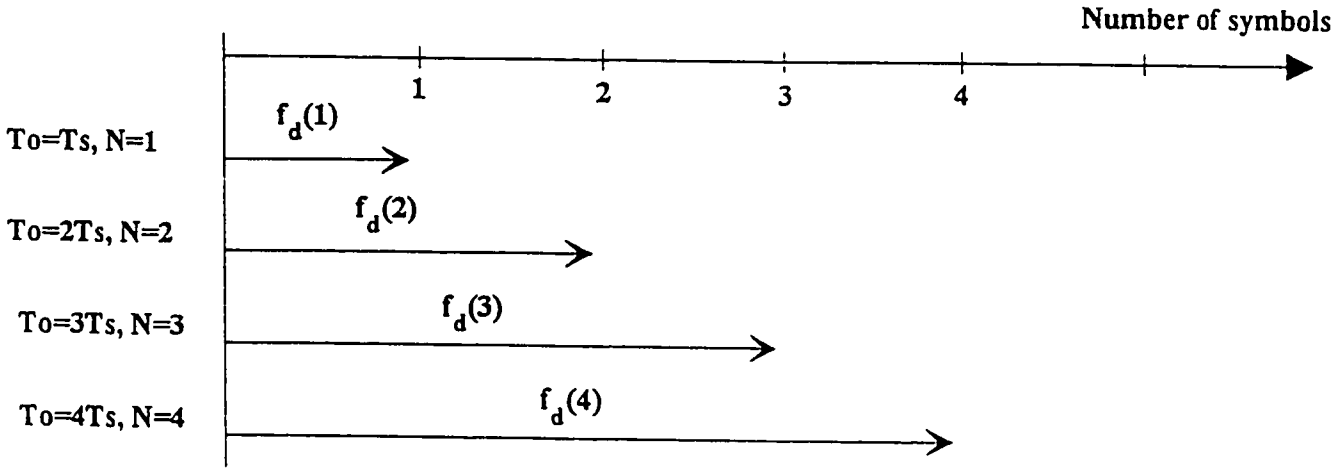
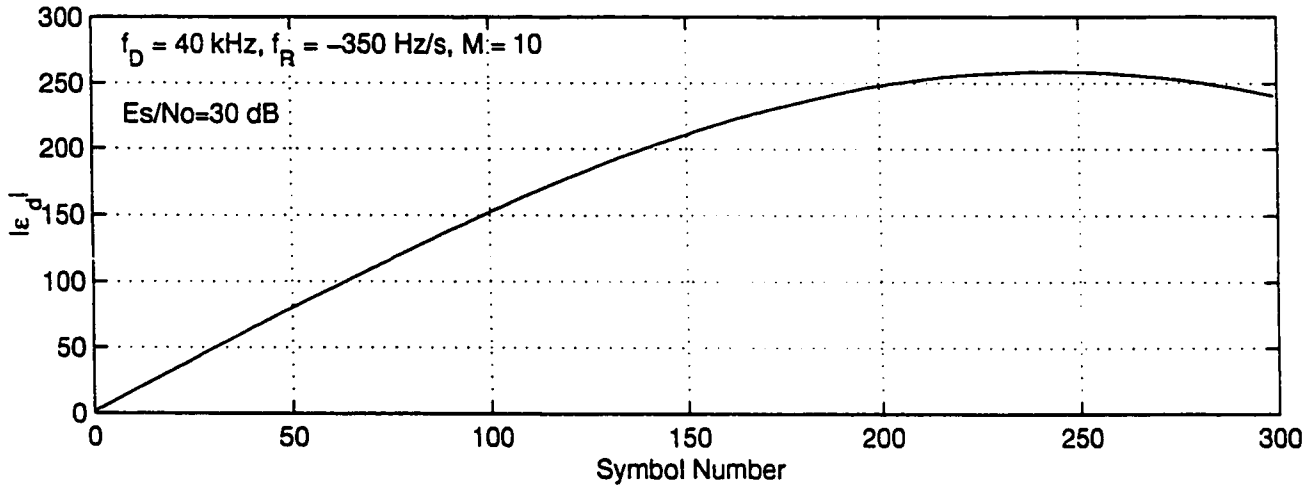


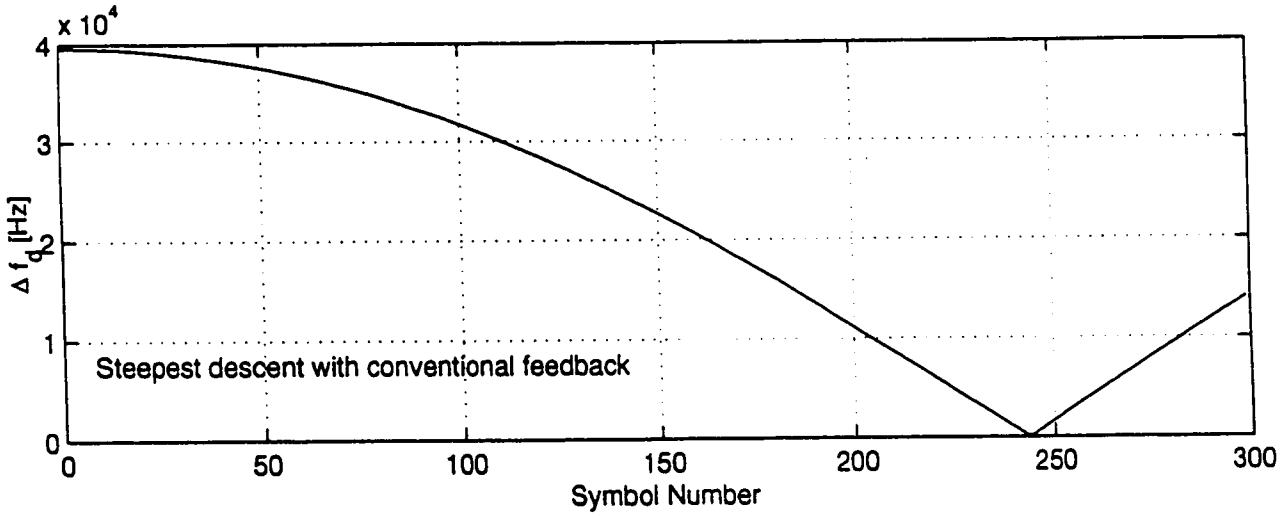
Figure 6.6(b): Adaptive approach

6.3 Tracking Analysis (Simulation)

Tracking analysis for the various modes of operation were investigated. The conventional steepest descent using both conventional and adaptive approach was tested, using f_D = 40 kHz and f_R = -350 Hz/s. As starting values, $f_d(1)$ = 400 Hz and $f_r(1)$ = 0 Hz/s were used.



(a) Frequency detector output $|\epsilon_d|$ versus symbol number

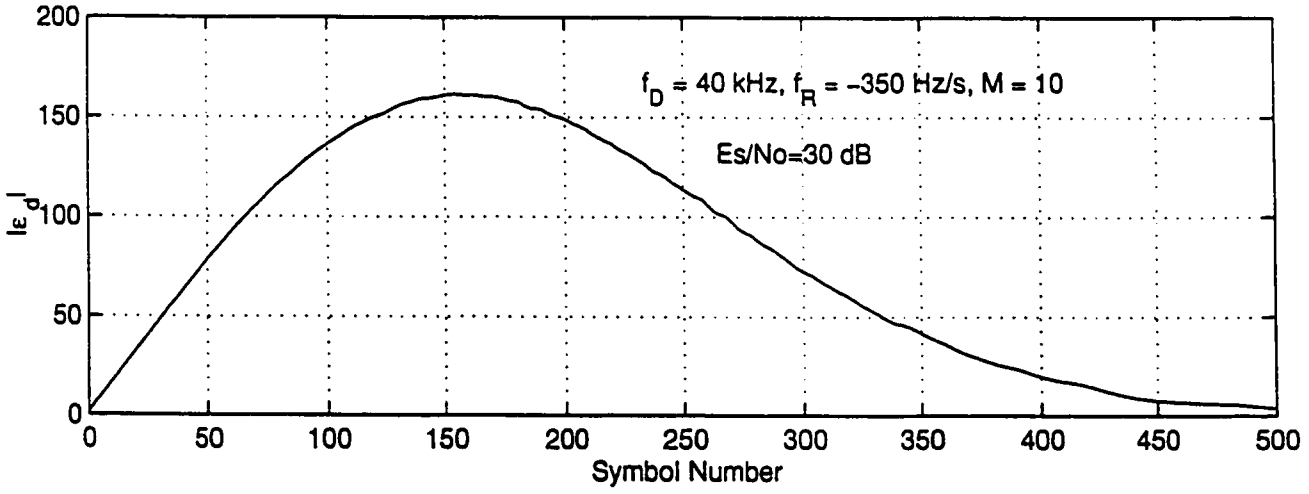


(b) Magnitude of frequency shift error $|\Delta f_d| = |f_D - f_d(n)|$ versus symbol number

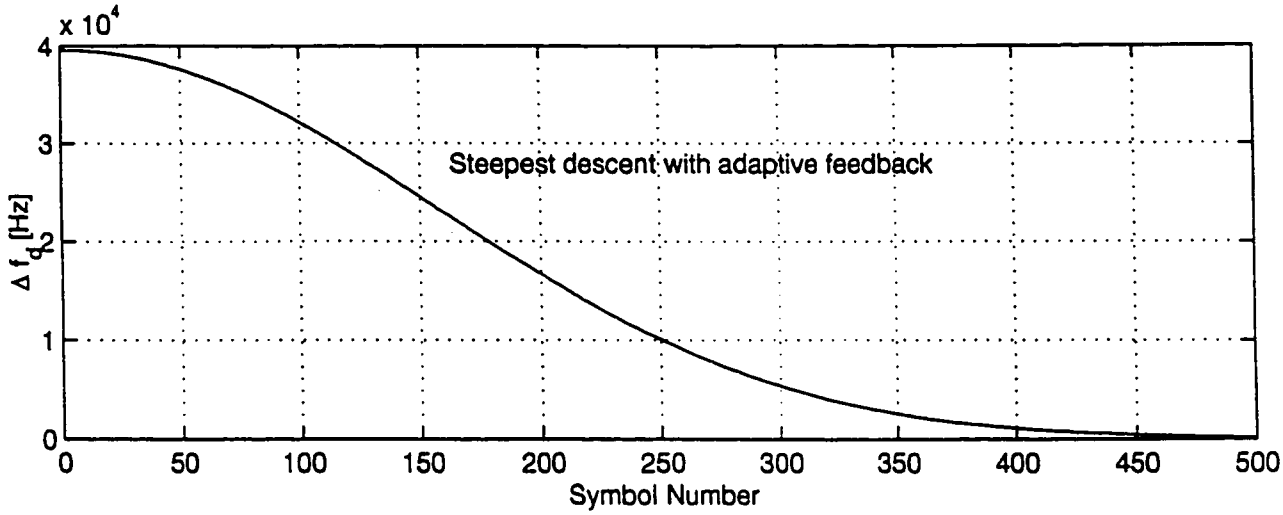
Figure 6.7: Steepest descent employing conventional feedback

Figure 6.7(a) shows the detector output $|\epsilon_d|$, while Figure 6.7(b) shows the magnitude of the frequency shift error $\Delta f_d = |f_D - f_d(n)|$ versus symbol number.

A fast estimate of the Doppler frequency shift f_d was obtained after 240 symbols using the steepest descent method, employing conventional feedback. The simulation then showed signs of divergence since Δf_d failed to remain close to zero, as a sign that carrier lock has been achieved. The solution obtained by applying the conventional steepest descent is therefore not reliable, and hence rejected. The symbol SNR was set to E_s/N_o = 30 dB to achieve this level of performance. In other words, for lower values of SNR, the simulation had to be stopped because it took too long to show any sign of a solution prospect. The adaptive steepest descent method produces carrier lock after 500 symbols and equilibrium is maintained beyond that point, as shown in Figure 6.8, more precisely 6.8 (b), where it depicts the frequency error of the detector.



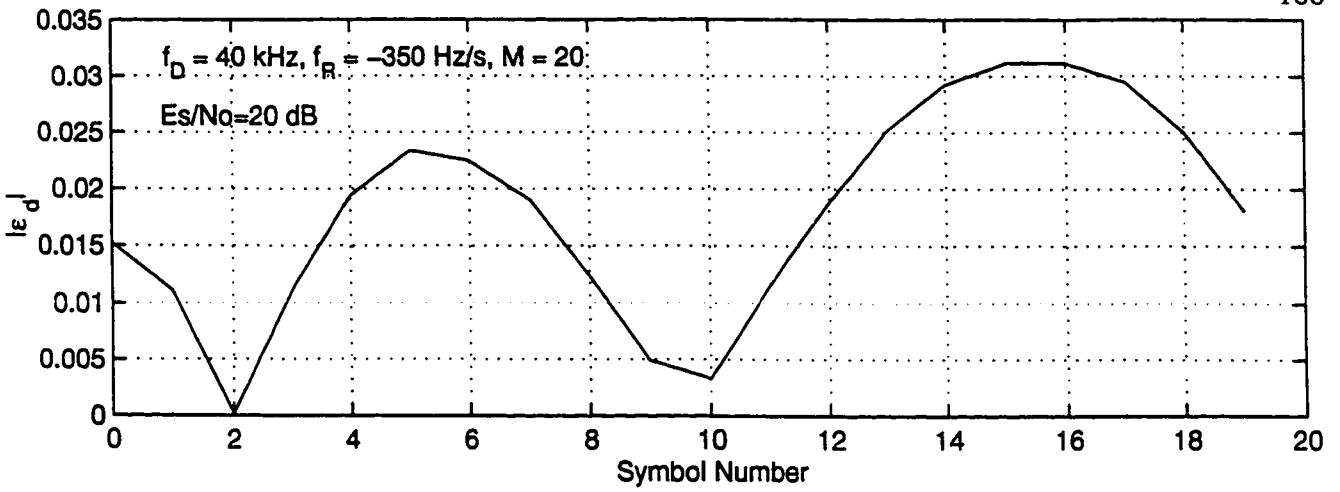
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



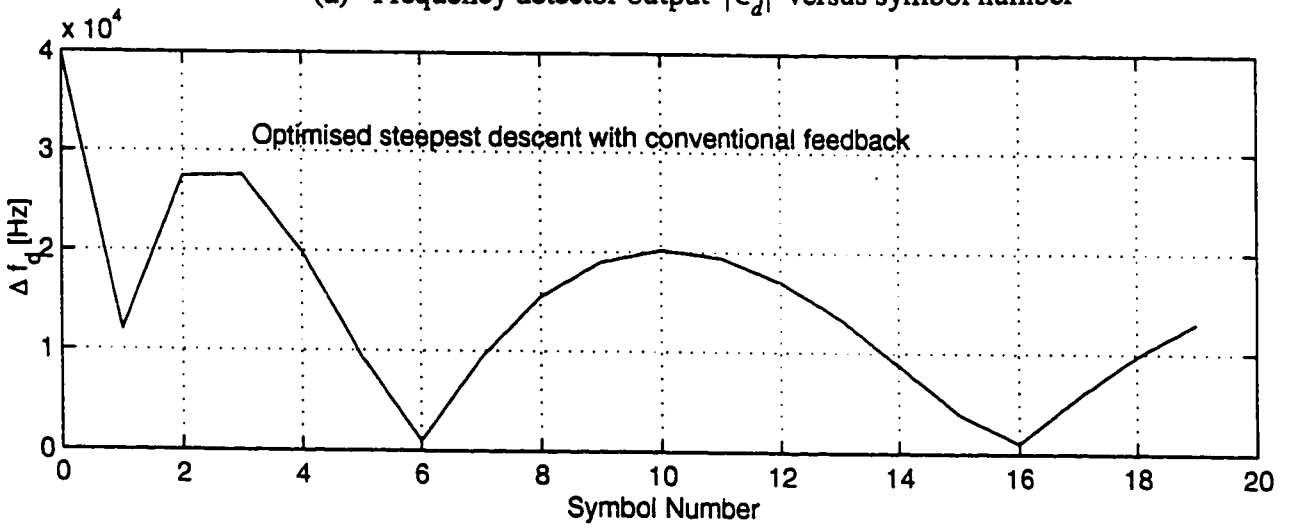
(b) Magnitude of frequency shift error $|\Delta f_d| = |f_D - f_d(n)|$ versus symbol number

Figure 6.8: Steepest descent employing adaptive feedback

The same conditions were applied to the optimized steepest descent strategy and the results are shown in Figures 6.9 and 6.10. Convergence towards a respectable solution for the optimized steepest descent is clearly shown by the progressive reduction in amplitude in the oscillations in Figure 6.9(b). On the other hand, the adaptive optimized steepest descent provides a stable carrier lock after only 6 symbols as shown in Figure 6.10(b). Therefore, we conclude that the adaptive technique has once again outperformed the conventional method. Furthermore, the adaptive optimized steepest descent is the best tracking method thus far. The symbol SNR was set to E_s/N_o = 20 dB and higher values have shown faster convergence.



(a) Frequency detector output $|\epsilon_d|$ versus symbol number

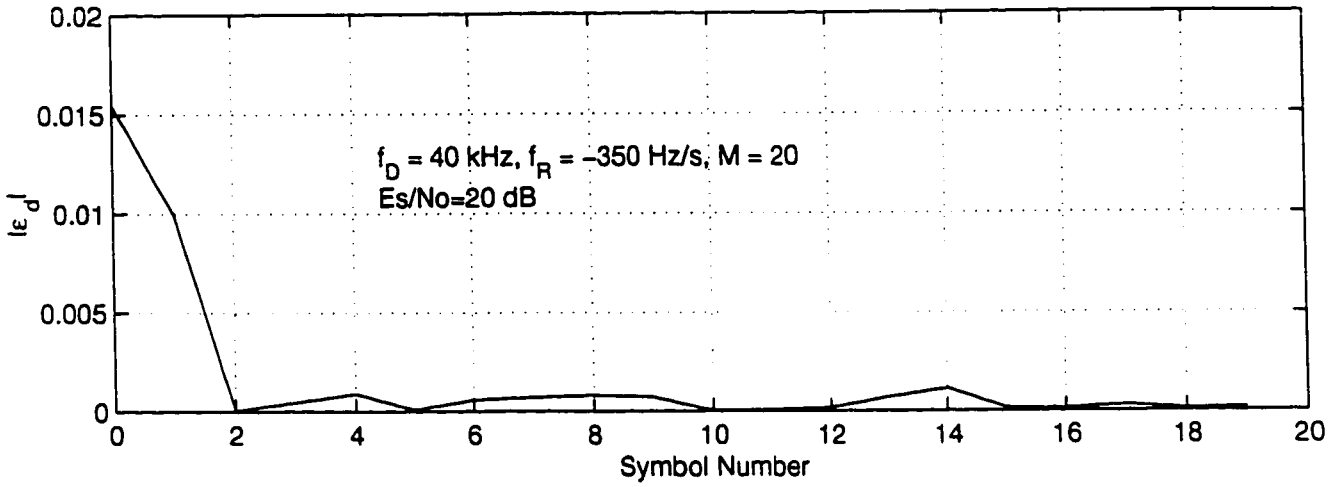


(b) Magnitude of frequency shift error $|\Delta f_d| = |f_D - f_d(n)|$ versus symbol number

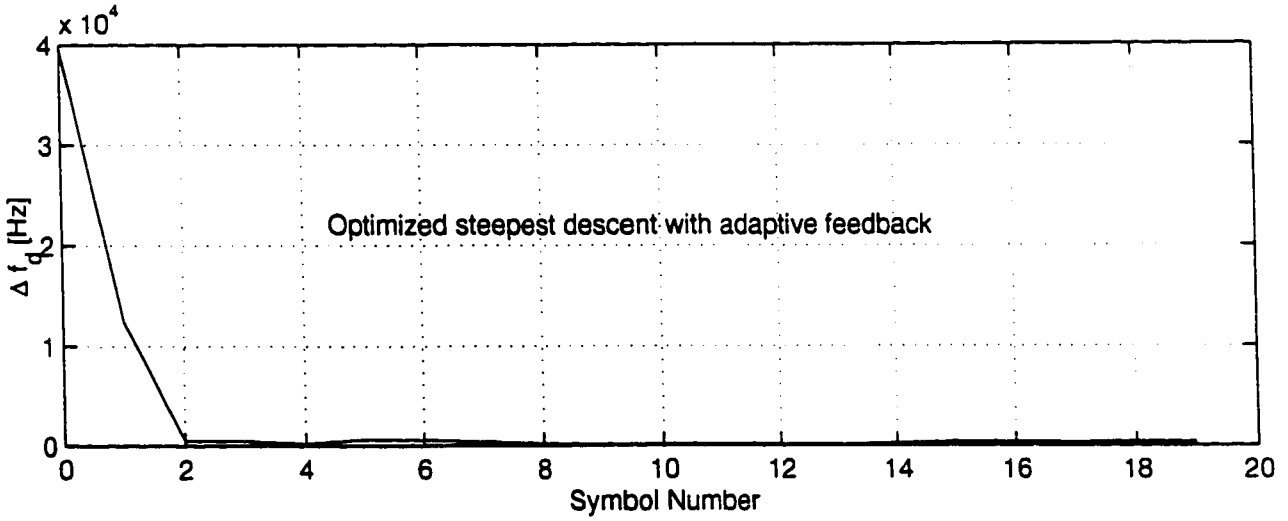
Figure 6.9: Optimized steepest descent employing conventional feedback

To minimize the number of graphs, only $|\epsilon_d|$ and $\Delta f_d = |f_D - f_d|$ were used for comparisons, and to reduce computation effort, the over-sampling factor was set to $M = 20$. Higher over-sampling values could have been used for the optimized steepest descent, since a solution was obtained after only a few symbols. Besides, for a fair comparison of various

tracking algorithms, the same over-sampling factor must be used throughout all simulations. Values $M > 20$ for the steepest descent case, showed to be too computationally demanding for a personal computer.



