

A Hybrid Precoding and Signal Detection Framework for Future Wireless Systems

by

Kevin Bradley Dsouza

B.Tech., National Institute of Technology Karnataka, India, 2017

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2018

© Kevin Bradley Dsouza 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

A Hybrid Precoding and Signal Detection Framework for Future Wireless Systems

submitted by *Kevin Bradley Dsouza* in partial fulfillment of the requirements for the degree of *Master of Applied Science in Electrical and Computer Engineering*.

Examining Committee:

Vijay K. Bhargava

Supervisor

Vincent Wong

Supervisory Committee Member

Sudip Shekhar

Supervisory Committee Member

Abstract

With energy efficiency and spectrum management being major concerns in future wireless systems, this thesis primarily focuses on the precoding and signal detection capabilities of next generation wireless transceivers.

In the first part of the thesis, we present a parallel framework to make hybrid precoding competitive in fast-fading environments. To enumerate, (i) a low-complexity algorithm which exploits the block diagonal phase-only nature of the analog precoder in a partially connected structure is proposed to arrive at a hybrid precoding solution for a multi-carrier single-user system using orthogonal frequency division multiplexing (OFDM), (ii) the original problem is broken down into independent subproblems of finding the magnitude and the phase components which are solvable in parallel, (iii) a per-RF chain power constraint is introduced instead of the sum power constraint over all antennas, which is much more practical in real systems, (iv) an alternating version of this scheme is proposed for increased spectral-efficiency gains, (v) wideband PCS architecture is critiqued for its applicability in future wireless systems and possible alternatives are discussed.

In the second part of the thesis, we present a signal detection and time-frequency localization framework for smart transceivers. Although deep learning techniques for image analysis have been advancing at a breakneck pace for the past few years, their application to RF data has been relatively less explored. To enumerate our contributions, (i) we present a modification of an existing state-of-the-art object classification technique called Faster-RCNN (FRCNN) [1] for detection and time-frequency localization of the signal in a spectrogram of a wideband RF capture, (ii) insights into the design choices pertaining to the

variables such as short-time Fourier transform (STFT) parameters, spectrogram and anchor sizes and network thresholds are discussed, (iii) synthetic data as per the recently proposed WiFi High Throughput (WiFi-HT) protocol [2] is generated and a mean average precision (mAP) of up to 0.9 is achieved when the model is trained and tested on positive signal to noise ratio (SNR) values, (iv) certain drawbacks of the model with respect to low SNR levels and disparate signal sizes are brought to light and possible solutions are discussed.

Lay Summary

Future wireless systems are poised for the adoption of high frequency bands in conjunction with multiple antenna systems for increased speed of communication. However, the adoption of these technologies comes with its own set of design challenges. One such challenge is to limit the consumption of energy with increasing complexity. While future wireless transceivers need to be energy efficient, they should also aid in security applications. As RF communication becomes pervasive for control and data transmission, from unmanned aerial vehicles (UAVs) to internet of things (IoT) devices, detection of the presence of a wireless device by passive sensing becomes paramount for security purposes. The two concerns mentioned above will be the main focal points of this thesis.

Preface

The following publications have resulted from the research presented in this thesis:

- K. B. Dsouza, K. N. R. S. V. Prasad, and V. K. Bhargava, “Hybrid Precoding with Partially Connected Structure for mmWave Massive MIMO OFDM: A Parallel Framework and Feasibility Analysis” *IEEE Transactions on Wireless Communications*, 2018. (Linked to Chapter 2)
- K. B. Dsouza*, K. N. R. S. V. Prasad*, V. K. Bhargava, “A Deep Learning Framework for Signal Detection and Time-Frequency Localization in Wideband RF Systems” to be submitted to *IEEE Journal on Selected Areas in Communications*. (Linked to Chapter 3)

Statement of Authorship

I am the primary author for the first publication and a joint primary author of the second publication listed above. I have been responsible to develop original ideas, derive mathematical solutions, and generate simulation results for these publications. Prof. Vijay K. Bhargava, my research supervisor, shared with me his vast knowledge in the field and was an instrument of constant support. My colleague, friend and co-author of both the above mentioned works, K.N.R. Surya Prasad, provided valuable guidance and insights in identifying the research problems, developing solution methodologies, and documenting the results.

Some of the simulation results were obtained using MATLAB, the computing environment by MathWorks [3] and CVX, a convex optimization software developed by Grant et al. [4].

Table of Contents

Abstract	iii
Lay Summary	v
Preface	vi
Table of Contents	vii
List of Tables	x
List of Figures	xi
Mathematical Notations	xiii
List of Abbreviations	xiv
Acknowledgements	xvii
Dedication	xviii
1 Motivation	1
1.1 The Promise of 5G Wireless Systems	1
1.2 Need for Hybrid-Precoding	1
1.3 Need for Automatic Signal Detection	2
1.4 Outline of the Thesis	3

2	A Per-RF Chain Hybrid Precoding Approach for Wideband PCS Systems	4
2.1	Introduction	4
2.2	System Model	6
2.2.1	Channel Model	8
2.3	Magnitude-Phase Approach	9
2.3.1	Magnitude Sub-Problem	12
2.3.2	Phase Sub-Problem	16
2.3.3	Par-ArgMod Algorithm	19
2.4	Joint Formulation and Alt-ArgMod	20
2.5	Complexity Analysis	21
2.6	Simulation Results	23
2.6.1	Spectral-Efficiency Evaluation	23
2.6.2	Power Analysis	26
2.6.3	Run-Time Evaluation	28
2.6.4	Realistic Scenarios	29
2.7	Critique of Wideband PCS	30
2.8	Conclusion and Possible Future Work	32
 3	 Signal Detection and Time-Frequency Localization Using Deep Learning	 35
3.1	Introduction	35
3.2	Framework for signal detection and time-frequency localization	38
3.3	Faster RCNN Architecture	41
3.3.1	Base Network	42
3.3.2	Region Proposal Network	43
3.3.3	Detector Network	45
3.3.4	Loss Functions	46
3.3.5	Training procedure	48

