



Circuits and Systems

Mekelweg 4,
2628 CD Delft
The Netherlands
<http://ens.ewi.tudelft.nl/>

CAS-2009-1394584

M.Sc. Thesis

Compressive Sampling for PPM and FSK Modulated Signals

Shahzad Sarwar Gishkori

Abstract

Efficiency of the Analog to Digital Converters (ADCs) has always been an issue of concern, especially, when it comes to sampling wide band signals which require extremely high sampling rates. As the systems with wide band signals are gaining the front position in digital communications, the need to find ways to reduce the sampling rates of ADCs but still maintaining exact reconstruction, is becoming ever more resurgent. In this thesis we present the utilization of a newly discovered technique to reduce the sampling rates much below the Nyquist rates. We explore the combination of Compressive Sampling (CS) with Pulse Position Modulation (PPM) and Frequency Shift Keying (FSK) modulation schemes. CS has been suggested for 'sparse signals' [1], [2]. Sparsity helps represent the signal in much less dimensions. We evaluate the suitability of this technique for PPM and FSK modulated signals in multipath fading environments so as to reduce the complexity at the receiver side. We detect the signals without having to first estimate the channel. We present scenarios where we can achieve the most from the combination of CS and PPM/FSK. We extend this frame work to Ultra Wide Band (UWB) applications. We also give theoretical expressions for the error probabilities of our signal models. The real challenge, with regard to using CS, is to reconstruct the signal from its reduced dimensions. In this respect, we have opted for the Orthogonal Matching Pursuit (OMP) algorithm for most of the cases, nonetheless, we elaborate on other available reconstruction methods as well.



Compressive Sampling for PPM and FSK Modulated Signals

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

TELECOMMUNICATIONS

by

Shahzad Sarwar Gishkori
born in D. G. Khan, Pakistan

This work was performed in:

Circuits and Systems Group
Department of Electrical Engineering
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology



Delft University of Technology

Copyright © 2009 Circuits and Systems Group
All rights reserved.

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
ELECTRICAL ENGINEERING

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled “**Compressive Sampling for PPM and FSK Modulated Signals**” by **Shahzad Sarwar Gishkori** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: August 31, 2009

Chairman:

Prof.dr.ir. Alle-Jan Van der Veen, Technische Universiteit Delft

Advisors:

Dr.ir. Geert Leus, Technische Universiteit Delft

Prof.dr.ir. Hakan Deliç, Boğaziçi University, Istanbul

Committee Members:

Dr.ir. Homayoun Nikookar, Technische Universiteit Delft

Abstract

Efficiency of the Analog to Digital Converters (ADCs) has always been an issue of concern, especially, when it comes to sampling wide band signals which require extremely high sampling rates. As the systems with wide band signals are gaining the front position in digital communications, the need to find ways to reduce the sampling rates of ADCs but still maintaining exact reconstruction, is becoming ever more resurgent. In this thesis we present the utilization of a newly discovered technique to reduce the sampling rates much below the Nyquist rates. We explore the combination of Compressive Sampling (CS) with Pulse Position Modulation (PPM) and Frequency Shift Keying (FSK) modulation schemes. CS has been suggested for 'sparse signals' [1], [2]. Sparsity helps represent the signal in much less dimensions. We evaluate the suitability of this technique for PPM and FSK modulated signals in multipath fading environments so as to reduce the complexity at the receiver side. We detect the signals without having to first estimate the channel. We present scenarios where we can achieve the most from the combination of CS and PPM/FSK. We extend this framework to Ultra Wide Band (UWB) applications. We also give theoretical expressions for the error probabilities of our signal models. The real challenge, with regard to using CS, is to reconstruct the signal from its reduced dimensions. In this respect, we have opted for the Orthogonal Matching Pursuit (OMP) algorithm for most of the cases, nonetheless, we elaborate on other available reconstruction methods as well.

Acknowledgments

First and foremost, I would like to thank God, The Most Compassionate and The Most Merciful, for giving me the ability to finish my thesis in time. Despite my short comings, He has always bestowed upon me His blessings and mercy. Verily He is the Source of all strength.

I would like to thank my advisors Dr. Geert Leus and Prof.dr. Hakan Deliç. I don't have words to express my gratitude for these two great men who made me grow in my thesis from a scratch. I troubled them, every now and then, with even minor questions but they never played it down and responded with full commitment. From Geert, I learnt how to think critically and to combine multiple aspects to make my work meaningful. From Hakan, I learnt never to forget minute details in research and the difference between technical and non-technical writing. It will not be wrong to say that without them, this thesis would not have seen the light of day. I would also like to thank my professor, Dr. Alle-Jan Van der Veen for giving me the opportunity to work in this group and for being a role model for all the young researchers. I would like to thank Dr.ir. Homayoun Nikookar for being a member of my thesis committee.

I would like to thank my siblings and my father back home, for giving me the vital moral support and encouragement which kept me going through all the ebb and flow of this work. Though, miles away, their best wishes and prayers have always surrounded me.

Finally, I would like to thank my late mother. She was always the happiest person on earth for even very small achievements of her son. I miss her every day and every moment of my life. This little work is dedicated to her.

Shahzad Sarwar Gishkori
Delft, The Netherlands
August 31, 2009

Contents

Abstract	v
Acknowledgments	vii
Abbreviations	xv
1 Introduction	1
1.1 Motivation	1
1.2 Effects of ADC Sampling Rates	2
1.3 Outline and Contributions	4
2 Background Discussion	7
2.1 Sparsity	7
2.2 Basic Theory of CS	7
3 Signal Model	9
3.1 General Signal Model	9
3.2 General CS Model	11
3.2.1 AIC	11
4 CS for PPM	13
4.1 PPM Scheme	13
4.2 PPM Signal Model	14
4.3 PPM Detection Rule	14
4.4 Quasi-Synchronous 2-PPM	16
4.5 Fully-Synchronous 2-PPM	18
4.6 PPM Simulation	20
4.7 Conclusions	24
5 CS for FSK	25
5.1 FSK Scheme	25
5.2 FSK Signal Model	25
5.3 FSK Detection Rule	27
5.4 N-FSK (single carrier)	28
5.5 N-FSK (double carrier)	29
5.6 FSK Simulation	30
5.7 Signal Reconstruction	34
5.8 Conclusions	37
6 Analysis of CS	39
6.1 CS Analysis	39
6.2 Conclusions	42

7	CS for UWB	43
7.1	Signal Model	43
7.2	Signal Model according to Mengali et al [3]	45
7.2.1	Simulation Results of [3]	47
7.3	Analysis of the Signal Model of [3]	49
7.4	Our Theoretical Model	52
7.4.1	Simulation results for our Theoretical Model	54
7.5	CS for UWB	55
7.6	CS for UWB with Multiple Access	58
7.6.1	Signal Model	58
7.6.2	CS Simulations for Multiple Access UWB	60
7.7	Conclusions	61
8	Conclusions and Future Work	63
8.1	Conclusions	63
8.2	Suggestions for Future Work	64
A	Analog to Information Converter	65
A.1	Random Sampling	65
A.2	Random Filtering	66
A.3	Pseudo-random Demodulation	66

List of Figures

1.1	ADC, ENOB versus sampling rate [4]	3
1.2	ADC, Power versus sampling rate [4]	3
3.1	CS Signal Model	11
4.1	Illustration of a received 2-PPM symbol (absolute squared) along with the channel spread	13
4.2	Matrix representation of CS for 2-PPM	14
4.3	Reconstruction and Detection Scheme (block-diagram)	15
4.4	2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, reconstruction algorithm = OMP	20
4.5	2-PPM BER for fully-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, reconstruction algorithm = OMP	21
4.6	2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = Gaussian random matrix, Reconstruction Algorithm: OMP	22
4.7	2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = Gaussian random matrix, Reconstruction Algorithm: BP	22
4.8	2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, N=16, M = 1:N, SNR = 0dB, Reconstruction Algorithm: OMP	23
4.9	2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, N=16, M = 1:N, SNR = 10dB, Reconstruction Algorithm: OMP	23
5.1	Matrix representation of CS for FSK (I)	26
5.2	Matrix representation of CS for FSK (II)	26
5.3	8-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, N=8, M=4, measurement matrix = random matrix, Reconstruction Algorithm: OMP	31
5.4	16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, N=16, M = [1:N], measurement matrix = random matrix, Reconstruction Algorithm: OMP	31
5.5	16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, N=16, M = [1:N], measurement matrix = random matrix, Reconstruction Algorithm: OMP	32
5.6	16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, N=16, M = [1:N], measurement matrix = Identity matrix, Reconstruction Algorithm: OMP	33

5.7	16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=16$, $M = [1:N]$, measurement matrix = Identity matrix, Reconstruction Algorithm: OMP	33
5.8	8-FSK, Double Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=8$, $M = 6$, measurement matrix = Gaussian matrix, Reconstruction Algorithm: OMP	34
5.9	8-FSK, Double Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=8$, $M = 6$, measurement matrix = Identity matrix, Reconstruction Algorithm: OMP	35
6.1	2-PPM, Quasi-Synchronous, Channel-taps (L) = 3, Sparsity (K) = 3, Gaussian fading channel with noise, $M = N/2$, measurement matrix = random matrix, Reconstruction Algorithm: OMP	40
6.2	Lower bounds on sparsity, [5] (black for [6], red for [7] and green for [5])	41
7.1	2-PPM UWB, Multiple frames Symbol, $N_f = 2$	43
7.2	2-PPM UWB, BER performance for Multiple frames for CM1, for large integration interval	47
7.3	2-PPM UWB, BER performance for Multiple frames for CM3, for large integration interval	48
7.4	2-PPM UWB, BER performance for Multiple frames for CM1, for small integration interval	48
7.5	2-PPM UWB, BER performance for Multiple frames for CM3, for small integration interval	49
7.6	2-PPM UWB, BER performance using (??) on Multiple frames	52
7.7	2-PPM UWB, Comparison with small integration interval	55
7.8	2-PPM UWB, Comparison with large integration interval	56
7.9	2-PPM UWB, CS with reconstruction algorithm = OMP, measurement matrix = DFT matrix	56
7.10	2-PPM UWB, CS with reconstruction algorithm = OMP, Effect of M/N variations	57
7.11	2-PPM UWB, CS with reconstruction algorithm = OMP, measurement matrix = Gaussian matrix	58
7.12	2-PPM UWB, CS with reconstruction algorithm = OMP, Effect of M/N variations, measurement matrix = Gaussian matrix	59
7.13	UWB symbol with Multiple Access	60
7.14	2-PPM UWB , comparison between single user and multiple users, CS with reconstruction algorithm = OMP, measurement matrix = DFT matrix	61
A.1	AIC implementation with random sampling, [8]	65
A.2	AIC implementation with random filtering (I), [9]	66
A.3	AIC implementation with random filtering (II), [9]	66
A.4	AIC implementation with pseudo-random demodulation, [10]	67

List of Tables

6.1	Sparsity according to [11]	42
6.2	Sparsity according to [6]	42
7.1	IEEE Channel Parameters [12]	44
7.2	Effect of Averaging on Multiframe Signal Model	53

Abbreviations

ADC	Analog to Digital Converter
AIC	Analog to Information Converter
BER	Bit Error Rate
BP	Basis Pursuit
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FHSS	Frequency Hopping Spread Spectrum
FSK	Frequency Shift Keying
FTB-OMP	Flexible Tree-search-Based OMP
GML	Generalized Maximum Likelihood
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
LOS	Line of Sight
MA	Multiple Access
MP	Matching Pursuit
NLOS	Non Line of Sight
OMP	Orthogonal Matching Pursuit
PPM	Pulse Position Modulation
RIP	Restricted Isometry Property
Seq-OMP	Sequential OMP
SNR	Signal to Noise Ratio
SS	Spread Spectrum
UWB	Ultra Wide Band

In this thesis we explore an alternative way of reducing the sampling rate for wideband signals. Sampling rate is one of the key features of any communications system. Its reduction can have far reaching effects on the overall performance of the communications system.

This chapter provides the motivation for the thesis along with the outline and our contributions for that matter.

1.1 Motivation

Digital communications is witnessing a phenomenal growth in applications which involve signals of very high bandwidth, for example, Ultra Wide Band (UWB) and Spread Spectrum (SS) signals, etc. Due to this, PPM and FSK modulation schemes have gained quite an importance in the realization of such systems. Efforts are being made by the research community to find ways to reduce the overall system complexity in this regard. CS with PPM and FSK yields one step forward in this direction.

At the receiver side, a big hurdle is the efficiency of the Analog to Digital Converter (ADC). According to the classical Shannon-Nyquist-Whittaker-Kotelnikov sampling theorem [13], [14], a band limited signal $x(t)$ (or $X(\omega) = 0, |\omega| > \omega_m$) can be determined completely from its samples $x(nT)$ if $T \leq \pi/\omega_m$. So the sampling rate should be at least twice the highest frequency. Therefore, if the bandwidth of the signal is too high, ADCs can be heavily stressed. It could take 'decades' before the ADCs, with current technology, can be fast and precise enough for the present day high bandwidth applications [10]. Furthermore, it has been described in [13] that most of the signals with large bandwidth have a small rate of information. This property of wideband signals makes them sparse in information and has led to methods of sampling based on the amount of information (or the rate of innovation). The combination of sparsity with finite rate of innovation has been described in [15], primarily for the non-discrete domain. CS offers more flexible options to deal with sparse signals in terms of the location of the information and the non-uniformity of measurements as we shall elaborate upon in subsequent sections. Therefore we take advantage of the sparsity of the signals which are PPM and FSK modulated through CS.

PPM and FSK signals are sparse in the time and frequency domains respectively. Both \mathcal{M} -ary PPM and \mathcal{M} -ary FSK can be represented in the geometric form of \mathcal{M} -

dimensional orthogonal vectors as in [16],

$$\begin{aligned}
\mathbf{a}_0 &= \left(\sqrt{S}, 0, 0, \dots, 0 \right)^T \\
\mathbf{a}_1 &= \left(0, \sqrt{S}, 0, \dots, 0 \right)^T \\
&\vdots \\
\mathbf{a}_{M-1} &= \left(0, 0, \dots, 0, \sqrt{S} \right)^T,
\end{aligned} \tag{1.1}$$

where S represents signal energy. So there are many zeros in the signal but only one significant component. CS aims at exploiting such structures of the signals.

1.2 Effects of ADC Sampling Rates

Sampling rates (f_s) have a direct bearing on the Performance of ADCs besides resolution (i.e., Effective Number of Bits ($ENOB$)). The two most widely used figures-of-merit P and F are defined as in [4],

$$\begin{aligned}
P &= 2^{ENOB} f_s \\
F &= \frac{2^{ENOB} f_s}{P_{diss}}
\end{aligned} \tag{1.2}$$

where $ENOB$ is also called *SNR bits* as in [17]. The performance trend of past-present-and-future for different ADCs has been given in [4] and [17]. Figure 1.1 gives the relationship between ADC resolution and sampling rates for different ADC architectures. It is clear that flash ADCs give the highest sampling rates (in the range of Giga samples per second)but lowest resolution. Sigma-Delta give highest resolution but the sampling rates are very low. The pipelined structure ADCs can be considered to give very good over all performance.

Furthermore, its is clear from (1.2) that power dissipation is inversely proportional to the performance of ADCs. The relationship between the power dissipated and resolution along with sampling frequency is almost linear. A derivation of this relationship has been presented in [18] with the assumptions that 1. power is consumed only at the sample-and-hold block of the ADC, 2. the input signal supplies the power to charge the sample-and-hold capacitance . It is given as,

$$P_{min} = k T f_s 10^{(6N+1.76)/10} \text{ [W]} \tag{1.3}$$

where N is stated number of bits of resolution, k is Boltzmann's constant and T is temperature.

The slope between consumed power and sampling frequency according to (1.3) is quite linear but in practical (or non-ideal) situations it is even steeper. Figure 1.2 shows the plot for power consumed versus sampling frequency for different ADC architectures. It shows that the power consumed for higher sampling rates is extremely high. The flash structures consume highest power though the number of resolution bits is very

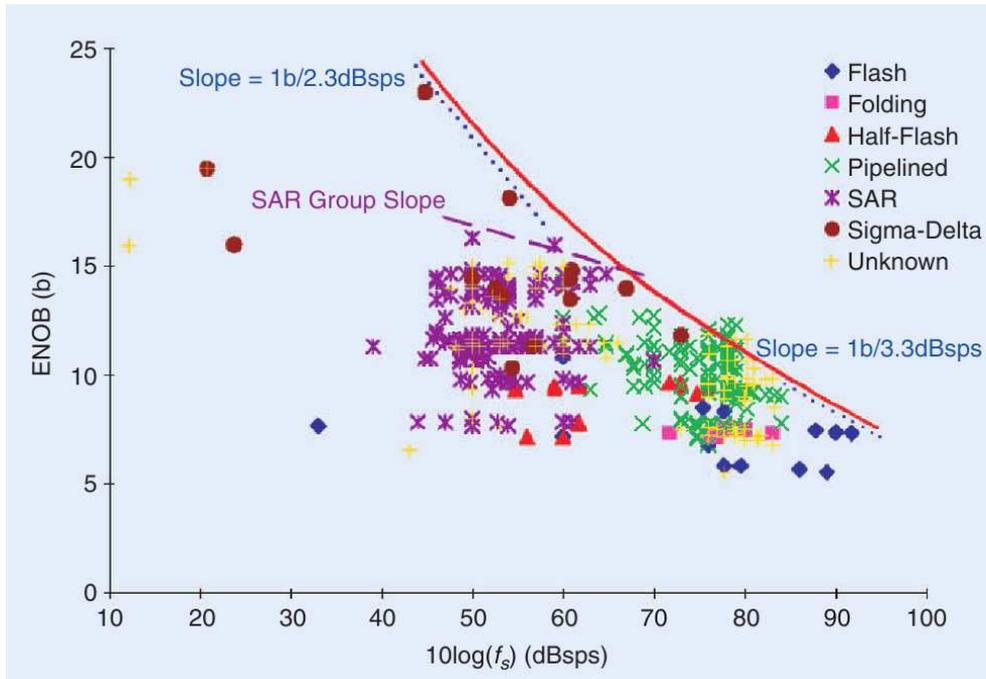


Figure 1.1: ADC, ENOB versus sampling rate [4]

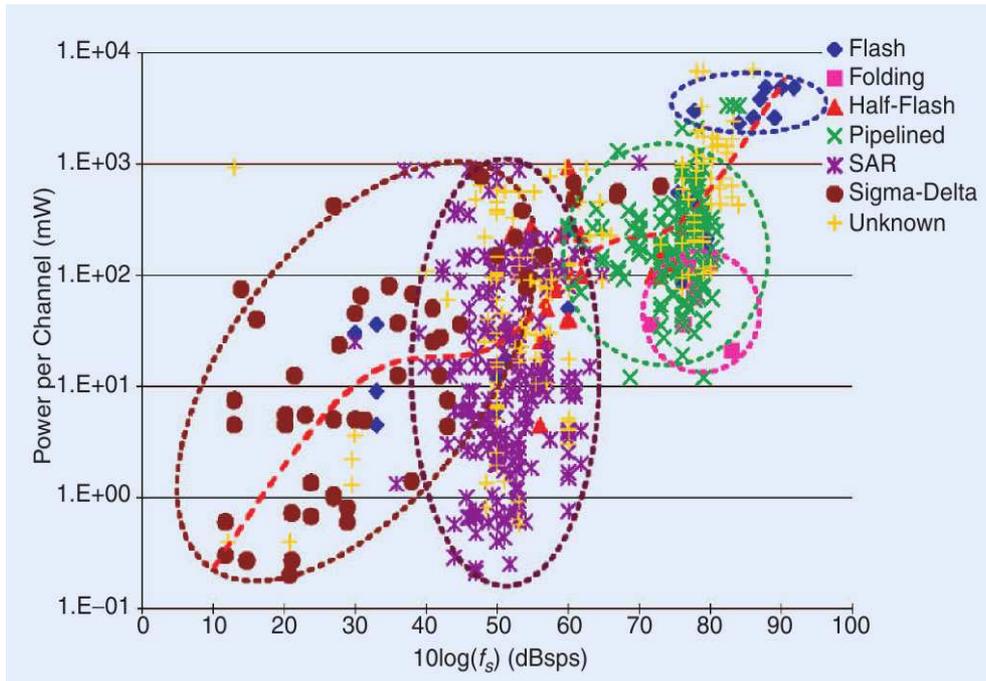


Figure 1.2: ADC, Power versus sampling rate [4]

small (see Figure 1.1). Along with the sampling rate and number of resolution bits, a lot of power is consumed in the comparator processes. In the case of flash ADCs, number of comparisons are around $2^N f_s$. Given the purely parallel structure of flash

ADCs, power consumed is highest. On the other hand, we can see from Figure 1.2 that the sigma-delta ADCs consume least power with very good resolution (see Figure 1.1) but at the same time low sampling rates. Both these architectures have to 'sacrifice' alot of power to give a 'balanced performance'.

So if the signals are wideband, very high sampling rates would be required as per the requirement of Shannon-Nyquist-Whittaker-Kotelnikov sampling theorem. This would imply, in view of the above discussion, very high power consumption. Our goal in this thesis is to reduce the sampling rates in order to increase the efficiency of ADCs.

1.3 Outline and Contributions

In this thesis we have utilized CS to reduce sampling rates for the wideband signals. We explicitly explored the effect of CS on PPM and FSK modulated signals. We have provided theoretical analysis for these schemes and simulations thereof. We have come out with a detailed analysis of the scenarios where CS can be utilized to its maixmum potential. We have contributed to the understanding of measurement matrices in terms of their structures and utilization. We have highlighted upon the bounds on the order of sparsity for better reconstruction of the signals. We have also given the applicability of CS for UWB signals. We have also provided the effect of multiuser interference on the performance of CS for UWB signals. We have considered non coherent detection to reduce complexity on the receiver side.

Chapter 2: Background Discussion

In this chapter we provide the necessary background material relating to our thesis. We briefly describe the basic theory of compressive sampling. We focus mostly on the key points which directly affect the performance in terms of its application.

Chapter 3: Signal Model

We provide general signal model in this chapter. It includes the improtant notations and terminologies which shall be used subsequently in the thesis. We also present the CS model.

Chapter 4: CS for PPM

In this chapter we provide the application of CS on PPM modulated signals. We present the theoretical analysis of PPM signals. We present the important issues relating to the application of CS for PPM.

Chapter 5: CS for FSK

In this chapter we provide the application of CS on FSK modulated signals. We present the theoretical analysis of FSK signals both for single tone and dual tone FSK. We present the important issues relating to the application of CS for FSK.

Chapter 6: CS Analysis

In this chapter we analyse our findings regarding application of CS for PPM and FSK signals. We focus on the measurement matrices and the order of sparsity for better reconstruction of the signal resulting in improved performance of CS.

Chapter 7: CS for UWB

In this chapter we present the application of CS for UWB signals. We provide our theoretical analysis for the signal model and comparison with an existing model. We extend the application of CS for the UWB signal with multiple access.

Background Discussion

In this chapter we provide the necessary background material relating to our thesis. We briefly describe the basic theory of compressive sampling. We focus mostly on the key points which directly affect the performance in terms of its application.