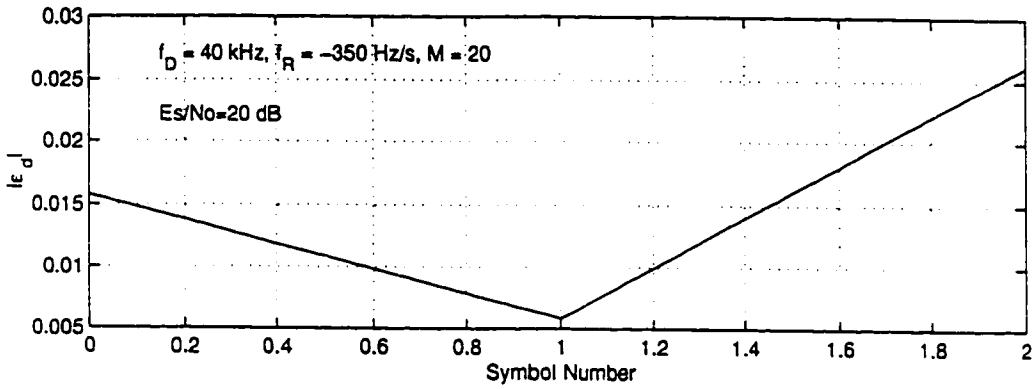
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



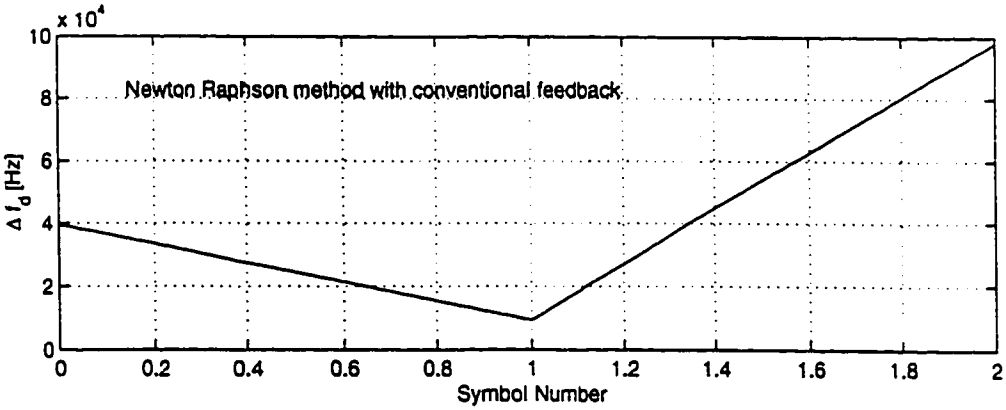
(b) Magnitude of frequency shift error $|\Delta f_d| = |f_D - f_d(n)|$ versus symbol number

Figure 6.10: Optimized steepest descent employing adaptive feedback

The Newton Raphson method is a special case since the objective function (5.35), that is, the likelihood function is quadratic. It will therefore produce a solution after the first symbol and hence, does not require an adaptive approach. However, application for more than one symbol may diverge as shown in Figure 6.11. Simulation results show that the Doppler shift is reduced from 40 kHz to 10 kHz after only the first symbol, and this is shown in Figure 6.11(b).



(a) Frequency detector output $|\epsilon_d|$ versus symbol number



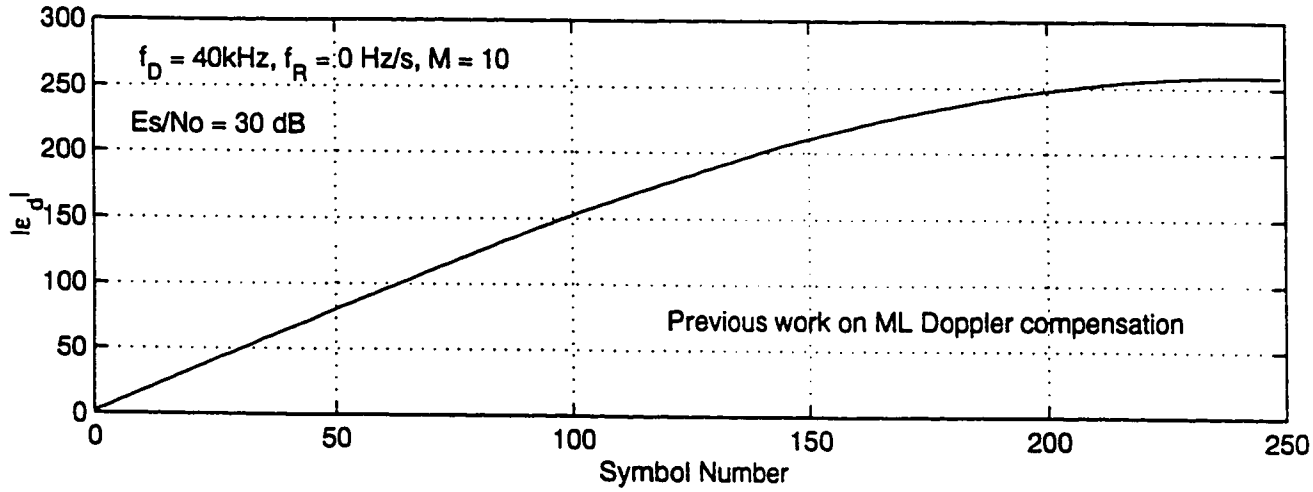
(b) Magnitude of frequency shift error $|\Delta f_d| = |f_D - f_d(n)|$ versus symbol number

Figure 6.11: Newton Raphson Method

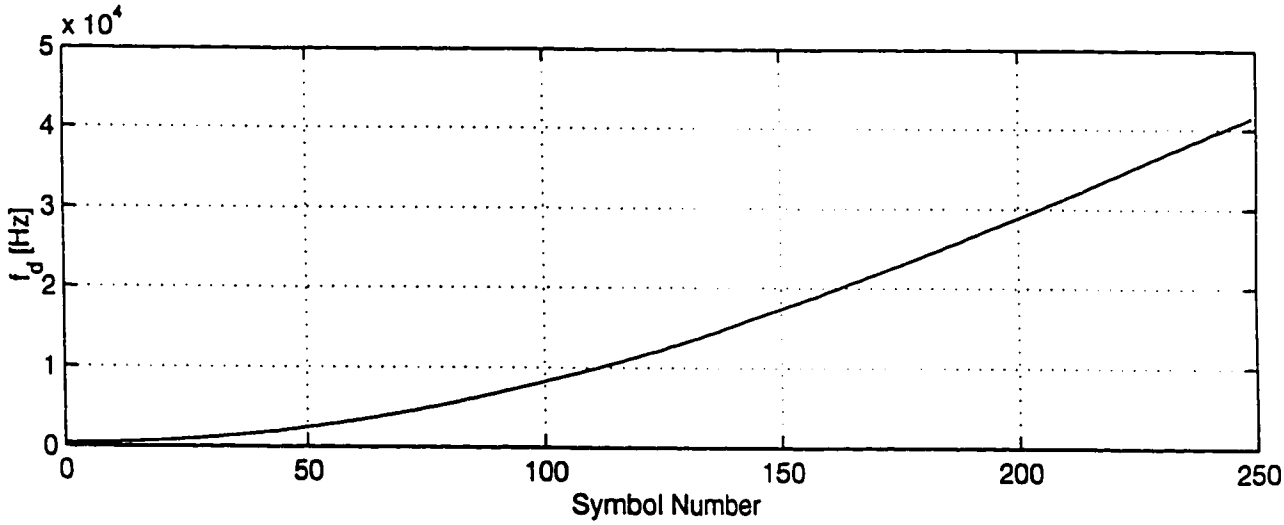
6.4 Doppler Offset (Shift) and Doppler Change (Rate)

One way to test how well a proposed detector performs is to compare it with those designed by others. We have chosen the detector developed by A. N. D'Andrea and U. Mengali since it had the best performance as shown earlier in Table 4.2 [90]. In section 4.8.4 we discussed the short comings of current ML frequency detectors. It is under those conditions, that is when $f_D = 0$ Hz and $f_R \neq 0$ Hz/s, we now compare the proposed ML QPSK Doppler compensator employing the steepest descent with conventional feedback and the adaptive optimized steepest descent methods to one developed by A. N. D'Andrea *et al.* The primary reason for choosing conventional steepest descent was because it employed a similar feedback technique to theirs.

First we compare their performance under Doppler conditions, $f_D \neq 0$ Hz and $f_R = 0$ Hz/s. More precisely, when $f_D = 40$ kHz and $f_R = 0$ Hz/s and the results from the simulation are shown below in Figure 6.12. In particular, Figure 6.12(a) depicts the frequency detector output $|\epsilon_d|$ versus symbol number and Figure 6.12(b), the estimated Doppler frequency shift $f_d(n)$ versus symbol number. The latter basically shows the estimated Doppler frequency of the detector and probably the best variable for comparison.



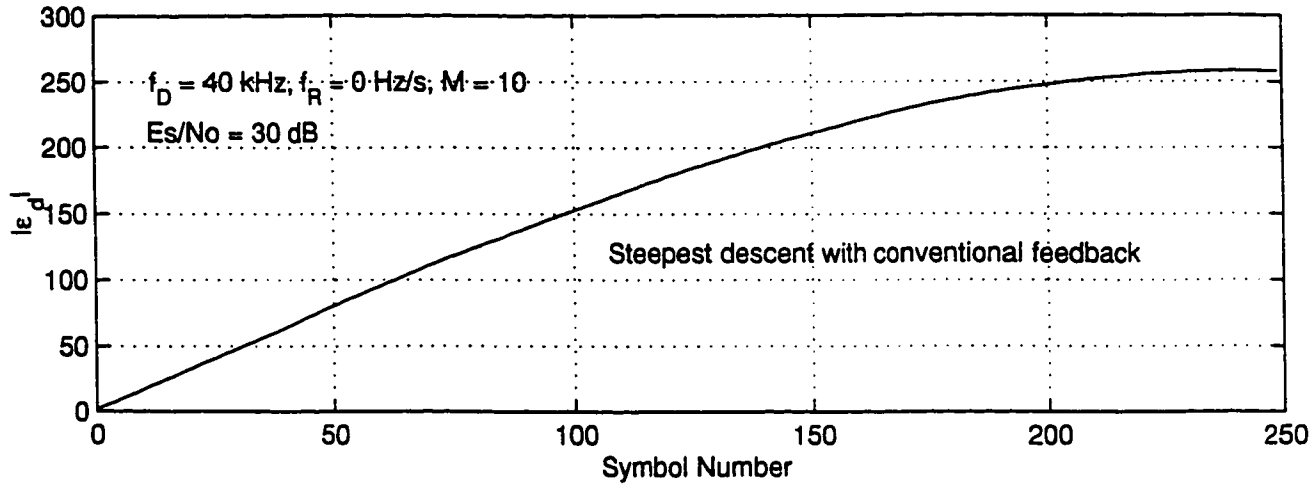
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



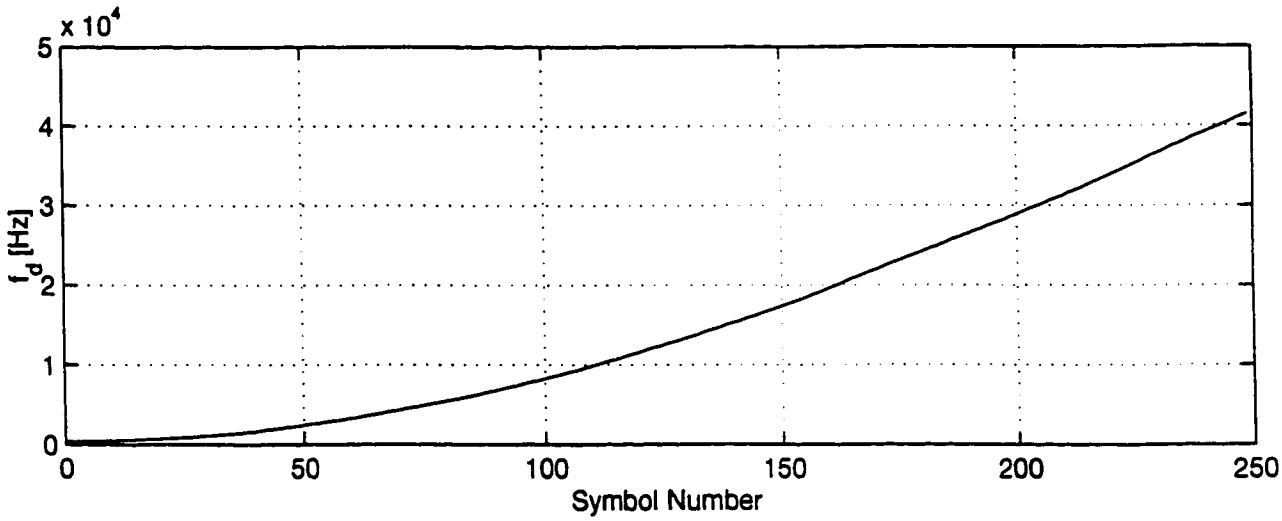
(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.12: Previous work on Doppler compensation

Under the same conditions on the QPSK signal, we make use of the proposed ML QPSK Doppler compensator employing steepest descent and using conventional feedback, we obtain results which are shown in Figure 6.13.



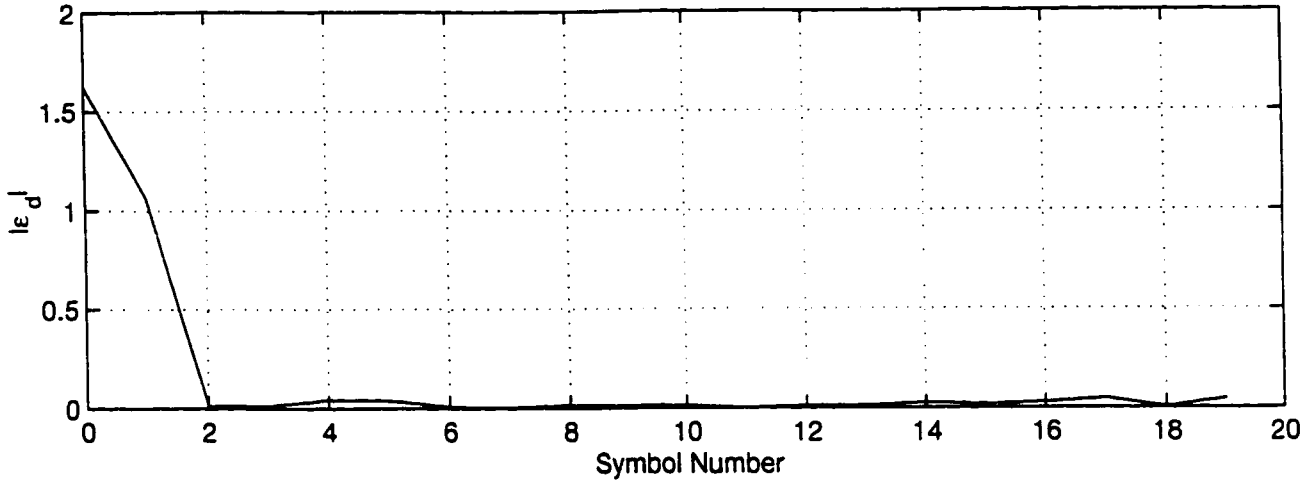
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



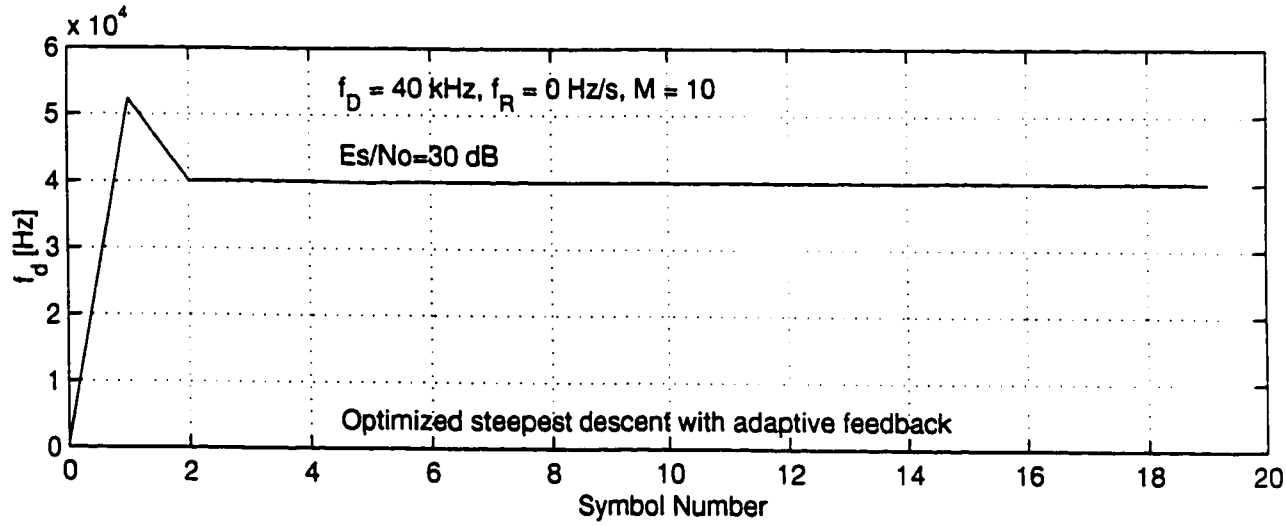
(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.13. Proposed steepest descent employing conventional feedback

The results looks identical, implying that they both have the same performance. Next we compare Figure 6.12 to the proposed ML QPSK Doppler compensator employing the adaptive optimized steepest descent which is shown in Figure 6.14.



(a) Frequency detector output $|\epsilon_d|$ versus symbol number

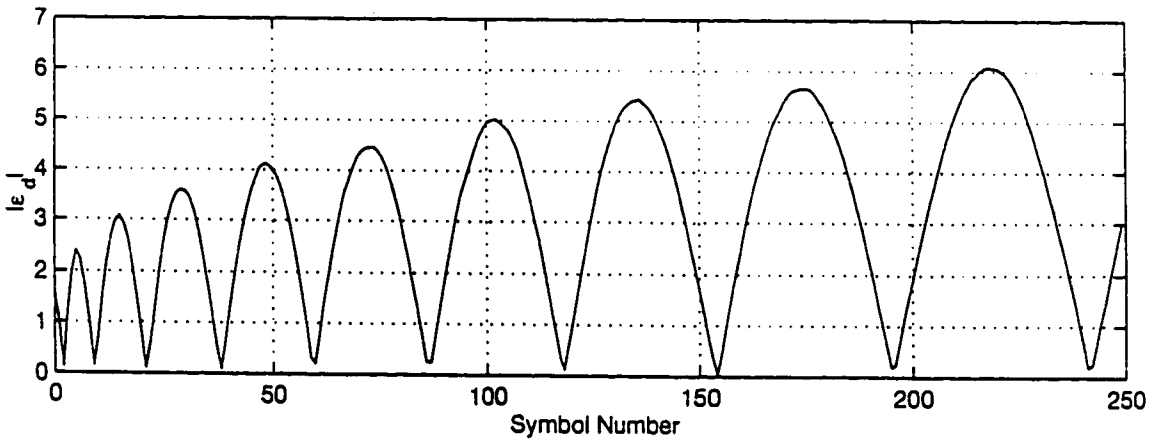


(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

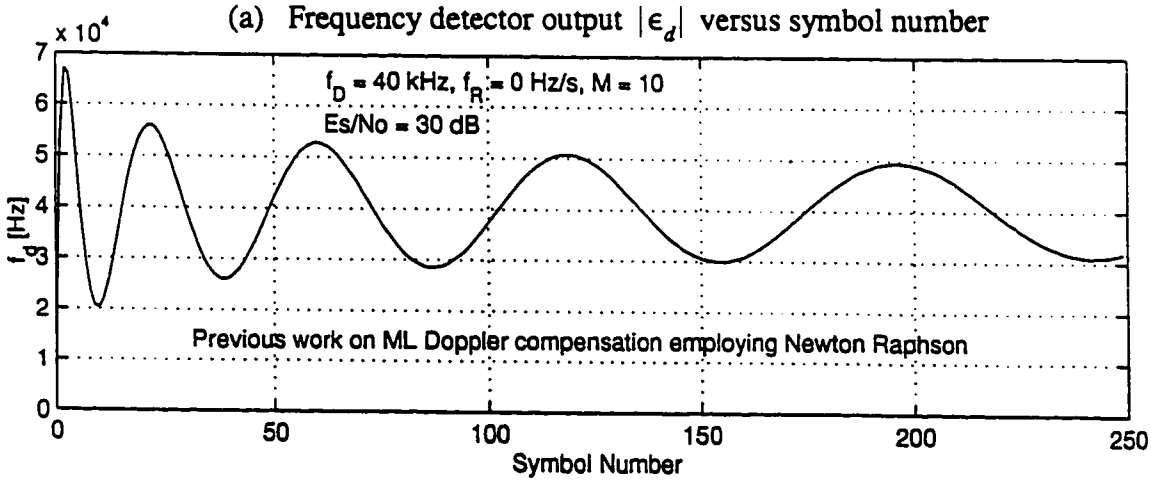
Figure 6.14: Proposed optimized steepest descent employing adaptive feedback

This method obtains a solution much faster than either methods presented earlier in Figure 6.12 and 6.13. The adaptive optimized steepest descent achieves carrier lock after 6 symbols, that is over a 100 fold time saving and maintains lockup.