Table of Contents

3.4	Design Choices to Adopt FRCNN for Signal Detection and Time-Frequency	
	Localization	48
3.4.1	Choice of STFT parameters	48
3.4.2	Spectrogram size	49
3.4.3	Choice of base network	50
3.4.4	Anchor box sizes and aspect ratios	50
3.4.5	RPN max and min overlap	51
3.4.6	Detector network min and max overlap	52
3.5	Numerical Studies	52
3.5.1	Dataset for training and testing	53
3.5.2	Spectrogram generation	53
3.5.3	Numerical thresholds for the FRCNN model	54
3.5.4	Training performance evaluation	54
3.5.5	Prediction Performance evaluation	57
3.6	Conclusion and Possible Future Work	64
4	Conclusion	66
	Bibliography	67
 Appendix		
A	Proof of results in Chapter 2	76
A.1	Optimality condition for wideband systems PCS	76
A.2	Results with varying number of streams and subcarriers	76

List of Tables

2.1	Complexity comparison.	22
2.2	Run time comparison	28
3.1	mAP with different number of anchors.	59

List of Figures

2.1	Hybrid precoding partially connected scheme in mmWave OFDM systems . . .	6
2.2	Hybrid precoding in OFDM as tensor factorization.	10
2.3	Frontal slice of the tensor \mathbf{F}_{opt}	10
2.4	Decomposition of the frontal slice of the tensor \mathbf{F}_{opt} in Fig. 2.3 into the product of an RF precoder \mathbf{F}_{RF} and a baseband precoder \mathbf{F}_{BB}	11
2.5	Spectral efficiency achieved by precoding algorithms when $N_s = N_{RF}^t = N_{RF}^r = 4$	24
2.6	Spectral efficiency achieved by Par-ArgMod algorithm (a,b) with varying N_{RF} and (c,d) with varying SNR for different values of N_{RF}	25
2.7	Power level achieved by different RF-chains (a,b) with SDR-AltMin and (c,d) with Alt-ArgMod.	27
2.8	Spectral efficiency achieved by Par-ArgMod using different channel models when $N_s = N_{RF}^t = N_{RF}^r = 4$	29
2.9	Spectral efficiency achieved by Optimal and hybrid precoders with increasing N_t	31
3.1	Proposed framework for signal detection and time-frequency localization . . .	38
3.2	The time content, frequency content, and spectrogram of an example wide-band RF capture, when the capture duration is 633 ms, center frequency is 2.4 GHz, wideband bandwidth is 56 MHz, and sampling rate is 56 MHz. . .	39
3.3	Faster RCNN Architecture.	42

List of Figures

3.4	Training loss convergence for the RPN classification, RPN regression, Detector classification, and Detector regression tasks in the FRCNN model.	55
3.5	Targets for the RPN and the inputs for the Detector. The red boxes are the computed targets and the RoIs where as the blue ones indicate the GT signals.	56
3.6	Test images of the trained model. The red boxes are the predicted bounding boxes where as the blue ones indicate the GT signals.	57
3.7	mAP for different base networks.	58
3.8	mAP with different RPN and Detector thresholds.	60
3.9	mAP vs SNR.	62
3.10	Test images of model trained on disparate anchor sizes. The red boxes are the predicted bounding boxes where as the blue ones indicate the GT signals.	63
A.1	Spectral efficiency at SNR = 0 dB, with $N_{RF}^t = N_{RF}^r = 4$	77

Mathematical Notations

The following notations are used in this thesis. Boldface lowercase and uppercase letters, for example, \mathbf{a} and \mathbf{A} , refer to vectors and matrices respectively. The notation $\mathbf{A}_{i,j}$ refers to the entry on the row i and column j of matrix \mathbf{A} . \mathbf{A}^H denotes the complex conjugate transpose of matrix \mathbf{A} . $|\mathbf{C}|$ denotes the modulus of the complex number \mathbf{C} and $\arg(\mathbf{C})$ denotes its argument. $\|\mathbf{A}\|_F$ denotes the Frobenius norm of \mathbf{A} . \mathbf{A}^\dagger denotes the Moore-Penrose pseudo inverse of \mathbf{A} , i.e, $\mathbf{A}^\dagger = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$. A vector with superscript of $*$, for example, \mathbf{x}^* , refers to the ground truth and a subscript of a , for example, \mathbf{x}_a , refers to the anchor box.

List of Abbreviations

3GPP	:	3rd Generation Partnership Project
5G	:	Fifth Generation
AB	:	Anchor Box
ADC	:	Analog to Digital Converter
AED	:	Audio Event Detection
BN	:	Base Network
BS	:	Base Station
CDL	:	Clustered Delay Line
CNN	:	Convolutional Neural Network
CRNN	:	Convolutional Recurrent Neural Network
CRP	:	Chinese Restaurant Process
CSI	:	Channel State Information
CWD	:	Choi-Williams Distribution
DFT	:	Discrete Fourier Transform
FCS	:	Fully Connected Structure
FFT	:	Fast Fourier transform
FRCNN	:	Faster-RCNN
GT	:	Ground Truth
HT	:	High Throughput
IFFT	:	Inverse Fast Fourier Transform
IoT	:	Internet of Things