2.1 Sparsity

Sparsity is a relevant term. In terms of image theory, a sparse structure may mean having few large coefficients and many small coefficients. So the image can be approximated by these large coefficients. In terms of analog signals, it may mean that the signal is constituted by few of its basis functions at a unit interval of time, out of a large dictionary of possible basis functions. An other way of putting it is that we can define a signal in terms of such basis functions which give it a sparse structure but still we need to know the basis functions to reconstruct the signal.

2.2 Basic Theory of CS

Compressive Sampling can be thought of as a concept to reduce the number of measurements required to approximate a signal without any loss of important information. So the signal is transformed into a compressed representation. Therefore, it can be processed with less resources and then finally reconstructed to its original form. This theory works under certain assumptions viz a viz the class of signals and the transform operator. Let $\mathbf{x}_{N \times 1}$ be a discrete-time signal (in vector representation), which is sparse in some basis (e.g., Fourier, Wavelet, etc.). If $\Psi_{N \times N}$ represents the matrix containing the basis vectors of \mathbf{x} and $\mathbf{s}_{N \times 1}$ its respective coefficients then we can write,

$$\mathbf{x} = \Psi \mathbf{s}. \quad (2.1)$$

Since \mathbf{x} is sparse in its basis then \mathbf{s} will have very few non-zero coefficients. Therefore \mathbf{x} can be represented by M linear measurements with $M \ll N$. Let $\Phi_{M \times N}$ be the transform operator: $\mathbb{R}^N \mapsto \mathbb{R}^M$, with M linear functionals as its rows. Then \mathbf{x} can be transformed into a new representation $\mathbf{y}_{M \times 1}$ i.e.,

$$\mathbf{y} = \Phi \Psi \mathbf{s}. \quad (2.2)$$

Here we shall attribute the basis matrix Ψ as the sparsity matrix and Φ as the measurement (or the sparsifying) matrix. The sparsity matrix consists of orthogonal basis vectors and \mathbf{s} contains the coefficients of the basis functions which constitute the signal \mathbf{x} . In sparse signals, \mathbf{s} contains very few large coefficients and many zeroes or negligibly small valued coefficients (obeying e.g., a power-law decay condition [19]). It is

this sparse structure of the signals which motivates their compressed representation. The key role played here is by the measurement matrix Φ . It is a special matrix with certain requirements. It does not only have to reduce the dimensions of the signal but should also allow for its reconstruction. Therefore, as a first condition, it has to satisfy the restricted isometry property (RIP) [2] and as our signals are mostly sparse in some orthonormal basis, RIP should hold for $\Theta := \Phi\Psi$ [20]. Secondly, its rows should be incoherent with the basis matrix [21].

Let x_i constitute the elements of \mathbf{x} , and let $T = \text{supp}\{i : x_i \neq 0\}$ (i.e., support of \mathbf{x}), with $K = |T|$ (i.e., cardinality of T) then the signal is said to be K -sparse (i.e., $K = \|\mathbf{x}\|_{l_0}$). RIP can then formally be written as,

$$(1 - \delta_K)\|\mathbf{s}\|_{l_2} \leq \|\Theta\mathbf{s}\|_{l_2} \leq (1 + \delta_K)\|\mathbf{s}\|_{l_2} \quad (2.3)$$

where δ_K is a sparsity constant of order K , $0 < \delta_K < 1$. If Θ_T represents a submatrix of Θ with columns $\theta_i : i \in T$, then RIP implies that the eigenvalues of $\Theta_T^t \Theta_T$ are in $[(1 - \delta_K)^2, (1 + \delta_K)^2]$ [22]. Furthermore RIP of order K implies that RIP also holds for sparsity less than K [22].

If the above conditions are met then one can reconstruct \mathbf{s} by solving the following minimization problem,

$$\begin{aligned} \min_{\mathbf{s}} \|\mathbf{s}\|_{l_1} &:= \sum_{i=1}^N |s_i| \\ \text{s.t. } \mathbf{y} &= \Theta\mathbf{s}, \end{aligned} \quad (2.4)$$

with

$$M = CK \log(N/K), \quad (2.5)$$

linear measurements where C is a positive constant ($C > 0$ and may range between 2 – 20) [1, 2]. In literature, a number of measurement matrices have been suggested which comply with the above mentioned requirements. For example, Gaussian, Bernoulli, Fourier ([1], [2], [20]) matrices and other random matrices based on predefined probabilities [23]. Though most of the literature on CS has been suggesting random measurement matrices, Candès *et al*, [21] strongly advocate the use of structured matrices e.g., Fourier matrices. We shall discuss this particular issue in subsequent sections.

Reconstruction with the help of (2.4) is known as Basis Pursuit (BP). This is not the only way to recover the signal but it is one of the stable methods for sure. In [24], it was demonstrated, both theoretically and empirically, that Orthogonal Matching Pursuit (OMP) [25] can recover a signal with almost similar bounds on the required linear measurements as that of (2.5). We also consider OMP as our primary reconstruction algorithm as it has been proven to be a faster alternative to BP [24].

Signal Model

In this chapter we define our general signal model. This consists of the transmitted and received signal structures along with the general CS model.

3.1 General Signal Model

Let $a(t)$ be the transmitted signal. $a(t)$ can either be PPM modulated or FSK modulated. The signal travels through the channel. Let $h'(t)$ represent the impulse response of the physical communications channel. At the receiver end, additive white Gaussian noise $n(t)$ gets added to the signal. So the received signal $x(t)$ can be represented as the combination of noise and the convolution of $a(t)$ and $h'(t)$. Assuming that $a'(t)$ represents the signal along with pulse shaping we can write,

$$x(t) = h'(t) * a'(t) + n(t) \quad (3.1)$$

where $*$ denotes the convolution.

We define $a(t)$ for the k -th information symbol as,

$$a_k(t) = \sum_{n=0}^{N-1} a_{n,k} \delta(t - nT) \quad (3.2)$$

where N is the total number of samples of the signal, sampled at the Nyquist rate. T is the sampling interval and equals inverse of twice the bandwidth, i.e., $T = 1/2B$ where B stands for the signal bandwidth. Furthermore, we can define $a_{n,k}$ separately for the case of PPM and FSK as,

$$a_{n,k}^{(PPM)} = \sum_{i=0}^{N_f-1} \delta[n - iN/N_f - k] \quad (3.3)$$

$$a_{n,k}^{(FSK)} = \exp[j2\pi kn/N] \quad (3.4)$$

where N_f are the number of frames per PPM symbol with a duration of T_f for an individual frame. We shall consider the scenarios where $N_f = 1$ and where $N_f > 1$.

Moreover, the physical channel $h'(t)$ can be defined as,

$$h'(t) = \sum_{z=0}^{Z-1} v_z \delta(t - T_z) \quad (3.5)$$

where v_z are the channel coefficients which we assume to be independently identically distributed, Z is the total number of multipaths with inter-path arrival time, T_z . We

also define the maximum channel delay spread $T_{m ds} := ZT$ where T is the sampling interval.

Furthermore, we consider the notion of composite channel $h(t)$ as the convolution of physical channel $h'(t)$ and the received pulse shape $p(t)$, i.e.,

$$h(t) = h'(t) * p(t) \quad (3.6)$$

Then we can define $h(t)$ as,

$$h(t) = \sum_{l=0}^{L-1} h_l p(t - T_l) \quad (3.7)$$

where h_l are the channel coefficients which we assume to be independently identically distributed, L is the total number of multipaths with inter-path arrival time, T_l . We also define the maximum channel delay spread $T_{m ds} := LT$ where T is the sampling interval.

For the sake of a general model, we now evaluate (3.1) with the composite channel $h(t)$ instead of the physical channel $h'(t)$. The basic principle will still remain same in either case. Given the fact that the channel has a Finite Impulse Response (FIR) of LT , (3.1) can be written as,

$$\begin{aligned} x(t) &= h(t) * a(t) + n(t) \\ &= \int_0^{LT} h(\tau) a(t - \tau) d\tau + n(t) \\ &= \int_0^{LT} h(\tau) \sum_{i=0}^{N-1} a_i \delta(t - \tau - iT) d\tau + n(t) \\ &= \sum_{i=0}^{N-1} a_i \int_0^{LT} h(\tau) \delta(t - \tau - iT) d\tau + n(t) \\ &= \sum_{i=0}^{N-1} a_i h(t - iT) + n(t) \end{aligned} \quad (3.8)$$

Defining $x_n = x(nT)$, $h_n = h(nT)$ and $n_n = n(nT)$, we can write (3.8) as,

$$\begin{aligned} x_n &= \sum_{i=0}^{N-1} a_i h_{n-i} + n_n \\ &= \sum_{l=0}^{L-1} h_l a_{n-l} + n_n \end{aligned} \quad (3.9)$$

To make every received frame independent, we consider a guard interval of length $L - 1$ with every frame. We do zero padding in the case of PPM and add a cyclic prefix

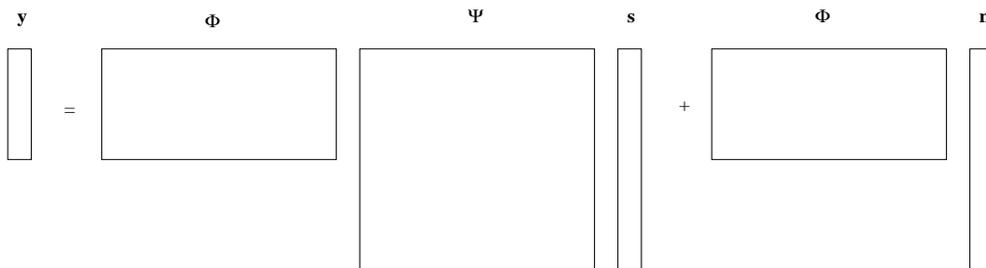


Figure 3.1: CS Signal Model

in the case of FSK, both of length $L - 1$. Thus we can write (3.9) in the matrix-vector notation as,

$$\mathbf{x} = \mathbf{H}\mathbf{a} + \mathbf{n}. \quad (3.10)$$

Here \mathbf{H} represents the channel convolution matrix, vectors \mathbf{x} , \mathbf{a} and \mathbf{n} contain all elements of x_n , a_n and n_n .

3.2 General CS Model

At the receiver side, we first compress the signal and then sample at reduced rate. Given \mathbf{x} to be the received signal, we can express the compressed signal \mathbf{y} in terms of the basis functions matrix Ψ and coefficient vector \mathbf{s} of \mathbf{x} as

$$\mathbf{y} = \Phi\Psi\mathbf{s} + \Phi\mathbf{n} \quad (3.11)$$

where Φ is the measurement matrix as defined in Section 2.2. Figure 3.1 gives an illustration of our CS model.

3.2.1 AIC

We carry out the compression of the analog signals through Analog to Information Converter (AIC). A number of different implementations of AICs have been suggested in literature. Appendix A describes the details of these implementations. So the received signal shall be first processed by the AIC to give it a compressed representation and sampled at sub-Nyquist. Then, we reconstruct the signal through CS reconstruction algorithms e.g., BP and OMP. Later on, we apply our detection rule over the reconstructed signal.

In this chapter we focus on the utilization of CS for PPM modulated signals. We present the PPM signal models for quasi-synchronous and fully-synchronous cases, theoretical expressions for the probability of error and simulation results.

4.1 PPM Scheme

Pulse Position Modulation (PPM) is an \mathcal{M} -ary orthogonal modulation scheme [26]. PPM is the most common time-based modulation technique used in Ultra Wide Band (UWB) communications [27]. Different symbols are realized by shifting a pulse to specific positions in time within the specified symbol duration. A simple example of 2-PPM with the symbol duration T_f divided equally into two pulse positions is shown in Figure 4.1. PPM is advantageous because of its simplicity and the ease of controlling delays [27] but the disadvantage (in the sense of receiver ADCs) is the large bandwidth associated with it. CS can ease the stress (because of PPM) on ADCs. Our signal model follows next.

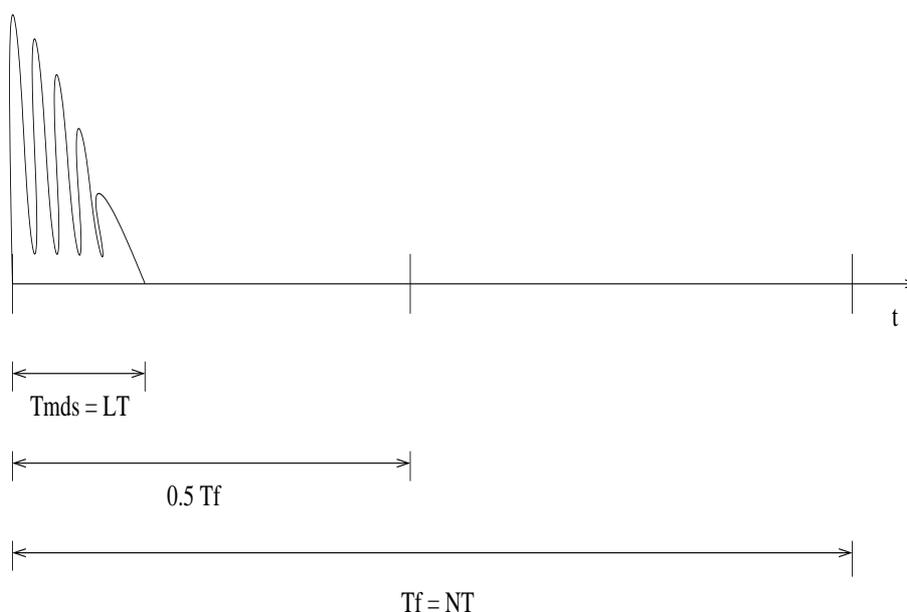


Figure 4.1: Illustration of a received 2-PPM symbol (absolute squared) along with the channel spread

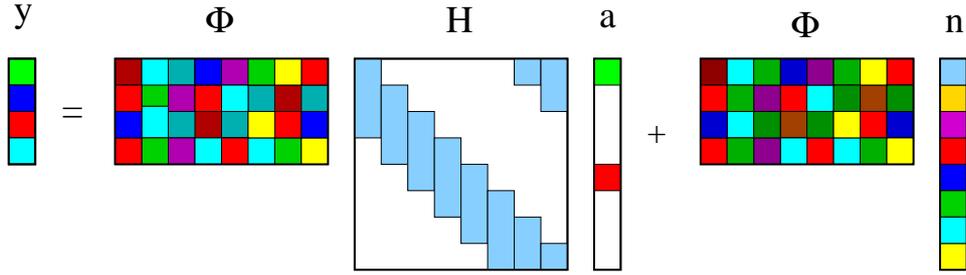


Figure 4.2: Matrix representation of CS for 2-PPM

4.2 PPM Signal Model

Let $a(t)$ represent the chain of transmitted PPM symbols. For the purpose of illustration we present the case of 2-PPM symbols. Extension to M-PPM is quite straightforward. Let $h(t)$ be the multipath channel impulse response (see (3.6)). We assume that the maximum channel delay spread $T_{m ds} < T_f/2$ (see Figure 4.1), so each pulse is contained within its pulse position and does not spill over to other pulse positions. h_l represents the l -th channel path coefficient. We assume that the channel coefficient h_l is i.i.d Gaussian random variables i.e., $h_l \sim \mathcal{N}(0, 1)$. $h(t)$ includes the pulse shaping filters both at the transmitter and the receiver side. So in our model, $a(t)$ would basically consist of Diracs at the beginning of every pulse and as it travels through the channel, the shape and length of the pulse would be the result of its convolution with the channel. If $n(t)$ describes the additive white Gaussian noise i.e., $n(t) \sim \mathcal{N}(0, \sigma^2)$, we can write the received signal $x(t)$ in the matrix form as in (4.1), i.e.,

$$\mathbf{x} = \mathbf{H}\mathbf{a} + \mathbf{n}. \quad (4.1)$$

Although we shall process the signal in the analog form, for reasons of understanding and simulations, we shall describe it in a discrete form sampled at the Nyquist rate. So after compressive sampling (i.e., taking M linear measurements of the signal), we can write the received signal as,

$$\mathbf{y} = \Phi\mathbf{H}\mathbf{a} + \Phi\mathbf{n} \quad (4.2)$$

Figure 4.2 gives an illustration of (4.2). In terms of (2.1) and (2.2) we can write (4.2) as;

$$\mathbf{y} = \Phi\Psi\mathbf{s} + \Phi\mathbf{n}. \quad (4.3)$$

where,

$$\mathbf{s} = \mathbf{H}\mathbf{a} \text{ and } \Psi = \mathbf{I}. \quad (4.4)$$

4.3 PPM Detection Rule

A block diagram of the reconstruction and detection process for PPM is given in Figure 4.3. It shows mainly, the AIC, reconstruction block and the decision block. We

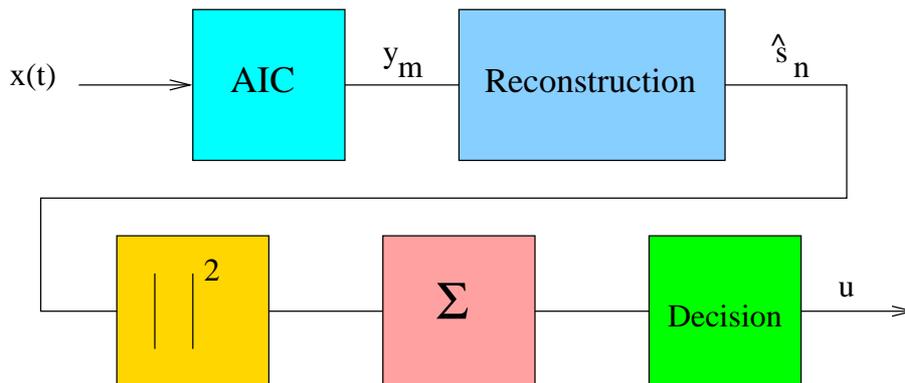


Figure 4.3: Reconstruction and Detection Scheme (block-diagram)

shall talk about the reconstruction process (i.e., from \mathbf{y} to $\hat{\mathbf{s}}$ where $\hat{\mathbf{s}}$ is the reconstructed vector) in detail in a later section (jointly for PPM and FSK), but here we describe the detection rule for PPM.

Since our goal is to reduce the overall system complexity, we have used noncoherent receivers for our system model. The advantage of noncoherent detection is the reduction in complexity and power consumption [3]. Our detection process is more like a generalized maximum likelihood detector. From the reconstructed signal $\hat{\mathbf{s}}$, our decision is based on the estimate as to which pulse position contains more energy than the rest of the positions. We process every frame (i.e., symbol) separately. We take the absolute square of the signal and integrate it over the different pulse position intervals. The interval which contains more energy than the rest (within the symbol) is our decision. In this way we do not need to estimate the channel. Instead, we are collecting the energy of the multipath components of the signal to increase the detection probability of the actual transmitted pulse.

As an example, we can consider the case of 2-PPM (Figure 4.1). If N is the total number of Nyquist rate samples of the symbol, then the decision rule is based on the following relation,

$$\begin{array}{c} 0 \\ \|\hat{\mathbf{s}}_{1:\frac{N}{2}}\|_2^2 \geq \|\hat{\mathbf{s}}_{\frac{N}{2}+1:N}\|_2^2 \\ 1 \end{array} \quad (4.5)$$

where $\hat{\mathbf{s}}_{i:j}$ denotes the subvector of $\hat{\mathbf{s}}$ from position i to j with $i, j \in \mathbb{Z}^+$. The relation (4.5) solves for bit 0 or bit 1. From (4.1)-(4.4), we can see from (4.5), that a decision on $\hat{\mathbf{s}}$ leads to that of \mathbf{a} . Since the channel effect is limited to the duration of the pulse, multiple paths in fact contribute to a correct decision. Furthermore, the noise \mathbf{n} can affect the decision in two ways. If we have perfect synchronization, i.e., we only compare the intervals which might contain the received pulse (i.e., within $T_{m ds}$, see Figure 4.1), in both of the pulse positions. In case of quasi-synchronous communications, we compare the whole pulse interval as described in (4.5). As noise is assumed to be an i.i.d Gaussian random sequence, in the former case (perfectly synchronous), only the noise

from the exact pulse interval contributes to the pulse estimate, whereas in the latter (i.e., quasi-synchronous case), extra noise also becomes part of the pulse energy. We shall experiment with both of these cases separately.

We define our average Signal to Noise Ratio (SNR) as,

$$\text{SNR} = \frac{E\|\mathbf{a}\|_2^2 E\|\mathbf{h}\|_2^2}{E\|\mathbf{n}\|_2^2} \quad (4.6)$$

where $E\|\mathbf{a}\|_2^2$ is the average signal energy (let it be denoted by $S := E\|\mathbf{a}\|_2^2$), $E\|\mathbf{h}\|_2^2$ is the average channel energy (given that the channel coefficients are zero mean i.i.d Gaussian random variables with unit variance, it is equal to L) and $E\|\mathbf{n}\|_2^2$ is the average noise energy (for i.i.d Gaussian $\sim \mathcal{N}(0, \sigma^2)$ noise variables, it is equal to $N\sigma^2$, where N is the Nyquist rate sampling factor). So (4.6) represents the average signal to noise ratio. For the case of PPM, let it be denoted by \mathcal{Y}_{PPM} ,

$$\mathcal{Y}_{PPM} = \frac{SL}{N\sigma^2}. \quad (4.7)$$

We shall express the SNR in terms of dB through out all the simulations. Our Bit Error Rate (BER) is in fact an indicator of how well Compressed Sensing has performed. To compare the performance of CS, we have derived the probability of error for our system model without CS to give us a theoretical reference. In the next section we present these mathematical expressions both for quasi-synchronous and the fully synchronous cases.

4.4 Quasi-Synchronous 2-PPM

Let $h_i \sim \mathcal{N}(0, 1)$ be the channel coefficients, $n_i \sim \mathcal{N}(0, \sigma^2)$ be the independent Gaussian distributed noise random variables and S be the signal energy. We can write $(h_i + n_i) \sim \mathcal{N}(0, 1 + \sigma^2)$ and $(\sqrt{S}h_i + n_i) \sim \mathcal{N}(0, S + \sigma^2)$ where \sqrt{S} is a deterministic value.