ML Doppler compensators developed by [90] which are currently used can be modified to utilize Newton Raphson strategy, minimizing an objective function of only one variable, since their method estimates only the Doppler frequency shift f_d. The simulation results of this method is shown in Figure 6.15 and is referred to as the revised method. Oscillation in the solution was present but the error did appear to converge to zero eventually.



(a) Frequency detector output $|\epsilon_d|$ versus symbol number



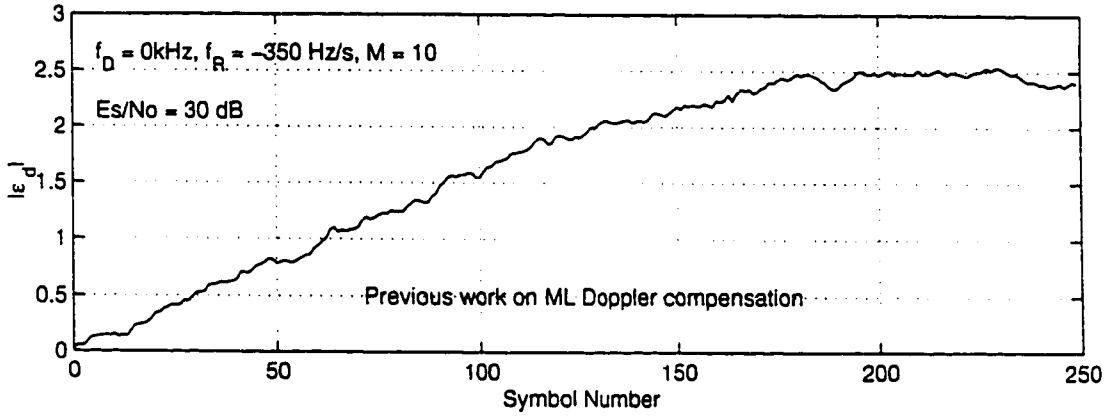
(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.15: Previous work on Doppler compensation (revised)

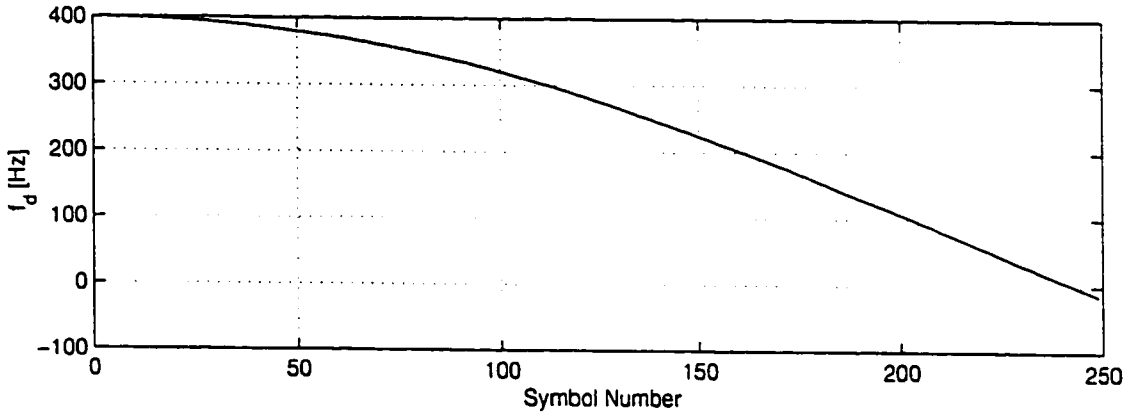
Here the estimation of f_d oscillates around 40 kHz, which is the actual value of f_D.

Now we investigate the case when the Doppler conditions are f_D = 0 Hz and f_R = -350 Hz/s.

Figure 6.16 shows the results using previous work from [90] and it was deduced that a good estimate of f_d was obtained after 240 symbols.



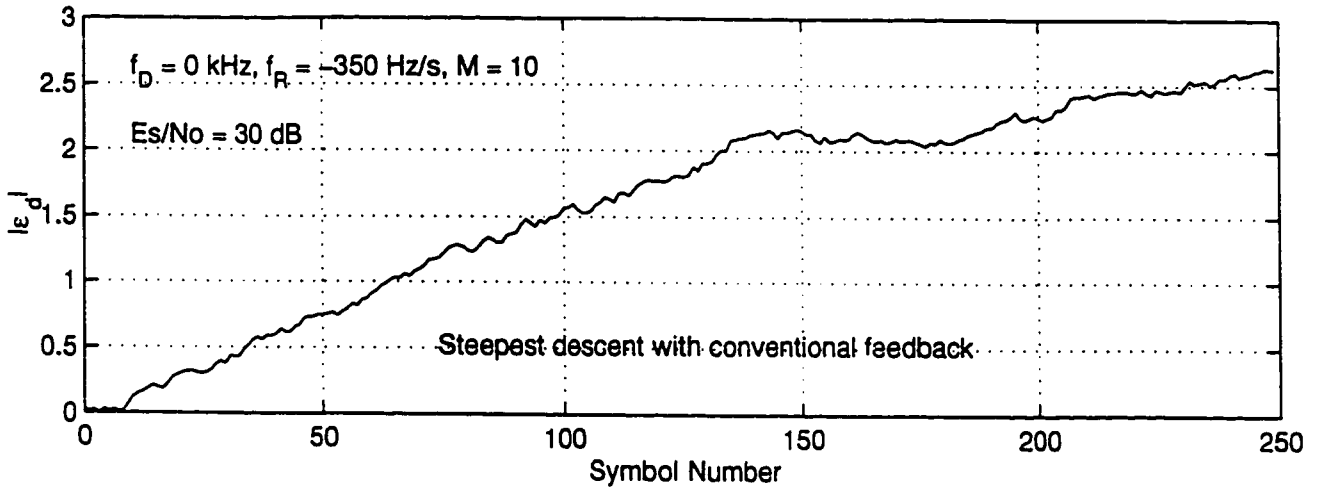
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



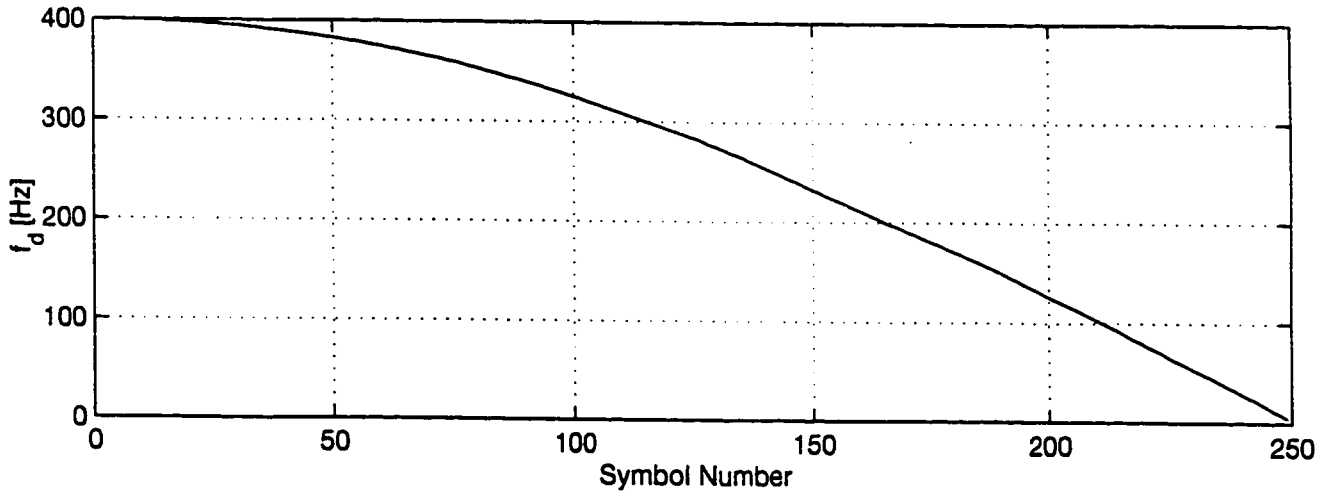
(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.16: Previous work on Doppler compensation

Next we use the steepest descent employing conventional feedback method and results are shown in Figure 17. A solution is obtained after 210 symbols, hence a slightly faster method and both solutions, Figure 16 and 17, showed signs of divergence.



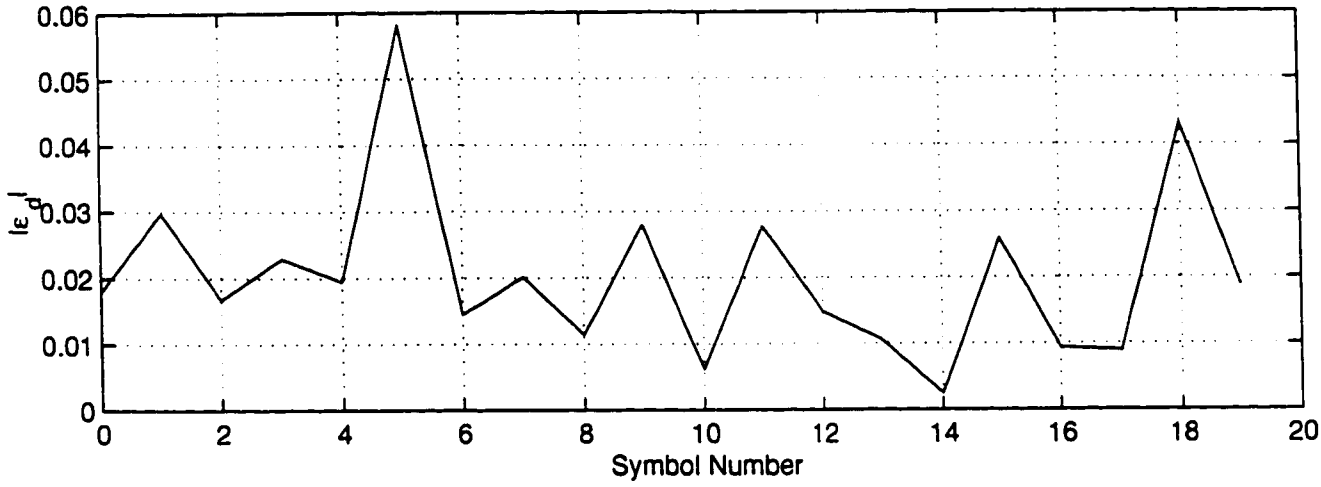
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



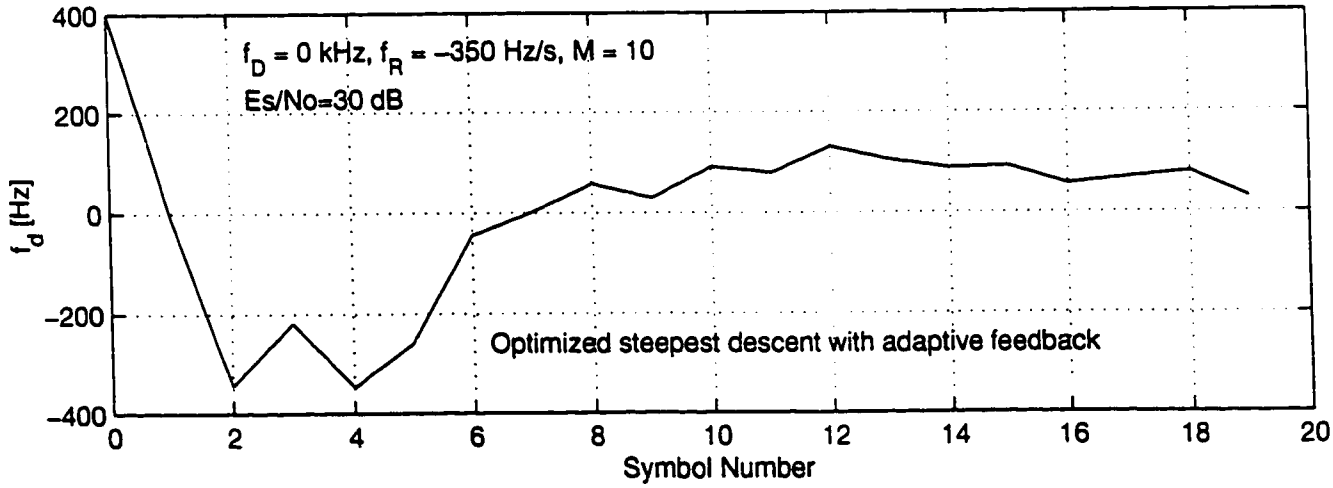
(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.17. Proposed steepest descent employing conventional feedback

Simulation of the adaptive optimized steepest descent method produced the results which are shown Figure 6.18. It is clear that the detector locks up after 7 symbols and it seems that f_d oscillates around $f_d = 0$ Hz, which is the actual value.



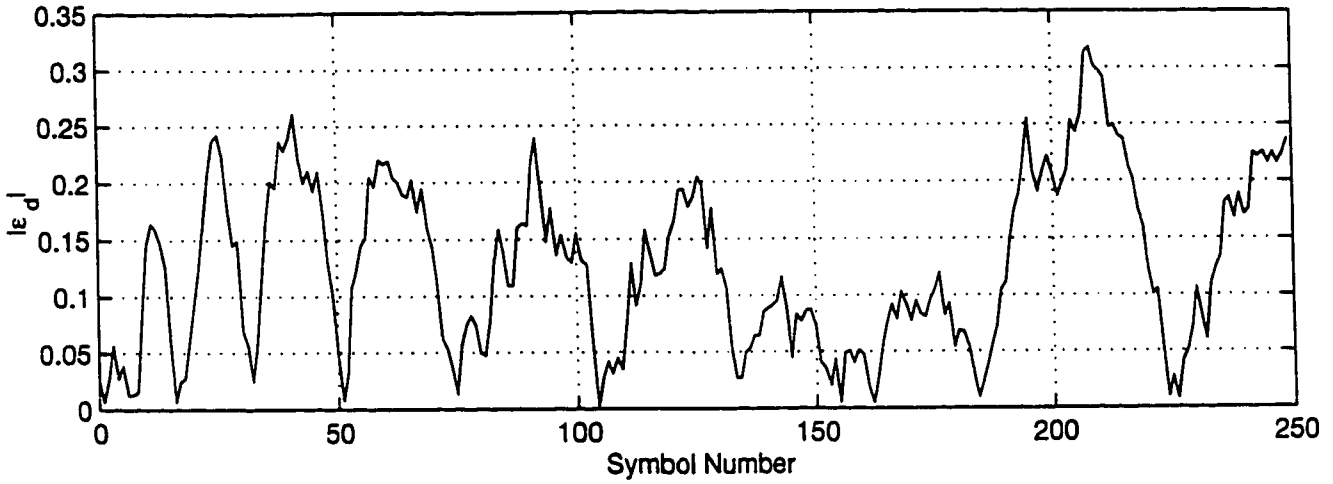
(a) Frequency detector output $|\epsilon_d|$ versus symbol number



(b) Estimated Doppler frequency shift $f_d(n)$ versus symbol number

Figure 6.18: Proposed optimized steepest descent employing adaptive feedback

Finally, the revised method of previous work by [90] on Doppler compensation is illustrated in Figure 19. The resultant plot in Figure 19(b) looks like a partially damped oscillation on f_d around 0 Hz, since the amplitude of the oscillations showed signs of attenuation as more symbols are observed.



(a) Frequency detector output $|\epsilon_d|$ versus symbol number

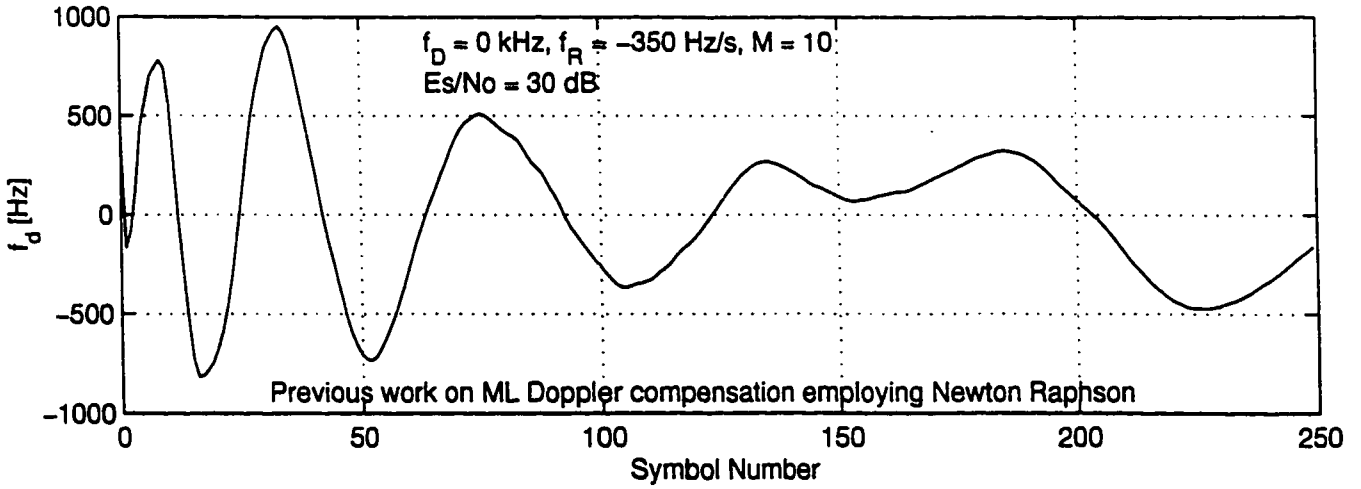


Figure 6.17(a,b): Previous work on Doppler compensation (revised)

6.8 Summary

It has been shown that the s-curve is dependent on the symbol SNR and QPSK signal pulse $h(t)$. Theoretically, both the expectation of the carrier synchronizer frequency and frequency rate feedback loop errors the have been shown to be non-affected by the presence of AWGN, where the former is the s-curve when only the first symbol is considered. Simulations have shown to agree with theory, but for low symbol SNR, the effects of noise are noticeable. This was shown by the ability of the s-curve to cross the origin. Representation of the noise process on a per symbol basis is dependent on the over-sample factor. The value of $M = 50$ was chosen since any increase in M showed no significant improvement in the s-curve. The theoretical useful range of the ML QPSK Doppler compensator was found to be $-R_s \leq \Delta f_d \leq R_s$, but its practical useful range was shown to be $-0.4R_s \leq \Delta f_d \leq 0.4R_s$.

Theoretical error analysis for the Doppler compensator at equilibrium has shown that the QPSK signal shape must be real and must satisfy Nyquist criterion for the feedback loop error in the compensator to settle at zero after lockup.

The grid search has shown to be adequate as a coarse Doppler estimator scheme but it was found to be highly computationally dependent. Therefore, it may not be suitable for any real time applications. In sections 6.1.3 and 6.1.8 it was confirmed that signals with higher symbol SNR would achieve shorter lockup times.

In order to have a thorough investigation of the proposed ML Doppler compensation, three optimization methods used for implementing the compensator were explored and these include; steepest descent, optimized steepest descent; and Newton Raphson method. Simulation results have shown that the adaptive optimized steepest descent was capable of locking up in only a few symbols (over 100 fold improvement), thereby reducing the lockup time considerably. Furthermore, it was the only method that was stable after lockup.

Comparison of the implementation of the ML QPSK Doppler compensator using adaptive optimized steepest descent method to current Doppler compensation methods, have shown that the former has a remarkable performance in terms of faster lockup and stability.

CHAPTER 7

Conclusions and Contributions

7.1 Contributions of this Thesis

The major contributions made by this thesis include:

1. The development of a model for generating Doppler frequency shift and Doppler rate curves for an LEO satellite system for a user located anywhere on earth at any satellite cross track angle.
2. The construction of a model for the phase error introduced by Doppler frequency shift and Doppler rate.
3. The derivation of two ML solutions namely, the grid search and ML QPSK Doppler compensator for LEO satellite systems.

4. The development of a carrier synchronizer structure to implement the ML QPSK Doppler compensator.
5. A validation of the grid search technique through computer simulations for LEO satellite systems.
6. A verification of the ML QPSK Doppler compensator s-curve (discriminator characteristic) via theory and computer simulations.
7. A comparison of the performances of the steepest descent, optimized steepest descent and Newton Raphson, employing conventional and adaptive feedback mode of operation, by using tracking analysis.
8. Performance comparison to current methods of Doppler compensation for LEO satellite systems.

7.2 Conclusions

In this thesis we have examined the Doppler shift problem associated with typical LEO satellite communication systems for mobile users. The Doppler shift was found to be rapidly time varying and was beyond the QPSK acquisition range fundamental limit for low data rate applications. In addition, the effects of Doppler shift and Doppler rate are more severe in the case for an overhead pass satellite.

By careful investigation of the Doppler curves obtained for LEO satellites at various cross track angles, it was apparent that all of them had the same s-shape characteristic. The

novel technique was developed and later refined for improved performance for overhead pass satellites, as a technique to reduce Doppler shift to tolerable levels, where conventional carrier synchronization may then be employed. Simulation showed that the novel technique performed better for satellites at maximum range orbits, that is, for satellites at larger cross track angles. However, the modified novel technique was found to be best suited for overhead pass satellites. The novel technique can be implemented with very little hardware since it only involves sweeping the local oscillator frequency.