List of Abbreviations

IoU	: Intersection over Union
ISM	: Industrial Scientific and Medical
mAP	: Mean Average Precision
MIMO	: Multiple-Input Multiple-Output
MSE	: Mean Square Error
NMS	: Non-Max Suppression
NYU	: NewYork University
OFDM	: Orthogonal Frequency Division Multiplexing
OMP	: Orthogonal matching pursuit
PCS	: Partially Connected Structure
PSD	: Power Spectral Density
PZF	: Phased-Zero-Forcing
QCQP	: Quadratic Constraint Quadratic Programming
QP	: Quadratic Programming
RF	: Radio Frequency
RoI	: Region of Interest
RPN	: Region Proposal Network
RSR	: Ratio of Successful Recognition
SDP	: Semi-Definite Programming
SIC	: Successive Interference Cancellation
SNR	: Signal to Noise Ratio
STFT	: Short Time Fourier Transform
SV	: Saleh-Valenzuela
SVD	: Singular Value Decomposition
TCSLs	: Time Clusters and Spatial Lobes
UAV	: Unmanned Aerial Vehicle
UMi-Sc	: Urban Microcell - Street Canyon

List of Abbreviations

- URA : Uniform Rectangular Array
USPA : Uniform Square Planar Array
VGA : Variable Gain Amplifier

Acknowledgements

I would like to thank my supervisor Professor Vijay K. Bhargava for his patience, knowledge, and generous financial support. I thank him for providing me with an excellent research atmosphere in his lab. This thesis would not have been possible without his support, guidance, and encouragement.

I would like to sincerely thank Professor Lutz Lampe and Professor Vincent Wong for their informative courses on wireless systems and convex optimization, which provided me the background and knowledge I needed to venture into these topics. The second part of this thesis was majorly inspired by collaboration with Skycope Technologies Inc. in association with Mitacs Canada. I am grateful to my colleagues at Skycope for their mentorship and Mitacs for funding the research.

My colleague and friend, K. N. R. Surya Prasad, provided invaluable insights throughout the duration of this research and helped refine the final manuscripts to a large extent. I'm indebted to him for his kindness and support.

Dedication

To my mom, who taught me that affection is not for the weak, my dad, who constantly reaffirmed that inquisitiveness is always appreciated, and my sister, who demonstrates that resilience is a virtue.

Chapter 1

Motivation

1.1 The Promise of 5G Wireless Systems

Millimeter-Wave (mmWave) is a key technology that will play a pivotal role in 5G wireless communication systems [5]. The high capacity requirements of next-generation systems make adoption of higher frequency bands such as mmWave inevitable. Recent measurement studies of frequency bands from 5 GHz to 100 GHz in New York City have shown the feasibility of using mmWave technology for cellular communication [6]. The main concern with mmWave is the high path loss and the absorption by atmospheric elements at such high frequencies [7]. On the upside, the small wavelengths at these frequencies make viable the accommodation of high number of antenna elements at the base station (BS) empowering massive MIMO systems. This combination of mmWave and massive MIMO will provide spatial multiplexing and beamforming gains to make up for the detrimental high frequency effects. Further, the spatial diversity offered by this system can be exploited for reliable communication.

1.2 Need for Hybrid-Precoding

Conventional MIMO systems use digital precoders in the baseband which can manipulate both magnitude and phase of signals. This fully digital precoder requires RF chains comprising of analog-to-digital converters (ADCs) and signal mixers, equal in number to the number of base station (BS) antennas. The power consumption and cost of these devices make it prohibitive to implement digital precoders for mmWave massive MIMO systems.

These constraints give rise to a new type of hybrid precoding architecture that addresses these issues by restricting the number of RF chains connecting the baseband precoder to the analog precoder [8]. The baseband precoder is converted to a low-dimensional one and the analog precoder is made high-dimensional, because of which low-cost phase shifters are used instead of variable gain amplifiers (VGAs) [6]. The phase shifter network that connects the RF chains to the antennas determines whether the system is fully or partially connected. The fully connected structure (FCS) connects each RF chain to all the antennas, delivering full beamforming gain. The PCS significantly reduces the number of phase shifters to be used by connecting each RF chain to only a subset of the available antennas, sacrificing some beamforming gain in the process. In this thesis, we adopt the PCS design, propose a low-complexity algorithm respecting per-RF chain power constraints and investigate its performance in a single-user mmWave wideband system.

1.3 Need for Automatic Signal Detection

With the emergence of the IoT, we are currently witnessing a steep surge in the number of wireless devices around us. In future wireless systems, with the IoT devices, UAVs and smart sensors-actuators coexisting alongside traditional mobile phones and access points, it becomes imperative to distinguish between these devices from a privacy and security point of view. Technology that can detect and differentiate such heterogeneous wireless signals and localize their time-frequency span can be commercialized into products for wireless security and spectrum management. From the wireless security perspective, commercial products can be built to make ad-hoc security decisions such as, sending emergency alerts on the potential presence of an unexpected wireless device, employing techniques to jam signals from the detected device, and also estimating the geolocation of the detected device from the incumbent signals. From the spectrum management perspective, commercial products can be built to dynamically share spectrum among the vast number and variety of heterogeneous

devices in the IoT space. Knowing *a priori* which time-frequency resources are underutilized and which ones have minimum interference, smart spectrum allocations can be made in densely populated scenarios. In this thesis, we develop a deep learning framework to passively detect transmitting devices that are transmitting time-frequency localized content in the wideband RF spectrum of interest. We also estimate the time and frequency span of each detected wireless transmission.

1.4 Outline of the Thesis

In Chapter 2, we introduce a per-RF chain hybrid precoding approach for wideband PCS systems. The original joint formulation is broken down into independent magnitude and phase formulations, which is discussed in Sections 2.3, 2.3.1, 2.3.2. The Par-ArgMod algorithm is presented in Section 2.3.3. Following this, a joint formulation leading to an alternating approach is discussed in Section 2.4. Complexity analysis and simulation results are presented in Sections 2.5 and 2.6. Wideband PCS systems are critiqued in Section 2.7 and finally Section 2.8 concludes with a few insights and avenues for future work.