Let the received vector be represented as

$$\mathbf{x} = \left[(\sqrt{S}h_1 + n_1), (\sqrt{S}h_2 + n_2), \dots, (\sqrt{S}h_L + n_L), (n_{L+1}), \dots, (n_N) \right]^T. \quad (4.8)$$

If $U_1 = \|\mathbf{x}(1 : \frac{N}{2})\|^2$ and $U_2 = \|\mathbf{x}(\frac{N}{2} + 1 : N)\|^2$, then in light of (4.5), the probability of error for the case when the pulse is transmitted in the first half of T_f (see Figure 4.1), can be written as,

$$P_e = P_r(U_1 < U_2 | U_1). \quad (4.9)$$

The decision variable U_1 can be written as

$$U_1 = \sum_{i=1}^{\frac{N}{2}} |\sqrt{S}h_i + n_i|^2. \quad (4.10)$$

Given L channel taps and the condition that $L < \frac{N}{2}$, we can write (4.10) as

$$U_1 = \sum_{i=1}^L |\sqrt{S}h_i + n_i|^2 + \sum_{i=1}^{\frac{N}{2}-L} |n_i|^2. \quad (4.11)$$

Let

$$\begin{aligned} X_1 &= \sum_{i=1}^L |\sqrt{S}h_i + n_i|^2, \\ X_2 &= \sum_{i=1}^{\frac{N}{2}-L} |n_i|^2. \end{aligned} \quad (4.12)$$

X_1 and X_2 represent independent Chi-square random variables. The probability density functions (pdfs) of these variables can be written as [26]

$$p_{X_1}(x_1) = \frac{x_1^{\frac{n_1}{2}-1}}{\sigma_1^{n_1} 2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{x_1}{2\sigma_1^2}}, \quad n_1 = L, \quad \sigma_1^2 = S + \sigma^2 \quad (4.13)$$

$$p_{X_2}(x_2) = \frac{x_2^{\frac{n_2}{2}-1}}{\sigma_2^{n_2} 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{x_2}{2\sigma_2^2}}, \quad n_2 = \frac{N}{2} - L, \quad \sigma_2^2 = \sigma^2 \quad (4.14)$$

As a consequence U_1 is a random variable which is the sum of two Chi-square random variables i.e.,

$$U_1 = X_1 + X_2 \quad (4.15)$$

and the pdf of random variable U_1 is the convolution of (4.13) and (4.14). i.e.,

$$p_{U_1}(u_1) = \int_0^\infty p_{X_1}(x_1) p_{X_2}(u_1 - x_1) dx_1. \quad (4.16)$$

Using [28, Eq. (5.26)], a solution of (4.16) can be written as,

$$\begin{aligned} p_{U_1}(u_1) &= \frac{1}{2\sigma_1\sigma_2\Gamma\left(\frac{n_1+n_2}{2}\right)} \left[\frac{u_1}{2\sigma_1^2}\right]^{\frac{n_1-1}{2}} \left[\frac{u_1}{2\sigma_2^2}\right]^{\frac{n_2-1}{2}} \\ &\times e^{-\frac{u_1}{2\sigma_2^2}} {}_1F_1\left[\frac{n_1}{2}, \frac{n_1+n_2}{2}; \frac{(\sigma_2^2 - \sigma_1^2)^2}{2\sigma_1^2\sigma_2^2} u_1\right], \end{aligned} \quad (4.17)$$

where ${}_1F_1[\cdot, \cdot; \cdot]$ is the Kummer confluent hypergeometric function which is defined in [29, Eq. (9.210.1)] as,

$${}_1F_1[\alpha, \beta; z] = 1 + \frac{\alpha z}{\beta 1!} + \frac{\alpha(\alpha+1)z^2}{\beta(\beta+1)2!} + \frac{\alpha(\alpha+1)(\alpha+2)z^3}{\beta(\beta+1)(\beta+2)3!} + \dots \quad (4.18)$$

It can be seen that (4.18) is an ever increasing function. Since the largest value representable by Matlab is around $1.79e308$, it is not possible to get an accurate value for ${}_1F_1[\alpha, \beta; z]$ using this software especially for higher values of z .

The decision variable U_2 can be written as,

$$U_2 = \sum_{i=\frac{N}{2}+1}^N |n_i|^2. \quad (4.19)$$

So we can also write U_2 as a Chi-square distributed random variable with pdf,

$$p_{U_2}(u_2) = \frac{u_2^{\frac{n'}{2}-1}}{\sigma_w^{n'} 2^{\frac{n'}{2}} \Gamma\left(\frac{n'}{2}\right)} e^{-\frac{u_2}{2\sigma_w^2}} : n' = \frac{N}{2}, \sigma_w^2 = \sigma^2 \quad (4.20)$$

The probability of correct decision is:

$$P_r(U_2 < u_1 | U_1 = u_1) = \int_0^{u_1} p_{U_2}(u_2) du_2 \quad (4.21)$$

Using (4.20), we can write,

$$P_r(U_2 < u_1 | U_1 = u_1) = \int_0^{u_1} \frac{u_2^{\frac{n'}{2}-1}}{\sigma_w^{n'} 2^{\frac{n'}{2}} \Gamma\left(\frac{n'}{2}\right)} e^{-\frac{u_2}{2\sigma_w^2}} du_2 \quad (4.22)$$

$$P_r(U_2 < u_1 | U_1 = u_1) = \frac{\gamma\left(\frac{n'}{2}, \frac{u_1}{2\sigma_w^2}\right)}{\Gamma\left(\frac{n'}{2}\right)} \quad (4.23)$$

where $\gamma(.,.)$ is the lower-incomplete-gamma function and $\Gamma(.)$ is the gamma function such that $\gamma(n, u) = \int_0^u t^{n-1} e^{-t} dt$ and $\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt$, [29].

Since U_1 is a random variable with pdf given by (4.16), we need to average it out in (4.23) to get the probability of error. Thus,

$$\begin{aligned} P_e &= 1 - \int_0^\infty P_r(U_2 < u_1 | U_1 = u_1) p_{U_1}(u_1) du_1 \\ &= 1 - \int_0^\infty \frac{\gamma\left(\frac{n'}{2}, \frac{u_1}{2\sigma_w^2}\right)}{\Gamma\left(\frac{n'}{2}\right)} p_{U_1}(u_1) du_1 \end{aligned} \quad (4.24)$$

Since (4.24) is not in closed form, we solve it numerically for the purpose of simulations. Clearly it is computationally intensive.

4.5 Fully-Synchronous 2-PPM

In the case of full synchronization, we can write (4.10) as

$$U_1 = \sum_{i=1}^L |\sqrt{S} h_i + n_i|^2 \quad (4.25)$$

So U_1 is a Chi-square random variable. Similar to (4.19), U_2 can be written as

$$U_2 = \sum_{i=\frac{N}{2}+1}^{\frac{N}{2}+L} |n_i|^2. \quad (4.26)$$

The pdf for U_1 is given by

$$p_{U_1}(u_1) = \frac{u_1^{\frac{n_1}{2}-1}}{\sigma_1^{n_1} 2^{\frac{n_1}{2}} \Gamma(\frac{n_1}{2})} e^{-\frac{u_1}{2\sigma_1^2}} : n_1 = L, \sigma_1^2 = S + \sigma^2. \quad (4.27)$$

Similarly the pdf for U_2 is given by

$$p_{U_2}(u_2) = \frac{u_2^{\frac{n_2}{2}-1}}{\sigma_2^{n_2} 2^{\frac{n_2}{2}} \Gamma(\frac{n_2}{2})} e^{-\frac{u_2}{2\sigma_2^2}} : n_2 = L, \sigma_2^2 = \sigma^2 \quad (4.28)$$

The probability of a correct decision given that pulse is transmitted in the first half of the symbol, can be written as,

$$P_r(U_2 < u_1 | U_1 = u_1) = \int_0^{u_1} p_{U_2}(u_2) du_2, \quad (4.29)$$

which, similar to (4.22) and (4.23), can be simplified to,

$$P_r(U_2 < u_1 | U_1 = u_1) = \frac{\gamma(\frac{L}{2}, \frac{u_1}{2\sigma_2^2})}{\Gamma(\frac{L}{2})}. \quad (4.30)$$

Since U_1 is a Chi-square random variable, we need to average it out in (4.30). Thus the probability of error is;

$$\begin{aligned} P_e &= 1 - \int_0^\infty P_r(U_2 < u_1 | U_1 = u_1) p_{U_1}(u_1) du_1 \\ &= 1 - \int_0^\infty \frac{\gamma(\frac{L}{2}, \frac{u_1}{2\sigma_2^2})}{\Gamma(\frac{L}{2})} p_{U_1}(u_1) du_1 \end{aligned} \quad (4.31)$$

By using [29, Eq. (6.455.2)], we can have (4.31) reduced to the following closed form,

$$P_e = 1 - \frac{2\Gamma(L)}{L[\Gamma(\frac{L}{2})]^2} \left[\frac{\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} \right]^L {}_2F_1\left(1, L; \frac{L}{2} + 1; \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right) \quad (4.32)$$

where ${}_2F_1(., .; .; .)$ is the Gaussian hypergeometric function and is defined by [29, Eq. (9.14.2)] as,

$$\begin{aligned} {}_2F_1[\alpha, \beta; \kappa; z] &= 1 + \frac{\alpha.\beta}{\kappa.1} z + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\kappa(\kappa+1).1.2} z^2 \\ &+ \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{\kappa(\kappa+1)(\kappa+2).1.2.3} z^3 + \dots \end{aligned} \quad (4.33)$$

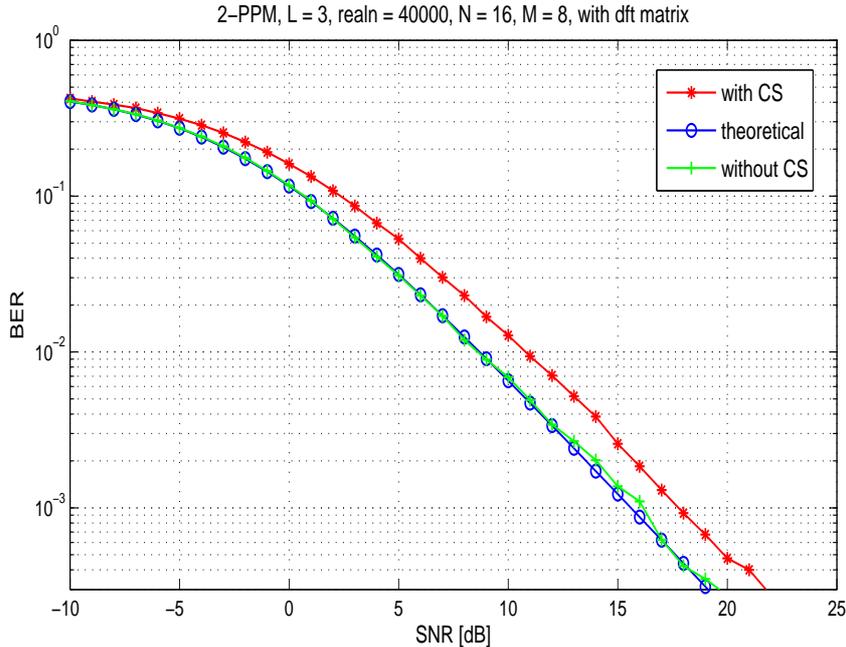


Figure 4.4: 2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, reconstruction algorithm = OMP

4.6 PPM Simulation

We consider 2-PPM. The symbol was represented by 16 components taken at the Nyquist rate (so $N = 16$). The channel coefficients were generated as real Gaussian random variables and three taps were used (i.e., $L = 3$). The noise variance was calculated from (4.7). In general we took compression ratio (i.e., M/N) around 50 percent (i.e., $M = 8$ in our case). Figure 4.4 describes the BER for the case of quasi-synchronous PPM where the measurement matrix used was the Fourier matrix (i.e., DFT matrix). Figure 4.4 displays three curves for Gaussian fading channel with white Gaussian noise, the BER obtained after using compressed sensing, the BER without CS (i.e., noncoherent detection of the received signal without processing through CS block, see Figure 4.3) and the theoretical BER plot for 2-PPM (according to (4.24)).

We see in Figure 4.4 that the simulated curve without CS closely follows the theoretical plot. The difference between the plot after CS and the reference plot is only around 2 dB. So for this quasi-synchronous setup, CS is performing very well. Figure 4.5 shows the BER simulation results for the fully synchronous case. The simulation parameters are the same as before. Here we see that the BER performance has improved by around 2 dB but the performance of CS is similar. The reason is that although we have changed our decision criterion, for CS, the reconstruction is still based on the same signal vector.

Figure 4.6 shows some results for quasi-synchronous PPM where we have used a Gaussian random matrix ($\sim \mathcal{N}(0, \frac{1}{\sqrt{M}})$) as measurement matrix instead of a DFT matrix. The result, as can be seen, is quite surprising. We see that the BER performance

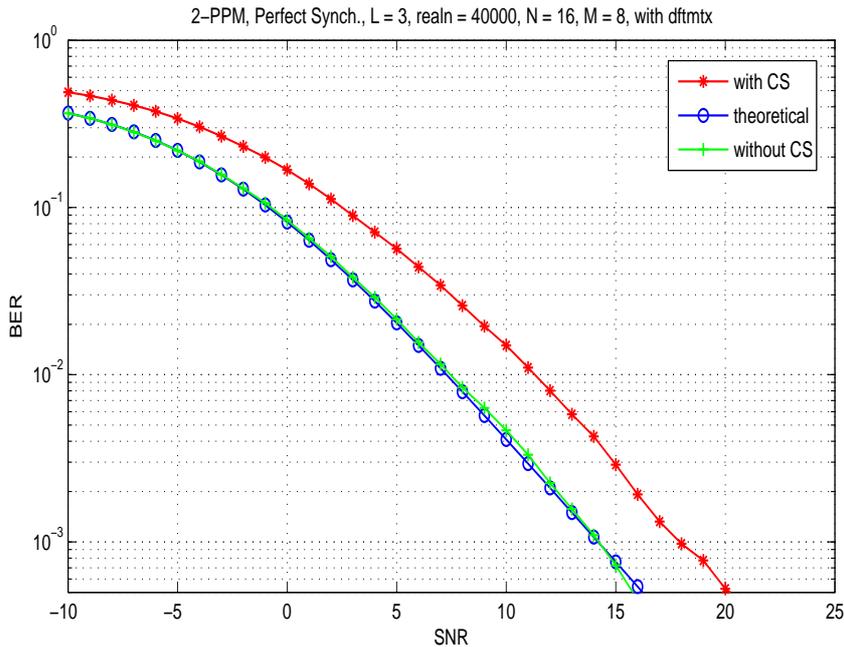


Figure 4.5: 2-PPM BER for fully-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, reconstruction algorithm = OMP

of CS saturates after 10 dB and that the margin of error is large even before that. Though all the simulation conditions are the same as for Figure 4.4 except for the measurement matrix, the performance is quite different. This means that CS with a random Gaussian measurement matrix and OMP as the recovery algorithm, could not reconstruct a 3-sparse signal (since $L = 3$, we have $K = 3$).

In the CS literature, Basis Pursuit has been suggested as a stable reconstruction method, so we used the l_1 -magic recovery program of [30] where a primal-dual algorithm was used to solve the standard BP problem as given in (2.4). Figure 4.7 shows the result of this algorithm. We see that even this does not help much and we still have the same kind of result. This shows that the problem may be lying in the choice of the measurement matrix. We shall talk about the behavior of these matrices in later sections.

Since the probability of error is an indicator for the performance of CS, we now present some results for the BER corresponding to different compression ratios for an SNR value. Figure 4.8 shows the curves for such a setup where the compression ratio varies from 0.1 to 1 for 0 dB SNR. We see that the CS curve catches the theoretical curve very quickly as we increase the compression ratio. Similarly, Figure 4.9 shows the case when SNR = 10 dB. Again we see that the CS curve approaches the theoretical limit at a compression ratio of 50 to 60 percent. Figure 4.8 and Figure 4.9 demonstrate that CS can achieve a very good performance result with a compression ratio of around 60 percent.

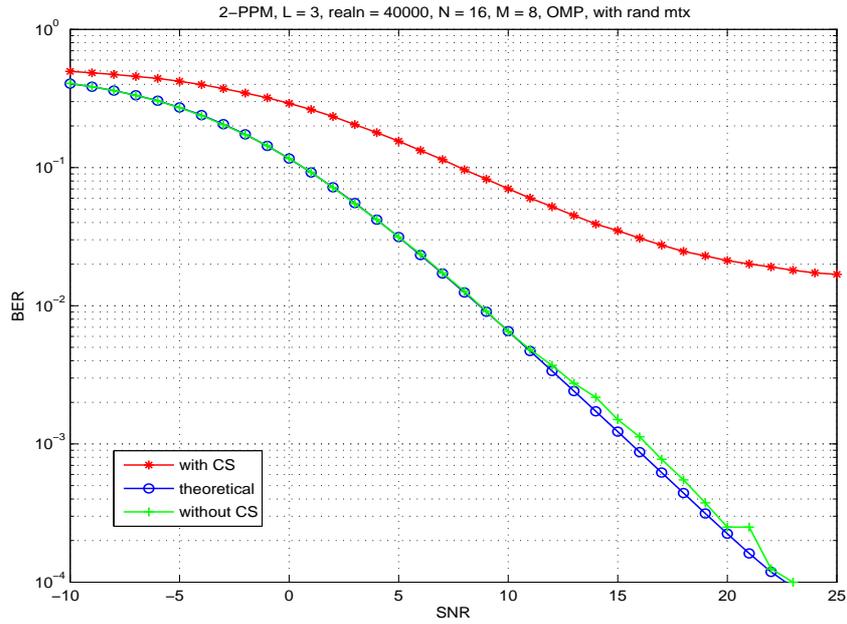


Figure 4.6: 2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = Gaussian random matrix, Reconstruction Algorithm: OMP

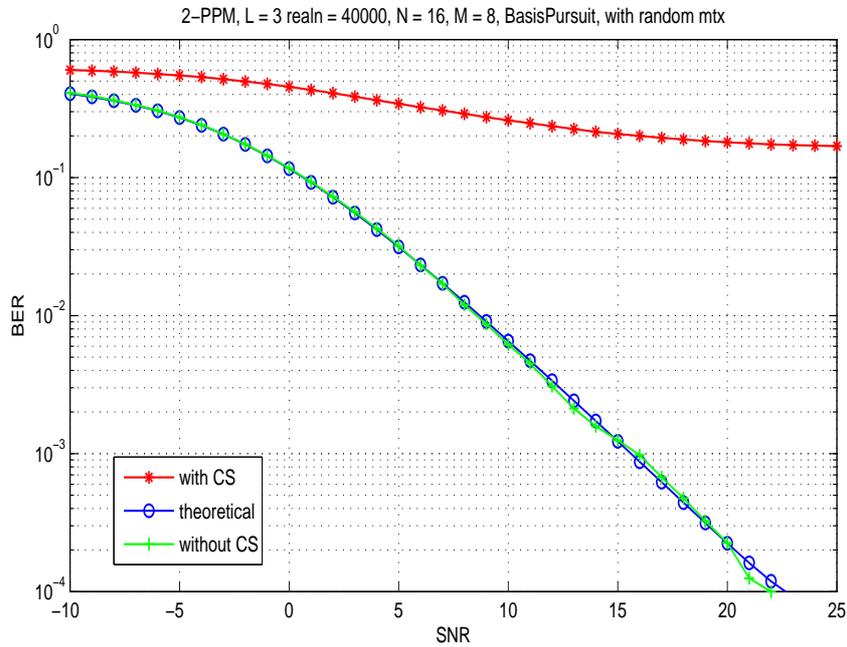


Figure 4.7: 2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = Gaussian random matrix, Reconstruction Algorithm: BP

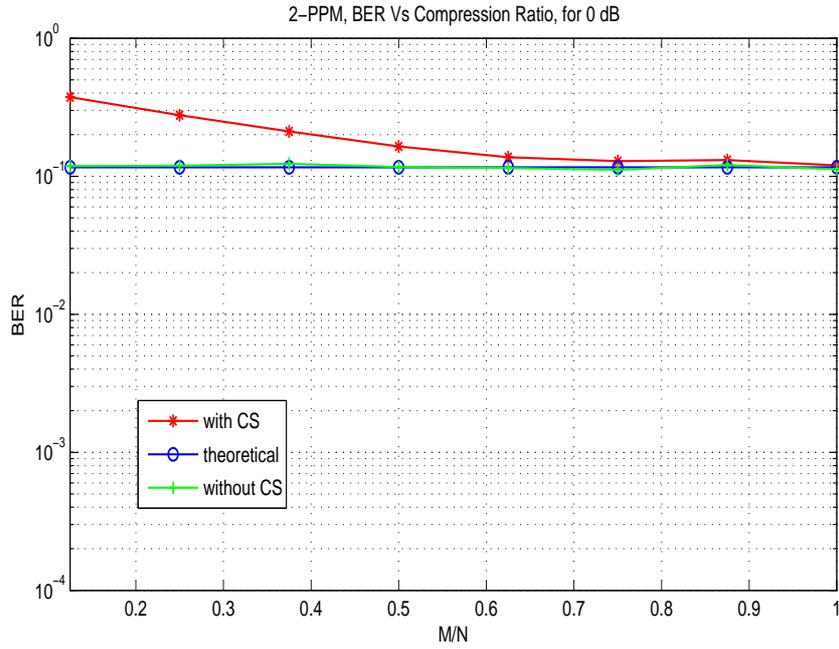


Figure 4.8: 2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, $N=16$, $M = 1:N$, SNR = 0dB, Reconstruction Algorithm: OMP

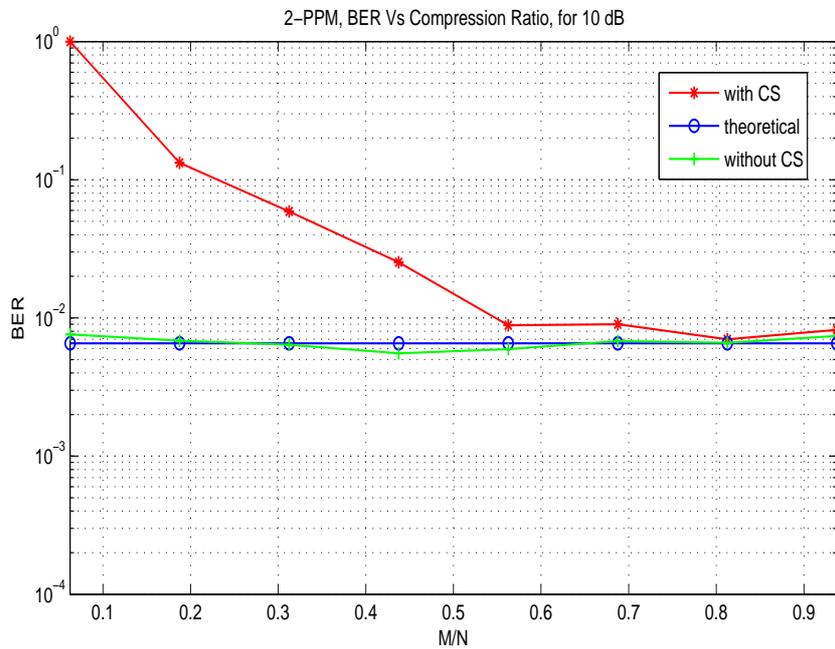


Figure 4.9: 2-PPM BER for quasi-synchronous scheme in a Gaussian fading channel with noise, Channel-taps (L) = 3, measurement matrix = DFT matrix, $N=16$, $M = 1:N$, SNR = 10dB, Reconstruction Algorithm: OMP

4.7 Conclusions

In this chapter we focused on the utilization of CS for PPM modulated signals. We saw that CS gives very good results for PPM modulated signals. We found that the measurement matrix plays a very important role in the reconstruction of the signal and finally the performance of CS. We presented the PPM signal models for quasi-synchronous and fully-synchronous cases along with their theoretical expressions for the probability of error. Our theoretical analysis portrays the exact behavior of our signal models.