We also modelled the phase error introduced by Doppler shift and Doppler rate. ML solutions for estimating Doppler shift and Doppler rate produced two implementable solutions. The grid search was found to be a quick and dirty method of obtaining a rough estimate of Doppler, whereas, the ML QPSK Doppler compensator was a more accurate estimation process, since it involved tracking the Doppler shift and Doppler rate. Tracking was made possible by employing a feedback loop. Three methods were suggested and compared; steepest descent, optimized steepest descent, and Newton Raphson method. The latter always produced a rough estimate after one symbol observation, since the log-likelihood function was quadratic in nature. The first two were capable of either conventional or adaptive feedback mode. Both the steepest descent and optimized steepest descent were simulated employing conventional and adaptive feedback operation. Simulations show that adaptive optimized steepest descent had superior performance, by locking in carrier after only a few symbols and maintaining equilibrium. This method showed close to 100 fold lockup time improvement.

The adaptive optimized steepest descent was then compared to a current carrier synchronizer. Under only Doppler frequency shift (Doppler rate assumed zero), the adaptive optimized steepest descent method was found to be better. The same conclusion was made when there was only Doppler rate and no Doppler frequency shift. Hence, our final conclusion is that the adaptive optimized steepest descent implementation of the ML QPSK Doppler compensator is a superior Doppler resistant carrier synchronizer than current methods for low data rate LEO satellite communication systems. Its rapid lockup characteristics will enable these systems to be more efficient, by increasing their capacity (ability to handle more users at any given time).

7.3 Future Work

Future work in several aspects are suggested as follows.

1. The study has assumed ideal rectangular data pulses, the latter simply meaning wideband QPSK. Raised cosine with various roll-offs, band limited QPSK and even offset QPSK should also be examined.
2. Investigate the effects of various data pulse shapes $h(t)$ to achieve a wider frequency acquisition range for the Doppler compensator.
3. Equilibrium analysis should be explored in more detail to understand the stability of the carrier synchronizer at lockup.

4. Conduct a study of using grid search or Newton Raphson as the initial estimator.

5. Investigate the performance of the ML QPSK Doppler compensator in fading satellite channels.

6. The study assumed random data since the search was for a non-data aided Doppler compensator. Performance using known data patterns or periodic sequences should also be explored.

APPENDIX A

Domestic Satellite Communications for Islands and Developing Countries

A.1 Background

There are many important aspects in custom designing an economical and efficient thin-route domestic communication satellite system for remote island and developing countries. Factors such as the service requirement, geographical location and the availability of satellite slots must be considered. As an example of a country consisting of scattered islands, we shall consider the Republic of Seychelles.

A.1.1 Thin-Route Communication System

Countries consisting of sparsely populated islands have certain requirements for a basic telecommunication service which might consist of 2 to 6 telephone channels, telex, facsimile, low baud rate data communications, at least one television channel for daily news, weather forecasts and rural education. This is referred to as thin-route communications and up until now, several of these systems have been implemented around the world including undersea cables to radio links in the HF/VHF/UHF bands and microwave, and satellite links [2, 117].

Satellite communications is oftenly viewed as a solution for rural telecommunication problems but this impression is not necessarily correct. One must realize that the most appropriate rural transmission system depends on many factors including: year round weather conditions; geographical distribution of population; terrain; technical expertise available in the country; quality of service; cost-effectiveness of the system; availability and reliability of commercial power; compatibility with existing infrastructure, that is, interfacing with PSTN; and, most importantly, its financial capabilities [1].

A.1.2 Why Choose a Satellite System?

The geographical location of the Seychelles is depicted in the Figure A1(a) and its territorial waters in Figure A1(b). Due to the geography of Seychelles, consisting of 115

islands with 447 km^2 total land area and a population of only 69000, it is obvious that the data traffic will always be thin-route. The islands are sparsely populated and spread over 1.3 million km^2 of the Indian Ocean. This corresponds to 0.03% land to sea area ratio. Seychelles consists of two types of tropical islands. The granitic islands which are mountainous and the coralline islands which are very flat. Since the islands are clustered in groups which are separated by several hundred kilometres, a satellite system is well suited for this application.

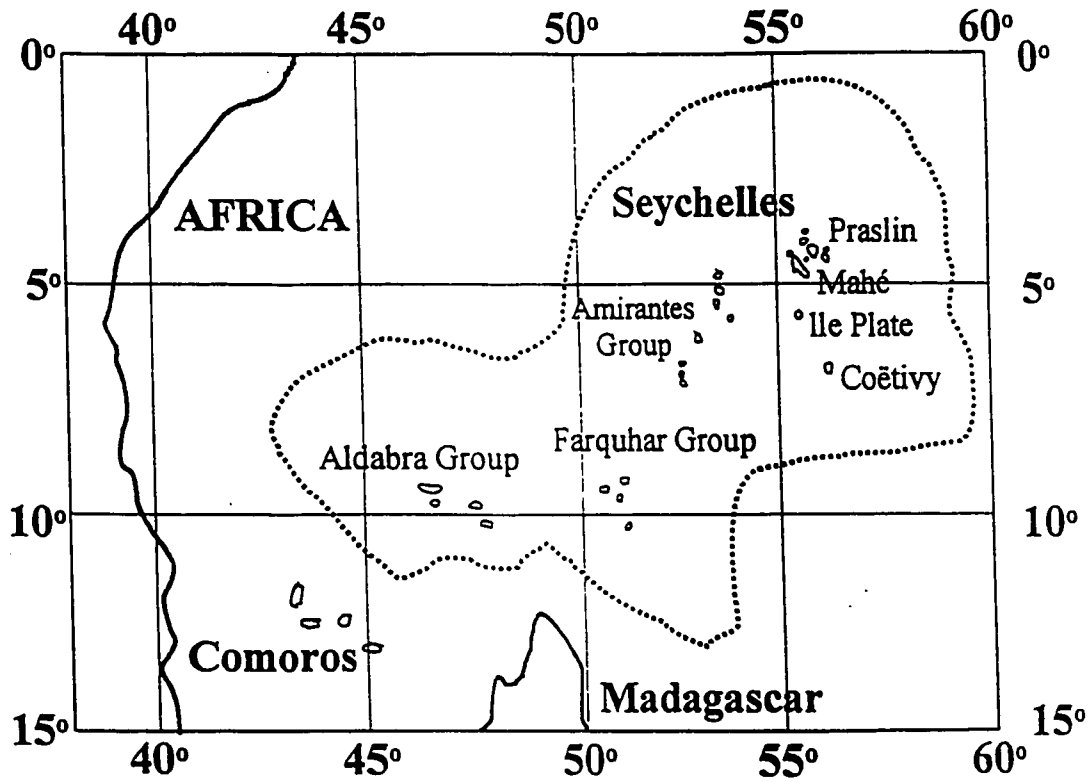


Figure A.1(a): The Seychelles geographical location

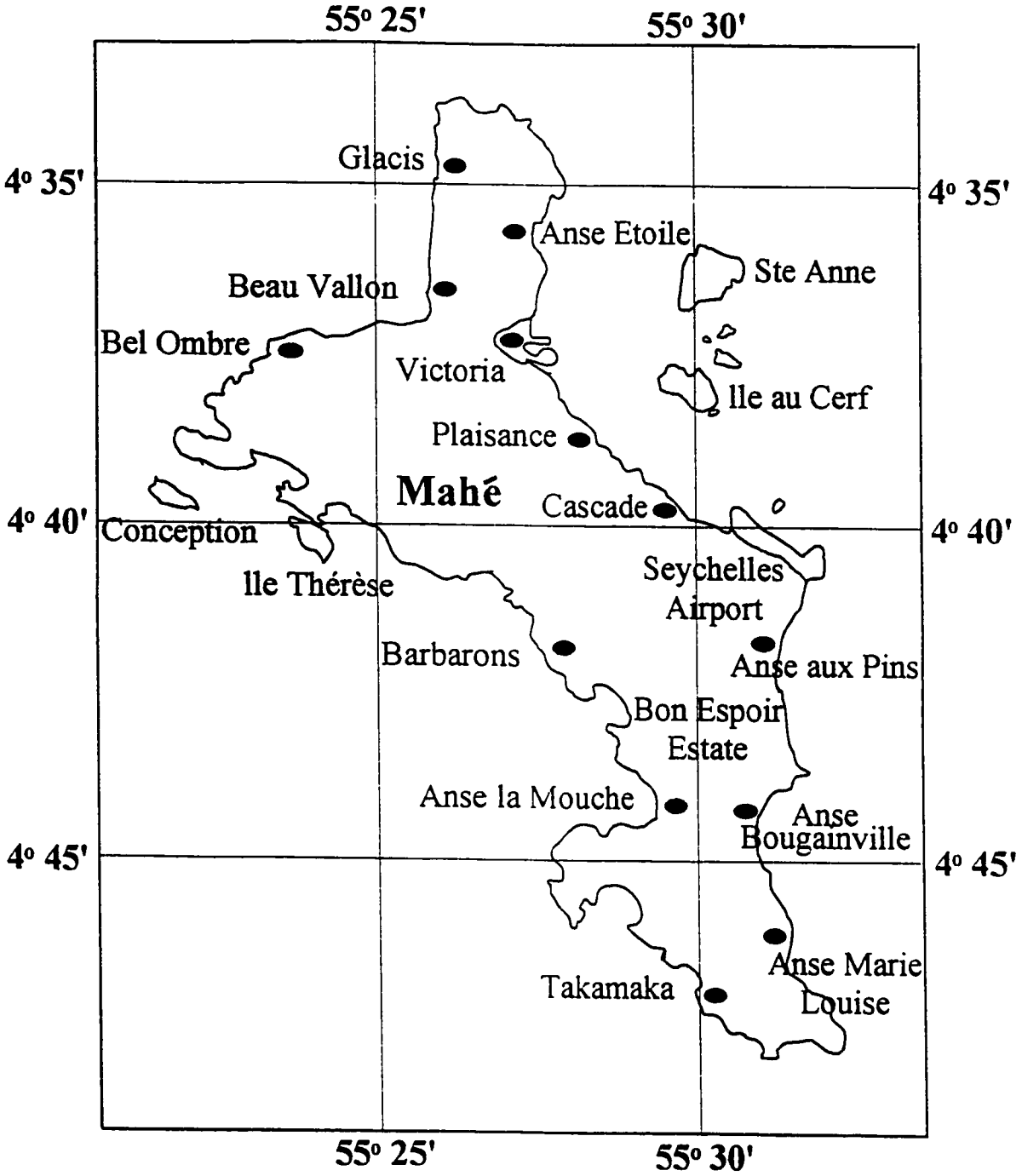


Figure A.1(b): Seychelles Islands and territorial waters

High grade commercial electricity produced by the government power station is available on only the three largest islands. The remaining islands depend on small low power gasoline generators, wind power and sometimes, solar power making it important to implement low power consumption satellite user terminals. In addition, the high monsoon and trade-winds in the Indian Ocean Region place a constraint on the size of the antenna, hence there is a demand for the use of very small aperture terminals (VSAT). Tropical islands in the Indian Ocean Region experience heavy rainfall from May to November and rain fade is of great concern when selecting the satellite frequency bands. Although radiocommunication sector of the international telecommunication union (ITU-R) (formerly international radio consultative committee (CCIR)) attenuation prediction models are available for temperate regions, little has been reported for tropical or equatorial regions. In the design of a reliable and economical satellite system, one of the key design parameter, is the rain fade [2, 118-128].

A.1.3 Satellite Slots

Many countries without a domestic satellite system in service have one or more available satellite slots in the space segment allocated on the Clarke Orbit by the international telecommunication union (ITU). Countries deciding whether or not to introduce a domestic satellite system are at the planning stage, which involves comparing the cost of leasing transponders from international satellite consortia, such as INTELSAT and

INMARSAT, to the expense of launching of a new domestic satellite system custom designed to tailor-fit their needs. Some countries eventually lease their slots to major international satellite organizations.

A.2 Domestic Satellite Systems

In this section satellite systems for India, Peru, South Africa, Indonesia, Japan, and Seychelles are now examined as candidates for domestic satellite communication systems serving islands and developing countries. The systems investigated are generally employed for domestic and international communications and they are often multi-purpose, providing communications, meteorology, military, surveillance, navigation and guidance, search and rescue, distress communication, disaster warning and various other purposes.

A.2.1 India

India lying between 8° and 37° north latitude, and 60° and 100° east longitude with about 80% of its population living in rural areas [129], employs the INTELSAT and INMARSAT facilities for its international telecommunication needs [130-131]. Domestically, the Indian National Satellite (INSAT) handles the extensive use of long-distance public telecommunications on thick as well as thin trunk routes to remote areas and islands, direct broadcast satellite (DBS) TV and radio broadcasting, and meteorological

services. The INSAT system is a multi-purpose domestic satellite system specifically intended for India's specific requirements and has been demonstrated to be very cost effective.

After performing extensive satellite communication research with the application technology satellite-6 (ATS-6) satellite of national aeronautics and space administration (NASA) from 1975 to 1976, the Indian authorities formulated their plans for the design of a domestic satellite system. The objective of the research was to use ATS-6 satellite as a test bed for the development of domestic satellite television broadcasting, to determine optimal system parameters, and to gain experience in the development, testing and managing a satellite system [130, 132]. The first generation Indian National Satellite system (INSAT-1) began operation in October 1983 consisting of three multi-purpose, axis-stabilized geostationary satellites designed to meet the anticipated service requirements of India well into the early 1990's [130].

A pictorial view of the INSAT-1 system is shown in Figure A2. INSAT-1A, launched in April 1982 and stationed at 74° east longitude, was the very first of the INSAT-1 satellite series but failed to deploy its solar array and therefore was never operational [133]. INSAT-1B, INSAT-1C and INSAT-1D employ an S-band slot at 2.5 GHz and C-band slots with the uplink at 6 GHz the downlink at 4 GHz (6/4 GHz). Both C and S frequency bands are used for communication purposes, with C-band used mainly for telephony and data traffic while S-band is used for direct TV broadcast and meterological services [134]. The satellites are stationed at 74°, 93.5° and 83° east longitude, respectively, and were built by Ford

Aerospace and Communications Corporation (FACC) of USA [130]. INSAT-1B, INSAT-1C and INSAT-1D were launched in August 1983, July 1988 and June 1990, respectively.

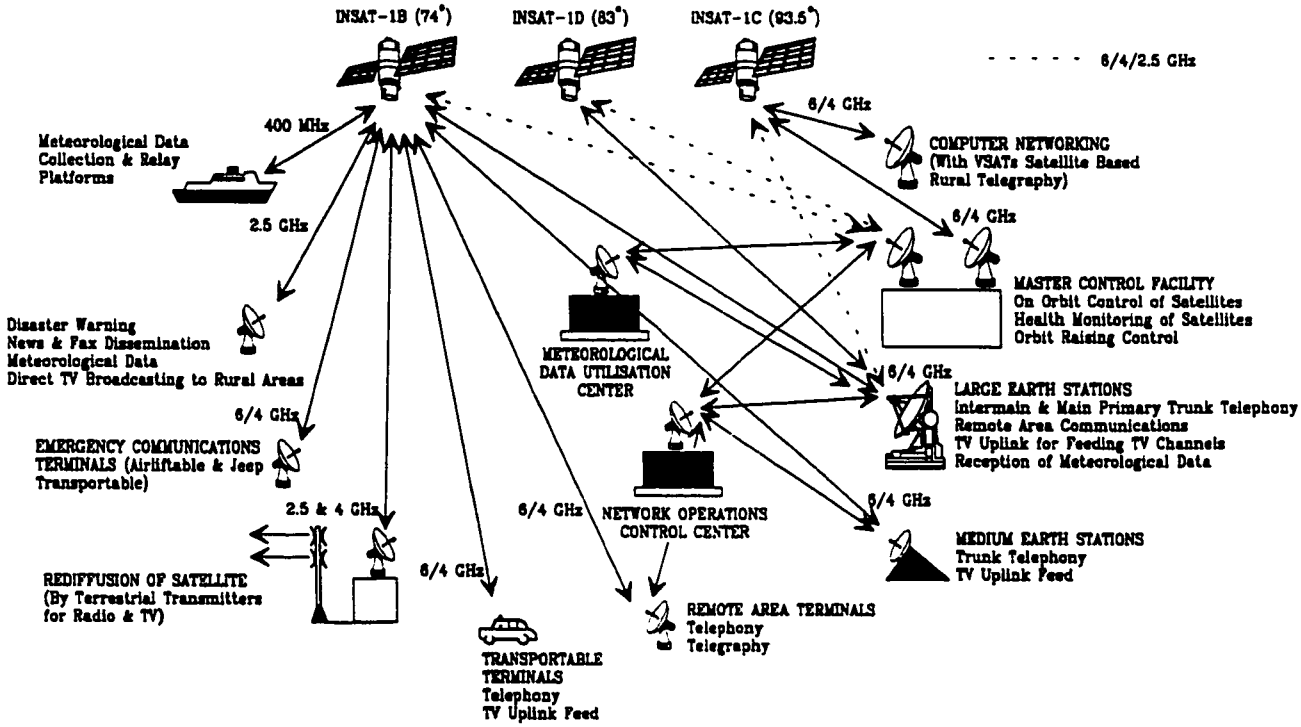


Figure A2: Space and ground segment of the INSAT-1 system

To cope with the increase in service requirements of the early 1990's, the second generation of satellites INSAT-2 was introduced to provide more capacity than its predecessor and operate in the same C and S frequency bands. Besides C and S bands, INSAT-2C and INSAT-2D provide dedicated Ku-band coverage for India but these satellites do not have any meteorological payload [135]. A pictorial view of the INSAT-2 system is

illustrated in Figure A3.

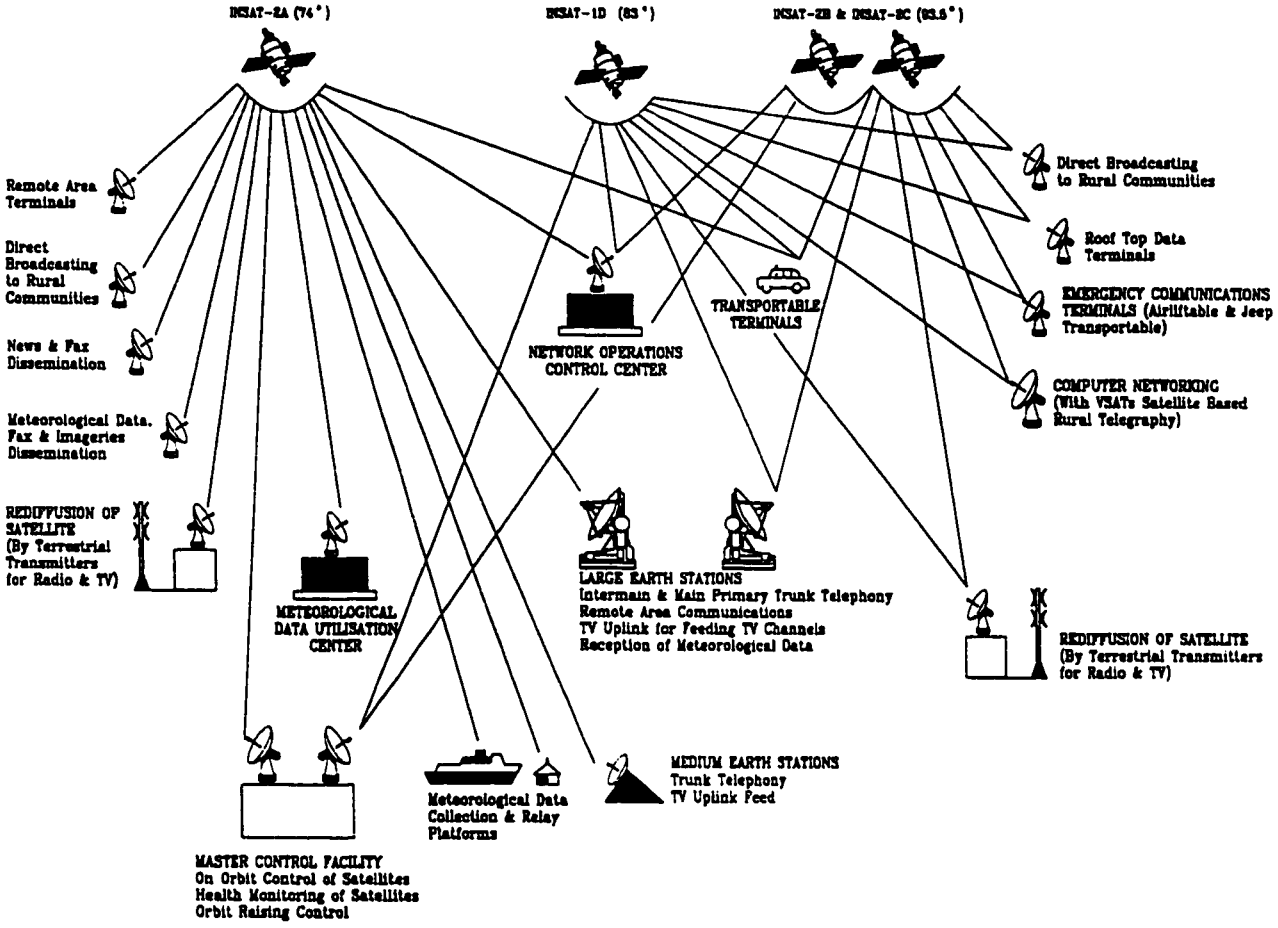


Figure A3: Space and ground segment of the INSAT-2 system

India has three geosynchronous orbit slots at 74°, 83° and 93.5° east longitude and plans to eventually collocate two satellites per slot to achieve a total of six satellites [136]. INSAT-2A was launched in July 1992 and INSAT-2B in July 1993 and stationed at 74° and 93.5° east longitude, respectively. In December 1995, INSAT-2C was launched and

stationed at 93.5° east longitude. The complete system will consist of a total of four multi-purpose satellites, of which INSAT-2B and INSAT-2C are placed very close together (collocated) and appear to operate as a single high capacity satellite. This technique becomes necessary when slots are not available in the geostationary orbit and India was the first country to implement it. Later this year INSAT-2D will be launched and collocated with INSAT-1D, where the former will eventually replace the latter. Launch dates and orbital slot locations of INSAT satellite series are shown in Table A1. While the basic configuration of the two satellite systems is similar, only INSAT-2 is capable of detecting emergency distress radio signals on land or anywhere in the Indian Ocean from the 406.025 MHz emergency location transmitters (ELT). The INSAT-2 satellite series has more capacity with 18 C-band transponders as compared to INSAT-1 series which has only 12 C-band transponders [133]. The multiple access technique utilized in INSAT-2 is TDMA and the size of the user's antennas range from small hand-held sets to roof top VSAT's. There are more than 500 rural earth stations on the mainland and eight others on the nearby islands [135]. It is also interesting to note that INSAT-2E will be launched sometime in 1997 and INTELSAT plans to lease 11 C-band transponders [137].