In Chapter 3, we introduce a signal detection and time-frequency localization framework using deep learning. The overall framework and the FRCNN architecture is discussed in Sections 3.2 and 3.3. The various design choices adopted are presented in Section 3.4. This is substantiated by the numerical studies elaborated in Section 3.5 and finally Section 3.6 concludes with the insights obtained and possible avenues for future work.

In Chapter 4 we provide a few concluding remarks, consolidate the insights obtained and motivate future research directions.

Chapter 2

A Per-RF Chain Hybrid Precoding Approach for Wideband PCS Systems

2.1 Introduction

With massive MIMO systems poised to take over next generation wireless systems, hybrid precoding will become indispensable in these systems from the energy-efficiency point of view. Several authors have investigated the hybrid precoding problem in the recent past [9]. Orthogonal matching pursuit (OMP) algorithm is widely adopted and gives satisfactory performance [8], [10]. A spatially sparse precoding scheme is proposed in [8] which exploits the sparse-scattering structure of the mmWave channel to formulate the precoder design as a matrix reconstruction problem constrained by sparsity. A phased-zero-forcing (PZF) scheme is proposed in [11], where a low-dimensional zero-forcing method is implemented on the equivalent channel obtained by the product of RF precoder and the actual channel matrix. The above mentioned schemes cater to the FCS, which consumes high power due to the number of phase shifters used [12]. PCS is seen as a promising candidate towards this end and can provide good performance with higher energy efficiency [13].

Partially connected architectures are looked into in [14]-[18], [19]. The concept of successive interference cancellation (SIC) is used to derive the hybrid precoder in [16]. This work assumes that the baseband precoding matrix is diagonal, implying that the power allocation is only done for different data streams, and the number of RF chains should be equal to that of the data streams. This leaves the analog-precoder to provide the beamforming

gain, which might not be an optimum strategy [16], [17]. Codebook-based design of hybrid precoders for the narrowband systems is discussed in [14]. This codebook-based method searches exhaustively over the analog and baseband precoders without a given design criterion. The above mentioned methods are developed for narrowband systems and they do not consider multi-carrier transmission. Alternating minimization is used as the design criterion in [20] to optimize the analog and digital precoder iteratively. However, the method in [20] is computationally intensive and may not be the best strategy for fast fading channels.

All the above mentioned works consider a sum-power constraint on the baseband precoder. In practical hybrid precoding systems, each RF chain is equipped with its own amplifier and therefore it would be more natural to consider a per-RF chain power constraint. There are works that use the per-antenna power constraint to perform hybrid precoding, however, most of the works like [21] and [22] use the FCS and their schemes are not extendable to PCS. PCS design with per-antenna power constraint is considered in [23], but, this design is limited to narrowband systems and the extension of their scheme to wideband systems is not straightforward. Therefore, to the best of our knowledge, there is no existing work that addresses the problem of hybrid precoding with PCS in wideband systems employing per-RF chain power constraints.

In this chapter we tackle the two main shortcomings in the existing literature. Firstly we make our design adhere to the practical per-RF chain power constraint. Secondly we separate the magnitude and phase formulations to arrive at a much faster precoding solution than existing state of the art in [20] which has the potential to be competitive in fast fading channels. Following this, we recognize that spectral efficiency gains can be achieved by alternating and extend the proposed algorithm to alternate between the magnitudes and phases of the precoders. This alternating version is faster and has better convergence than state of the art alternating minimization in [20] because of the per-RF chain optimization approach along with minimal performance loss at higher SNR. In the coming section we discuss the system and channel model used in this work and also introduce the metric that

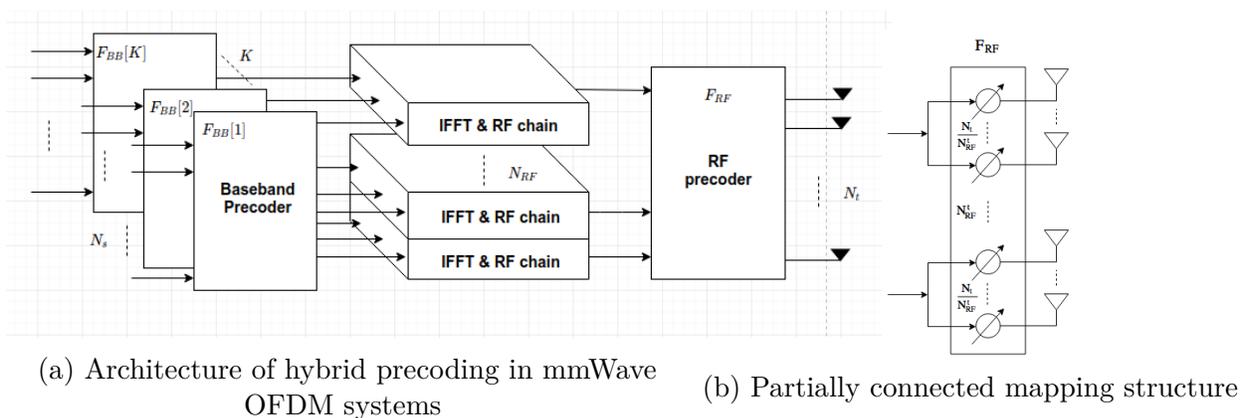


Figure 2.1: Hybrid precoding partially connected scheme in mmWave OFDM systems

will be used to evaluate performance.

2.2 System Model

We consider a single-user mmWave OFDM system as shown in Fig. 2.1a, where N_s data streams are sent over each sub-carrier by N_t transmit antennas and received by N_r receive antennas. The number of RF chains at the transmitter and the receiver for each subcarrier are denoted as N_{RF}^t and N_{RF}^r , respectively.