In this chapter we focus on the utilization of CS for FSK modulated signals. We present the FSK signal models for single-tone and dual-tone cases, theoretical expressions for the probability of error and simulation results. We shall also describe our reconstruction algorithm.

5.1 FSK Scheme

\mathcal{M} -ary Frequency Shift Keying (\mathcal{M} -FSK) is a Spread-Spectrum (SS) modulation scheme, mostly used in Frequency Hopping Spread Spectrum (FHSS). Each \mathcal{M} -FSK symbol consists of \mathcal{M} different frequency slots. These slots contain equal-energy orthogonal signal waveforms [26]. So \mathcal{M} -FSK consists of \mathcal{M} basis functions and one basis function is transmitted at one time [31].

5.2 FSK Signal Model

We define the FSK signal \mathbf{a} in the frequency domain and transmit it in the time domain by taking its normalized inverse Discrete Fourier Transform (IDFT), i.e., $\mathbf{F}^H \mathbf{a}$. If \mathbf{F} represents the $N \times N$ normalized Discrete Fourier Transform (DFT) matrix then,

$$\mathbf{F}^H = (\mathbf{F})^{-1}, \quad (5.1)$$

is the normalized inverse Discrete Fourier Transform (IDFT) matrix, where $(\cdot)^H$ stands for the Hermitian of a matrix. Different symbols of $\mathbf{a}_{N \times 1}$ can be realized by making any of its components equal to one and the rest to zero. The transmitted signal convolves with the multipath fading channel and at the receiver, it suffers from the additive noise. As in the case of PPM, we consider the channel to consist of L paths and we assume the uniformly spaced channel coefficients to be Rayleigh faded with unit variance. The noise is assumed to be i.i.d Gaussian i.e., $n(t) \sim \mathcal{N}(0, \sigma^2)$. We can write the received signal as,

$$\mathbf{x} = \mathbf{H} \mathbf{F}^H \mathbf{a} + \mathbf{n}, \quad (5.2)$$

where $\mathbf{H}_{N \times N}$ represents the channel convolution matrix and $\mathbf{n}_{N \times 1}$, the noise vector. After compressed sensing, we can write (5.2) as;

$$\mathbf{y} = \mathbf{\Phi} \mathbf{H} \mathbf{F}^H \mathbf{a} + \mathbf{\Phi} \mathbf{n} \quad (5.3)$$

where $\mathbf{\Phi}_{M \times N}$ represents the measurement matrix and $\mathbf{y}_{M \times 1}$, the compressed signal vector as described before. Figure 5.1 gives an illustration of (5.3). Also, by using the fact that $\mathbf{F}^H \mathbf{F} = \mathbf{I}$, we can write (5.3) as,

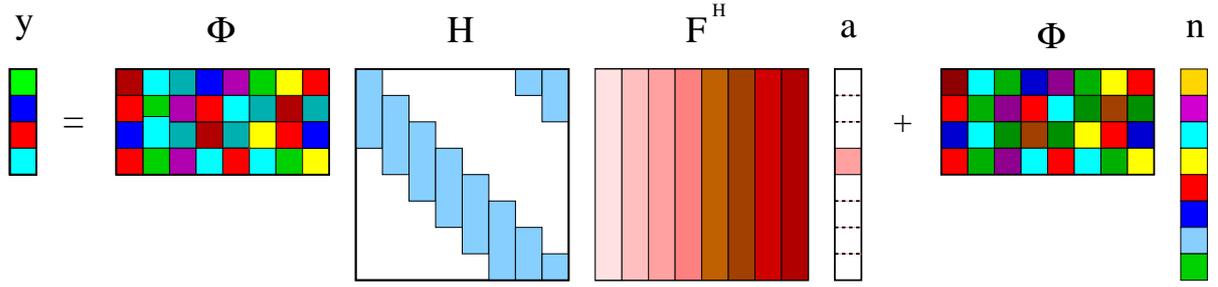


Figure 5.1: Matrix representation of CS for FSK (I)

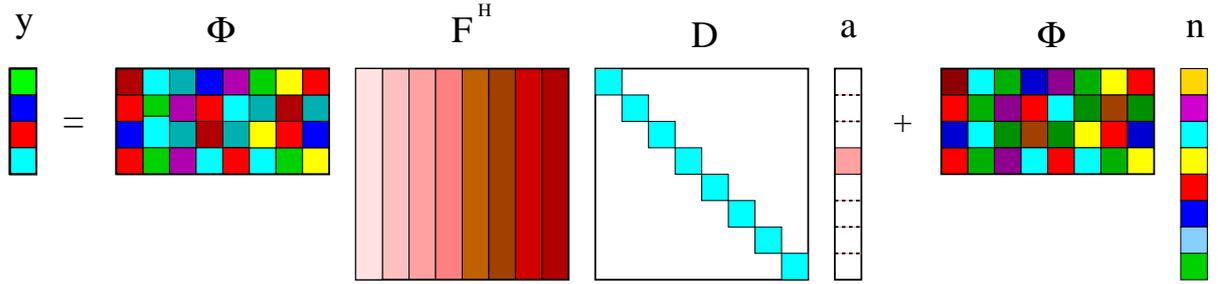


Figure 5.2: Matrix representation of CS for FSK (II)

$$\mathbf{y} = \Phi \mathbf{F}^H \mathbf{D} \mathbf{a} + \Phi \mathbf{n} \quad (5.4)$$

where $\mathbf{D}_{N \times N}$ is the diagonal matrix containing the eigenvalues of the channel matrix (since \mathbf{H} is a circular matrix), i.e.,

$$\mathbf{D} = \mathbf{F} \mathbf{H} \mathbf{F}^H \quad (5.5)$$

Figure 5.2 gives an illustration of (5.4). In terms of our model equations (2.1) and (2.2) we can write (5.4) as;

$$\mathbf{y} = \Phi \Psi \mathbf{s} + \Phi \mathbf{n} \quad (5.6)$$

where,

$$\begin{aligned} \Psi &= \mathbf{F}^H \\ \mathbf{s} &= \mathbf{D} \mathbf{a} \end{aligned} \quad (5.7)$$

Furthermore, we define the Signal to Noise Ratio (SNR) for the FSK modulated signals, as in (4.6). So the average SNR for N-FSK can be written as,

$$\gamma_{FSK} = \frac{SL}{k\sigma^2} : k = \log N \quad (5.8)$$

where S is the average signal energy, L is the average channel energy, σ^2 is the noise variance and k is the number of bits.

5.3 FSK Detection Rule

For FSK we use noncoherent energy detection. Noncoherent detection excludes the phase estimation from the FSK detection process and hence further simplifies the detection process. From (5.7), we can see that a decision on the vector \mathbf{s} is in fact the decision on the transmitted signal. This is how we have also excluded the process of channel estimation. Furthermore, although there are L channel taps, because of (5.5), the vector \mathbf{s} would be one-sparse if every FSK symbol is represented by a single tone (i.e., frequency signal). So the CS reconstruction process becomes much easier and fewer linear measurements are needed to estimate the signal.

The probability of error is calculated from the reconstructed signal vector \mathbf{s} . To make a decision of which carrier was transmitted, we estimate the energy for each component of \mathbf{s} . So if s_i represents an element of \mathbf{s} (after reconstruction) then we can write the decision variable u as,

$$u = \arg \max_{i \in \{1 \dots N\}} \|s_i\|_2^2 \quad (5.9)$$

Moreover, the case when two carriers have been transmitted, we call it as Dual tone FSK. Dual tone FSK has been presented in literature (e.g., [32]) as a solution for limited frequency bands to improve multiple access performance in frequency-hopped spread spectrum communications. The benefit of dual tone is in terms of the ability to increase multiple access but at the same time, the tone energy is half that of the conventional \mathcal{M} -FSK. Therefore, the benefit of increase in multiple access comes at the cost of degradation in noise performance. So dual-tone trades noise performance for multiple access enhancement [32].

Out of a total of N frequency bands (where $\mathcal{M} = N$ in this case), G unique combinations are possible where each combination consists of two frequency bands. So we can say as in [32],

$$G \leq \frac{(N)(N-1)}{2} \quad (5.10)$$

Thus the decision will be based on the maximum energy of any two unique combinations of elements of \mathbf{s} , i.e.,

$$\{u_1, u_2\} = \arg \max_{\substack{i,j \in \{1 \dots N\} \\ i \neq j}} \sum (\|s_i\|_2^2 + \|s_j\|_2^2) \quad (5.11)$$

But as is clear from (5.11), we cannot use all these combinations since these combinations can potentially produce dependent variables. For the purpose of simplicity, we only consider independent combinations. So we assume that out of only $N/2$ combinations one is transmitted in a consecutive fashion and we compare the energies of these many combinations to assess the combination with maximum energy. So we can say,

$$\{u_1, u_2\} = \arg \max_{(i,j) \in \{(1,2), (3,4), \dots, (N-1, N)\}} \sum (\|s_i\|_2^2 + \|s_j\|_2^2) \quad (5.12)$$

In the next section we present the theoretical expression for the probability of error for N-FSK for the case when only one carrier is transmitted per symbol and when two carriers are transmitted (i.e., dual tone).

5.4 N-FSK (single carrier)

Let U_1 represent the decision variable when first carrier in the FSK symbol was transmitted. So following the procedure of [26, section 14.4.3], we can write

$$U_1 = \left[|\sqrt{S}h e^{-j\phi} + n'_1| \right]^2 \quad (5.13)$$

where h is the respective diagonalized channel coefficient value (see (5.5)). ϕ is the respective carrier phase, S is the carrier energy and n'_1 is the noise component. Then we can write the pdf for the Chi-square distributed random variable U_1 as [26],

$$p_{U_1}(u_1) = \frac{u_1^{\frac{n_1}{2}-1}}{\sigma_1^{n_1} 2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{u_1}{2\sigma_1^2}}, \quad n_1 = 2, \quad \sigma_1^2 = SL + \sigma^2 \quad (5.14)$$

Similarly, the decision variable for any other element of \mathbf{s} , e.g., $U_i : i \neq 1$, can be written as,

$$U_i = [|n'_i|]^2, \quad (5.15)$$

Its respective pdf can be written as,

$$p_{U_i}(u_i) = \frac{u_i^{\frac{n_2}{2}-1}}{\sigma_2^{n_2} 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{u_i}{2\sigma_2^2}}, \quad n_2 = 2, \quad \sigma_2^2 = \sigma^2 \quad (5.16)$$

We can write (5.14) and (5.16) in the following simplified form,

$$p_{U_1}(u_1) = \frac{e^{-\frac{u_1}{2\sigma_1^2}}}{2\sigma_1^2} \quad (5.17)$$

$$p_{U_i}(u_i) = \frac{e^{-\frac{u_i}{2\sigma_2^2}}}{2\sigma_2^2} \quad (5.18)$$

The probability of correct decision given that the first carrier was transmitted is,

$$P_r(U_i < u_1 | U_1 = u_1) = \int_0^{u_1} p_{U_i}(u_i) du_i \quad (5.19)$$

Using (5.18),

$$P_r(U_i < u_1 | U_1 = u_1) = \int_0^{u_1} \frac{e^{-\frac{u_i}{2\sigma_2^2}}}{2\sigma_2^2} du_i \quad (5.20)$$

$$P_r(U_i < u_1 | U_1 = u_1) = 1 - e^{-\frac{u_1}{2\sigma_2^2}} \quad (5.21)$$

If all the decision variables U_i , where $i = 2, \dots, N$ are identically and independently distributed as (5.16), so we can write the joint pdf as,

$$P_r(U_2 < U_1, U_3 < U_1, \dots, U_N < U_1) = \left[1 - e^{-\frac{u_1}{2\sigma_2^2}}\right]^{N-1} \quad (5.22)$$

Probability of error can be calculated by averaging out U_1 ,

$$\begin{aligned} P_e &= 1 - \int_0^\infty \left[1 - e^{-\frac{u_1}{2\sigma_2^2}}\right]^{N-1} p_{U_1}(u_1) du_1 \\ &= 1 - \frac{1}{2\sigma_1^2} \int_0^\infty \left[1 - e^{-\frac{u_1}{2\sigma_2^2}}\right]^{N-1} e^{-\frac{u_1}{2\sigma_1^2}} du_1 \end{aligned} \quad (5.23)$$

Using [29, Eq. (3.312)], we were able to get (5.23) in the following closed form,

$$P_e = 1 - \frac{\sigma_2^2}{\sigma_1^2} B\left(\frac{\sigma_2^2}{\sigma_1^2}, N\right) \quad (5.24)$$

where $B(.,.)$ is the beta function such that $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt$ [29].

5.5 N-FSK (double carrier)

If two carriers, s_i and s_j , have been transmitted then according to our detection rule we need to find any two unique combinations of the vector \mathbf{s} which give us the maximum energy according to (5.11). Probability of error would be the probability of such a wrong decision. Similar to Section 5.4, our decision variable will have same pdf as (5.14) with twice its degrees of freedom, i.e.,

$$p_{U_1}(u_1) = \frac{u_1^{\frac{n_1}{2}-1}}{\sigma_1^{n_1} 2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{u_1}{2\sigma_1^2}}, \quad n_1 = 4, \quad \sigma_1^2 = SL + \sigma^2 \quad (5.25)$$

which reduces to the following simplified form,

$$p_{U_1}(u_1) = \frac{u_1 e^{-\frac{u_1}{2\sigma_1^2}}}{4\sigma_1^4} \quad (5.26)$$

The pdf for the second decision variable U_i will be similar to (5.16) with two times its degrees of freedom, i.e.,

$$p_{U_i}(u_i) = \frac{u_i^{\frac{n_2}{2}-1}}{\sigma_2^{n_2} 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{u_i}{2\sigma_2^2}}, \quad n_2 = 4, \quad \sigma_2^2 = \sigma^2 \quad (5.27)$$

which can be simplified to,

$$p_{U_i}(u_i) = \frac{u_i e^{-\frac{u_i}{2\sigma_2^2}}}{4\sigma_2^4} \quad (5.28)$$

The probability of correct decision given that the carrier of decision variable U_1 were transmitted is,

$$\begin{aligned} P_r(U_i < u_1 | U_1 = u_1) &= \int_0^{u_1} p_{U_i}(u_i) du_i \\ &= \frac{1}{4\sigma_2^4} \int_0^{u_1} u_i e^{-\frac{u_i}{2\sigma_2^2}} du_i \end{aligned} \quad (5.29)$$

Using [29, Eq. (3.351.1)], (5.29) simplifies to,

$$P_r(U_i < u_1 | U_1 = u_1) = 1 - e^{-\frac{u_1}{2\sigma_2^2}} \left(1 + \frac{u_1}{2\sigma_2^2}\right) \quad (5.30)$$

The decision variable U_1 , has to be compared with every other two unique combinations of the elements of \mathbf{s} . So if all other variables are independently distributed and there are G such variables then we can write the joint pdf as,

$$P_r(U_2 < U_1, U_3 < U_1, \dots, U_G < U_1) = \left[1 - e^{-\frac{u_1}{2\sigma_2^2}} \left(1 + \frac{u_1}{2\sigma_2^2}\right)\right]^G \quad (5.31)$$

And the probability of error can be written as,

$$P_e = 1 - \int_0^\infty \left[1 - e^{-\frac{u_1}{2\sigma_2^2}} \left(1 + \frac{u_1}{2\sigma_2^2}\right)\right]^G p(u_1) du_1 \quad (5.32)$$

Using (5.26), we can write (5.32) as,

$$P_e = 1 - \frac{1}{4\sigma_1^4} \int_0^\infty \left[u_1^{\frac{1}{G}} e^{-a_1 u_1} - u_1^{\frac{1}{G}} e^{-a_2 u_1} - \frac{u_1^{\frac{G+1}{G}}}{2\sigma_2^2} e^{-a_2 u_1} \right]^G du_1 \quad (5.33)$$

where,

$$\begin{aligned} a_1 &= \frac{1}{2\sigma_1^2 G} \\ a_2 &= \left(\frac{1}{2\sigma_2^2} + \frac{1}{2\sigma_1^2 G} \right) \end{aligned} \quad (5.34)$$

5.6 FSK Simulation

For the simulation of FSK, we consider first, 8-FSK. so vector \mathbf{a} consists of 8 components representing eight subcarriers in frequency domain (so $N = 8$) and let $M = 4$ (i.e., 50

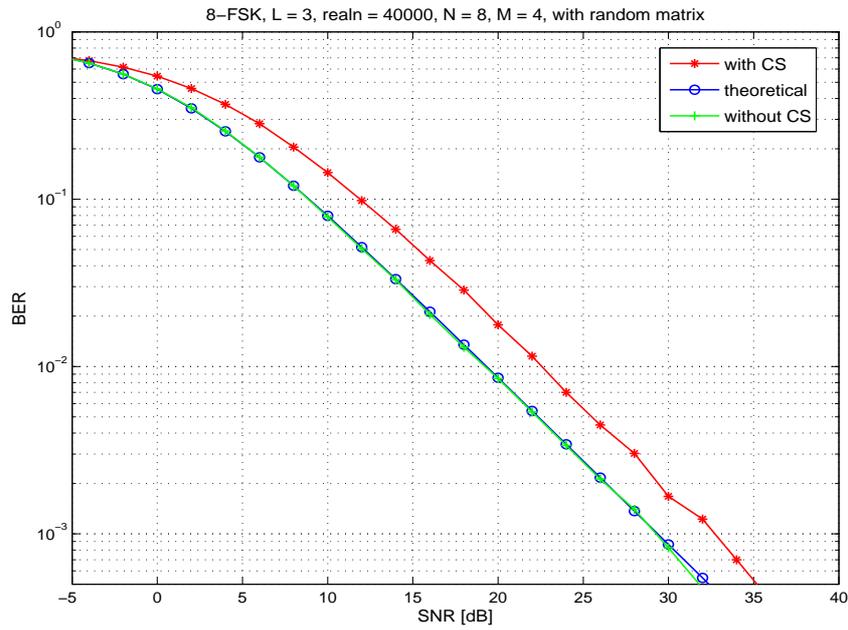


Figure 5.3: 8-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=8$, $M=4$, measurement matrix = random matrix, Reconstruction Algorithm: OMP

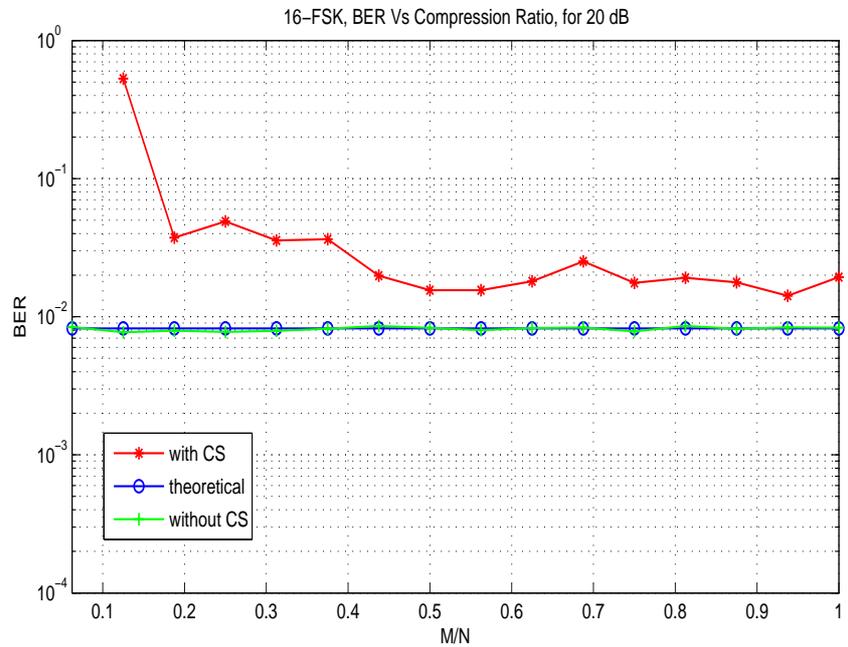


Figure 5.4: 16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=16$, $M=[1:N]$, measurement matrix = random matrix, Reconstruction Algorithm: OMP

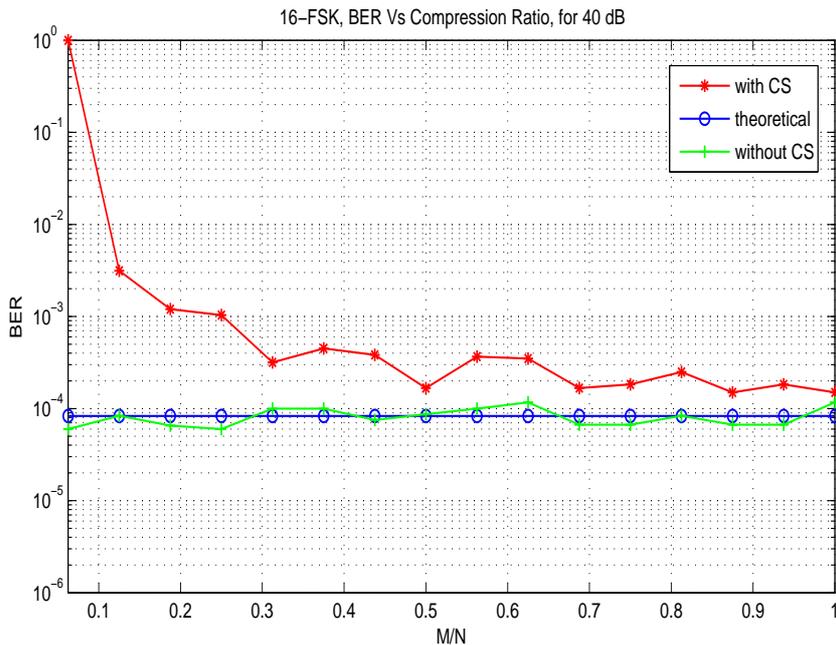


Figure 5.5: 16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=16$, $M = [1:N]$, measurement matrix = random matrix, Reconstruction Algorithm: OMP

percent compression). We considered complex Rayleigh fading channel. Noise variance was calculated according to (5.8). Figure 5.3 shows the simulation results for 40,000 realizations of noise and channel when 3 channel taps were used. The figure shows three curves, similar to the case of PPM. One curve with using CS and the other without CS while the third one is the theoretical FSK curve for a Rayleigh fading channel according to (5.24). We see that the plot without CS follows the theoretical curve closely. The difference between the plot with CS and without CS is around three dBs. This a very good performance with 50 percent compression ratio.

Furthermore, Figure 5.4 and Figure 5.5, show simulation results for probability of error as a function of compression ratio for 20 and 40 dBs. For this scenario, we considered $N = 16$, $M = 1 : N$ and measurement matrix as Gaussian matrix. We see that for a compression ratio of above 40 percent, results of CS are very close to the reference plots. In these figures we observe some fluctuations in CS plots. This can be attributed to the randomness of the Gaussian matrix. To clear this point we do the same simulations with the only difference of a structured matrix (here, Identity matrix) as our measurement matrix. Figure 5.6 and Figure 5.7 show the simulation results for probability of error as a function of compression ratio for 20 and 40 dBs with Identity measurement matrix. The results show more smoother performance that the Gaussian matrix. Also, above a compression ratio of 40 percent, CS results are comparable to the theoretical plot. Therefore, we can say that for single tone FSK, CS gives good results both for random as well as for structured measurement matrices.