There are several types of VSAT networks currently operating in India [129, 135]. Even though VSAT's can be used for any kind of digital transmission within the limitation of data speed, it is primarily used for data applications such as e-mail, text messaging, telex, facsimile, remote computer access, and electronic money order [135]. The reason why data communications in particular messaging needs, are more vital compared to voice services,

is from the point of view of cost, affordability, market, and technology [131].

Satellite	Launch Date	Orbit Location (East Longitude)
INSAT-1A	April 1982	74° (Failed)
INSAT-1B	August 1983	74°
INSAT-1C	July 1988	93.5°
INSAT-1D	June 1990	83°
INSAT-2A	July 1992	74°
INSAT-2B	July 1993	93.5°
INSAT-2C	December 1993	93.5°
INSAT-2D	Late 1996	83°

Table A1: Launch dates and orbital slots of INSAT

A.2.2 Peru

Geographically, Peru is divided into three different areas, including the coastal area which is desert, the mountains and the jungle. The Peruvian telecommunications authorities have been more concerned in improving and extending thin-route communications service primarily to the inhabitated mountainous and jungle regions of Peru by using a domestic satellite communication system [138-139]. Thus, reliable high quality transmission media for international and domestic telephony, and television channels are now readily available to the small towns and villages which are remotely scattered in the mountains and jungle

areas of Peru. These services are not economically achievable by other means, such as HF/VHF/UHF or microwave links or cable, due to the distances and nature of the geographical coverage involved.

The Peruvian telecommunication authority leases ¼ of a 36 MHz transponder (9 MHz bandwidth) from INTELSAT IVA which is stationed at 30° west longitude above the Atlantic Ocean. This provides a maximum of 200 telephone channels, thus allowing communication links between small towns and villages.

The satellite system utilizes C-band (6/4 GHz) and coverage is provided by a global beam with the telephone channel system being based on single channel per carrier (SCPC) pre-assigned (PA). In PA, also known as the fixed assignment scheme, each user will always employ the same channels of which others may not utilize. Such a technique is ideal for high capacity users, but is inappropriate for low traffic links between remote villages as those encountered in many developing nations. For small towns and villages, a thin route approach is far more efficient where demand assignment (DA) can be implemented.

Demand assignment scheme offers a solution where channels are pooled; hence, channels are allocated upon the user's demand and when no longer required, returned to the common pool. Contrary to PA, demand assignment of channels is never permanently assigned between any two stations. Consequently, DA channels have a larger traffic-carrying capability than PA, and hence can serve a larger number of users, therefore, the DA scheme is best suited for thin and medium traffic, especially for satellite voice circuits.

SCPC is a FDMA technique oftenly used for thin-route applications. There have

been plans to modify the PA FDMA system into a demand assigned multiple access (DAMA) system in order to handle the growth of traffic. This is a more effective and economical technique to handle relatively higher data traffic than normal thin-route cases since it makes better utilization of the available number of channels.

A.2.3 South Africa

In the Republic of South Africa, a small number of radio links are used for communication purposes to nearby islands in the Atlantic and Indian Ocean. Initially, most of the international data traffic was carried by undersea cable which terminates in Portugal, but this resource is completely utilized.

The Republic is a major participant in the INTELSAT consortium making use of the latest developments in INTELSAT services utilizing both C-band and Ku-band (14/11 GHz). At the present time, the Republic has expanded and is now utilizing three INTELSAT earth stations and its submarine cable to handle all of its international traffic which is mainly composed of telephone, data, telex and television [5]. The submarine cable handles about 360 two-way voice grade circuits while INTELSAT compliments the system by providing an additional 1,000 voice grade circuits.

South Africa has implemented a unique system referred to as the "phone shop", which provides voice communications services to disadvantaged rural communities. A phone shop consists of up to 10 GSM fixed cellular transceivers in a building, where callers can make

private calls similar to public telephone booths [117]. Although South Africa is technologically far more advanced than the rest of Africa, problems still arise in terms of the transmission of essential health information on a national basis. Where the medical profession is concerned, contact by fax or phone to peers in developed countries is often prohibitively expensive and the lack of adequate information often means planning without facts or making decisions based on information that is several years out of date. In 1994, store-and-forward satellites were launched into low earth polar orbit at an altitude of 800 km. The orbital period is approximately 100 minutes and each ground station is within the satellite's footprint twice a day for at least 10 minutes per satellite. This satellite system referred to as HealthNet consists of around 500 ground stations located at medical institutions and are capable of exchanging low volume health related information at no cost since the system does not depend on international telecommunication links [140].

A.2.4 Indonesia

Indonesia is the world's largest archipelago consisting of 17,508 islands geographically situated between 6° north and 11° south in latitude, and 95° and 141° east in longitude corresponding to 5,000 km in the east-west direction and 2,000 km in the north-south direction. The total sea surface area is 10 million km^2 and more than 5,000 islands are inhabited [141]. In the early 1970's, inter-island communication relied on HF, troposcatter radio links and cables, while communication within islands was achieved by terrestrial

microwave systems. Unfortunately, due to increased present day capacity requirements, this system now has limitations. With such a large number of islands, it is excessively time consuming and an extremely expensive task to interconnect all the islands by either undersea cable or microwave links. It is therefore not surprising that the Indonesian authorities decided very early to introduced the PALAPA satellite system [142-143].

Indonesia was the first developing country to custom design a satellite system for its own domestic purposes. The first generation Indonesian satellite system PALAPA-A began operation in July 1976 with the launch of PALAPA-A1. In March 1977, a second satellite named PALAPA-A2 was launched, thus completing the system configuration. PALAPA-A1 is stationed at 83° east longitude while the PALAPA-A2 is stationed at 77° east longitude. The PALAPA-A satellite series are used to carry telephony, thin-route data traffic, television, educational programs, and specialized services to large corporations. These satellites are spin stabilized and have a combined total of 12 C-band transponders providing a capacity of either 6,000 voice circuits or 12 television channels or a combination. However, due to the excess capacity of the two combined satellites, it has enabled Indonesia to lease transponders to neighbouring countries including, the Philippines, and Thailand.

Later, due to the rapid growth of data traffic through the PALAPA-A system, a second generation satellite system called PALAPA-B came into service in 1983. PALAPA-B, also known as the B-series satellite, has twice the capacity of its predecessor, that is a total of 24 C-band transponders having a capacity of 12,000 voice circuits or 24 television channels [142]. PALAPA-B1 was launched in June 1983 and stationed at 108° east

longitude. Then, in February of the following year, PALAPA-B2 was launched and to be stationed at 113° east longitude. Originally, the latter failed to reach the geostationary orbit due to a problem in the launcher. It was later retrieved and a replacement PALAPA-B3, also known as PALAPA-B2P was later launched in March 1987. Table A2 shows the launch dates and orbital slot locations of PALAPA satellite series. With the increased capacity, the lease to neighbouring countries was extended to Malaysia, Korea, and Singapore [142, 144]. Both the PALAPA-A and PALAPA-B satellite systems make use of (6/4 GHz) C-band. In April 1990, PALAPA-B1 had drifted to 134° but not removed from the the belt while PALAPA-B2R, which was the retrieved PALAPA-B2 and modernized, was launched and stationed at 108° east longitude.

Satellite	Launch Date	Orbit Location (East Longitude)
PALAPA-A1	July 1976	83°
PALAPA-A2	March 1977	77°
PALAPA-B1	June 1983	134°(Drifted)
PALAPA-B2	February 1984	113° (Failed)
PALAPA-B2P	March 1987	113°
PALAPA-B2R	April 1990	108°
PALAPA-B4	May 1992	118°
PALAPA-C1	February 1996	113°

Table A2: Launch dates and orbital slots of PALAPA

The last of the second generation satellites, namely the PALAPA-B4 was launched

in May 1992 and positioned at 118° east longitude. The design of the third generation satellites PALAPA-C is well under way and scheduled for operation by 1999. PALAPA-C1 has 34 transponders of which 30 are C-band and 4 are Ku-band. It was successfully launched in February 1996 and stationed at 113° east longitude. The broadcast footprint will range from the Middle East to New Zealand and Japan. Specifically, the C-band will have Australasia coverage [141, 144]. All the PALAPA satellites were built by Hughes Space and Communications Group [144].

The ground segment of PALAPA-B consists of approximately 200 earth stations with antennas ranging from 4.5 to 10 m in diameter. It is estimated that there are over 1,000 television receive only (TVRO's) receiving video signals from the PALAPA satellite system and most of the small earth terminals, of which some are at least 2 m in diameter, are assembled in Indonesia [132].

A.2.5 Japan

The Nippon Telegraph and Telephone Public Corporation (NTT) is the pioneer in the use of K-band (30/20 GHz) for domestic satellite communication since their successful development in high speed digital transmission and high frequency devices. From 1967 to 1968, various TDMA experiments were carried out using the ATS-1 satellite [145]. Then, a medium capacity communication satellite referred to as CS, was put into geostationary orbit at 135° east longitude in December 1977 for experimental purposes. In 1983, the

Japanese launched their first generation domestic satellites using their own rockets. The Japanese communication satellite, namely the CS-2 was specifically designed for medium capacity communication and utilizes both C-band (6/4 GHz) and K-band (30/20 GHz). The CS-2 plays an important role in providing the Japanese with reliable communication links between the Japanese mainland and remote islands. A brief overview of the Japanese satellite system is demonstrated in Figure A4. Specifically, the CS-2 consists of two spin-stabilized geostationary satellites, namely CS-2a and CS-2b, where the latter is the standby satellite. The CS-2a satellite is stationed at 132° east longitude while the CS-2b is stationed at 136° east longitude and both were built in Japan [146].

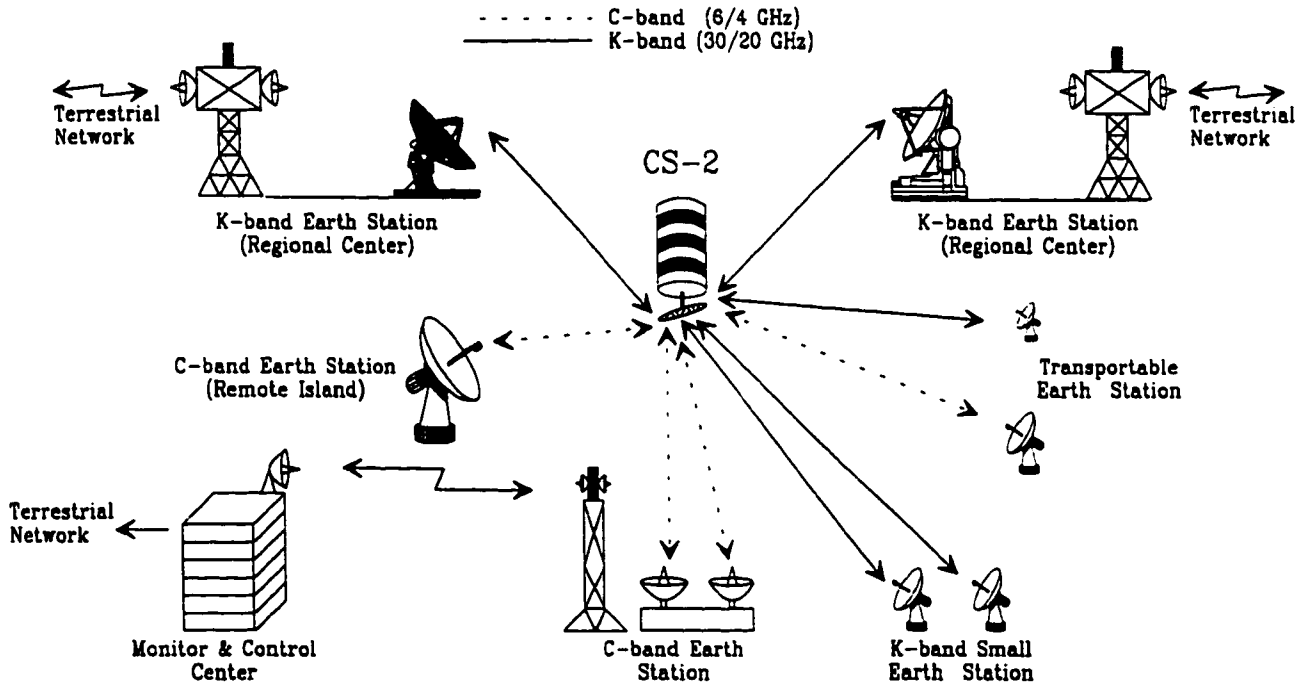


Figure A4: Japanese domestic satellite communication system

The C-band antenna beam covers the whole Japanese territory, including the mainland and the remote islands, whereas the mainland is served by K-band. As a consequence, satellite links between the mainland and the remote islands are provided only by C-band while satellite links within the mainland can be established by both C-band and K-band. Interference in C-band on the mainland especially in the cities, is very high compared to the remote islands and this is due to widespread terrestrial microwave radio relay systems. The main reasons for the usage of the K-band on the mainland only are: the need for high capacity traffic, small and compact high performance portable terminals and, in addition, this frequency band is not extensively used for terrestrial communication links [145].

Each satellite has six K-band transponder units of which only four are used and two C-band transponder units which are fully employed. The K-band TDMA link provides a maximum of 2,000 telephone channels and 2 television channels between the regional centers on the mainland, whereas the C-band TDMA link, is capable of providing only 192 telephone channels and 2 television channels between the mainland and the remote islands [124].

The adoption of K-band has made it possible for the employment of high performance, compact transportable earth terminals and high capacity digital transmission due to its wide bandwidth. K-band suffers considerably from high rain fade [2, 118-128]; therefore, in order to improve the satellite communication links during rainy periods, utilization of high performance hardware, proper selection of modulation and demodulation

techniques and finally, optimization of the transmission channel design had to be carefully implemented. In 1969 extensive research in radio propagation characteristics, site diversity effect, and cross polarization discrimination variation of K-band were conducted at several locations in Japan. From the results of this research, the Japanese have implemented an estimation method for space propagation characteristics which is based on rain fade distribution in Japan [147]. Another technique utilized is the simultaneous use and mutual switching of C-band and K-band as a countermeasure against rain fade [125]. A conceptional diagram of the configuration of the system is shown in Figure A5.

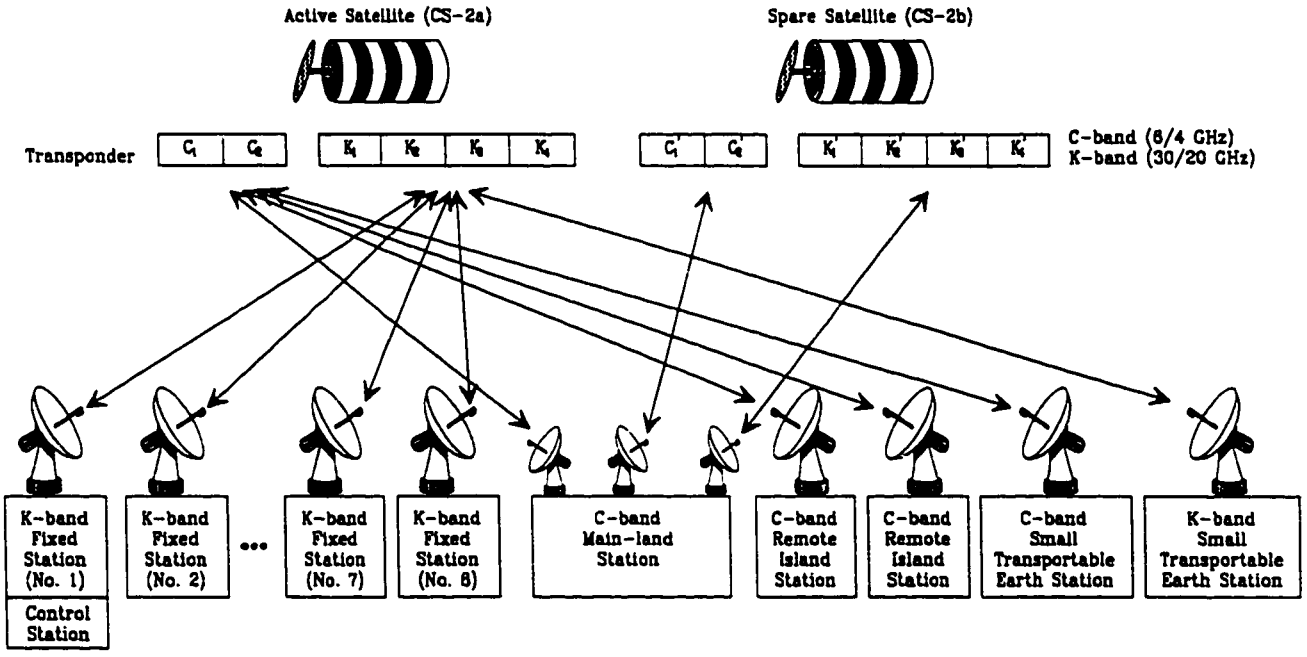


Figure A5: Configuration of the Japanese domestic satellite system

The overall ground communication network consists of the following three systems: K-band fixed earth stations, C-band remote island stations, and K-band and C-band small transportable earth stations. The ground segment of the C-band remote island communication system is composed of a control station on the mainland and three remote island fixed earth stations [6, 147-148]. In addition, C-band and K-band small transportable earth station (STES) includes a 3 m and 2.7 m in diameter Cassegrain antennas, respectively, and makes use FDMA [145-147]. The whole system, excluding the truck, weighs about 6,000 kg. A primary advantage of this system is that it can easily set up emergency or temporary links due to it transportability [124, 146, 149]. This system can only establish point-to-point links between a transportable earth station and a specified fixed earth station. C-band satellite communication links can be made all over the Japanese mainland and the remote islands, but the K-band system is only limited to the mainland because of coverage of the 30/20 GHz satellite beam.

Due to the growing demand of satellite communication services in Japan and also the need for a new test-bed satellite for assisting in the development of higher level communication technologies, the CS-3 series satellites were launched to replace the CS-2 satellites. CS-3a was launched in February 1988 while CS-3b were launched in September 1988 and stationed at 132° and 135° east longitude, respectively. Each CS-3 satellite had a total of 2 C-band and 10 K-band transponders. Launch dates, orbital slots and total number of transponders of various Japanese communications satellites are tabulated below in Table A3.

In April 1985, deregulation of satellite communication systems was introduced to Japan and this permitted private companies to obtain licences to own and operate such systems.

Satellite	Launch Date	Orbit Location (East Longitude)	Transponder/Band
CS	Dec. 1977	135°	C and K
CS-2 (a/b)	a: Feb. 1983 b: Aug. 1983	132° 136°	C: 2 K: 6
CS-3 (a/b)	a: Feb. 1988 b: Sep. 1988	132° 136°	C:2 K: 10
JCSAT (1/2/3)	1: Mar. 1989 2: Jan. 1990 3: Sep. 1995	150° 154° 128°	Ku: 32
SuperBird (A/B)	A1: Jun. 1989 B1: Feb. 1992 A2: Dec. 1992	158° 162° 158°	Ku: 23 K: 2
N-STAR (A/B)	A: Early 1995 B: Mid. 1995	132° 136°	S, C, Ku and K

Table A3: Launch dates, orbital slots and total number of transponders

Two private companies, Japan Communications Satellite Company, Inc. (JCSAT) and Space Communications Corporation (SCC), launched satellites referred to as JCSAT-1 and SCC SuperBird-A, respectively. JCSAT-1 was launched in March 1989 and stationed at 150° east longitude while SCC SuperBird-A was launched in June 1989 and stationed at 158° east longitude. Both satellites carried Ku-band transponders and were used mostly for VSAT networks since the beam widths of these satellites are extremely narrow as they have to cover

only Japan, the G/T and, EIRP are very favorable with values exceeding 10 dB/K and 49 dBW, respectively, which are significantly higher than most other domestic satellite systems around the world [150].

Due to an increased demand of the VSAT service in Japan, in January 1990 JCSAT-2 was launched and stationed at 154° east longitude and this provided JCSAT with a total of 32 Ku-band and 10 K-band transponders. Later that same year, in December, SCC SuperBird-A1 failed due to a propellant leak. Therefore, a second satellite called SuperBird-B1 was launched in February 1992 and stationed at 162° east longitude. Furthermore, a replacement for the SuperBird-A, namely SuperBird-A2 was launched in December of that same year thus, giving SCC access to a total of 23 Ku-band and 2 K-band transponders [150]. In June 1992, there were 21,000 VSATs in operation and that accounted for 70% of all earth stations in Japan. Most of these VSATs have a diameter of 0.75 m [150]. In September 1995, JCSAT-3 was successfully launched and stationed at 128° east longitude.