The hybrid precoder at the transmitter comprises of an $N_{RF}^t \times N_s$ digital baseband precoder \mathbf{F}_{BB} and an $N_t \times N_{RF}^t$ analog RF precoder \mathbf{F}_{RF} . The transmitted signal therefore can be written as $\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}$, where \mathbf{s} is the $N_s \times 1$ symbol vector such that $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s}\mathbf{I}_{N_s}$. The baseband precoding is performed in the frequency domain for each subcarrier followed by an inverse fast Fourier transform (IFFT) operation that consolidates the signals of all the subcarriers. Now, as the analog precoding is a post-IFFT operation, all the signals have to share a common analog precoder [24]. Therefore, the received signal on each subcarrier k can be expressed as

$$\mathbf{y}[k] = \sqrt{\rho}\mathbf{W}_{BB}^H[k]\mathbf{W}_{RF}^H\mathbf{H}[k]\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\mathbf{s} + \mathbf{W}_{BB}^H[k]\mathbf{W}_{RF}^H\mathbf{n}, \quad (2.1)$$

2.2. System Model

where $k \in [0, K - 1]$ is the subcarrier index, ρ stands for the average received power, $\mathbf{W}_{BB}[k]$ is the $N_{RF}^r \times N_s$ digital baseband decoder for the k th subcarrier, \mathbf{W}_{RF} is the $N_r \times N_{RF}^r$ analog decoder shared across all the subcarriers at the receiver, $\mathbf{H}[k]$ is the channel matrix for the k th subcarrier, \mathbf{F}_{RF} is the $N_t \times N_{RF}^t$ shared analog RF precoder, $\mathbf{F}_{BB}[k]$ is the $N_{RF}^t \times N_s$ digital baseband precoder for the k th subcarrier, K is the total number of subcarriers, and \mathbf{n} is the additive white noise vector, the elements of which are independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and variance σ_n^2 . We assume that perfect channel state information (CSI) is available at both the transmitter and the receiver. In practical setups, CSI can be efficiently obtained by channel estimation at the receiver using an adaptive compressed sensing approach with discrete Fourier transform (DFT)-based codebook design [25].

With transmission of Gaussian symbols, the achievable spectral efficiency is given by

$$\begin{aligned}
 R[k] = \log \det(\mathbf{I}_{N_s} + \frac{\rho}{\sigma_n^2 N_s} (\mathbf{W}_{RF} \mathbf{W}_{BB}[k])^\dagger \mathbf{H}[k] \mathbf{F}_{RF} \mathbf{F}_{BB}[k] \\
 \times \mathbf{F}_{BB}^H[k] \mathbf{F}_{RF}^H \mathbf{H}^H[k] (\mathbf{W}_{RF} \mathbf{W}_{BB}[k]))
 \end{aligned} \tag{2.2}$$

The phase shifter only implementation of the analog precoder confines the values of its elements to satisfy the constant modulus constraint given by $|(\mathbf{F}_{RF})_{i,j}| = \frac{1}{\sqrt{N_t}}$ and $|(\mathbf{W}_{RF})_{i,j}| = \frac{1}{\sqrt{N_r}}$. The nature of the phase shifter network that connects the RF chains to the antennas makes this network partially-connected. Fig. 2.1b illustrates the partially connected mapping structure considered in our work, where each RF chain is connected to only a subset of antennas, $N_t/N_{RF}^t = M_t$ at the transmitter end. A similar structure is followed on the receiver side, with the number of antennas per RF chain being $M_r = N_r/N_{RF}^r$. Unless otherwise specified, just the use of M assumes $M = M_t = M_r$ and the use of N_{RF} assumes $N_{RF} = N_{RF}^t = N_{RF}^r$.

2.2.1 Channel Model

In this work, we consider the clustered channel model which is known to characterize the mmWave channels very well [26], [27]. Specifically, considering the theoretical Saleh-Valenzuela (SV) model [6], the channel matrix for the k th subcarrier in the frequency domain is given following [24] as

$$\mathbf{H}[k] = \beta \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{il} \mathbf{a}_r(\phi_{il}^r, \theta_{il}^r) \mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)^H e^{-j2\pi ik/K}, \quad (2.3)$$

where $\beta = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}}$ is the normalization factor, N_{cl} and N_{ray} represent the number of clusters and number of rays in each cluster, K is the total number of subcarriers, and α_{il} is the gain of the l th ray in the i th propagation cluster. It is assumed that the α_{il} terms are i.i.d across the N_{ray} rays in each cluster i and that they follow a complex Gaussian distribution $\mathcal{CN}(0, \sigma_{\alpha,i}^2)$. The variance terms $\sigma_{\alpha,i}^2$ across the N_{cl} clusters satisfy the condition $\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \gamma$, which is the normalization factor used to satisfy the constraint $\mathbb{E}[\|\mathbf{H}\|_F^2] = N_t N_r$. In (2.3), $\mathbf{a}_r(\phi_{il}^r, \theta_{il}^r)$ and $\mathbf{a}_t(\phi_{il}^t, \theta_{il}^t)$ refer to the array response vectors for the receiver and the transmitter respectively, where ϕ_{il}^r and θ_{il}^r represent the angles of azimuth and elevation for arrival at the receiver and ϕ_{il}^t and θ_{il}^t represent likewise for departure at the transmitter. For the antenna array architecture, we consider a uniform square planar array (USPA) at the transmitter with $\sqrt{N_t} \times \sqrt{N_t}$ antenna elements. Hence, the array response vector corresponding to the l th ray in the i th cluster for the transmitter is given as

$$\begin{aligned} \mathbf{a}_t(\phi_{il}^t, \theta_{il}^t) = & \frac{1}{\sqrt{N_t}} [1, \dots, e^{j\frac{2\pi}{\lambda} d(p \sin \phi_{il}^t \sin \theta_{il}^t + q \cos \theta_{il}^t)}, \\ & \dots, e^{j\frac{2\pi}{\lambda} d((\sqrt{N_t}-1) \sin \phi_{il}^t \sin \theta_{il}^t + (\sqrt{N_t}-1) \cos \theta_{il}^t)}]^T \end{aligned} \quad (2.4)$$

where d and λ are the antenna spacing and signal wavelength, and $0 \leq p < \sqrt{N_t}$ and $0 \leq q < \sqrt{N_t}$ are the antenna indices in the 2D plane. A USPA with a similar array response vector is used at the receiver with $\sqrt{N_r} \times \sqrt{N_r}$ antenna elements and same antenna spacing.