For the dual tone FSK, i.e., the sparsity level is of order 2, the simulation results are shown in Figure 5.8 and 5.9 with the Gaussian and identity matrix as the measure-

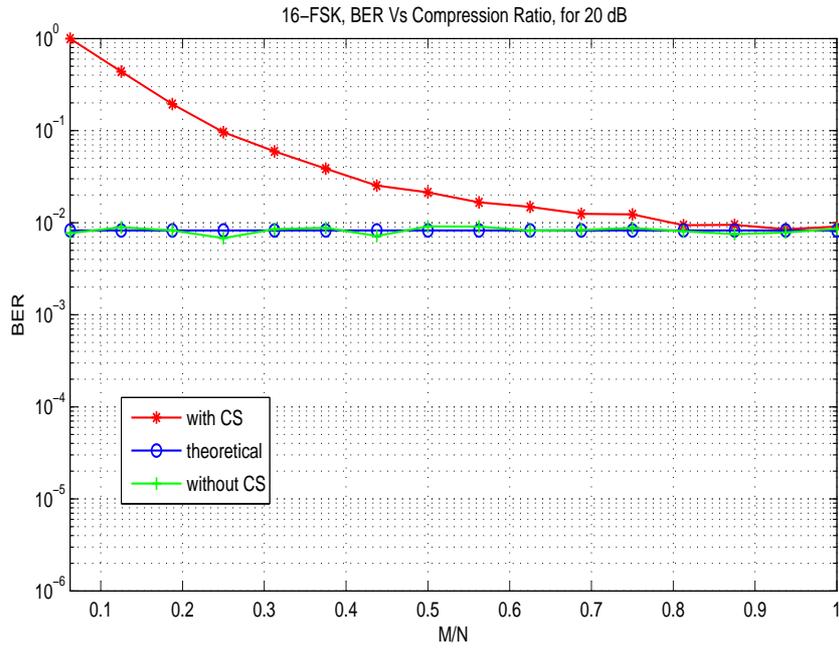


Figure 5.6: 16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=16$, $M = [1:N]$, measurement matrix = Identity matrix, Reconstruction Algorithm: OMP

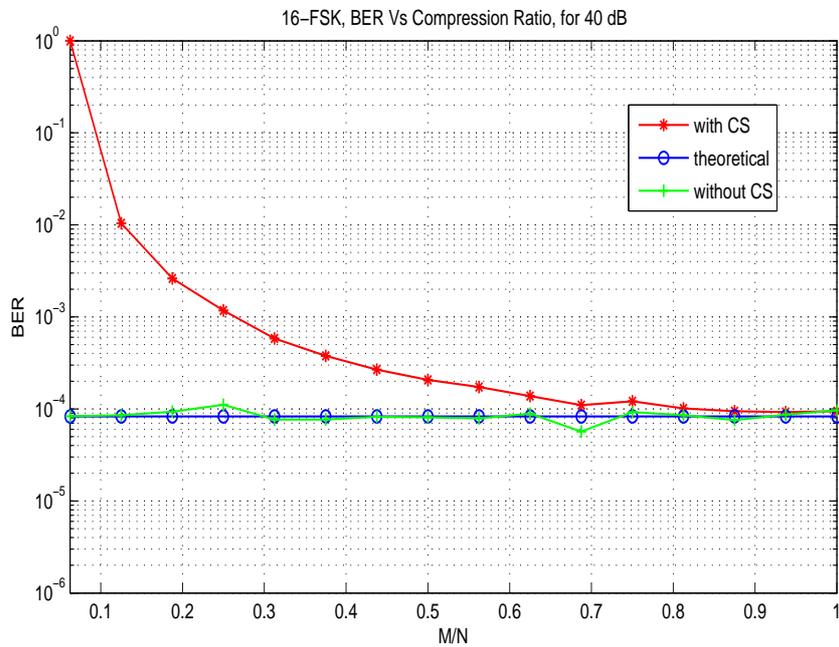


Figure 5.7: 16-FSK, Single Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=16$, $M = [1:N]$, measurement matrix = Identity matrix, Reconstruction Algorithm: OMP

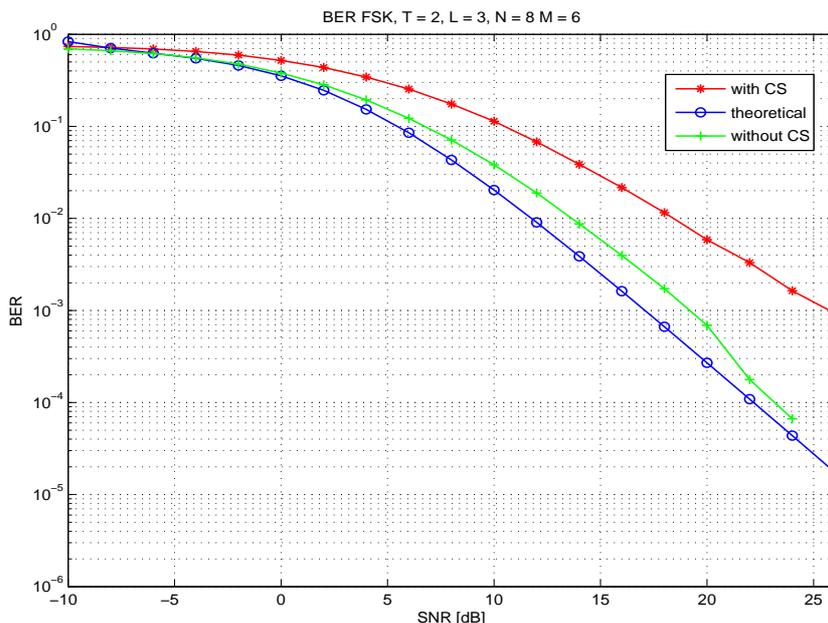


Figure 5.8: 8-FSK, Double Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=8$, $M = 6$, measurement matrix = Gaussian matrix, Reconstruction Algorithm: OMP

ment matrices respectively. We can see from the simulation results that the Gaussian matrix is not giving us good performance in comparison to the structured matrix. The performance of Gaussian matrix is similar to that of PPM for sparsity order of more than one as explained in previous sections.

5.7 Signal Reconstruction

For signal reconstruction, we have tried a number of different algorithms. These algorithms range from Basis Pursuit (BP) [30] and Orthogonal Matching Pursuit (OMP) [25] to Matching Pursuit (MP) [33], Flexible Tree-search-Based OMP (FTB-OMP) [34] and Sequential OMP (Seq-OMP) [35]. Though BP is the optimal choice for signal recovery but we found OMP feasible from a practical point of view. BP is more powerful than OMP in situations where the measurement matrices used are Gaussian or Bernoulli matrices [25]. The Greedy pursuit (i.e., OMP) is advantageous in terms of computational cost especially when the signal vector is highly sparse [25]. OMP did better than BP, in terms of computational time, for our simulations (the results can be seen in Figure 4.6 and Figure 4.7). Also OMP is easier to implement than BP [25]. Furthermore, the recovery performance of OMP is better than MP, FTB-OMP and Seq-OMP for our system setup. Here we briefly describe the our slightly modified OMP.

The signal is approximated by projecting the received signal onto its basis functions. OMP is an improved form of MP, in the sense that the projections onto basis

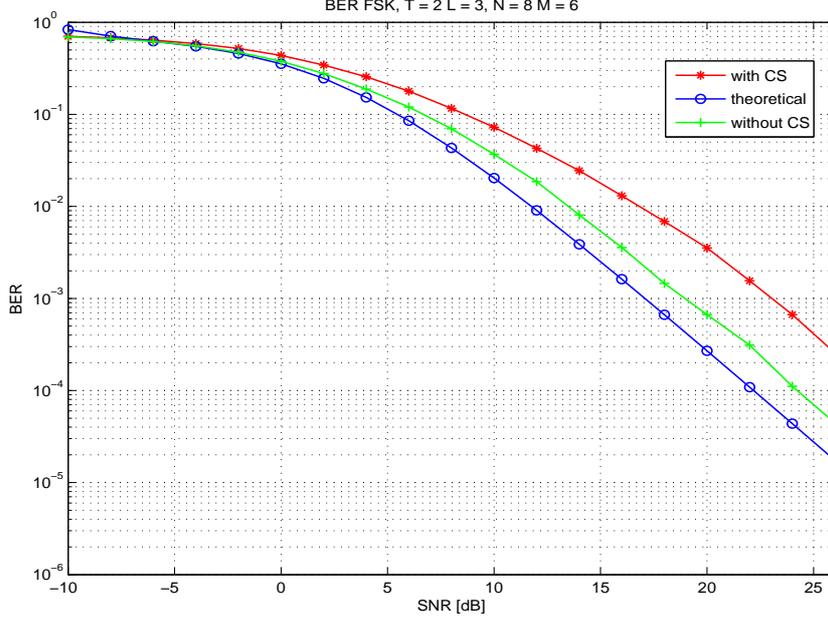


Figure 5.9: 8-FSK, Double Carrier, Channel-taps (L) = 3, Rayleigh fading channel with noise, $N=8$, $M = 6$, measurement matrix = Identity matrix, Reconstruction Algorithm: OMP

functions are optimized. The received signal is projected onto a basis matrix. The aim is to find the basis vectors which best represent the signal. A set of basis vectors is selected based on the correlation with the signal. The basis vector which gives the least residual becomes part of the signal representation. Every new selection depends on the previously selected basis functions as well. Every basis vector is selected once. This process is iterated a number of times as per requirement for the algorithm to converge.

In our case, the basis matrix used for reconstruction is the product of the sparsity matrix and the measurement matrix i.e. $\Theta_{M \times N}$. So we can write it as;

$$\Theta = \Phi\Psi \quad (5.35)$$

The $[M \times 1]$ dimensional vectors, θ_i where $i = 1, 2, 3, \dots, N$, constitute the matrix Θ , i.e.,

$$\Theta = [\theta_1 \theta_2 \theta_3 \dots \theta_N] \quad (5.36)$$

Let Ω be the set of vectors which best represent the signal. The set will have a new index of the suitable vector with every new iteration. Also Ω is an empty set at the time of initialization, i.e., $\Omega_0 = \{\emptyset\}$.

Now we present the steps involved in the reconstruction algorithm;

1. The received signal \mathbf{y} itself is considered as the initial residual error \mathbf{r} . i.e. $\mathbf{r}_0 = \mathbf{y}$

2. At the k -th iteration \mathbf{r}_k is projected onto all columns of Θ and the maximum correlation value is selected to shortlist a number of basis vectors.;

$$c_{k+1} = \max |\langle \mathbf{r}_k, \boldsymbol{\theta}_i \rangle| : i = 1, 2, 3, \dots, N - k \quad (5.37)$$

Let,

$$\Lambda_{k+1} = \{i : |\langle \mathbf{r}_k, \boldsymbol{\theta}_i \rangle| \geq \alpha c_{k+1}; i = 1, 2, 3, \dots, N - k\} \quad (5.38)$$

where the set Λ_{k+1} contains the indices of the candidate basis vectors depending on the value of α which ranges from 0 to 1.

3. Among the vectors indexed by Λ_{k+1} , the best candidate is the one which (in union with the already selected most suitable basis functions Ω_k) gives the minimum residual after being projected upon by the current residual;

$$\lambda_{k+1} = \arg \min_{\lambda \in \Lambda} \{ \|\mathbf{r}_k - \mathbf{P}_{span\{\boldsymbol{\theta}_t : t \in \Omega_k\}} \mathbf{r}\|_2 \} \quad (5.39)$$

4. λ_{k+1} is made part of Ω_k so;

$$\Omega_{k+1} = \Omega_k \cup \lambda_{k+1} \quad (5.40)$$

5. The residual is updated;

$$\mathbf{r}_{k+1} = \mathbf{y} - \mathbf{P}_{span\{\boldsymbol{\theta}_t : t \in \Omega_{k+1}\}} \mathbf{y} \quad (5.41)$$

6. We remove vector $\boldsymbol{\theta}_{k+1}$ from the matrix Θ and record its actual position in the original basis matrix. Then repeat from step 2 and continue till a certain level of residual error (which is a fraction of the received signal energy) and/or maximum no. of allowed iterations is achieved.

7. In the end the signal is approximated. The recorded positions of the used basis functions correspond to the components of \mathbf{s} that have to be assigned values. These values are generated by the following equation.

$$\mathbf{z} = \Theta^\dagger \mathbf{y} \quad (5.42)$$

where Θ represents the reduced form of matrix Θ , containing vectors indexed by Ω . Also, Θ^\dagger represents Moore-Penrose pseudo inverse of Θ , i.e.,

$$\Theta^\dagger = (\Theta^H \Theta)^{-1} \Theta^H. \quad (5.43)$$

5.8 Conclusions

In this chapter we focused on the utilization of CS for FSK modulated signals. We saw that CS gives very good results for FSK modulated signals. We presented the FSK signal models for single-carrier and dual-carrier cases along with their theoretical expressions for the probability of error. Our theoretical analysis portrays the exact behavior of our signal models. Furthermore, the role of measurement matrix became vivid especially for the case of dual-carrier FSK. Overall, CS is still valid for the case of FSK. Finally we presented our reconstruction algorithm.

In this chapter we present an analysis of CS for the PPM and FSK signals.

6.1 CS Analysis

We presented different scenarios for the application of CS for PPM and FSK signal. We used both random and structured measurement matrices. We used lower values of N (i.e., 8, 16) and hence even lower values of linear measurements M . We found that for the sparsity of order one (i.e., $K = 1$), both kinds of measurement matrices performed well. When the order of sparsity was increased from one to two or three, the random matrices performed poorly. Even then, the structured matrices had still acceptable results. This phenomenon points to the fact that the random matrices do not fulfill the basic requirements of RIP and/or non coherence.

For incoherence, it is required that the rows (ϕ_k) of the measurement matrix Φ should not be able to sparsely represent the columns (ψ_j) of the basis matrix Ψ . In other words, we can say that the rows of the measurement matrix should not be concentrated but rather 'spread out' in the Ψ domain [21]. If $\mu(\Theta)$ represents the mutual coherence between Φ and Ψ then, following [21], we can write,

$$\mu(\Theta) = \max_{k,j} | \langle \phi_k, \psi_j \rangle | : \Theta = \Phi\Psi \quad (6.1)$$

Lack of incoherence can have direct impact upon the required number of measurements [21]. Furthermore, we require that the measurement matrix should be ideally orthonormal. Since the rows of measurement matrix are much less than the columns, this may not be possible. So Restricted Isometry Property (RIP) gives us a less strict condition. According to (2.3), the measurement matrix should not amplify the signal vector beyond a specific value related to δ_K (where δ_K is determined for a specific sparsity order).

It has been shown in [2], [1] and [20] that given the number of measurement as in (2.5), the Gaussian matrix fulfills RIP with high probability. Recently, [36] has determined that for a Gaussian measurement matrix to have sufficient RIP, $\delta_{2K} < \sqrt{2} - 1$ in (2.3). But the most important thing is that the number of measurements in (2.5) are sufficient for the 'asymptotic' case. Where the value of N ranges from several hundreds to several thousands. Even then these measurements are not the most optimum value but rather sufficient for good reconstruction. We carried out simulations to prove this point and we found it to hold true (see Figure 6.1).

To put the above conditions plainly, we require at least two conditions to be fulfilled for a measurement matrix to fully recover the sparse vector.

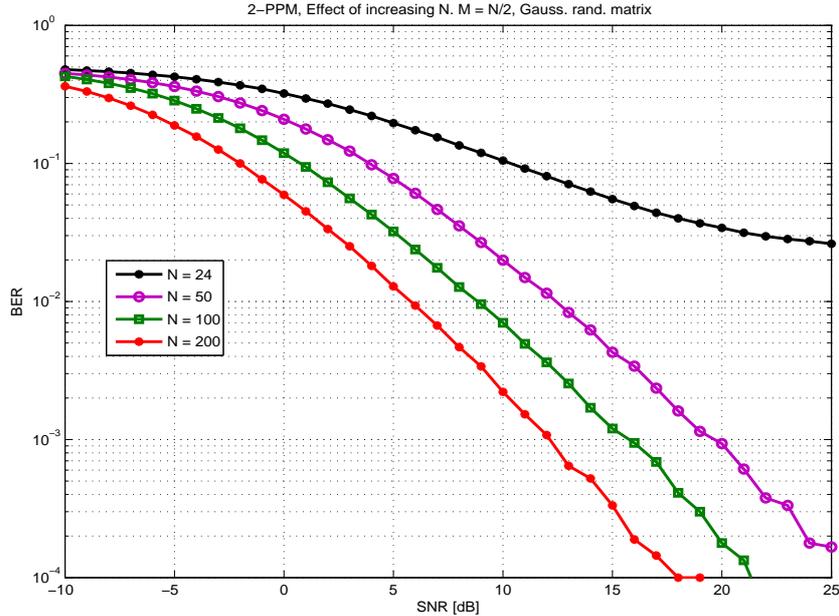


Figure 6.1: 2-PPM, Quasi-Synchronous, Channel-taps (L) = 3, Sparsity (K) = 3, Gaussian fading channel with noise, $M = N/2$, measurement matrix = random matrix, Reconstruction Algorithm: OMP

1. Every collection of the K columns of the measurement matrix must form a full rank matrix to be invertible so that we can get K -nonzero values.
2. Every K columns of measurement matrix must span a different subspace than the rest of the columns so that the positions of the K nonzero elements are known.

We can see that the above two properties are readily fulfilled by both Fourier and Identity (structured) matrices by selecting at least their first M rows. Whereas, for Gaussian matrix to fulfill these properties, it has to have high dimensional matrix. Though Gaussian matrix has many properties which made it an element of choice for researchers but these properties are better utilized in the asymptotic sense for Compressed Sensing.

The performance of CS, is heavily dependent on the measurement matrix used. RIP is one way of characterizing the measurement matrix as a sufficient condition. There are other ways to describe the conditions for the measurement matrix so that (2.4) can give sparse approximation with an overwhelming probability. In [6], a K -neighborly polytope characterization for the measurement matrix is given. Similarly, in [11], a null-space characterization of the measurement matrix is given. Both these characterizations are sufficient and necessary conditions and assume $N \rightarrow \infty$. Similarly, a geometric functional analysis for the measurement matrices has been done in [7]. A comparison of the different methods of analysis for the random measurement matrices has been carried out in [5]. For the sake of insight and then their relation to our proposition, we define below some new notations.

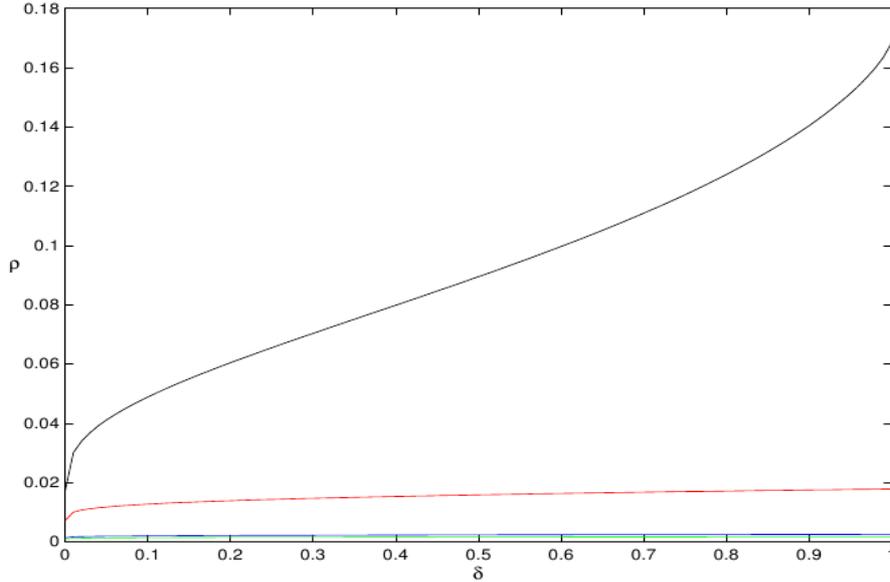


Figure 6.2: Lower bounds on sparsity, [5] (black for [6], red for [7] and green for [5])

Let $\rho = \frac{K}{M}$ and $\zeta = \frac{M}{N}$ (i.e., the compression ratio). Furthermore, let $\rho(\zeta)$ be a function of the compression ratio. Then, in terms of [6], [7] and [5], we can write a generalized expression for the bound on ρ , so that (2.4) can recover the sparse signal with high probability,

$$\rho \leq (1 - \epsilon)\rho(\zeta) \quad (6.2)$$

where $\epsilon > 0$ is a small constant. In [5], a graphical description for the bounds on ρ as a function of ζ for all the above cited methods has been given (as shown in Figure 6.2 where the variable δ stands for ζ in our case). There we see that the bounds given by [7] and [5] are the tightest and guarantee better results than [6].

Here we describe some numerical results relating to these bounds for clear intuition. From the discussion above, we can write,

$$K \leq \rho M = \rho \zeta N = (1 - \epsilon)\rho(\zeta)\zeta N. \quad (6.3)$$

(6.3) shows that the number of nonzero elements that can be recovered, grow linearly with N for a fixed compression ratio (while $N \rightarrow \infty$). This relation can indirectly be interpreted in terms of the minimum value of N needed to guarantee the sparse recovery of order K with high probability.

Considering $\epsilon = 0.011$ and $\zeta = 0.5$, according to [6], the relation between K and N can be calculated as, $K \leq (0.044708)N$. Similarly, this relationship, according to [7], amounts to $K \leq (0.01483)N$ and [11] puts it at $K \leq (0.0016)N$. We tabulate below the exact numerical values of N for a required order of sparsity K for the two extreme bounds. Tables 6.1 and 6.2 show these results. From Table 6.2, it is clear that for a sparsity of order 3, N should be greater than 67. These results vindicate our proposition regarding the dimensions of the random measurement matrices. Figure 6.1

shows the effect of increasing values of N on the probability of error which in turn is an indicator for the recoverability of CS with the Gaussian measurement matrix.

Table 6.1: Sparsity according to [11]

$K = 0.0016N$	
K	N
1	625
3	1875
5	3125

Table 6.2: Sparsity according to [6]

$K = 0.044708N$	
K	N
1	22
3	67
5	112

6.2 Conclusions

In this chapter we analyzed the performance of CS in terms of sparsity of the signals. We saw that the compression ratio is not an enough parameter for the true performance of CS. It is also vital that the both M and N should be asymptotic in nature, especially for the random measurement matrices to work. In the literature, this aspect has not been really emphasized to its true degree of import.

In this chapter we shall analyse the applicability of CS for Ultra Wide Band (UWB) based on PPM. First, we present the signal models, theoretical expressions, a comparison of our models with [3]. Secondly, we shall present the application of CS to PPM based UWB.