A domestic mobile satellite communications system was introduced in Japan in 1995 by the launching of N-STAR-A and N-STAR-B in early and mid 1995, respectively. The N-STAR satellites are three axis stabilized and are stationed at the same geostationary slots as the CS-3(a/b). The N-STAR satellites are of the multibeam type and they each produce four fixed beams to cover the Japanese region as illustrated in Figure A6 [151]. The mobile communication service is achieved by using the S-band whereas fixed communications uses C-band, Ku-band and K-band. The C-band is also used for the feeder link (satellite-base station link) for the mobile satellite communication service.

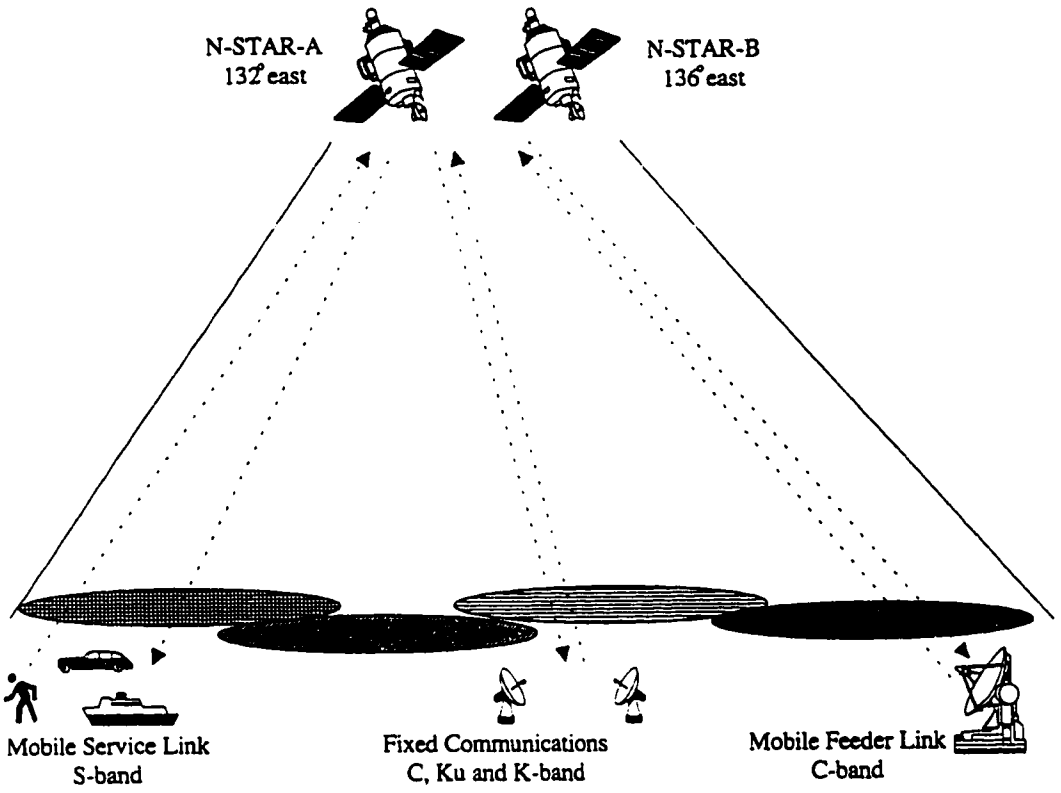


Figure A6: Overview of N-STAR mobile satellite communication system

In Japan, terrestrial analogue and digital cellular systems employs the 800 MHz band and is mostly used in the urban areas where demand for such services is high. The N-STAR mobile satellite system complements the terrestrial cellular system by extending the service to suburban and rural areas where the traffic is low and distributed over a large area. This integration of terrestrial and satellite mobile service has introduced dual-mode user terminals which can be used in both environments but has a preference to connect to the terrestrial system because the service is less costly. Once outside the cellular service area, the user terminal automatically connects to the satellite system [151].

A.2.6 Seychelles

In the past, Seychelles being in the maritime environment, has made extensive use of the Medium Frequency (MF)/ HF/VHF and UHF bands for normal voice and data links to the inner and outer islands. From 1983, the international traffic has been handled two INTELSAT satellites located at 60° and 63° east longitude. Currently, there are about 20,700 domestic telephone lines and this corresponds to a telephone density of around 30 telephones for every 100 inhabitants. This shows that it is feasible to provide telecommunication service to a reasonably high degree for island countries with a relatively small population.

The economy of the Seychelles depends heavily on the tourism industry with hotels, banks, financial institutions, information organizations, and other businesses relying on communications, both national and international. It is for these reasons that the country has tried its utmost to continually improve its telecommunications infrastructure.

Only C-band is being utilized and the two earth stations are owned by the private company Cable & Wireless. A standard B antenna, 9 m in diameter, is located in Bon Espoir at the top of a hill and a standard A antenna, 16 m in diameter, is located in downtown Victoria with both sites on the main island Mahé.

The telephone exchanges in Seychelles are interconnected by either cable including fibre optics, or digital microwave links. RUTEL, which stands for rural telecommunications and developed in Germany, is a system whose primary function is to connect scattered subscribers in rural or nearby islands to a local exchange via digital microwave links, using

dipole antennas. The frequency bands utilized are 2300-2330 MHz and 2400-2430 MHz and this system connects most of the small nearby inhabited islands (inner islands) to a central station on Mahé.

Due to commercial activity and development on the outer islands and the distance involved, VISTA which is a new low density telephony service offered by INTELSAT was introduced. This system employs antennas of 4.5 m in diameter, has now partially replaced the previously used MF/HF radio system and is part of a national telecommunication network. The control station for VISTA is also located on Mahé. Some of its desirable features are: low-cost earth stations; low power consumption; easy installation; operation; and, maintenance. It is interesting to note that the features of the VISTA system mentioned above were specifically introduced to benefit the developing world and thus the reason of its widespread acceptance [152].

The Seychelles telecommunication authorities have been given two satellites slots namely, 37.5° and 42.5° east longitude and are still in the primary stages of examining the possibility of initiating a national domestic satellite system [153]. They are now in the process of performing satellite frequency co-ordination according to ITU procedures with nearby satellites, examining interference and link budget calculations.

A.3 Discussion

The cost of a satellite transmission system is unconstrained by distance. In all of the

above mentioned cases, satellite systems are used only where it was found to be the most cost-effective method of telecommunications.

A.3.1 Next Generation Domestic Satellites

India has made great strides in the field of satellite communications for domestic purposes and search and rescue in the Indian Ocean. COSPAS-SARSAT provides SAR services to the entire world while INSAT-II is limited to the disc of the earth directly below the satellite and uses a geostationary satellite [154]. As far as meteorology is concerned, India maintains continuous meteorological earth observation and is the only system which provides weather observation, high resolution meteorology images, and severe weather/cyclone warnings in the Indian Ocean Region, since India's coastal regions are prone to cyclones. The utilization of INMARSAT will be extended to provide mobile satellite service (MSS) such as aeronautical, maritime and land communications. MSS is currently the fastest growing sector in the satellite communication field and will be available by the end of this decade through INSAT-3 [155].

Mobile data communications service and RDSS are to be introduced in the near future, where the former provides alphanumeric messaging with the use of a pager and the latter provides position/velocity determination, and is essential for navigational purposes [130]. The problem that India is facing now is the congestion of C-band traffic on the INSAT system. India is leasing a C-band transponder from INTELSAT as well [152]. The

expansion of communications has created the need for more data traffic handling, higher data rates, new services, smaller and inexpensive portable ground terminals which in return has made the Indian telecommunications authorities examine the usage of higher frequency bands and higher EIRP [156]. The frequency band which has initiated some interest is the Ku-band and extensive propagation studies, especially rain fades, have been conducted for this frequency band [127]. Furthermore, the earth station equipment cost for the Ku-band is slightly lower than for the C-band, making the former frequency band even more attractive.

In the near future, the digital satellite system, Iridium, will be put into operation by Motorola [8, 131, 154] offering radio-telephone service coverage to the entire world. The Indian government has proposed to set up subsidized solar-power telephone booths in remote villages to compliment their terrestrial telephone infrastructure. In India, from a point of view of cost, messaging services are in great demand and current studies have proposed an LEO satellite system as a cost effective solution. This system will deploy between 18 and 24 small LEO satellites at an altitude of 1000 km in three orbits inclined at 40° - 45° with respect to the equatorial plane. This system will guarantee coverage of the Indian region with a maximum time to forward a message being about 100 minutes [131].

Peru, like most third world countries, leases a fraction of a transponder from INTELSAT for their domestic communication purposes. However, if there is a need for more data traffic handling, these countries can easily lease more bandwidth from INTELSAT.

Indonesia at the present time, is in the planning stage for their third generation

domestic satellite system, which will employ the latest technology in satellite communications namely, three-axis stabilization or also known as body stabilization, 30 high power C-band and four Ku-band transponders, and perhaps spot beams [144]. It will be considered as a regional satellite system because of its extensive coverage of the Asia-Pacific region [141].

A.3.2 Future Trends in Domestic Satellite Communications

In both industrialized and developing countries, economic and social progress depends prominently on improved telecommunications and economic analyses have shown that investment in telecommunications is linked to economic growth. In an era where debts of developing countries have risen to very high levels, the purchase of any new goods must in some way generate a return. No sector is more successful in doing so than telecommunications, and no area of telecommunications holds more promise for the diversity of needs and budgets than satellite communications.

Since many countries are now applying for their satellite slots, the time has come for such countries to investigate the possibility of launching their own domestic satellites custom tailored for their needs. Embarking on such a project of this sort depends on the country's financial resources and the capital return for the services the satellite system will provide. The location of the country also plays an important role. For example a country strategically located, as in the case of Indonesia, could lease excess transponder bandwidth to its

neighbours (ASEAN) [157].

Countries located near the desert regions of the world have the advantage of using the high frequency bands especially the Ku-band and the K-band, because the lack of long rainfall periods. These higher frequency bands are a requirement for today's high data rates.

Tropical countries have the worst scenario as they have a rainy season for up to six months a year. Therefore, they encounter severe rain fade problems in the high frequencies and are forced to use C-band and S-band [2, 128]. Frequencies above 10 GHz are affected by various forms of hydrometeors including rain, snow, hail, fog, cloud and ice. The most serious impairment to communication systems is caused by rain and this is due mainly to its large dielectric constant which produces a high displacement current and hence absorption. Unfortunately, the present ITU-R rain fade prediction model is not as accurate in tropical regions as it has been found to be in temperate regions [121-123, 158]. It has been shown that the model actually underestimates the rain fade [118-123, 128]. In the design of a reliable and economical domestic satellite communication system, fairly accurate estimates of rain fade are required. Hence there is a need for accurate information on the propagation characteristics for frequencies above 10 GHz which can only be obtained by performing on-site rain measurement experiment. It has also been proposed that inter-satellite links may be use as a means of avoiding rain fades in a communication system [158].

For remote island operation, important factors include: reliability; simple installation; compactness (small and light weight); high performance; low power consumption; intelligent equipment, and cost effective terminals that are maintenance free or require little

maintenance; and, invulnerability to natural disasters. The capability of setting up of temporary links in the case of natural disasters using portable terminals offers improvements in overall communication network reliability [159].

Some of the leading trends in the satellite communication field include higher space segment capacity, longer satellite orbit life-time, increased antenna pointing accuracy, and higher transmitted power. Table A4 shows the trend of INTELSAT satellite characteristics [160].

	INTELSAT V/VA	INTELSAT VI	INTELSAT VII
C-band Maximum EIRP (dBW)	29.0	31.0	36.5
Ku-band Maximum EIRP (dBW)	44.4	44.7	46.7
Orbital Life-Time (Years)	7	13	15

Table A4: Trend in INTELSAT satellite characteristics

Notice the trend in increased satellite maximum EIRP and satellite orbit life-time. Coverage by global beam areas such as vast ocean surfaces having a sparse user population requiring thin-route telecommunication service results in a wastage of satellite radiated power. Conversely, in making use of shaped antenna spot beams to cover the areas with dense user population will make efficient use of the satellite radiated power. Conventional EIRP of global beam of today's modern satellites is around 36.5 dBW (INTELSAT VII) but C-band spot beams give an EIRP in excess of 40 dBW. Similarly, Ku-band spot beams can achieve more than 55 dBW [154]. With these typical high EIRP levels, the use of even smaller user terminals has become possible.

Japan is currently the leader in research and development of advanced technology commercial domestic communication satellites and has launched several advanced experimental satellites rather than using those of today's production capabilities [161-162]. Three experimental communication satellites projects, currently under way, are the engineering test satellites (ETS-V and ETS-VI) and communication and broadcasting engineering test satellite (COMETS). ETS-V was launched in August 1987 at 150° east longitude and is being used as a test bed to provide mobile satellite communications for ships, aircraft and land vehicles. L-band provides the service link whereas C-band is used for the feeder link. Furthermore, ETS-V delivers educational programs and medical treatment communications to Asia-Pacific countries through low-cost user terminals.

ETS-VI extends the mobile communications to handheld user terminals and portable VSATs in the S-band (2.3/2.1 GHz), and provides high quality satellite broadcasting services such as high definition television (HDTV) by using K-band multibeams. The millimetre wave bands (43/38 GHz) have been selected by considering atmospheric attenuation allowable in personal satellite communications and availability of millimetre wave devices. This frequency band will be exploited to develop high data rate inter-satellite link technology and to investigate its usage as the new frequency band for PCS. Experimental inter-satellite link for tracking, and data relay between LEO satellites and earth stations are being done using the S-band and optical satellite communications.

COMETS is scheduled for launch in February 1997 and is the most ambitious Japanese satellite project. The purpose of COMETS is to perform further tests on HDTV and

PCS service using the K-band, inter-satellite links using S-band and K-band and fixed satellite service using millimetre wave bands (46.9/43.8 GHz) [163].

The role of experimental satellites in the evolution of operational domestic systems is well established and requires a reasonable amount of resources. A good test bed in the form of an experimental satellite and earth station as a control facility can be used to demonstrate the hardware and concepts of the systems for trying out new services. India and Japan have undertaken this route and many others are following [154].

The trend over the last decade clearly has been towards larger, more powerful and thus higher EIRP satellites in geostationary orbit but developing countries are interested in more economical systems. There is some thought that small, cost-effective, reliable, high quality satellites could play a major role in future domestic satellite systems. Certainly small LEO satellite systems can achieve the same high EIRP levels as bigger geostationary satellites therefore, permitting the use of low power terminal. This is one of the major reasons for the proposal of several LEO systems [8, 131, 154].

A.4 Conclusions

This chapter examines domestic satellite communication systems for six different countries: India; Peru; South Africa; Indonesia; Japan; and, Seychelles. Their needs range from domestic and international voice and data circuits to DBS TV, meteorological services, SAR, and mobile radio links. Fibre optics communications physically ties the user to the

system, but combining satellite, terrestrial, and fibre optics as a hybrid system, is a better suited option to fully digitalize communication systems in developing nations, since they compliment each other. Depending on the type and quantity of services requirements of these countries, multi-purpose domestic satellite communication implementations include customized satellites as in the case of India, Indonesia and Japan, to leasing transponder bandwidth from INTELSAT.

The requirement for developing counties is thin-route communications, but with the ever increasing traffic demands and the introduction of new services, one has to compare the cost-effectiveness of either having a national satellite or simply leasing more transponder bandwidth. For those countries having their own domestic satellite systems, future generation satellite will have to be designed primarily to handle more capacity and offer new services such as DBS HDTV, MSS, PCS, RDSS and SAR. Countries without a domestic satellite can request a slot on the Clarke Orbit from ITU and look into the possibility of launching their own satellites instead of leasing transponders from INTELSAT. In instances where a particular country has no more satellite slots available, collocating two satellites in one slot is a technique for increasing the capability to handle more traffic or to introduce new services.

When the traffic grows, either to an increase in traffic density and/or number of earth stations, domestic satellite systems employing PA FDMA-SCPC and may consider re-equipping with DA FDMA-SCPA. The only cost involved will be on re-equipping since there is no increase in the cost of leasing the same space segment.

It clearly stands out that the completely exhausted MF, HF, VHF and UHF bands in the terrestrial systems and the congestion of C-band, has caused many island countries to look into the possibility of using higher frequency bands, namely Ku-band, K-band and also millimetre wave band. The benefits of these frequency bands are considerable reduction in size of the earth stations or portable terminals, easier realization of higher EIRP spot beams, and the ability to handle higher data rates [156].

In the case of tropical islands, frequencies in K-band and higher will suffer from severe rain fades. Since rain fade is an important parameter in the design of a reliable domestic satellite system, several techniques have being developed to combat it.

APPENDIX B

Trigonometric Formulas for Spherical Triangles

B.1 Spherical Triangles

For any spherical triangle ABC such as the one depicted in Figure B.1, whose side lengths a, b and c are measured by the circle arcs subtended at the center of the sphere we have the sine law given by

$$\frac{\sin(A)}{\sin(a)} = \frac{\sin(B)}{\sin(b)} + \frac{\sin(C)}{\sin(c)} \qquad \textbf{(B.1)}$$

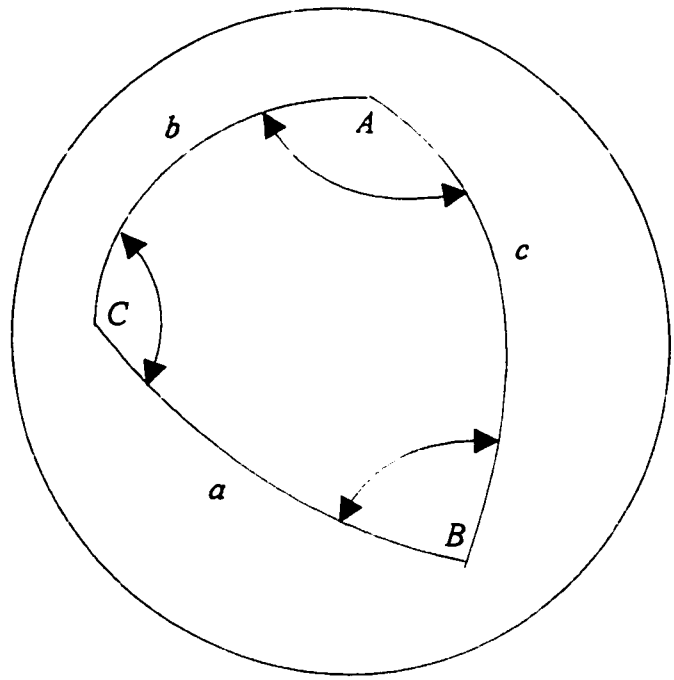


Figure B.1: Spherical triangles

The cosine law for the sides are given by

$$\cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A) \tag{B.2}$$

and the cosine law for the angles are

$$\cos(A) = -\cos(B)\cos(C) + \sin(B)\sin(C)\cos(a) \tag{B.3}$$

APPENDIX C

Doppler Estimation Using Satellite and User Location

C.1 Practical Method of Doppler Estimation

The slant range is a measure of distance of the satellite from the user terminal. In Figure C.1 we consider the satellite location as (X_s, Y_s, Z_s) and user terminal location as (X_u, Y_u, Z_u). The coordinate origin coincides with the center of the earth. The slant range s, using this coordinate system is given by

$$s = \sqrt{(X_s - X_u)^2 + (Y_s - Y_u)^2 + (Z_s - Z_u)^2} \qquad \textbf{(C.1)}$$

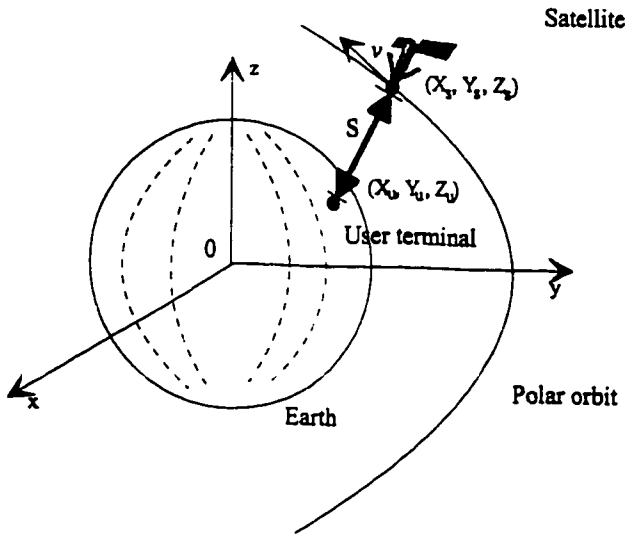


Figure C.1: Satellite slant range estimation

Therefore, knowing the location of the satellite and user terminal allows us to calculate the slant range. Satellite location can be found by using a real time model for the constellation of Iridium satellites. Conversely, the user location is roughly known to within a city block.