The results of this work can be extended to a uniform rectangular array (URA), the response vector for which can be found in [28]. This channel model is extended to conform to practical constraints as given in [26] to obtain the New York University (NYU) model and as given in [27] to obtain the 3rd Generation Partnership Project (3GPP) model for evaluation in realistic scenarios.

2.3 Magnitude-Phase Approach

One approach to solve the hybrid precoding problem is to make our design as close to the unconstrained fully digital precoder as possible with respect to a cost function [29]. This means that we have to minimize the Euclidean distance between the two designs in order to maximize performance, as noted in [8], [20], [30]. This poses to us an interesting problem of tensor factorization with certain constraints. The digital precoder has to deal with power constraints and the analog precoder faces constant modulus constraint because of the phase shifter only approach. Also, the analog precoder is shared by all the subcarriers, which means it has to be factored out of the tensor. We know that the fully digital optimal precoder and decoder are the first N_s columns of the unitary matrices \mathbf{V} and \mathbf{U} of the channel matrix \mathbf{H} respectively. We can obtain \mathbf{V} and \mathbf{U} from the singular-value decomposition (SVD) of the channel matrix, i.e, $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$.

From here on, we mainly focus on the precoder design only. We emphasize that the same design can be extended to the decoder as well because of the similarity in their mathematical formulations (c.f. [8] for example). For OFDM systems, hybrid precoding is essentially a tensor factorization problem as shown in Fig 2.2. The data tensor \mathbf{F}_{opt} is decomposed into a factor matrix \mathbf{F}_{RF} and a core tensor \mathbf{F}_{BB} according to the Tucker-1 decomposition model [29]. The factor matrix is shared by all the subcarriers introducing an inherent suboptimality in the decomposition approach.

We transform the above mentioned tensor factorization problem into two subproblems of

2.3. Magnitude-Phase Approach

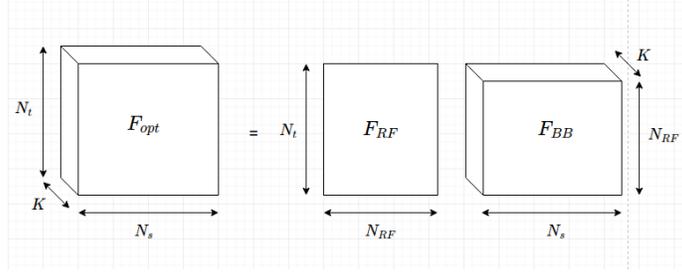


Figure 2.2: Hybrid precoding in OFDM as tensor factorization.

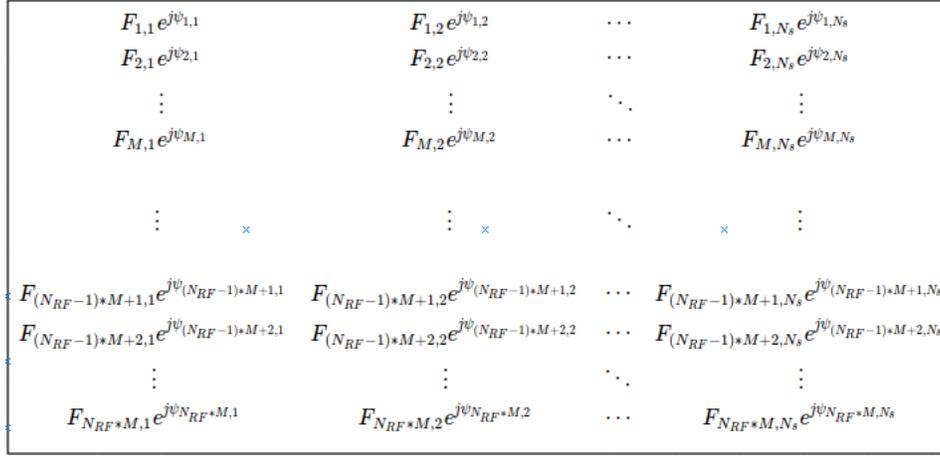


Figure 2.3: Frontal slice of the tensor \mathbf{F}_{opt} .

estimating the magnitudes and phases of the elements of $\mathbf{F}_{BB}[k]$ and \mathbf{F}_{RF} . Each element of \mathbf{F}_{opt} can be written as

$$\mathbf{F}_{opt_{i,j,k}} = F_{i,j,k} e^{j\psi_{i,j,k}}. \quad (2.5)$$

A frontal slice of the tensor \mathbf{F}_{opt} , representing a subcarrier is illustrated in Fig. 2.3, where, $F_{i,j,k} = |\mathbf{F}_{opt_{i,j,k}}|$ and $\psi_{i,j,k} = \arg(\mathbf{F}_{opt_{i,j,k}})$. Since \mathbf{F}_{RF} is the analog precoder, we can express each non-zero element in \mathbf{F}_{RF} as $\mathbf{F}_{RF_{i,j}} = e^{j\theta_{i,j}}$. Note that we do not include the subscript k for the $\theta_{i,j}$ terms in \mathbf{F}_{RF} because the analog precoder is shared across all the subcarriers. For the partially-connected precoder architecture under study (c.f. Fig. 2.1b), \mathbf{F}_{RF} takes a block-diagonal structure because each RF chain is only connected to a small number M_t out of the total number N_t of antennas at the transmitter. This is also illustrated in Fig. 2.4a, where, $|\mathbf{F}_{RF_{i,j}}| = 1/\sqrt{N_t}$ for all the non-zero elements of \mathbf{F}_{RF} and $\theta_{i,j} = \arg(\mathbf{F}_{RF_{i,j}})$.