7.1 Signal Model

We use 2-PPM for our UWB transmitted signal. Here we consider a single-user transmission. We shall use Gaussian as well as, more importantly, IEEE 802.15.3a [12] channel models. Now we mention some of the notations that will be used in our model. Some of these have been borrowed from [3]. The PPM signal interval is denoted by T_{sig} . If one symbol contains multiple frames (i.e. N_f), each with frame time T_f , then $T_{sig} = N_f T_f$. So in this case one signal interval contains multiple copies of the same symbol. If $N_f = 1$ then $T_{sig} = T_f$. $h(t)$ represents the channel impulse response and T_{mds} represents the maximum delay spread of the channel. Figure 7.1 gives a view of a PPM symbol consisting of two frames.

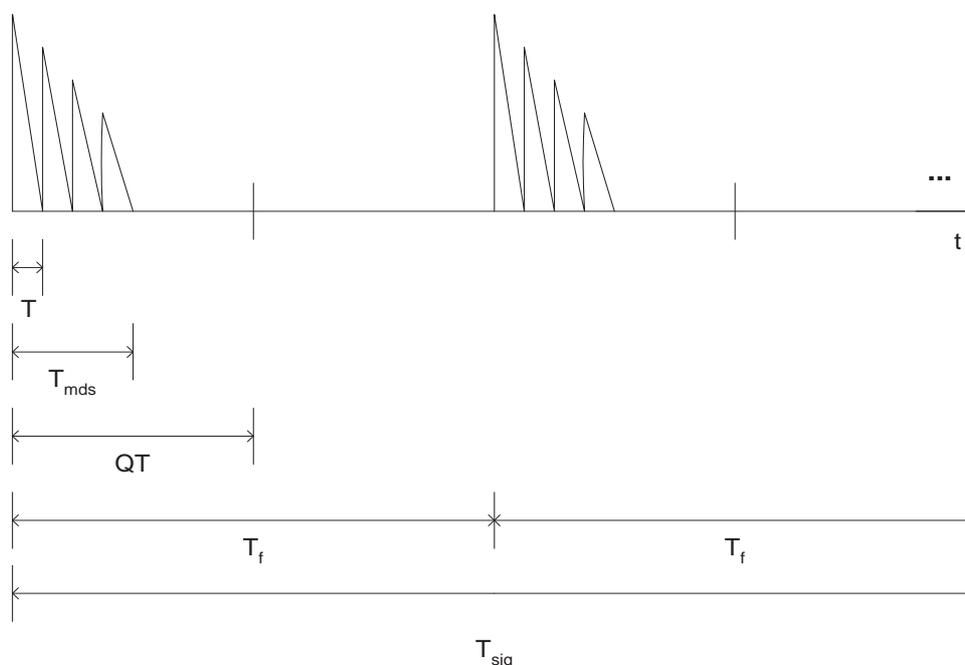


Figure 7.1: 2-PPM UWB, Multiple frames Symbol, $N_f = 2$

Our E_b/N_0 is defined as below.

$$\frac{E_b}{N_0} = \frac{E_s}{\sigma^2/B} \quad (7.1)$$

while

$$E_s = N_f E_h$$

E_s is the symbol energy (which in our case is equal to the bit energy i.e., E_b), E_h is the channel energy B is the receiver bandwidth and σ^2 is the noise variance. The channel energy can be defined as [3],

$$E_h = \int_{-\infty}^{+\infty} h^2(t) dt \quad (7.2)$$

$$\simeq \int_0^{\Delta} h^2(t) dt$$

where Δ is the integration interval. It is our design parameter and $\Delta = LT$. We have assumed that the channel delay spread $T_{m ds}$ is less than the pulse interval $QT := T_f/2$ (see Figure 7.1), but in the case where there is no perfect channel synchronization, the integration interval can make the difference. Its prudent choice is important to capture actual signal energy while keeping excessive noise accumulation at bay.

The IEEE 802.15.3a channel model [12] is based on the S-V model [37]. It can briefly be described as follows (for one realization),

$$h(t) = X \sum_{i=0}^I \sum_{k=0}^K \alpha_{k,i} \delta(t - T_i - \tau_{k,i}) \quad (7.3)$$

$$P_r(T_i|T_{i-1}) = \Lambda \exp[-\Lambda(T_i - T_{i-1})], \quad i > 0$$

$$P_r(\tau_{k,l}|\tau_{(k-1),i}) = \lambda \exp[\tau_{(k)} - \tau_{(k-1),i}], \quad k > 0$$

where $\alpha_{k,i}$ is the multipath gain coefficient, T_i is the delay of the i -th cluster, $\tau_{(k),i}$ is the delay of k -th multipath component in relation to the i -th cluster arrival time T_i and X caters for the shadowing effect which is distributed log-normally. Λ is the cluster arrival rate and λ is the ray arrival rate within each cluster. IEEE 802.15.3a considers four channel models i.e., CM1-CM4. We shall use CM1 and CM3 which correspond to Line of Sight (LOS) and Non Line of Sight (NLOS) signal transmission situations. Given Γ and γ as the cluster and ray decay factors respectively, Table 7.1 briefly describes the values for channel parameters.

Table 7.1: IEEE Channel Parameters [12]

Model Parameters	CM 1	CM 3
Λ [1/nsec]	0.0233	0.0667
λ [1/nsec]	2.5	2.1
Γ	7.1	14.00
γ	4.3	7.9

Finally the pulse that we use for the purpose of simulation is the second derivative of a Gaussian pulse with unit energy. We construct the pulse such that the pulse duration is 1 nsec. The mathematical expression for the second derivative of a Gaussian pulse is given below,

$$\begin{aligned}
p(t) &= P \frac{-2}{\tau'} \left(1 - \frac{2t^2}{\tau'^2} \right) \exp \left[\frac{-t^2}{\tau'^2} \right] \\
\text{s.t.} & \\
P &= \tau' \sqrt{\frac{\tau' \mathbf{E}}{3\sqrt{\pi/2}}}
\end{aligned} \tag{7.4}$$

where \mathbf{E} is the pulse energy as described in [27] (it has been taken as unity) and τ' is the time scaling factor (it has been adjusted so that the complete pulse is contained within 1 nsec).

7.2 Signal Model according to Mengali et al [3]

In [3], an \mathcal{M} -PPM noncoherent receiver for UWB has been presented. The signal is transmitted in multiple frames. The motivation for a multiple frame transmission has been attributed to the FCC limits on the signal power spectral density. Repeating a pulse N_f times, reduces the energy of an individual pulse for a constant symbol energy.

Since the UWB received pulses do not have a fixed shape and are a result of a superposition of many multipath components, the received waveform has the following form,

$$r(t) = \sum_{k=-\infty}^{\infty} h(t - kT_f - a_k T_f / \mathcal{M} - \tau) + n(t) \tag{7.5}$$

where $h(t)$ is the unknown channel impulse response (only the maximum delay spread is known which is equal to T_{mfs}). a_k is the data symbol taken from the alphabet $\{0, 1, \dots, \mathcal{M} - 1\}$ ($\mathcal{M} = 2$ in the case of 2-PPM). $n(t)$ is the white Gaussian noise with power spectral density $N_0/2$ in the interval $-B \leq f \leq B$ (B is the bandwidth of the rectangular low-pass filter at the receiver). τ is the time offset between the transmit and receive clocks.

If the symbol is transmitted in N_f frames then the received signal for the k -th symbol, with full synchronization, can be written as,

$$r(t) = \sum_{j=0}^{N_f-1} h(t - (j + kN_f)T_f - a_k T_f / \mathcal{M}) + n(t) \tag{7.6}$$

For the sampled version of (7.6) with sampling interval $T := 1/2B$, $r[i] := r(iT)$, $n[i] := n(iT)$, all samples are stacked into vectors $\mathbf{r}^{(j)}$ and $\mathbf{n}^{(j)}$. Assuming that the

channel remains fixed during the symbol, (7.6) can be written in the following vector form,

$$\mathbf{r}^{(j)} = \mathbf{u}(a_k, \mathbf{h}) + \mathbf{n}^{(j)} \quad (7.7)$$

where $\mathbf{u}(a_k, \mathbf{h})$ is the concatenation of \mathcal{M} subvectors. For the case of 2-PPM, it is a concatenation of two vectors where each vector will have half of the samples of one frame. If the received signal vectors corresponding to each frame are stacked in one big vector \mathbf{r} (i.e., $\mathbf{r} := [\mathbf{r}^{(0)T} \ \mathbf{r}^{(1)T} \ \dots \ \mathbf{r}^{(N_f-1)T}]^T$) then the pdf of the received signal can be written as,

$$p(\mathbf{r}|a_k, \mathbf{h}) = C \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=0}^{N_f-1} \|\mathbf{r}^{(j)} - \mathbf{u}(a_k, \mathbf{h})\|_2^2 \right\} \quad (7.8)$$

where C is a positive constant and $\sigma_n^2 = N_0 B$ is noise variance. Using generalized maximum likelihood (GML) criterion, (7.8) can be maximized by minimizing its summation term which equals the following cost function,

$$\Lambda(a_k, \mathbf{h}) = \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} (h_l^2 - 2h_l r[(j + kN_f)\mathcal{M}Q + a_k Q + l]) \quad (7.9)$$

where Q is the oversampling term and $\mathcal{M}Q = T_f/T$. Next, varying \mathbf{h} while keeping a_k fixed,

$$h_l = \frac{1}{N_f} \sum_{j=0}^{N_f-1} r[(j + kN_f)\mathcal{M}Q + a_k Q + l] : \quad 0 \leq l \leq L-1 \quad (7.10)$$

which leads to their decision variable, i.e.,

$$\hat{a}_k = \arg \max_{m \in [0, \mathcal{M}-1]} \sum_{j=0}^{N_f-1} \sum_{i=i_{k,j,m}}^{i_{k,j,m}+L-1} r^2[i] \quad (7.11)$$

where $i_{k,j,m} := [(j + kN_f)\mathcal{M} + m]Q$ is the first sample in the m -th slot of the j -th frame. (7.11) can be interpreted as maximizing the sum of the squares of the samples (of the received signal) at every pulse location over all frames. We shall explore in next sections if this conclusion can be derived from (7.9) and (7.10). Furthermore, using a Gaussian approximation of the decision variable they have suggested the following theoretical expression for the probability of error,

$$P_e = Q \left(\left[2 \left(\frac{N_0}{E_s} \right) + 2N_f B \Delta \left(\frac{N_0}{E_s} \right)^2 \right]^{-1/2} \right) \quad (7.12)$$

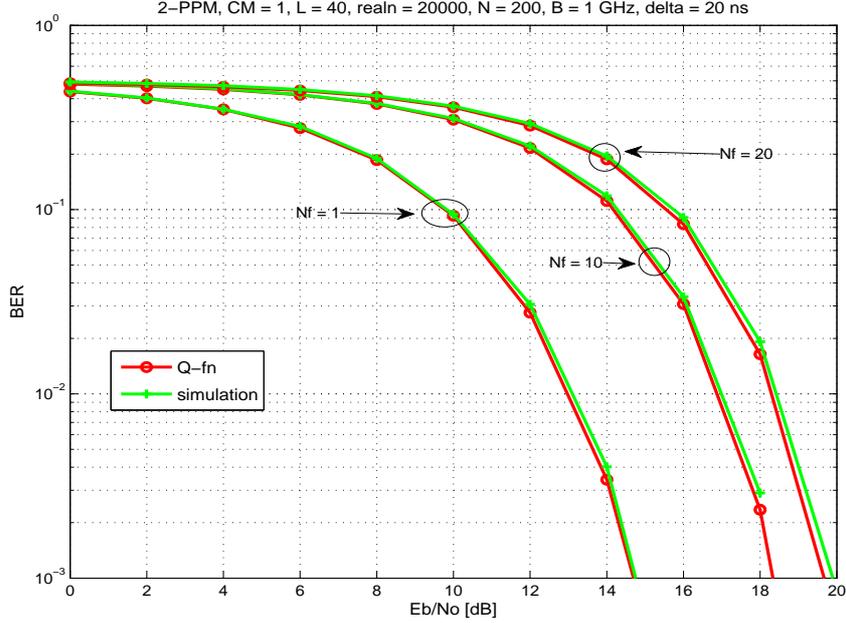


Figure 7.2: 2-PPM UWB, BER performance for Multiple frames for CM1, for large integration interval

7.2.1 Simulation Results of [3]

Figures 7.2-7.5 show the simulation results according to the methodology of [3] with different number of frames both for the CM1 and CM3 channel models. All these four figures show simulation results for one, ten and twenty frames respectively. The frame time T_f has been considered as 100 nsec for these simulations which is sufficient to analyse the behavior of this signal model. So each pulse interval QT is at most 50 nsec. The plots depict results of signal simulation as well as the Q-function according to (7.12).

Figure 7.2 shows the multi-frame simulation results for CM1 with an integration interval (Δ) of 20 nsec. This corresponds to 40 channel taps (L) with the receiver bandwidth (B) of 1 GHz and a sampling interval (T) of 0.5 nsec. Figure 7.3 shows the multi-frame simulation results for CM3 with and integration interval (Δ) of 45 nsec. This corresponds to 180 channel taps (L) with the receiver bandwidth (B) of 2 GHz and sampling interval (T) of 0.25 nsec. We see that in both of these figures, simulation closely follows the Q-function. But it is also very suspicious that the overall BER performance degrades with an increasing number of frames.

Figure 7.4 shows the multi-frame simulation results for CM1 with an integration interval (Δ) of 10 nsec. This corresponds to 20 channel taps (L) with the receiver bandwidth (B) of 1 GHz and sampling interval (T) of 0.5 nsec. We see that the simulation deviates from the Q-function and this deviation keeps increasing with increasing number of frames. Figure 7.5 shows the multi-frame simulation results for CM3 with and integration interval (Δ) of 40 nsec. This corresponds to 80 channel taps (L) with

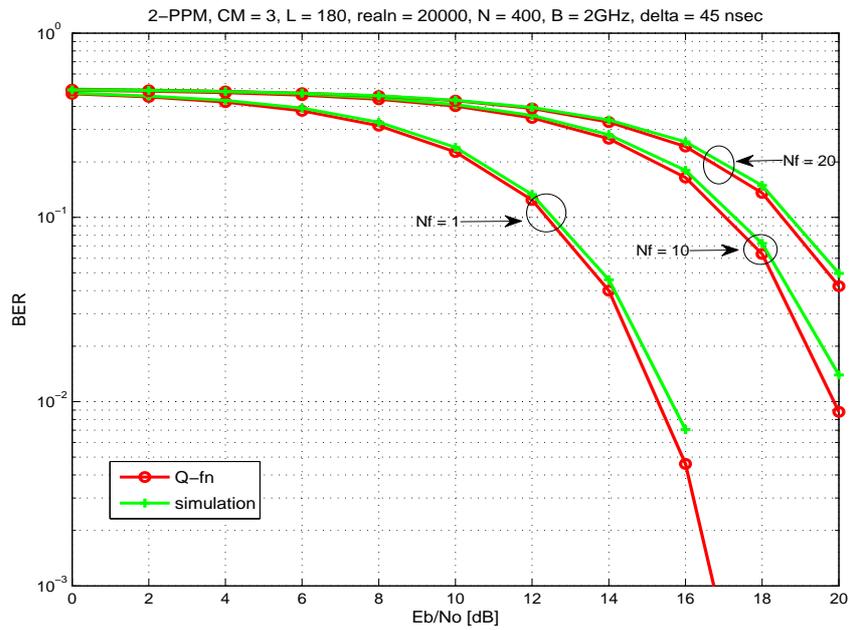


Figure 7.3: 2-PPM UWB, BER performance for Multiple frames for CM3, for large integration interval

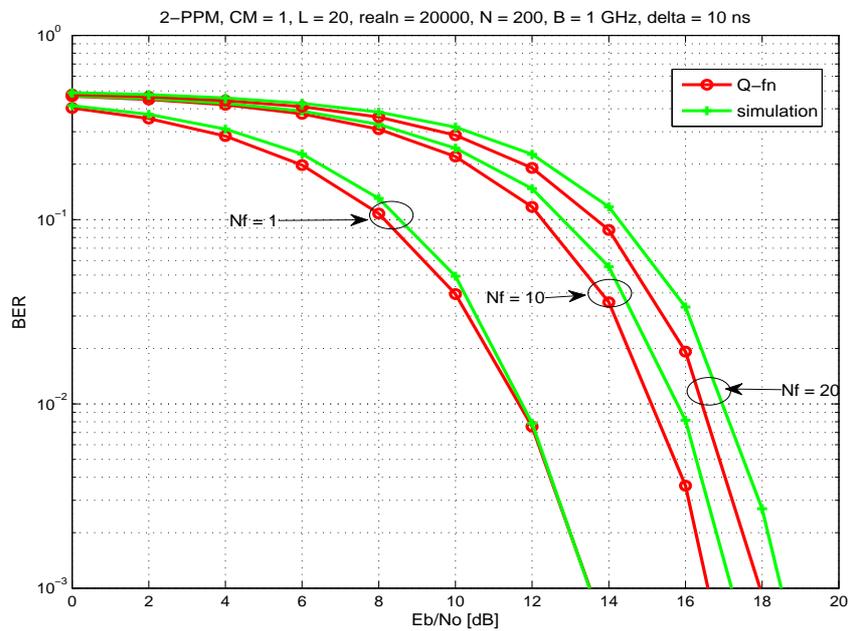


Figure 7.4: 2-PPM UWB, BER performance for Multiple frames for CM1, for small integration interval

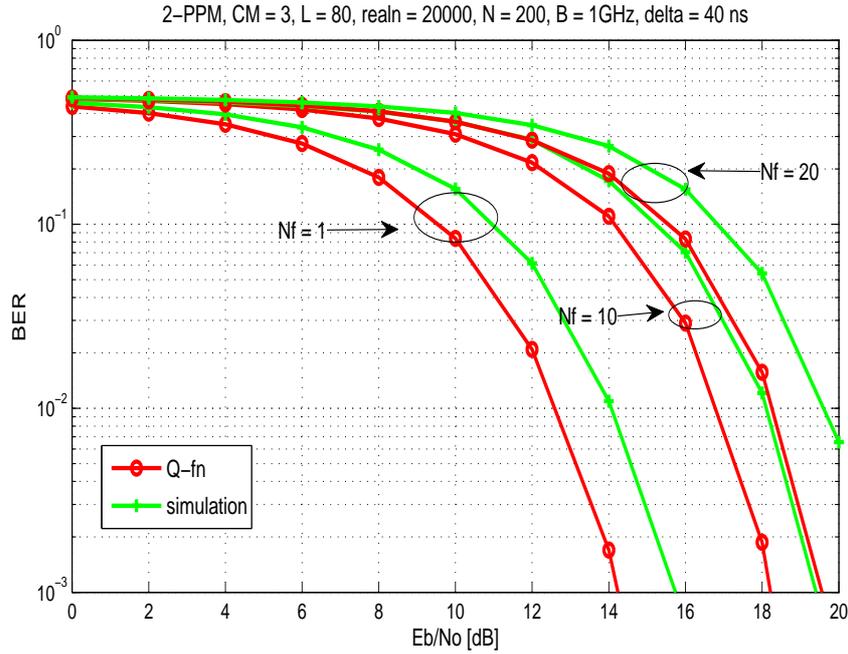


Figure 7.5: 2-PPM UWB, BER performance for Multiple frames for CM3, for small integration interval

the receiver bandwidth (B) of 1 GHz and sampling interval (T) of 0.5 nsec. We see that the simulation deviates from the Q-function which is even more wider than the case of CM1. Furthermore, the overall BER performance is also degraded with increasing number of frames.

These simulation results show,

1. The Q-function assumption of (7.12) is valid only for the case of a large integration interval when a large number of channel components are available.
2. The Q-function assumption of (7.12) is not valid for a small integration interval and the deviation from the actual performance is larger for CM3 than for CM1.
3. The overall BER performance degrades with an increasing number of frames which is a questionable result.

In the next section we shall analyse the model of [3] and provide our suggestions and/or improvements in this regard.

7.3 Analysis of the Signal Model of [3]

We now analyze the detection and decision model of [3], especially from (7.8) to (7.11). We carry out the in between steps of these equations to be better able to assess the conclusions reached by [3] in justification of the multi-frame methodology.