It has been shown earlier in chapter 3, that for the case of an overhead pass satellite, the slant range s is given by (3.2). Rewriting (3.2) to obtain the coverage angle Φ we have

$$\Phi = \cos^{-1}\left[\frac{R_e^2 + (R_e + h)^2 - s^2}{2R_e(R_e + h)}\right] \qquad \text{(C.2)}$$

Making a small addition to the expression (2.2) we obtain

$$\frac{R_e + h}{\sin(\gamma + 90^o)} = \frac{R_e}{\sin(\delta)} = \frac{s}{\sin(\Phi)} \qquad \text{(C.3)}$$

Now rewriting the Doppler shift given by (3.10) into the form of (2.3) but for the overhead pass satellite case, we get

$$f_d = \frac{|v|\, f_c \cos(\Phi + \gamma)}{c} = \frac{|v|\, f_c \sin(\delta)}{c} \tag{C.4}$$

Using the latter part of (C.3) and (C.4) we obtain

$$f_d = \frac{|v|\, f_c\, R_e \sin(\Phi)}{c\, s} \tag{C.5}$$

Substituting (C.2) into (C.5) we obtain the estimate of Doppler shift using the satellite's altitude, slant range, and speed, earth's radius, and carrier frequency.

$$f_d = \frac{|v|\, f_c\, R_e}{c\, s} \sin\left(\cos^{-1}\left[\frac{R_e^2 + (R_e + h)^2 - s^2}{2R_e(R_e + h)} \right] \right) \tag{C.6}$$

Note that here we are not estimating the Doppler curve but rather the actual Doppler shift introduced by the relative radial velocity of the satellite with respect to the user terminal, which is both satellite and user terminal location dependent. For satellites at a cross track angle, the longitude of the satellite orbit at that instant in time is required.

C.2 Effects of Satellite Location Uncertainty

Consider an Iridium overhead pass satellite as shown in Figure C.2 at position P and at elevation angle of $\gamma = 45°$, carrier frequency $f_c = 1.618$ GHz, satellite altitude $h = 780.6$ km, earth's radius $R_e = 6377$ km, satellite absolute velocity $|v| = 7463$ m/s, speed of light c

= 3E8 m/s.

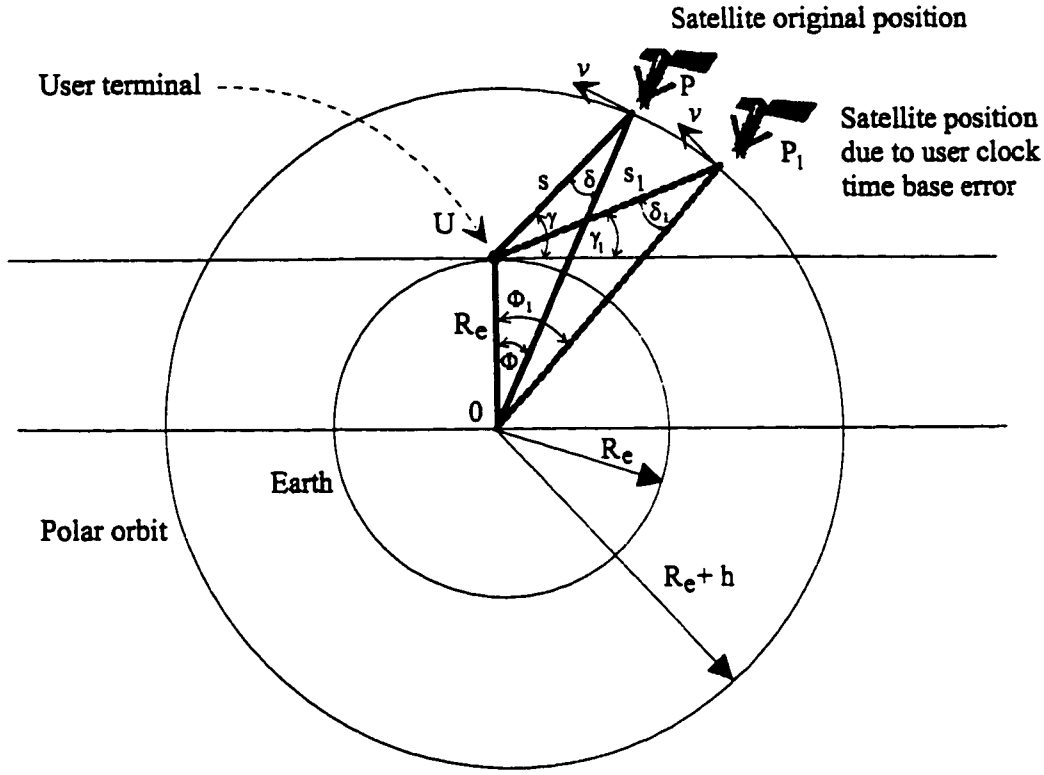


Figure C.2: Geometry of satellite location uncertainty

Applying cosine rule to the triangle ΔOUP we obtain

$$(R_e + h)^2 = R_e^2 + s^2 - 2R_e s \sin(\gamma) \qquad \text{(C.7)}$$

Rearranging (C.7) into standard quadratic form we get

$$s^2 - s2R_e\sin(\gamma) - 2R_e h - h^2 = 0 \qquad \text{(C.8)}$$

The satellite slant range can be found by solving the quadratic expression for s as follows

$$s = -R_e\sin(\gamma) + \sqrt{R_e^2 \sin^2(\gamma) + 2R_e h + h^2} \qquad \text{(C.9)}$$

The coverage angle can be found by using (C.2) and the Doppler shift by (C.5) or (C.6). Hence, we obtain s = 1049 km, Φ = 5.95°, and the nominal estimated Doppler shift f_d = 25.358 kHz.

Assume that the user terminal clock time base lags by ΔT = 1 μs. Then, the new coverage angle is given by (angles in radians)

$$\Phi_1 = \Phi + \frac{|v|\,\Delta T}{(R_e + h)} \qquad \textbf{(C.10)}$$

This gives us a new coverage angle Φ_1 = 5.95°. Clearly an error in the user terminal clock time base does not contribute any significant error into the location of the satellite. Therefore, the satellite slant range as well as the estimated Doppler shift are not affected.

C.3 Effects of User Location Uncertainty

If a user terminal location has an uncertainty of ± A meters, then the satellite coverage angle, as shown in Figure C.3 is given by (angles in radians)

$$\Phi_2 = \Phi - \frac{A}{R_e} \qquad \textbf{(C.11)}$$

We first consider A = 10 km and we find that Φ_2 = 5.86°, s_2 = 1042 km, and $f_d(2)$ = 25.145 kHz. Therefore, the deviation from the nominal estimated Doppler shift is $\Delta f_d(2)$ = 213 Hz.

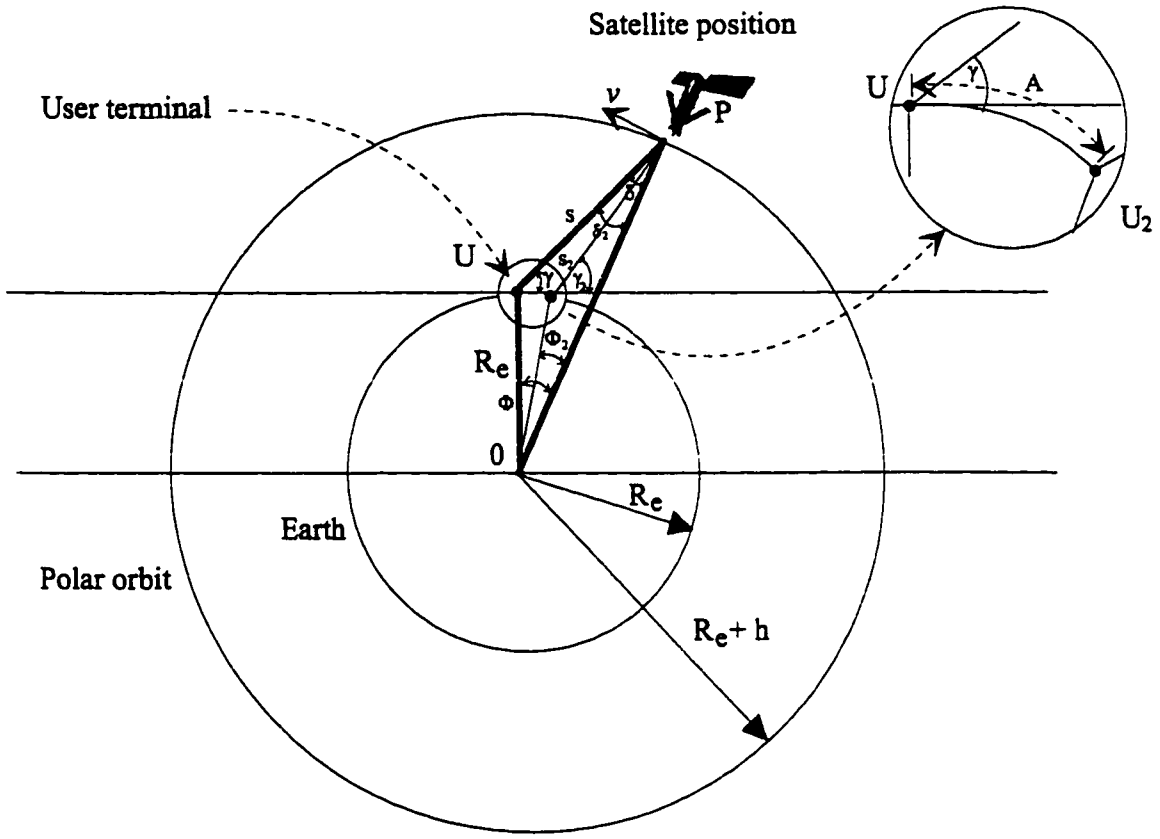


Figure C.3: Geometry of user terminal uncertainty

In the second case we let A = 100 km and obtain $\Phi_3 = 5.05°$, $s_3 = 982$ km, and $f_d(3) = 23.022$ kHz. Therefore, the deviation from the nominal estimated Doppler shift is $\Delta f_d(3) = 2.3$ kHz. Therefore, an uncertainty of ±10 km in the location of the user terminal produced a frequency deviation of around 200 Hz in the estimated Doppler shift. Similarly, for an uncertainty in user location of ±100 km, the frequency deviation increased to roughly 2.3 kHz.

In this appendix it has been shown that having knowledge of the satellite location via a real time model for satellite constellation and the user terminal location within a city block, allows us to calculate the nominal Doppler shift. Sensitivity analysis has shown that user terminal clock time base error does not contribute a significant error into the estimate of the

Doppler shift. However, accuracy of user location of ±10 km contributes a deviation of 200 Hz in the estimated Doppler shift and accuracy of user location of ±100 km, contributes a deviation of roughly 2.3 kHz. In conclusion, this method of Doppler shift estimation may be worthwhile considering as long as the methods of obtaining the location of the satellite and user terminal require very little processing overhead.

BIBLIOGRAPHY

[1] H. Chasia., "Satellites and Developing Countries," *IEEE Comm. Magazine*, vol. 33, no. 9, pp 110-112, September 1995.

[2] "Planning Rural Telecommunication Services," *Philips Telecommunications Review*, vol. 41, no. 2, pp 106-112, June 1983.

[3] R. Crowell, and R. Blois., "Satellite Telecommunications: Applications and Issues at the Regional and Global Levels," *IEEE Comm. Magazine*, vol. 33, no. 9, pp 108-109, September 1995.

[4] G. Maral, and M. Bousquet., *Satellite Communications Systems*, John Wiley & Sons, New York, USA, 1986.

[5] L. R. Erasmus., "Involvement of the Republic of South Africa in Satellite Communications," *Elektron*, (SAIEE Department of Posts and Telecommunications), June 1985.

[6] H. Shiota, and Y. Watanabe., "TDMA System for C-Band Satellite Communication with Remote Islands," *Review of the Electrical Communication Laboratories*, vol. 26, no. 1-2, pp 80-90, February 1978.

[7] R. J. Leopold, and A. Miller., "The Iridium™ Communications System," *1993 IEEE MTT-S International Microwave Symposium*, (Atlanta, GA, USA), pp 575-578, 14-18 June 1993.

[8] R. J. Leopold, and A. Miller., "The Iridium™ Communications," *IEEE Potentials*, vol. 12, no. 2, pp 6-9, April 1993.

[9] R. J. Leopold, A. Miller, and J. L. Grubb., "The Iridium Systems: A New Paradigm In Personal Communications," *Applied Microwave & Wireless*, vol. 5, no. 4, pp 68-76, Fall 1993.

[10] D. E. Sterling, and J. E. Hatlelid., "The Iridium™ System - A Revolutionary Satellite Communications System Developed with Innovative Applications of Technology," *IEEE, MILCOM'91*, pp 436-440, 1991.

[11] L. S. Kolsky, and P. L. Malet., "Application of Motorola Satellite Communications, Inc.," *FCC Application for the Iridium Low Earth Orbit Mobile Satellite System*, pp 1-119, December 1990.

[12] R. J. Leopold., "Low-Earth Orbit Global Cellular Communication Network," presented at the *Mobile Satellite Communication Conference*, (Aldelaide, Australia), pp 59-65, 23-24 August 1990.

[13] R. H. Williams., "Iridium Offers Contact to Any Point on Earth," *Signal*, vol. 45, no.6, pp 95-97, February 1991.

[14] G. R. McIntosh., "Iridium: A Global Satellite Cellular Network," *Telecommunication Journal of Australia*, vol. 41, no. 2, pp 63-66, 1991.

[15] R. J. Leopold., 'TUANZ'92: Communications for Competitive Advantage," *Conference and Trade Exhibition*, (Motorola Satellite Communications, 2501 South Price Road, Chandler, AZ 85248).

[16] P. A. Swan, and P. N. Cloutier., "Global Personal Communications this Decade with Iridium™," *Space Technology*, vol. 13, no. 4, pp 423-425, July 1993.

[17] D. C. Beste., "Design of Satellite Constellations for Optimal Continuous Coverage," *IEEE Trans. Aerospace and Elect. Sys.*, vol. AES-14, no. 3, pp 466-473, May 1978.

[18] F. Vatalaro, G. E. Corazza, C. Caini, and C. Ferrarelli., "Analysis of LEO, MEO, and GEO Global Mobile Satellite Systems in the Presence of Interference and Fading," *IEEE Journal on Select Areas Comm.*, vol. 13, no. 2, pp 291-300, February 1995.

[19] A. H. Ballard., "Rosette Constellations of Earth Satellites," *IEEE Trans. Aerospace and Elect. Sys.*, vol. AES-16, no. 5, pp 656-673, September 1980.

[20] J. Radzik, and G. Maral., "A Methodology for Rapidly Evaluating the Performance of Some Low Earth Orbit Satellite Systems," *IEEE Journal on Select Areas Comm.*, vol. 13, no. 2, pp 301-309, February 1995.

[21] J. Kaniyil, J. Takei, S. Shimamoto, Y. Onozato, T. Usui, I. Oka, and T. Kawabata., "A Global Message Network Employing Low Earth-Orbiting Satellites," *IEEE Journal Select Areas Comm.*, vol. 10, no. 2, pp 418-427, February 1992.

[22] D. I. Dalgleish., *An Introduction to Satellite Communications*, Peter Peregrinus Ltd., London, United Kingdom, 1989.

[23] M. Borota, R. J. Leopold, K. Johnson, and A. Miller., 'The Iridium™/SM1 Personal

Communication System," *Proceedings of 1994 Symposium VLSI Circuits*, (Honolulu, HI, USA), pp 1-4, 9-11 June 1994.

[24] J. J. Schuss, T. Carlson, R. Francois, P. Maloney, A. Rohwer, J. Upton, and L. Wardle., "Design of the Iridium Phased Array Antennas," *Proceedings of IEEE Antennas and Propagation Society International Symposium*, (Ann Arbor, MI, USA), pp 218-221, 28 June - 2 July 1993.

[25] J. L. Grubb., "The Traveller's Dream Come True," *IEEE Comm. Magazine*, vol. 29, pp 48-51, November 1991.

[26] J. Litva., *Electrical Engineering 742 Graduate course notes*, 1994.

[27] B. Bertiger., "Iridium - A Global Personal Communications System," pp 331-333, Motorola Inc., (Satellite Communications 2501 S. Price Road, Chandler, AZ 85248-2899).

[28] J. E. Hatlelid, and L. Casey., "The Iridium™ System Personal Communications Anytime, Anyplace," *Proceedings of Third International Mobile Satellite Conference*, IMSC'93 (Pasedena, California, USA), pp 285-290, 16-18 June 1993.

[29] F. Xiong., "Modem Techniques in Satellite Communications," *IEEE Comm. Magazine*, vol. 32, pp 84-98, August 1994.

[30] N. Levanon., *Radar Principles*, John Wiley & Sons, New York, USA, pp 1-18, 1988.

[31] M. Nelkon, and P. Parker., *Advance Level Physics*, Heinemann Educational Books, Hong Kong, pp 627-639, 1970.

[32] R. J. Leopold., "Low-Earth Orbiting Satellite System," Private communication, pp

287-299.

[33] W. L. Pritchard, and J. A. Sciulli., *Satellite Communication Systems Engineering*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1986.

[34] B. R. Bertiger, P. A. Swan, and R. J. Leopold., "Iridium is in the Works," *Aerospace America*, pp 40-42, February 1991.

[35] S. Haykin., *Adaptive Filtering Theory*, Second Edition, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1991.

[36] P. P. Ginsto., "Major Alternatives for a Global Non Geostationary Satellite Mobile Network," *CSELT Technical Reports*, vol. 22, pp 1-14, 1994.

[37] J. Wood., *Satellite Communications and DBS Systems*, Focal Press, Oxford, United Kingdom, 1992.

[38] L. W. Couch II., *Digital and Analog Communication Systems*, Fourth Edition, Macmillan Publishing Company, New York, USA, 1993.

[39] J. N. Pelton, and T. Iida., "21st Century Satellite Options for the Asia-Pacific: A Cooperative U.S.-Japanese Study," in *Proc. Pacific Telecommunications Council Fourteenth Annual Conference*, PTC'92 (Honolulu, HI, USA), pp 801-814, 12-15 January 1992.

[40] R. Frieden., "Assessing the Scope of Global Personal Communication Services," *Proceedings of 16th Annual Pacific Telecommunications Conference* (Honolulu, HI, USA), pp 357-362, 16-20 January 1994.

[41] A. J. Viterbi., "The Evolution of Digital Wireless Technology from Space

Exploration to Personal Communication Services," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp 638-644, August 1994.

[42] D. L. Schilling., "Wireless Communication Going Into the 21st Century," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp 645-652, August 1994.

[43] A. Swift., "Canadian Licenses Satellite Cellular Phone Project," *The Hamilton Spectator.*, vol. Business, pp C6, 23 July 1997.

[44] D. M. Jansky, and M. C. Jeruchim., *Communication Satellites in the Geostationary Orbit*, Artech House, Massachusetts, USA, pp 529-532, 1987.

[45] V. K. Bhargava, D. Haccoun, R. Matyas, and P. P. Nuspl., *Digital Communications by Satellite*, John Wiley & Sons, Washington D.C., USA, 1981.

[46] Tri T. Ha., *Digital Satellite Communications*, Macmillan Publishing Company, New York, USA, 1986.

[47] J. J. Spilker., *Digital Communications by Satellite*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1977.

[48] R. M. Gagliardi., *Satellite Communications*, Lifetime Learning Publications, Belmont, California, USA, 1984.

[49] W. L. Pritchard, H. G. Suyderhoud, and R. A. Nelson., *Satellite Communication Systems Engineering*, Second Edition, Prentice Hall, Englewood Cliffs, New Jersey, USA, pp 97-117, 1993.

[50] M. J. Miller, B. Vucetic, and L. Berry., *Satellite Communications (Mobile and Fixed Services)*, Kluwer Academic Publishers, Norwell, Massachusetts, USA, 1993.

[51] T. Pratt, and C. W. Bostian., *Satellite Communications*, John Wiley & Sons, New York, USA, pp 22-30, 1986.

[52] M. S. El-Hennawey, and C. R. Carter., "Total Information Processor for 121.5/243 MHz SARSAT Signals," *IEEE Trans. Aerospace and Elect. Sys.*, vol. AES-29, no. 3, pp 866-877, July 1993.

[53] A. P. Clark, A. Aftelak, M. Savjani, and C. Ang., "Doppler-Shift Correction by the DFT, for a QPSK Satellite Modem," *International Journal of Satellite Communications*, vol. 8, No. 5, pp 349-361, 1990.

[54] A. P. Clark, and A. Aftelak., "Doppler-Shift Correction for a UQPSK Satellite Modem," *International Journal of Satellite Communications*, vol. 8, no. 5, pp 337-347, September 1990.

[55] Y. Uno, M. Katayama, T. Yamazato, and A. Ogawa., "Carrier Regeneration in a Block Demodulator for Low Earth-Orbital Satellite Communication System," *The Fourth International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'93)*, (Yokohama, Japan), pp 458-462, 8-11 September 1993.

[56] D. K. Van Alphen, and W. C. Linsey., "Higher-Order Differential Modulation Techniques," *Proceedings of MILCOM'95*, (San Diego, CA, USA), vol. 3, pp 1008-1012, 5-8 November 1995.

[57] M. Katayama, A. Ogawa, and N. Morinaga., "Carrier Synchronization Under Doppler Shift of the Nongeostationary Satellite Communcation Systems," *Communications*

on the Move (ICCS/ISITA'92), (Singapore), vol. 2, pp 466-470, 16-20 November 1992.

[58] M. Ise, M. Katayama, T. Yamazato, and A. Ogawa., "A Block Demodulator for Low Earth-Orbit Satellite Communication Systems with Large Doppler Shift," *VTC96*.

[59] A. Kariwara., "Performance of Dual-Channel Demodulator for LEO Satellite CDMA Communications," *Proceedings of ICC'95*, (Seattle, WA, USA), vol. 3, pp 1875-1879, 18-22 June 1995.