2.3. Magnitude-Phase Approach

$$\begin{array}{cccc}
 e^{j\theta_1} & 0 & \dots & 0 \\
 e^{j\theta_2} & 0 & \dots & 0 \\
 \vdots & \vdots & \ddots & \vdots \\
 e^{j\theta_M} & 0 & \dots & 0 \\
 \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \dots & e^{j\theta_{(N_{RF}-1)+M+1}} \\
 0 & 0 & \dots & e^{j\theta_{(N_{RF}-1)+M+2}} \\
 \vdots & \vdots & \ddots & \vdots \\
 0 & 0 & \dots & e^{j\theta_{N_{RF}+M}}
 \end{array}$$

(a) Block diagonal structure of \mathbf{F}_{RF} .

$$\begin{array}{cccc}
 B_{1,1}e^{j\phi_{1,1}} & B_{1,2}e^{j\phi_{1,2}} & \dots & B_{1,N_s}e^{j\phi_{1,N_s}} \\
 B_{2,1}e^{j\phi_{2,1}} & B_{2,2}e^{j\phi_{2,2}} & \dots & B_{2,N_s}e^{j\phi_{2,N_s}} \\
 \vdots & \vdots & \ddots & \vdots \\
 B_{N_{RF},1}e^{j\phi_{N_{RF},1}} & B_{N_{RF},2}e^{j\phi_{N_{RF},2}} & \dots & B_{N_{RF},N_s}e^{j\phi_{N_{RF},N_s}}
 \end{array}$$

(b) Frontal slice of the tensor \mathbf{F}_{BB} .

Figure 2.4: Decomposition of the frontal slice of the tensor \mathbf{F}_{opt} in Fig. 2.3 into the product of an RF precoder \mathbf{F}_{RF} and a baseband precoder \mathbf{F}_{BB} .

Moving to the baseband precoder \mathbf{F}_{BB} , each of its elements can be written as

$$\mathbf{F}_{BB_{i,j,k}} = B_{i,j,k}e^{j\phi_{i,j,k}}, \quad (2.6)$$

where $B_{i,j,k} = |\mathbf{F}_{BB_{i,j,k}}|$ and $\phi_{i,j,k} = \arg(\mathbf{F}_{BB_{i,j,k}})$. A frontal slice of the tensor \mathbf{F}_{BB} , corresponding to the subcarrier k is illustrated in Fig. 2.4b. In order for our hybrid precoder to be as close to the optimal precoder as possible, we take the metric of mean squared error (MSE) following [16] and [20] and therefore to minimize the approximation involved adopt the least squares L2-norm as the design objective. The optimization problem for the q th RF chain turns out to be

$$\begin{aligned}
 & \underset{\theta, B, \phi}{\text{minimize}} && \sum_{i=(q-1)M_t+1}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K \|\mathbf{F}_{i,j,k}e^{j\psi_{i,j,k}} - CB_{q,j,k}e^{j(\theta_i+\phi_{q,j,k})}\|^2 \\
 & \text{subject to} && \|B_{q, :, k}\|^2 \leq \frac{N_s}{N_t}
 \end{aligned} \quad (2.7)$$

where $C = \frac{1}{\sqrt{N_t}}$, is the constant modulus of the non-zero terms of the analog precoder. Omitting the summations, subscripts and constraints for ease of representation, the norm

term can be simplified as follows

$$\begin{aligned}
\|Fe^{j\psi} - CBe^{j(\theta+\phi)}\|^2 &= [F \cos \psi - CB \cos(\theta + \phi)]^2 + [F \sin \psi - CB \sin(\theta + \phi)]^2 \\
&= F^2 \cos^2 \psi + C^2 B^2 \cos^2(\theta + \phi) - 2CFB \cos \psi \cos(\theta + \phi) + \\
&F^2 \sin^2 \psi + C^2 B^2 \sin^2(\theta + \phi) - 2CFB \sin \psi \sin(\theta + \phi) \\
&= F^2 + C^2 B^2 - 2CFB[\cos \psi \cos(\theta + \phi) + \sin \psi \sin(\theta + \phi)] \\
&= F^2 + C^2 B^2 - 2CFB \cos(\psi - (\theta + \phi))
\end{aligned}$$

Taking the limits of $\cos(x)$ into consideration, we get the inequality

$$F^2 + C^2 B^2 - 2CFB \leq F^2 + C^2 B^2 - 2CFB \cos(\psi - (\theta + \phi)) \quad (2.8)$$

Equation 2.8 will be used to pose the magnitude problem in the coming subsection. For now, plugging this simplified version back into (2.7), the original problem boils down to

$$\begin{aligned}
&\underset{\theta, B, \phi}{\text{minimize}} && \sum_{i=(q-1)M_t+1}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K F_{i,j,k}^2 + C^2 B_{q,j,k}^2 - 2CF_{i,j,k} B_{q,j,k} \cos(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k})) \\
&\text{subject to} && \|B_{q,:k}\|^2 \leq \frac{N_s}{N_t}, q = 1, \dots, N_{RF}^t
\end{aligned} \quad (2.9)$$

In the coming subsections we propose independent formulations for the magnitude and phase of the hybrid precoder. We also provide theoretical justifications on how the formulation given in (2.9), referred to from here on as the original formulation, can be relaxed into the proposed magnitude and phase formulations. Following this, a way to alternate between the magnitude and the phase formulations is discussed.