Starting from the maximum likelihood function, it is clear that to maximize (7.8), we need to minimize the summation term. The cost function can thus be written as,

$$\begin{aligned}
\Lambda(a_k, \mathbf{h}) &= \sum_{j=0}^{N_f-1} \|\mathbf{r}^{(j)} - \mathbf{u}(a_k, \mathbf{h})\|_2^2 \\
&= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} (h_l^2 - 2h_l r[(j + kN_f)\mathcal{M}Q + a_k Q + l]) \\
&= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} (h_l^2 - 2h_l r[P_{j,l}])
\end{aligned} \tag{7.13}$$

where $P_{j,l} = (j + kN_f)\mathcal{M}Q + a_k Q + l$ for notational simplicity. Now taking derivative w.r.t \mathbf{h} and keeping a_k fixed we obtain,

$$\frac{\partial \Lambda(a_k, \mathbf{h})}{\partial \mathbf{h}} = \frac{\partial}{\partial \mathbf{h}} \left\{ \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} (h_l^2 - 2h_l r[P_{j,l}]) \right\} \tag{7.14}$$

For the l -th component of \mathbf{h} we can write (7.14) as,

$$\frac{\partial \Lambda(a_k, h_l)}{\partial h_l} = \frac{\partial}{\partial h_l} \left\{ \sum_{j=0}^{N_f-1} (h_l^2 - 2h_l r[P_{j,l}]) \right\} \tag{7.15}$$

Now minimizing the cost function would mean setting its gradient w.r.t. h_l to zero, i.e.,

$$\begin{aligned}
0 &= \sum_{j=0}^{N_f-1} \left\{ (2\hat{h}_l - 2r[P_{j,l}]) \right\} \\
\sum_{j=0}^{N_f-1} \hat{h}_l &= \sum_{j=0}^{N_f-1} r[P_{j,l}] \\
N_f \hat{h}_l &= \sum_{j=0}^{N_f-1} r[P_{j,l}] \\
\hat{h}_l &= \frac{1}{N_f} \sum_{j=0}^{N_f-1} r[P_{j,l}]
\end{aligned} \tag{7.16}$$

where \hat{h}_l is the estimate for h_l . Rearranging (7.13) we obtain,

$$\begin{aligned}
\Lambda(a_k, \mathbf{h}) &= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} h_l^2 - \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} 2h_l r[P_{j,l}] \\
&= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} h_l^2 - \sum_{l=0}^{L-1} 2h_l \sum_{j=0}^{N_f-1} r[P_{j,l}]
\end{aligned} \tag{7.17}$$

Now substituting (7.16) in (7.17) we get,

$$\begin{aligned}
\Lambda(a_k, \hat{\mathbf{h}}) &= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2 - \sum_{l=0}^{L-1} 2\hat{h}_l N_f \hat{h}_l \\
&= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2 - 2N_f \sum_{l=0}^{L-1} \hat{h}_l^2 \\
&= \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2 - 2 \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2 \\
&= - \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2
\end{aligned} \tag{7.18}$$

So minimizing the cost function would mean,

$$\begin{aligned}
\min_{a_k} \Lambda(a_k, \hat{h}) &= \max_{a_k} \sum_{j=0}^{N_f-1} \sum_{l=0}^{L-1} \hat{h}_l^2 \\
&= \max_{a_k} N_f \sum_{l=0}^{L-1} \hat{h}_l^2
\end{aligned} \tag{7.19}$$

From (7.19), it is clear that the decision criterion can be in fact independent of the number of frames. Since each frame has a similar channel response, the only variation possible is because of the noise. So if we take an average channel response for all the frames then the performance will be similar whether their is one frame or multiple frames. Also now if we replace \hat{h}_l in (7.19) by its value from (7.16), we can write,

$$\begin{aligned}
\min_{a_k} \Lambda(a_k, \hat{h}) &= \max_{a_k} N_f \sum_{l=0}^{L-1} \left[\frac{1}{N_f} \sum_{j=0}^{N_f-1} r[P_{j,l}] \right]^2 \\
&= \max_{a_k} N_f \sum_{l=0}^{L-1} \left[\frac{1}{N_f} \sum_{j=0}^{N_f-1} r[(j + kN_f)\mathcal{M}Q + a_kQ + l] \right]^2
\end{aligned} \tag{7.20}$$

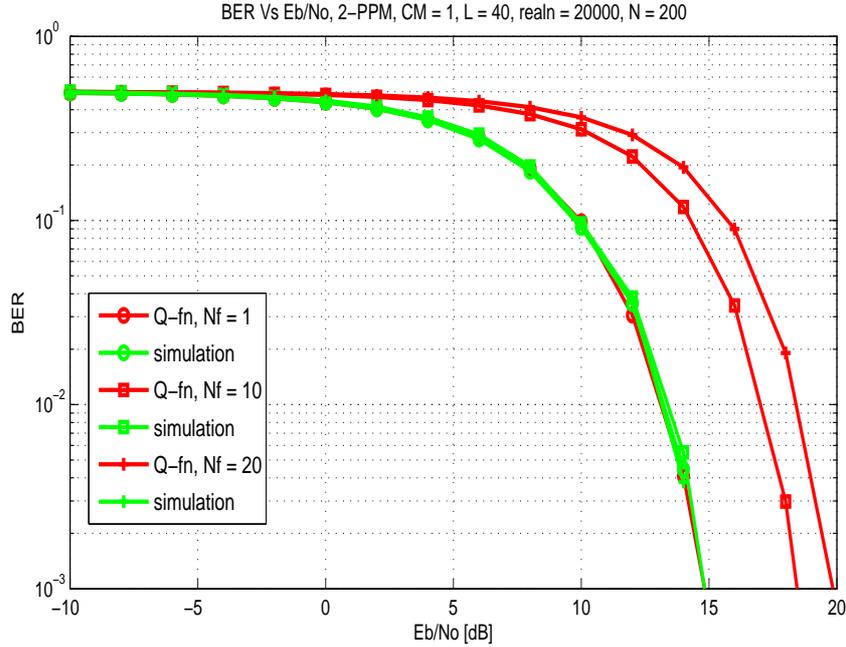


Figure 7.6: 2-PPM UWB, BER performance using (??) on Multiple frames

So from (7.20), we can see that if a decision on \hat{a}_k is made such that first we obtain the signal by taking an average over all the frames and then apply energy detection with respect to each pulse position to find the maximum, our performance will be better than (7.11).

We can validate our conclusion through simulations as well. Figure 7.6 shows the simulation result for CM1 with an integration interval (Δ) of 20 nsec. This corresponds to 40 channel taps (L) with a receiver bandwidth (B) of 1 GHz and a sampling interval (T) of 0.5 nsec. We can see that though the number of frames per PPM symbol varies from 1 to 20, the performance is similar to the case of a single frame per symbol. Similarly, Table 7.2 shows the simulated BER values for the case $N_f = 1, 10, 20$. For the E_b/N_0 range $[-10 : 2 : +10]$, we see very small changes in performance for a single frame per symbol and multiple frames per symbol.

Given these conclusions we shall, from now on, use only signals with a single frame per symbol and build our CS theory thereon. Our results for the application of CS on one frame may indirectly reflect its effect on multi-frame models as well.

7.4 Our Theoretical Model

As we can see from Figures 7.2-7.5 that the approximation of (7.12) is not valid for all situations of signal integration interval. Therefore we present our theoretical model to provide exact analysis of signal detection.

Since we are using 2-PPM signal, so we have two decision variables. The first decision variable will represent the energy contained in the signal duration of length

Table 7.2: Effect of Averaging on Multiframe Signal Model

E_b/N_0 [dB]	$N_f = 1$	$N_f = 10$	$N_f = 20$
-10	0.4923	0.5017	0.4969
-08	0.4891	0.4971	0.4935
-06	0.4838	0.4916	0.4884
-04	0.4749	0.4820	0.4801
-02	0.4603	0.4687	0.4679
00	0.4385	0.4774	0.4461
+02	0.4040	0.4148	0.4094
+04	0.3509	0.3619	0.3582
+06	0.2767	0.2982	0.2824
+08	0.1851	0.1966	0.1930
+10	0.0917	0.0965	0.0959

LT nsec (i.e., Δ , the integration interval). This can be represented as following,

$$U_1 = \sum_{i=1}^L |\sqrt{S}h_i + n_i|^2 \quad (7.21)$$

The second decision variable will have the energy contained in the second pulse interval position (see Figure 4.1 and 7.1). So we can write,

$$U_2 = \sum_{i=\frac{N}{2}+1}^{\frac{N}{2}+L} |n_i|^2. \quad (7.22)$$

We consider the channel modles according to CM1 and CM2. As the noise is zero mean white Gaussian, U_1 is a non-central chi-square distributed random variable. The channel components in (7.21) cause its noncentrality. The noncentrality parameter is given by,

$$s^2 = \sum_{i=1}^L Sh_i^2, \quad (7.23)$$

and its pdf is given by [26] as,

$$p_{U_1}(u_1) = \frac{1}{2\sigma^2} \left(\frac{u_1}{s^2}\right)^{(n_1-2)/4} \exp\left[\frac{-(s^2 + u_1)}{2\sigma^2}\right] \times I_{n_1/2-1}\left(\sqrt{u_1}\frac{s}{\sigma^2}\right) : n_1 = L, u_1 > 0 \quad (7.24)$$

with n_1 degrees of freedom. $I_\nu(z)$ is the modified Bessel function of the first kind as defined in [29, Eq. (8.445)], i.e.,

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(\nu + k + 1)} \left(\frac{z}{2}\right)^{\nu+2k}, \quad z \geq 0 \quad (7.25)$$

U_2 is a central chi-square distributed random variable since the only parameter involved is noise which is zero mean. The pdf of U_2 is given by [26] as,

$$p_{U_2}(u_2) = \frac{1}{\sigma^{n_2} 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} u_2^{(n_2-2)/2} \exp\left[\frac{-u_2}{2\sigma^2}\right] : n_2 = L, u_2 > 0 \quad (7.26)$$

with n_2 degrees of freedom.

To calculate the probability of error, we first consider a converse case. That is to say, we first find the probability that the energy due to only noise is less than the signal interval. So the probability of such a correct decision given that pulse is transmitted in the first half of the symbol, can be written as,

$$\begin{aligned} P_r(U_2 < u_1 | U_1 = u_1) &= \int_0^{u_1} p_{U_2}(u_2) du_2 \\ &= \int_0^{u_1} \frac{1}{\sigma^{n_2} 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} u_2^{(n_2-2)/2} \exp\left[\frac{-u_2}{2\sigma^2}\right] \end{aligned} \quad (7.27)$$

which can be simplified to,

$$P_r(U_2 < u_1 | U_1 = u_1) = \frac{\gamma\left(\frac{n_2}{2}, \frac{u_1}{2\sigma^2}\right)}{\Gamma\left(\frac{n_2}{2}\right)}. \quad (7.28)$$

where $\gamma(., .)$ is the lower-incomplete gamma function.

Since U_1 is a random variable, we need to average it out in (7.28). Thus the probability of error is;

$$\begin{aligned} P_e &= 1 - \int_0^\infty P_r(U_2 < u_1 | U_1 = u_1) p_{U_1}(u_1) du_1 \\ &= 1 - \int_0^\infty \frac{\gamma\left(\frac{n_2}{2}, \frac{u_1}{2\sigma^2}\right)}{\Gamma\left(\frac{n_2}{2}\right)} p_{U_1}(u_1) du_1 \end{aligned} \quad (7.29)$$

which simplifies to,

$$\begin{aligned} P_e &= 1 - \frac{1}{2\sigma^2 \Gamma\left(\frac{n_2}{2}\right) s^{(n_1-2)/2}} \exp\left[\frac{-s^2}{2\sigma^2}\right] \int_0^\infty \gamma\left(\frac{n_2}{2}, \frac{u_1}{2\sigma^2}\right) u_1^{(n_1-2)/4} \\ &\quad \times \exp\left[\frac{-u_1}{2\sigma^2}\right] I_{n_1/2-1}\left(\sqrt{u_1} \frac{s}{\sigma^2}\right) du_1. \end{aligned} \quad (7.30)$$

We can see that (7.30) is not in a closed form therefore it has to be computed numerically which is computationally intensive process. We carried it out in Matlab.

7.4.1 Simulation results for our Theoretical Model

Figures 7.7 and 7.8 show the simulation results of our theoretical model for two scenarios based on CM1. In case one, the integration interval is small i.e., $\Delta = 10$ nsec, which corresponds to 20 channel taps (L) with a receiver bandwidth (B) of 1 GHz and a

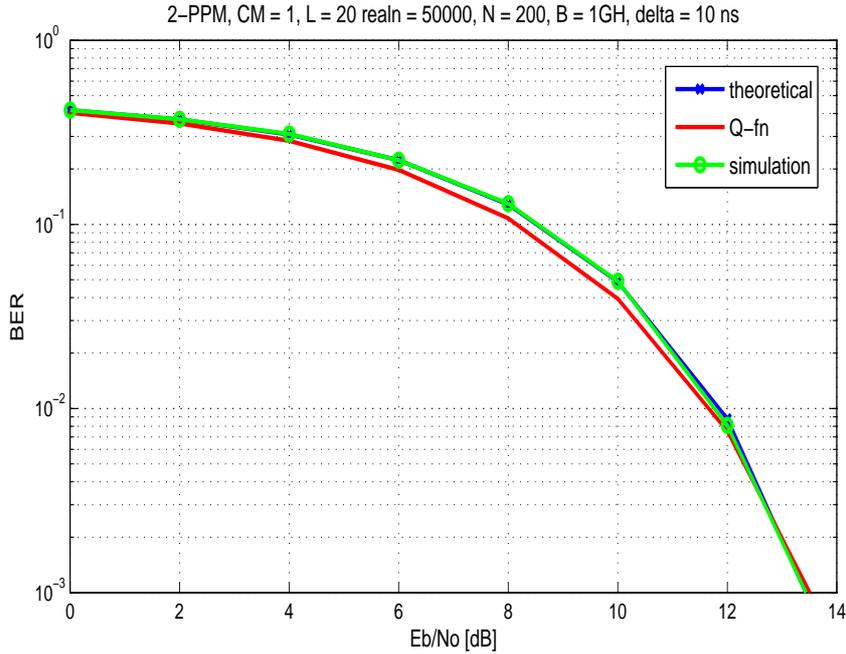


Figure 7.7: 2-PPM UWB, Comparison with small integration interval

sampling interval (T) of 0.5 nsec. In case two, the integration interval is large i.e., $\Delta = 20$ nsec, which corresponds to 80 channel taps (L) with a receiver bandwidth (B) of 2 GHz and a sampling interval (T) of 0.25 nsec.

In both cases we can see that our theoretical model is exact and it follows the simulation results very closely, whereas the Q-function proposed by [3] falls short of the exact values in the first case. In the second case, the Q-function is comparable with our theoretical model.

7.5 CS for UWB

Now we apply our CS theory as presented in Section 4.2. The major factor for CS to work is the sparsity of the signal. For the purpose of simulation, we consider first the case of a CM1 channel model. We take an integration interval of 10 nsec. This corresponds to 20 channel taps (L) with a receiver bandwidth (B) of 1 GHz and a sampling interval (T) of 0.5 nsec. So for a frame interval (T_f) of 100 nsec, the total signal vector has a size of $N = 200$. Using a compression ratio of 0.5, $M = 100$. For a given sparsity, our previous analysis tells us that the total signal vector size N must be sufficiently large for the Gaussian matrix to work as the measurement and consequently as reconstruction matrix. We see in this first case, that for a sparsity order $K = 20$, according to Table 6.1, $N = 12500$, and according to Table 6.2, $N = 448$, as a necessary condition for the Gaussian matrix to work as a feasible measurement matrix. But we see that in our case the available value of N falls short of both these bounds. Nonetheless, in terms of simulations, large values of N can put an intensive pressure on software like

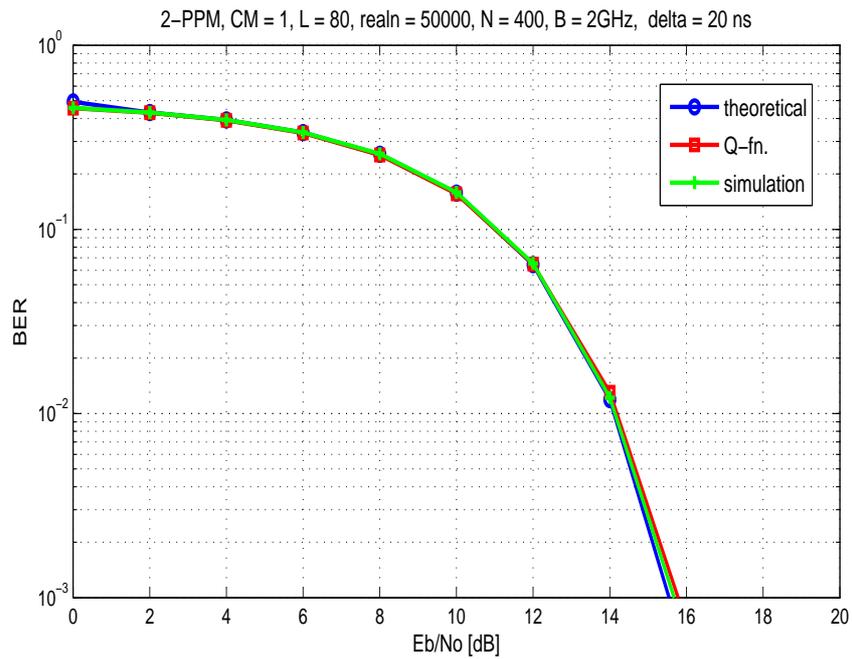


Figure 7.8: 2-PPM UWB, Comparison with large integration interval

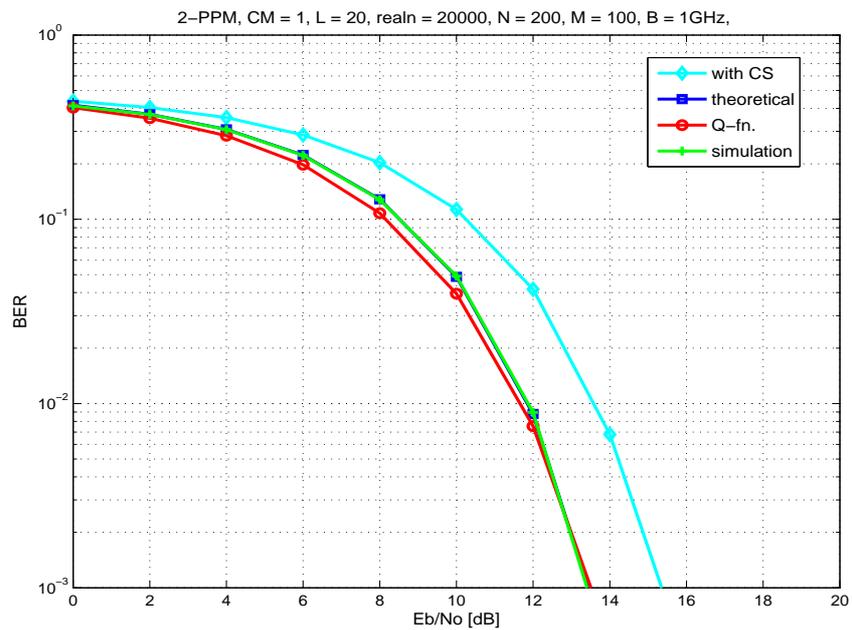


Figure 7.9: 2-PPM UWB, CS with reconstruction algorithm = OMP, measurement matrix = DFT matrix

Matlab which we are using for our simulation purpose, although in practice, this does not play a role.

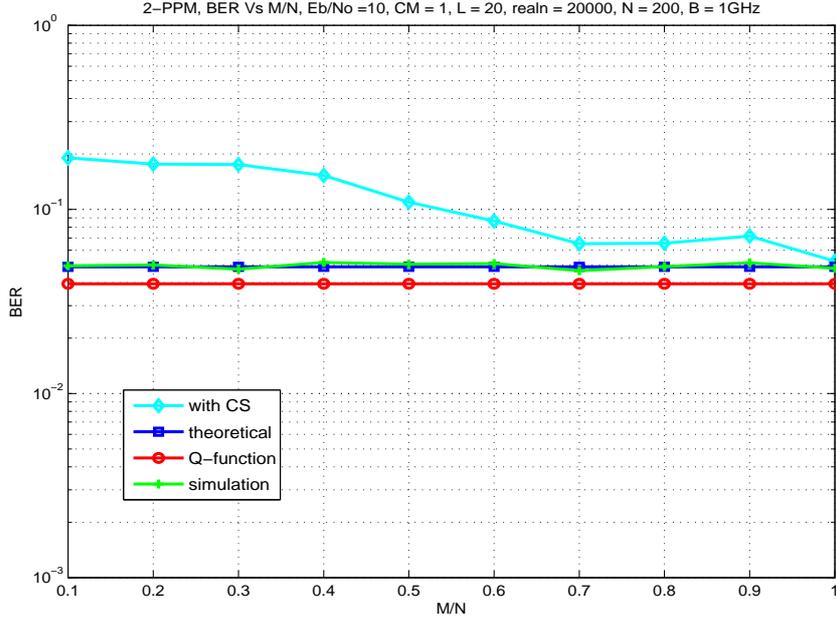


Figure 7.10: 2-PPM UWB, CS with reconstruction algorithm = OMP, Effect of M/N variations

Keeping in mind the above discussion, we chose to use Fourier matrix as our measurement matrix. Figure 7.9 shows the simulation results of the first case. The plots shown correspond to the Q-function, our theoretical model, simulation results without CS and the one after applying CS and then reconstructing thereof. We can see from the Figure,

- Our theoretical plot exactly follows the simulation results.
- The Q-function falls short of the simulation results.
- The application of CS has given good results. The difference between the performance of CS and the theoretical one is around 2 dB at a bit error rate of 10^{-3} .

So we are sampling at half of the Nyquist rate (since the compression ratio is 0.5) but the performance of CS is comparable with the Nyquist rate performance with a margin of loss of only around 2 dB. These are quite promising results. Figure 7.10 displays the effect of increasing the value of M on the performance of CS for an E_b/N_0 of 10 dB. We see that the CS curve closes on the references curves very quickly as the compression ratio (i.e., M/N) increases.

Furthermore, Figure 7.11 shows the result of applying CS on UWB for the case when Gaussian matrix is used as the measurement matrix, rest of the conditions are similar to Figure 7.9. Also Figure 7.12 shows the effects of varying compression ratio on BER performance similar to Figure 7.10 with the only difference of Gaussian matrix

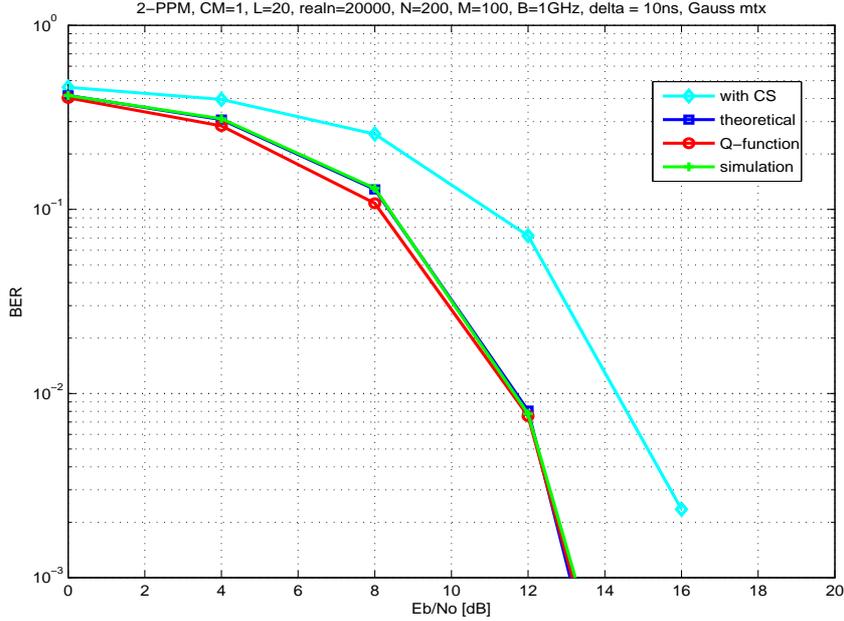


Figure 7.11: 2-PPM UWB, CS with reconstruction algorithm = OMP, measurement matrix = Gaussian matrix

as the measurement matrix. We can deduce from these figures that the structured matrix is performing better than the random matrix as the measurement matrix which vindicates our former claim on this issue as presented in section 6.1.

7.6 CS for UWB with Multiple Access

Multiple access in UWB was pioneered by Scholtz (e.g., [38]). Since then, a lot of progress has been made in realizing the vast potential of UWB for multiple users through time hopping and spread spectrum (e.g., [39, 40, 41, 42]). The capacity of multiple access in UWB is limited by the multiple users interference (MUI), multipath channels and noise. In this section we see how CS would work with UWB when we want to access a certain user among many users.