[60] D. E. Dodds, and M. Moher., "Spread Spectrum Synchronization for LEO Personal Communications Satellite System," *Proceedings of the 1995 Canadian Conference on Electrical and Computer Engineering, (CCGEI'95)*, (Montreal, Canada), vol. 1, pp 20-23, 1995.

[61] S. G. Glisic, T. Poutanen, and W. W. Wu., "New PN Code Acquisition Schemes for Low Earth Orbit Satellites," *Proceedings of IEEE 3rd International Symposium on Spread Spectrum Techniques and Applications, (ISSSTA'94)*, (Oulu, Finland), vol. 2, pp 608-613, 4-6 July 1994.

[62] A. P. Clark, and A. Aftelak., "Carrier Phase Synchronization in the Demodulation of UQPSK Signals," *IEE Proceedings (I Communications, Speech and Vision)*, vol. 136, no. 5, pp 351-360, October 1989.

[63] D. A. Jiraud., "Performance of Noncoherent FSK in a LEO Satellite Channel," *Proceedings of SPIE (Int. Soc. Opt. Eng.)*, (Philadelphia, PA, USA), vol. 2601, pp 115-122, 23-25 October 1995.

[64] P. G. Flikkema, and L. D. Davidson., "Performance Anaylsis of Spread Spectrum Acquisition Algorithm for Satellite Mobile Radio," *Proceedings of MILCOM'92*, vol. 1, pp 115-119, 11-14 October 1992.

[65] S. Haykin., *Communication Systems*, Second Edition, John Wiley & Sons, New York, USA, 1983.

[66] S. Benedetto, E. Biglieri, and V. Castellani., *Digital Transmission Theory*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1987.

[67] L. W. Couch., *Digital and Analog Communication Systems*, second Edition, Macmillan Publishing Company, New York, USA, 1987.

[68] S. Haykin., *Digital Communications*, John Wiley & Sons, New York, USA, 1988.

[69] S. Haykin., *An Introduction to Analog and Digital Communications*, John Wiley & Sons, New York, USA, 1989.

[70] E. Jones., *Digital Transmission*, McGraw-Hill Book Company, London, United Kingdon, 1993.

[71] J. Martin., *Communications Satellite Systems*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1978.

[72] W. C. Linsey, and M. K. Simon., *Telecommunications Systems Engineering*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1973.

[73] H. B. Killen., *Digital Communications with Fibre Optics and Satellite Applications*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 1988.

[74] R. Haeb, and H. Meyr., "A Systematic Approach to Carrier Recovery and Detection

of Digitally Phase Modulated Signals on Fading Channels," *IEEE Trans. on Comm.*, vol. 37, no. 7, pp 748-754, July 1989.

[75] M. Andronico, S. Casale, A. La Corte, and C. Pinnisi., "A New Algorithm for Fast Synchronization in Burst Mode PSK Demodulator," *Proceedings of ICC'95* (Seattle, WA, USA), vol. 3, pp 1641-1646, 18-22 June 1995.

[76] J. E. Ohlson, and R. J. Huff., "Multi-Frequency TDMA for Satellite Communications," *Proceedings ICC'83*, pp 892-896, 1983.

[77] U. Mengali, and M. Morelli., "Data-Aided Frequency Estimation for Burst Digital Transmission," *IEEE Trans. on Comm.*, vol. 45, no. 1, pp 23-25, January 1997.

[78] S. Bellini, C. Molinari, and G. Tartara., "Digital Frequency Estimation in Burst Mode QPSK Transmission," *IEEE Trans. on Comm.*, vol. 38, no. 7, pp 959-961, July 1990.

[79] M. Luise, and R. Reggiannini., "Carrier Frequency Recovery in All-Digital Modems for Burst-Mode Transmissions," *IEEE Trans. on Comm.*, vol. 43, no. 2/3/4, pp 1169-1178, February/March/April 1995.

[80] S. Hinedi, and W. C. Linsey., "Intersymbol Interference Effects on BPSK and QPSK Carrier Tracking Loops," *IEEE Trans. on Comm.*, vol. 38, no. 10, pp 1670-1676, October 1990.

[81] S. T. Kleinberg, and H. Chang., "Sideband False-Lock Performance of Squaring, Fourth-Power, and Quadriphase Costas Loops for NRZ Data Signals," *IEEE Trans. on Comm.*, vol. COM-28, no. 8, pp 1335-1345, August 1980.

[82] C. L. Weber, and W. K. Alem., "Demod-remod Coherent Tracking Receiver for

QPSK and SQPSK," *IEEE Trans. on Comm.*, vol. COM-28, no. 12, pp 1945-1949, December 1980.

[83] A. J. Viterbi, and A. M. Viterbi., "Nonlinear Estimation of PSK-Modulated Carrier Phase with Application to Burst Digital Transmission," *IEEE Trans. on Information Theory*, vol. IT-29, no. 4, pp 543-551, July 1983.

[84] C. R. Chan., "Improving Frequency Acquisition of a Costas Loop," *IEEE Trans. on Comm.*, vol. COM-25, no. 12, pp 1453-1459, December 1977.

[85] F. M. Gardner., "Properties of Frequency Difference Detectors," *IEEE Trans. on Comm.*, vol. COM-33, no. 2, pp 131-138, February 1985.

[86] H. Sari, and S. Moridi., "New Phase and Frequency Detectors for Carrier Recovery in PSK and QAM Systems," *IEEE Trans. on Comm.*, vol. COM-36, no. 9, pp 1035-1043, September 1988.

[87] D. G. Messerschmitt., "Frequency Detectors for PLL Acquisition in Timing and Carrier Recovery," *IEEE Trans. on Comm.*, vol. COM-27, no. 9, pp 1288-1295, September 1979.

[88] G. Karam, F. Daffara, and H. Sari., "Simplified Versions of the Maximum Likelihood Frequency Detector," *Proceedings of GLOBECOM'92*, (Orlando, FL, USA), vol. 1, pp 345-349, 6-9 December 1992.

[89] H. J. Clayton, P. Fines, and A. H. Aghvami., "Compariosn of Frequency Detectors for Low Earth Orbit Satellite Communication Systems," *Communications Research Group*, Electronic and Electrical Engineering Department, King's College, London.

[90] A. N. D'Andrea, and U. Mengali., "Performance of a Frequency Detector Based on the Maximum Likelihood Principle," *Proceedings of GLOBECOM'92*, (Orlando, FL, USA), vol. 1, pp 340-344, 6-9 December 1992.

[91] T. Alberty, and V. Hespelt., "A New Pattern Jitter Free Frequency Error Detector," *IEEE Trans. on Comm.*, vol. COM-37, no. 2, pp 159-163, February 1989.

[92] A. N. D'Andrea, and U. Mengali., "Performance of a Quadricorrelator Driven by Modulated Signals," *IEEE Trans. on Comm.*, vol. COM-38, no. 11, pp 1952-1957, November 1990.

[93] H. Sari, H. Houeix, and G. Karam., "Phase and Frequency Detectors for Clock Synchronization in High-Speed Optical Transmission Systems," *Proceedings of GLOBECOM'92*, (Orlando, FL, USA), vol. 1, pp 370-374, 6-9 December 1992.

[94] G. De Jonghe, and M. Moeneclaey., "Cycle Slipping Behaviour of NDA Feedforward Carrier Synchronizers for Time-Varying Frequency Non-Selective Fading Channels," *Proceedings of GLOBECOM'92*, (Orlando, FL, USA), vol. 1, pp 350-354, 6-9 December 1992.

[95] M. H. Meyers, and L. E. Franks., "Joint Carrier Phase and Symbol Timing Recovery for PAM Systems," *IEEE Trans. on Comm.*, vol. COM-28, no. 8, pp 1121-1129, August 1980.

[96] M. K. Simon, and D. Divsalar., "Maximum-Likelihood Block Detection of Noncoherent Continuous Phase Modulation," *IEEE Trans. on Comm.*, vol. COM-41, no. 1, pp 90-98, January 1993.

[97] L. E. Franks., "Carrier and Bit Synchronization in Data Communication - A Tutorial Review," *IEEE Trans. on Comm.*, vol. COM-28, no. 8, pp 1107-1121, August 1980.

[98] H. L. Van Trees., *Detection, Estimation, and Linear Modulation Theory*, John Wiley & Sons, New York, USA, 1968.

[99] J. G. Proakis., *Digital Communications*, Second Edition, McGraw-Hill Book Company, New York, USA, 1989.

[100] R. De Gaudenzi, V. Vanghi, and T. Garde., "Feed-Forward Carrier Phase Synchronizers for M-PSK Signals," *Proceedings of ICC'95*, (Seattle, WA, USA), vol. 1, pp 1055-1059, 18-22 June 1995.

[101] K. Goethals, and M. Moeneclaey., "ML Oriented NDA Symbol Synchronizers for Time-Varying Non-Selective Fading Channels," *Proceedings of GLOBECOM'92*, (Orlando, FL, USA), vol. 1, pp 355-359, 6-9 December.

[102] W. G. Cowley., "Phase Estimation for PSK Packets: Bounds and Algorithms," *IEEE Trans. on Comm.*, vol. COM-44, no. 1, pp 26-28, January 1996.

[103] T. Pollet, and M. Moeneclaey., "Effect of Carrier Phase Jitter on Single-Carrier and Multi-Carrier QAM Systems," *Proceedings of ICC'95*, (Seattle, WA, USA), vol. 1, pp 1046-1050, 18-22 June 1995.

[104] G. B. Thomas Jr., and R. L Finney, Calculus and Analytic Geometry, Seventh Edition, Addison-Wesley Publishing Company, New York, USA, 1988.

[105] H. Meyr, and G. Ascheid., Synchronization in Digitals Communications, vol. 1, John Wiley & Sons, New York, USA, 1983.

[106] P. Y. Kam., "Maximum Likelihood Carrier Phase Recovery for Linear Suppressed-Carrier Digital Data Modulations," *IEEE Trans. on Comm.*, vol. COM-34, no. 6, pp 522-527, June 1986.

[107] B. E. Paden., "A Matched Nonlinearity for Phase Estimation of a PSK-Modulated Carrier," *IEEE Trans. on Information theory.*, vol. IT-32, no. 3, pp 419-422, May 1986.

[108] M. P. Fitz., "Planar Filtered Techniques for Burst Mode Carrier Synchronization," *Proceedings of GLOBECOM'91*, vol. 1, pp 365-369, 1991.

[109] M. K. Simon., "On the Optimality of the MAP Estimation Loop for Carrier Phase Tracking BPSK and QPSK Signals," *IEEE Trans. on Comm.*, vol. COM-27, no. 1, pp 158-165, January 1979.

[110] S. Haykin., *Communication Systems*, Third Edition, John Wiley & Sons, New York, USA, 1994.

[111] J. W. Bandler., *Computer Engineering 3KB3 Course Notes*, McMaster University, 1991.

[112] A. Grace., *Optimization ToolBox*, The MathWorks Inc., December 1992.

[113] F. D. Natali, "AFC Tracking Algorithms," *IEEE Trans. on Comm.*, vol. COM-32, no. 8, pp 935-947, August 1984.

[114] A. N. D'Andrea, and U. Mengali, "Design of Quadricorrelators for Automatic Frequency Control Systems," *IEEE Trans. on Comm.*, vol. COM-41, no. 6, pp 988-997, June 1993.

[115] P. A. Wintz, and E. J. Luecke, "Performance of Optimum and Suboptimum Synchronizers," *IEEE Trans. on Comm.*, vol. COM-17, no. 3, pp 380-389, June 1969.

[116] F. Classen, H. Meyr, and P. Sehier, "Maximum Likelihood Open Loop Carrier Synchronizer for Digital Radio," Proceeding of ICC'93, (Geneva), pp 493-497, 1993.

[117] D. Baker, and M. Parsons., "Case Studies in Low Density Areas (Telecommunications)," *Electrical Communication*, vol. 1, pp 47-52, 1995.

[118] M. Juy, R. Maurel, and M. Rooryck., "Measurements of 11 GHz Attenuation Using The INTELSAT V Beacon in Indonesia," *International Journal of Satellite Communications*, vol. 8, pp 251-256, June 1990.

[119] A. Dissanayake, D. McCarthy, and J. Allnutt., "Radiometric Rain Attenuation Measurements at 11.6 GHz in Peru," *Electronic Letters*, vol. 26, no. 3, pp 168-170, February 1990.

[120] A. Dissanayake, D. McCarthy, J. Allnutt, and R. Shepherd., "11.6 GHz Rain Attenuation Measurements in Peru," *International Journal of Satellite Communications*, vol. 11, no. 4, pp 229-237, February 1990.

[121] P. de Maagt, G. Brussaard, and J. Allnutt., "Primary and Secondary Statistics of Rain Attenuation and Rain Rate Measured on Satellite-to-Earth Path in Indonesia," *Proceedings of the 9th International Conference on Antennas and Propagation, IEE Conference Publication*, vol. 2, no. 204, pp 123-126, 4-7 April 1995.

[122] G. Brussaard, D. Jaap, K. Liu, and J. Derksen., "Tropospheric Propagation Characteristics at Ku-band for Satellite to Ground and LOS Paths in Surabaya,

Indonesia," *IEICE Transactions on Communications*, vol. E76-B, no. 12, pp 1593-1597, December 1993.

[123] S. Sarkar, I. Ahmed, M. Prasad, H. Dutta, and B. Reddy., "Rain Attenuation on Earth Space Path Over Shillong," *IETE Technical Review*, vol. 11, no. 4, pp 239-242, July-August 1994.

[124] H. Fuketa, M. Kuramoto, and Y. Morihiro., "Medium Capacity Domestic Satellite Communication System," *Review of the Electrical Communication Laboratories*, vol. 28, no. 7-8, pp 502-518, August 1980.

[125] M. Shinji, S. Shimada, M. Koyama, and H. Kumazawa., "Systems Design Considerations of Borne antenna for Medium Capacity Domestic Satellite Communications," *Review of the Electrical Communication Laboratories*, vol. 24, no. 3-4, pp 274-289, April 1976.

[126] R. Tanaka., "30/20 GHz Domestic Satellite Communication System in the Public Communication Network of Japan: Design and Operation," *IEEE Proc.*, vol. 72, no. 11, pp 1637-1644, November 1984.

[127] N. Mohanty, M. L. Hasija, and S. Barathy., "Future Earth Station in 14/12 GHz for Domestic Satellite Communications," *IETE*, vol. 24, no. 8, pp 315-320, 1978.

[128] M. Juy, R. Maurel, and M. Rooryck., "Satellite Earth Path attenuation at 11 GHz in Indonesia," *Electronics Letters*, vol. 26, no. 17, pp 1404-1406, August 1990.

[129] M. Sharma., "MCPC VSATs For Rural Network," *Proceedings of the 10th International Conference on Digital Satellite Communications, IEE Conference*

Publication, no. 403/2, pp 484-490, 1995.

[130] J. Singh., "Status of Satellite Communications in India," *IETE Technical Review*, vol. 5, no. 2, pp 63-74, 1988.

[131] V. Thomas., "Satellite Systems for Global Personal Communications - Approaches, Trends and the Indian Scenario," *IETE Technical Review*, vol. 11, no. 5-6, pp 279-290, September-December 1994.

[132] P. McDougal., "VSAT and Developmental Communications," *Elsevier Science Publishers B.V, Space and Communication and Broadcasting 6*, pp 445-455, 1989.

[133] Donald Lewis., "Indian Tracking and Data Acquisition Facilities for Low Earth and Geosynchronous Satellites," *IEEE AES Systems Magazine*, vol. 8, no. 2, pp 19-27, February 1993.

[134] K. Karnik, and N. Raman., "Planning and Considerations for National Satellite-Based TV System for Education Services in a Developing Country," *42nd Congress of the International Astronautical Federation*, 5-11 October 1991.

[135] A. Ghose., "VSAT Networks in India - Specification and Application," *International Journal of Satellite Communications*, vol. 11, no. 4, pp 211-216, August 1993.

[136] D. Talapatra., "The Indian Space Program," *IEEE AES Systems Magazine*, vol. 8, no. 2, pp 211-216, February 1993.

[137] J. Asker., "Space Pacts Boost India," *Aviation Week and Space Technology*, vol. 142, no. 7, pp 59, February 1995.

[138] J. C. Leguyt, and A. Visser., "A Domestic Satellite System for Peru," *Philips*

Telecommunications Review, vol. 39, no. 2, pp 71-82, June 1981.

[139] J. C. Leguyt, and I. B. J. Rots., "A DAMA Satellite System for Thin-Route Telephone Communication," *Philips Telecommunications Review*, vol. 41, no. 2, pp 134-143, June 1983.

[140] G. Myers, T. Nkabinde, and D. Blaauw., "HealthLink: SatelLife and HealthNet in South Africa," *The Electronic Library*, vol. 13, no. 4, pp 293-298, August 1995.

[141] J. Parapak., "Satellite Communications' Role in National Development," *State Policy*, vol. 10, no. 3, pp 229-235, August 1994.

[142] Elsevier Science Publishers B. V. (North Holland)., "The PALAPA System," *Space and Communication and Broadcasting 6*, pp 127-129, 1988.

[143] F. Gao, and U. Rao., "Space and Education in Developing Countries," *Selected Proceedings of the 43rd International Astronautical Federation Congress*, vol. 30, pp 397-413, 28 August-5 September 1992.

[144] N. Davis., "Asia's Communications Market Booms," *Aerospace America*, vol. 33, no. 1, pp 34-38, January 1995.

[145] K. Miyauchi., "NTT's Domestic Satellite Communication System," *IEEE Comm. Magazine*, pp 25-34, September 1980.

[146] R. Tanaka., "Satellite Communications in Domestic Public Network," *Telecommunication Technologies*, vol. 20, pp 9-28, 1985.

[147] K. Miyauchi., "Domestic Satellite Communication System Research and Development," *Review of the Electrical Communication Laboratories*, vol. 28, no.

7-8, pp 489-497, August 1980.

[148] H. Fuketa, and Y. Watanabe., "TDMA System Outline for Domestic Satellite Communication," *Review of the Electrical Communication Laboratories*, vol. 26, no. 1-2, pp 55-62, February 1978.

[149] H. Shiota, S. Nakajima, and Y. Watanabe., "TDMA Equipment for 6/4 GHz Band Satellite Communication System," *Review of the Electrical Communication Laboratories*, vol. 28, no. 7-8, pp 659-666, August 1980.

[150] Akira Fujii., "VSATs in Japan," *International Journal of Satellite Communications*, vol. 11, no. 4, pp 217-222, August 1993.

[151] E. Hagiwara, K. Satoh, Y. Nishi, H. Komagata, and Y. Yasuda., "S-Band Satellite Communication System Using N-STAR Satellite," *AIAA International Communications Satellite Systems Conference 15th*, pp 673-675, 28 February - 3 March 1994.

[152] T. Michael., "Leased Services on The INTELSAT System Part 1: Domestic Service and International Television," *International Journal of Satellite Communications*, (INTELSAT, 490 L'Enfant Plaza S.W., Washington, D. C. 20024, U.S.A.), 1984.

[153] J. Pillay, and M. Laurence., *Reference Material on Seychelles Telecommunication Systems*, (Ministry of Finance and Communication, P.O. Box 313, Central Bank, Victoria, Seychelles).

[154] N. Pant., "Satellite Communication Technology and Applications 1995-2010," *Journal of Spacecraft Technology*, vol. 4, no. 2, pp 1-28, July 1994.

[155] K. Narayanan, and D. Venugopal., "INSAT - Mobile Satellite Communication Services," *1994 IEEE International Conference on Personal Wireless Communications: Conference Proceedings (ICPWC'94)*, pp 157-162, 1994.

[156] A. Sisodia, and K. Shankara., "Regenerative Communication Payload for Rural Development and Tele-Education in India," *Proceedings of 10th International Conference on Digital Satellite Communications, IEE Conference Publication*, no. 403/2, pp 539-550, 1995.

[157] J. N. Hurd., "Satellite Communications for the Pacific Islands: An Alternative to Isolation," *Public Service Satellite Consortium*, (Washington, D.C., U.S.A.), pp 89-91.

[158] T. Iida., "Satellite Communications Applications to Pacific Countries Above Ku Band," *Proceedings of Sixteenth NASA Propagation Experiments Meeting (NAPEX XVI) and Advance Communications Technology Satellite (ACTS) Propagation Studies Minworkshop*, pp 70-75, 29-30 May 1992.

[159] S. Jamshidi, and L. N. Nguyen., "INTELNET SERVICES: A Global Data Distribution and Collection Scheme," *International Journal of Satellite Communications*, vol. 4, pp 83-87, 1986.

[160] A. Sinha, B. Agrawal, and W. Wu., "Trends in Satellite Communications Technology, Techniques and Applications," *International Journal of Satellite Communications*, vol. 8, no. 4, pp 283-294, February 1990.

[161] B. Edelson, J. Pelton, C. Bostian, W. Brandon, V. Chan, E. Hager, R. Jennings, R.

Kwan, C. Miller, and L. Riley., "Satellite Communications Systems and Technology," pp 45-53, July 1993.

[162] G. Hyde, J. Pelton, B. Edelson, W. Brandon, T. Iida, and R. Collette., "World-Wide Satellite Communications Trend," *AIAA International Satellite Systems, 15th Conference*, pp 1182-1185, 28 February - 3 March 1994.

[163] T. Iida, H. Wakana, and N. Obara., "ETS-V, ETS-VI and COMETS Projects in Japan," *Proceedings of Sixteenth NASA Propagation Experiments Meeting (NAPEX XVI) and Advance Communications Technology Satellite (ACTS) Propagation Studies Minworkshop*, pp 64-69, 29-30 May 1992.