7.6.1 Signal Model

We consider binary PPM for time hopping spread spectrum. We transmit one symbol in N_f frames, each frame of duration T_f . Each frame is divided into N_c chips, with T_c as the duration of the chip. The transmitted pulse is position modulated within the chip between two positions of duration $T_m = T_c/2$ each. So each transmitted bit can be represented by these two positions, represented by b (first position for bit 0 (i.e., $b = 0$) and second in case of bit 1 (i.e., $b = 1$)). We also assume that the maximum channel delay spread $T_{m\text{ds}} < T_c/2$. The time hopping code c_j can take values between 0 and $N_c - 1$ at the most and the code is repeated after $N_f T_f$ for each user. Furthermore,

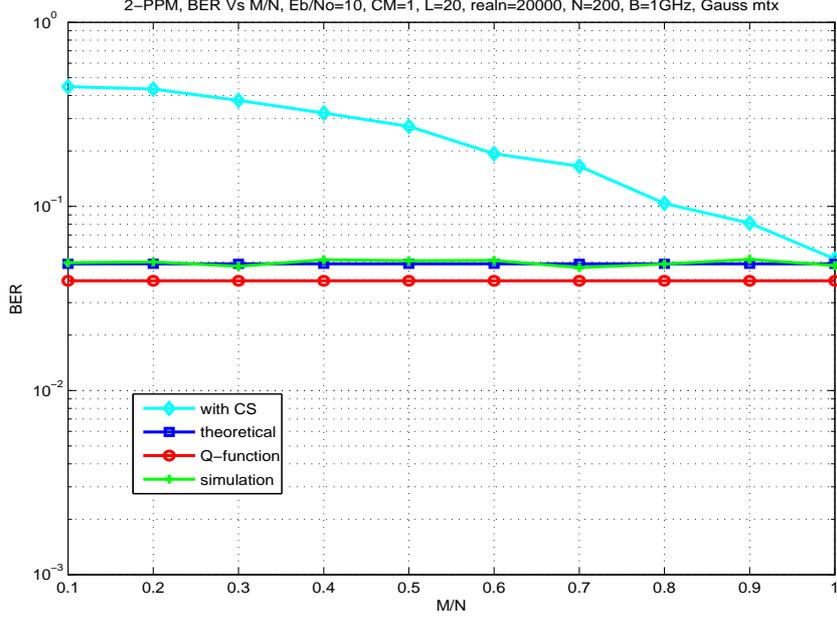


Figure 7.12: 2-PPM UWB, CS with reconstruction algorithm = OMP, Effect of M/N variations, measurement matrix = Gaussian matrix

the signal is amplitude modulated as a spreading code sequence i.e., $d_j \in \{\pm 1\}$. So we can write our transmitted UWB signal in a multiple access environment for the k -th user as,

$$a_{MA}^{(k)}(t) = \sum_{j=0}^{N_f-1} d_j^{(k)} p\left(t - jT_f - c_j^{(k)}T_c - b_k T_m\right) \quad (7.31)$$

where $p(t)$ is the second derivative of a Gaussian pulse as described in (7.4).

For detection, we again opt for noncoherent energy detection. We also assume that we are synchronized with the start of a symbol. Since in multiple access scheme signals from different users will reach the receiver for the k -th user, there can be small delay of arrival between different signals depending upon their distance and relative moment. For our case we assume this difference to be negligible. Furthermore, since the time hopping code and spreading sequence of a particular user is known at the receiver, the desired user's signal can be located. Then by using the spreading code, the averaging process would outstrip the undesirable users. Perfect orthogonal codes can be the best choice but in our simulation we did not use completely orthogonal codes.

So the detection process first consists of despreading then CS is applied and the received signal is sampled at the lower rate. Then the demodulation process is carried on the reconstructed signal to make the final decision.

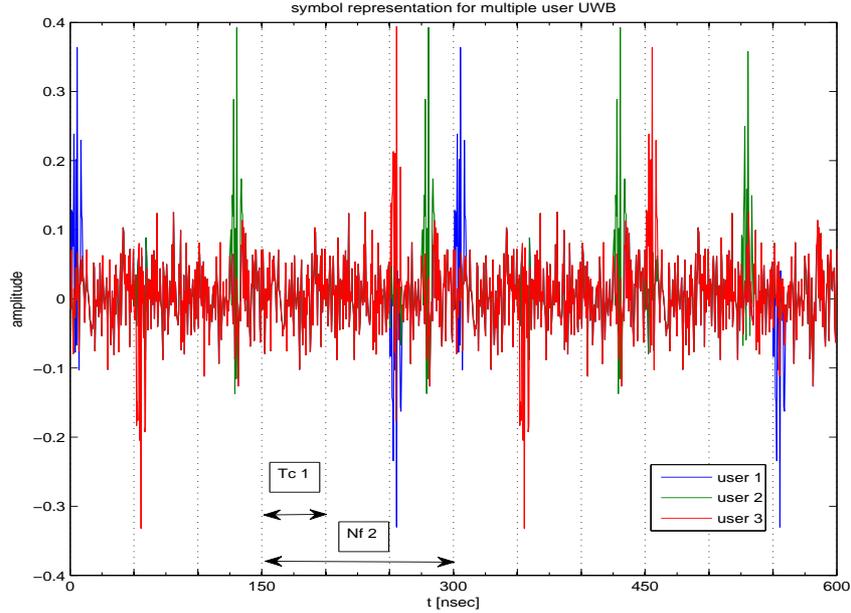


Figure 7.13: UWB symbol with Multiple Access

7.6.2 CS Simulations for Multiple Access UWB

Now we present the results of our simulations for CS with UWB in presence of multiple users interference, multiple paths and Gaussian noise.

We assume three users, i.e., $N_u = 3$, the frame interval $T_f = 150$ nsec, number of frames per symbol $N_f = 4$, number of chips per frame $N_c = 3$ with a chip interval $T_c = 50$ nsec. The time hopping codes for each user and their respective amplitude codes are given in the form of matrices below with rows representing the users and columns respective frames,

$$c = \begin{pmatrix} 1 & 3 & 1 & 3 \\ 3 & 3 & 3 & 2 \\ 2 & 3 & 2 & 1 \end{pmatrix}, \quad (7.32)$$

$$d = \begin{pmatrix} +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \end{pmatrix}. \quad (7.33)$$

Figure 7.13 shows the received UWB symbol for multiple access with a channel spread of 10 nsec which corresponds to $L = 20$ with a receiver filter bandwidth (B) of 1 GHz and a sampling interval (T) of 0.5 nsec. The signal vector on which CS has to work consists of 100 samples (i.e., N). We take the number of measurements M to be 50. Given a sparsity of order $K = 20$, Figure 7.14 shows the simulation results for CS on UWB with multiple access. We can see from the figure that CS is performing reasonably well. It also provides, for the sake of comparison, simulation

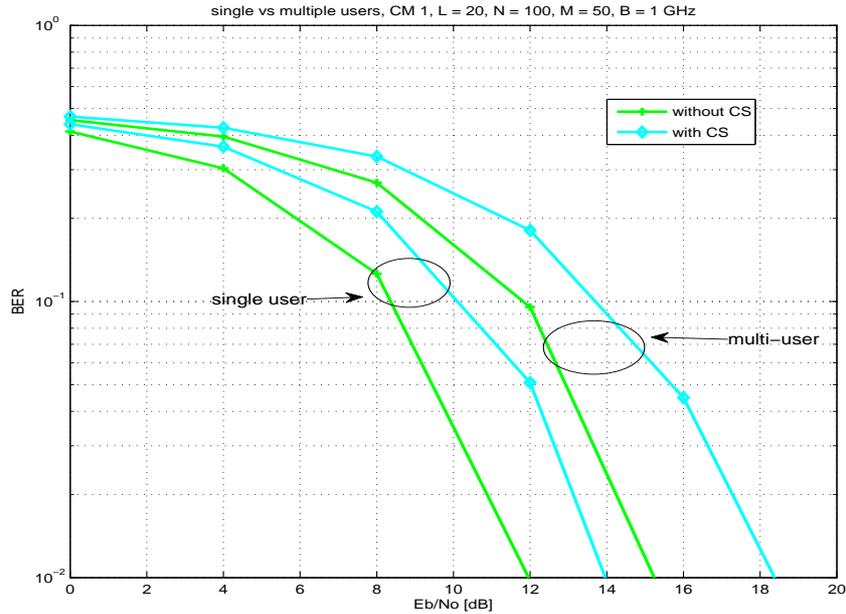


Figure 7.14: 2-PPM UWB , comparison between single user and multiple users, CS with reconstruction algorithm = OMP, measurement matrix = DFT matrix

results for the case of a single user. It can be seen that on the whole single user has better performance because of the absence of any interference. But the performance of CS is getting worse in case of multi-user in comparison to the case of a single user. The difference between the performance results with and without CS is around 2 dB in case of a single user and it is around 4 dB in case of the multiple users. The reason can be the increase in the order of sparsity because of multiple interference. So we can conclude that as the multiple users interference increases, CS may not be able to give a very good performance.

7.7 Conclusions

In this chapter we provided an application of PPM modulated signals. UWB signals are considered to be the most intensive in terms of sampling rates since the bandwidth is too large. We have successfully presented our solution through CS. Firstly, We elaborated upon the inconsistencies of the PPM based UWB model of [3] and then presented our improved theoretical model. Then we have implemented our CS methodology with achieving good results. We have also shown that CS can work even in the case of multiple access. We have seen here as well that the performance of CS is influenced by the level of sparsity of the signal.

Conclusions and Future Work

In this chapter we conclude our work briefly and provide some possible options for future research.

8.1 Conclusions

In this thesis we have explored CS for two major wideband modulation schemes, namely PPM and FSK. We have utilized CS to reduce sampling rates for these signals. We explicitly explored the effect of CS on UWB signals. We have provided theoretical analysis for these schemes and simulations thereof. We have come out with a detailed analysis of the scenarios where CS can be utilized to its maximum potential. We have contributed to the understanding of measurement matrices in terms of their structures and utilization. We have highlighted upon the bounds on the order of sparsity for better reconstruction of the signals. We have also given the applicability of CS for UWB signals. We have also provided the effect of multiuser interference on the performance of CS for UWB signals. We have considered non-coherent detection to reduce overall system complexity.

Chapter 3 provides the general signal model for FSK and PPM signals. It gives a unified approach to the description of multipath channel. It also gives the general CS model applied on wideband signals on the receiver side. In chapter 4 we have provided the application of CS on PPM modulated signals. We considered quasi-synchronous and fully-synchronous PPM signals. We have presented the theoretical analysis of PPM signals for both these cases. We present the important issues relating to the application of CS for PPM. These issues relate to the sparsity of the signals and also the measurement matrices. We found that the structured matrices contribute as good measurement matrices even when the signal is not largely sparse enough in comparison to the random matrices. In chapter 5 we provided the application of CS on FSK modulated signals. We considered single tone and dual tone FSK signals. We presented the theoretical analysis of FSK signals for both the cases. We presented the important issues relating to the application of CS for FSK in terms of sparsity and reconstruction. Here again we found that the structured matrices were a better choice when it comes to a reduction in the order of sparsity.

Our findings regarding structured and random matrices were analyzed in chapter 6. We gave reasons which may cause a random matrix not to work as a good measurement matrix. We have provided with explicit numerical values regarding the non-zero elements of a signal vector and the total number of basis functions. We saw that the CS theory with especially the random matrices as measurement matrices is feasible only when M and N are asymptotic in nature.

Last but not the least, we have provided the application of CS for UWB signals.

We have given our theoretical analysis for the signal model and comparison with an existing model. We have shown that our theory works for the IEEE 802.15.3a channel models. We extended the application of CS on UWB signals in the presence of multiple access. Even in this case we saw that CS gives promising results.

8.2 Suggestions for Future Work

CS has a very good potential of being applied in telecommunications. Below we present some of the possible dimensions to extend the application of CS in the field of wireless and data communications.

- One direct extension of our work can be the utilization of new reconstruction algorithms which can provide more tangible parameters to gauge their performance. Similarly a comprehensive analysis of available reconstruction algorithms can be carried out to provide the most suitable algorithms for practical implementations of CS.
- Since we have seen that sparsity plays an important role in the performance of CS, research can be carried out to ascertain the best estimate of the sparsity of the signal. This can have direct impact on the efficiency of reconstruction algorithms. Furthermore, if the signal is not fully sparse then some sparsifying methods can be applied to this end. It can also be analyzed what would be the cost of sparsifying the signals in terms of BER performance.
- In this research we have shown that the signal detection and reconstruction can be done without having to estimate the channel. In situations where the channel estimate is crucial, CS can still be applied to estimate the channel as well.
- Research can be carried out to ascertain the performance of CS in situations with practical abnormalities e.g., timing errors and amplifier non-linearities etc.

Analog to Information Converter



In this appendix we briefly elaborate upon some of the implementations of AICs. AICs aim at the practical realization of CS especially for the analog signals. We present here three suggested implementations of AICs.

- Random sampling
- Random Filtering
- Pseudo-random Demodulation

A.1 Random Sampling

These architectures have been suggested in [8, 43]. It is suitable for wideband signals which are sparse in a local Fourier representation e.g., Frequency hopping, slowly varying chirps from radar and geophysics, and many acoustic and audio signals. We consider one of its implementation as given in [8]. This architecture scheme consists of a parallel bank of low-rate ADCs with equal shifts between their starting conversion points. So a set of shifted samples is produced by each ADC. Switching between the ADCs is random. Figure A.1 shows an illustration of this scheme.

So we can look at this architecture as if implementing an identity matrix as the measurement matrix with its rows being knocked out randomly (for a certain signal length). This architecture can be implemented but it may increase the chip area and the second challenge might be that of the jitter effect which may ensue when controlling the switches.

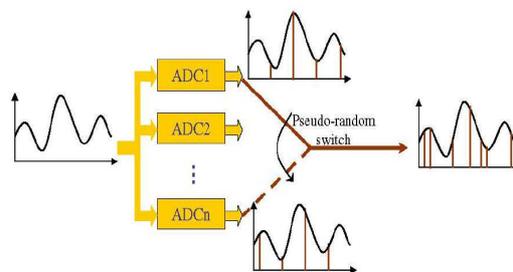


Figure A.1: AIC implementation with random sampling, [8]

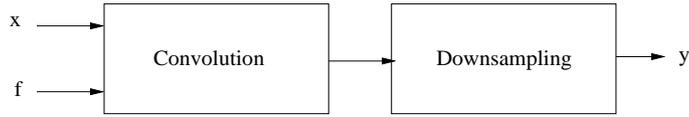


Figure A.2: AIC implementation with random filtering (I), [9]

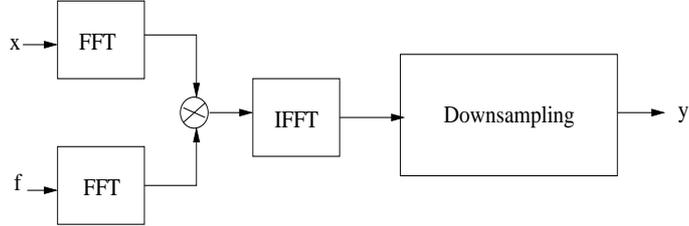


Figure A.3: AIC implementation with random filtering (II), [9]

A.2 Random Filtering

This architecture has been suggested in [9]. It is designed for a large class of compressible signals and can be generalized to streaming and continuous-time signals (e.g., signals sparse in time, frequency and wavelet domains). In this approach, measurements are obtained by convolving the signal (x) with a random-tap FIR filter (f) and then downsampling the filtered signal by a factor of $\lfloor \frac{N}{M} \rfloor$. The filter taps (of certain length) can be drawn from a normal distribution or from the Bernoulli distribution. Thus the measurement matrix will have a quasi-Toeplitz structure with each row containing shifted values of the filter taps. The amount of right shift equals $\lfloor \frac{N}{M} \rfloor$. Two implementations have been proposed for this architecture. One with obtaining measurements by performing linear convolution and downsampling simultaneously (see Figure A.2) and the other with first using the FFT for convolution and then downsampling the inverse-FFT of their product (see Figure A.3).

The first method can be useful for streaming and continuous-time signals. The second method can be faster than the first if filter has many taps but it needs the presence of an entire signal. This architecture has given promising results for signals which are sparse especially in time and frequency domains as per the simulation results of [9].

A.3 Pseudo-random Demodulation

This architecture has been suggested in [10, 44, 45]. In this scheme the signal acquisition process consists of mainly three components. These are demodulation, filtering and uniform sampling. The signal is modulated via a 'chipping sequence' of ± 1 s. This sequence alternates at a rate which is equal to or faster than the Nyquist frequency of the signal. The demodulation process spreads the frequency content of the signal so that it is not destroyed at the second stage of the acquisition process. Second stage consists of a lowpass filter. At the final stage, signal is sampled at reduced rate (M). Figure A.4 shows a schematic of this architecture.

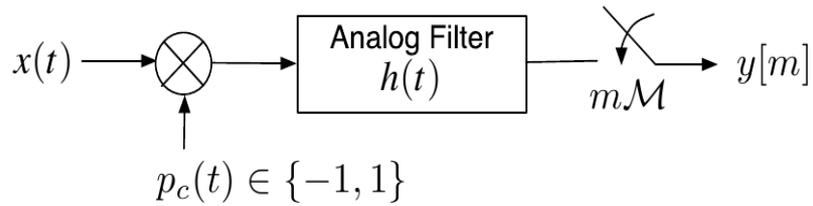


Figure A.4: AIC implementation with pseudo-random demodulation, [10]

The process can be summarized as below.

- The signal is demodulated by multiplying it with a high-rate pseudonoise sequence, which spreads the tones (i.e., frequency content) across the entire spectrum.
- A lowpass anti-aliasing filter is applied as an integrator.
- The signal is sampled at relatively low rate.

In ideal form, the system can be modeled as a measurement matrix Φ containing N/M pseudo-random ± 1 s on each row. The demodulator bypasses the need for a high-rate ADC but still allows the use of a low-rate ADC. This architecture can be applied to the signals which are sparse in both the time and frequency domains.

Bibliography

- [1] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, April 2006.
- [2] E. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [3] C. Carbonelli and U. Mengali, “M-ppm noncoherent receivers for uwb applications,” *Wireless Communications, IEEE Transactions*, vol. 5, no. 8, pp. 2285–2294, Aug. 2006.
- [4] B. Le, T. Rondeau, J. Reed, and C. Bostian, “Analog-to-digital converters,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 69–77, Nov. 2005.
- [5] J. D. Blanchard, C. Cartis, and J. Tanner, “Phase transitions for restricted isometry properties,” 2009, preprint.
- [6] D. Donoho and J. Tanner, “Neighborliness of randomly-projected simplices in high dimensions,” in *National Academy of Sciences*, vol. 102, no. 27, July 2005, pp. 9452–9457.
- [7] M. Rudelson and R. Vershynin, “On sparse reconstruction from fourier and gaussian measurements,” *Communications on Pure and Applied Mathematics*, vol. 61, pp. 1025–1045, August 2008.
- [8] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, “Random sampling for analog-to-information conversion of wideband signals,” in *IEEE Dallas/CAS Workshop*, Oct. 2006, pp. 119–122.
- [9] J. Tropp, M. Wakin, M. Duarte, D. Baron, and R. Baraniuk, “Random filters for compressive sampling and reconstruction,” in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 3, May 2006, pp. III–III.
- [10] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, “Analog-to-information conversion via random demodulation,” in *Design, Applications, Integration and Software, IEEE Dallas/CAS Workshop*, Oct. 2006, pp. 71–74.
- [11] M. Stojnic, W. Xu, and B. Hassibi, “Compressed sensing - probabilistic analysis of a null-space characterization,” in *Acoustics, Speech and Signal Processing, ICASSP. IEEE International Conference*, April 2008, pp. 3377–3380.
- [12] A. Molisch, J. Foerster, and M. Pendergrass, “Channel models for ultrawideband personal area networks,” *IEEE Wireless Communications*, vol. 10, no. 6, pp. 14–21, Dec. 2003.

- [13] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transaction on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [14] M. Unser, "Sampling-50 years after shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, Apr 2000.
- [15] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 31–40, March 2008.
- [16] J. G. Proakis and M. Salehi, *Fundamentals of Communication Systems*. Upper Saddle River, New jersey: Pearson Prentice Hall, 2005.
- [17] R. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, Apr 1999.
- [18] P. Kenington and L. Astier, "Power consumption of a/d converters for software radio applications," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 2, pp. 643–650, Mar 2000.
- [19] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [20] R. Baraniui, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 2008, 2007. [Online]. Available: <http://dsp.rice.edu/cs/jlcs-v03.pdf>
- [21] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007. [Online]. Available: <http://stacks.iop.org/0266-5611/23/969>
- [22] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k -term approximation," 2006, preprint.
- [23] D. Achlioptas, "Database-friendly random projections," in *PODS '01: Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. New York, NY, USA: ACM, 2001, pp. 274–281.
- [24] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [25] Y. C. Pati, R. Rezaifar, Y. C. P. R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, 1993, pp. 40–44.
- [26] J. G. Proakis, *Digital Communications*, 4th ed. Avenue of Americas, NY: McGraw-Hill, 2001.

- [27] M. Ghavami, L. B. Michael, and R. Kohno, *Ultra Wideband signals and systems in communication engineering*, 2nd ed. West Sussex, England: John Wiley Sons, 2007.
- [28] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. NY USA: Springer, 2006.
- [29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA: Academic Press, 1996.
- [30] E. Candès, “l1-magic: Recovery of sparse signals,” 2005. [Online]. Available: <http://www.acm.caltech.edu/l1magic/>
- [31] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [32] T. Gulliver and S. Ting, “Dual-tone mfsk for frequency-hopped spread-spectrum multiple access,” in *Communications, Computers and signal Processing, PACRIM. IEEE Pacific Rim Conference*, vol. 2, 2001, pp. 643–646.
- [33] S. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, Dec 1993.
- [34] G. Karabulut, L. Moura, D. Panario, and A. Yongacoglu, “Integrating flexible tree searches to orthogonal matching pursuit algorithm,” *Vision, Image and Signal Processing, IEE Proceedings -*, vol. 153, no. 5, pp. 538–548, Oct. 2006.
- [35] A. K. Fletcher, S. Rangan, and V. K. Goyal, “On-off random access channels: A compressed sensing framework,” *CoRR*, vol. abs/0903.1022, 2009.
- [36] E. J. Candès, “The restricted isometry property and its implications for compressed sensing,” *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.
- [37] A. Saleh and R. Valenzuela, “A statistical model for indoor multipath propagation,” *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, Feb 1987.
- [38] M. Win and R. Scholtz, “Impulse radio: how it works,” *IEEE Communications Letters*, vol. 2, no. 2, pp. 36–38, Feb 1998.
- [39] —, “Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications,” *IEEE Transactions on Communications*, vol. 48, no. 4, pp. 679–689, Apr 2000.
- [40] F. Ramirez-Mireles, “Performance of ultrawideband ssma using time hopping and m-ary ppm,” *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 6, pp. 1186–1196, Jun 2001.
- [41] L. Zhao and A. Haimovich, “Multi-user capacity of m-ary ppm ultra-wideband communications,” in *Ultra Wideband Systems and Technologies, 2002. Digest of Papers. 2002 IEEE Conference on*, 2002, pp. 175–179.

- [42] F. Ramirez-Mireles, “Performance of ultrawideband ssma using orthogonal ppm-th over dense multipath channels,” *Telecommunication Systems*, vol. 36, pp. 107–115, 2007.
- [43] T. Ragheb, S. Kirolos, J. Laska, A. Gilbert, M. Strauss, R. Baraniuk, and Y. Massoud, “Implementation models for analog-to-information conversion via random sampling,” in *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on*, Aug. 2007, pp. 325–328.
- [44] T. Ragheb, J. Laska, H. Nejati, S. Kirolos, R. Baraniuk, and Y. Massoud, “A prototype hardware for random demodulation based compressive analog-to-digital conversion,” in *Circuits and Systems, 2008. MWSCAS 2008. 51st Midwest Symposium on*, Aug. 2008, pp. 37–40.
- [45] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, “Beyond nyquist: Efficient sampling of sparse bandlimited signals,” *CoRR*, vol. abs/0902.0026, 2009. [Online]. Available: <http://arxiv.org/abs/0902.0026>