2.3.1 Magnitude Sub-Problem

From Fig. 2.2, we note that each row of the baseband precoder \mathbf{F}_{BB} is applied a different phase shift by the different non-zero entries of the analog precoder \mathbf{F}_{RF} to reach different

2.3. Magnitude-Phase Approach

rows of the optimal precoder. Taking all the K subcarriers into account, we can write the first $M_t K$ equations for the q th RF chain as

$$\begin{aligned} \mathbf{F}_{opt_{i,:},k} &\approx e^{j\theta_i} \mathbf{F}_{BB_{q,:},k}, \\ 1 \leq i \leq M_t, 1 \leq k \leq K, q &= \left\lceil \frac{i}{M_t} \right\rceil. \end{aligned} \quad (2.10)$$

Each of the $M_t K$ equations listed above can be written as two separate equations - one equating the magnitude of the LHS terms with that of the RHS terms and the other equating the phase of the LHS terms with that of the RHS terms. The same procedure can be followed for the next $M_t K$ equations and so on. From (2.6) and (2.10), we can mathematically equate the magnitudes in the hybrid precoder for subcarrier k and RF chain q as

$$\begin{aligned} F_{i,j,k} &\approx B_{q,j,k}, \\ 1 \leq i \leq M_t, 1 \leq j \leq N_s, q &= \left\lceil \frac{i}{M_t} \right\rceil. \end{aligned} \quad (2.11)$$

As we consider MSE as the optimization metric, we adopt the least squares L2-norm as the design objective and formulate the magnitude sub-problem for each sub-carrier k and the q th RF chain as

$$\begin{aligned} \underset{B_{q,:},k}{\text{minimize}} \quad & \sum_{i=(q-1)M_t+1}^{qM_t} \sum_{j=1}^{N_s} \|B_{q,j,k} - F_{i,j,k} \sqrt{N_t}\|^2 \\ \text{subject to} \quad & \|B_{q,:},k\|^2 \leq \frac{N_s}{N_t} \end{aligned} \quad (2.12)$$

Remark 1 (On the relation between the original formulation in (2.9) and the magnitude formulation in (2.12)). *From (2.8) and (2.9), we know that the lower bound to the original*

2.3. Magnitude-Phase Approach

formulation consists of only the magnitude terms and is given by

$$\begin{aligned}
 & \underset{B_{q,j,k}}{\text{minimize}} && \sum_{i=(q-1)M_t+1}^{qM_t} \sum_{j=1}^{N_s} \|F_{i,j,k} - B_{q,j,k}C\|^2 \\
 & \text{subject to} && \|B_{q,:k}\|^2 \leq \frac{N_s}{N_t}
 \end{aligned} \tag{2.13}$$

Note that the optimization problems in (2.12) and (2.13) are equivalent because $C = \frac{1}{\sqrt{N_t}}$. Consequently, we note that our magnitude formulation is trying to minimize a lower bound to the original formulation. An important aspect to note here is that by separating the magnitude formulation out of the original problem, we are able to do away with the dependence among the subcarriers. This allows us to solve the magnitude problem for each subcarrier independently. Since we operate with a total of K subcarriers, we should solve K similar problems as (2.12) - one for each subcarrier. This can be done in parallel because the magnitude subproblems do not have any inter-dependencies between sub-carriers.

In (2.12), we have considered a per RF chain power constraint for the baseband precoder. Beamforming with per-antenna power constraint has been investigated in [31]-[34], however, the philosophy in these works can be extended to a per-RF chain power constraint as well. Most of the works on beamforming adopt a sum-power constraint on the antennas and following [31], we believe that a per-RF chain power constraint is more realistic as in practical hybrid precoding systems each RF chain is equipped with its own power amplifier and is limited by the linearity of that amplifier. Having a per-RF chain power constraint would also aid in achieving equal power allocation among different RF chains so that all RF chains will be equally active at a given time [35]. This is further elaborated in section 2.6.2.

The per RF chain power constraint follows from the formulation of the regular precoding constraint which limits the Frobenius norm of the precoders, i.e.,

$$\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = N_s \tag{2.14}$$

Due to the special block diagonal structure of the analog precoder \mathbf{F}_{RF} (c.f. Fig. 2.4a), the power constraint can be rewritten as

$$\frac{N_t}{N_{RF}^t} \|\mathbf{F}_{BB}\|_F^2 = N_s. \quad (2.15)$$

Dividing the power equally among the N_{RF}^t RF chains, (2.15) transforms into the per RF chain power constraint given in (2.12).

To solve (2.12), we firstly relax the power constraint so as to obtain a power budget that maintains the convexity of the problem at hand. Upon doing so, we end up with a quadratically constrained quadratic programming (QCQP) problem which is convex and can be solved by standard convex optimization techniques [36]. Results show that the power limit is achieved by this formulation for each RF chain.

A similar formulation as given in (2.12) can be formed for each of the N_{RF}^t successive rows of the baseband precoder \mathbf{F}_{BB} (one per RF chain). That is, in order to cover the entire precoder, we need to solve a total N_{RF}^t magnitude problems per subcarrier. We can solve these N_{RF}^t problems simultaneously in parallel because of the block diagonal nature of the analog precoder \mathbf{F}_{RF} (c.f. Fig. 2.4a) as the partially connected nature of the hybrid precoder renders each RF chain to be independent of the others. We can visualize the optimal precoder as a group of N_{RF}^t tensors stacked on top of each other which allow us to deal with each of them separately. The resulting baseband precoder may not directly maximize the spectral efficiency [8], [24], but it makes for a good substitute because it helps simplify the given problem. The magnitudes of the baseband precoder for all the subcarriers are obtained by solving these independent set of N_{RF}^t formulations per subcarrier.

The system of equations in (2.11) is overdetermined because the number of independent equations ($M_t N_s$) is much higher than the number of variables (N_s) involved. Specifically, the magnitude equations in (2.11) reveal that, for a given subcarrier k , each of the rows of the magnitude matrix B of the baseband precoder tries to approximate a set of M_t rows of

the magnitude matrix F of the optimal precoder. For an optimal solution to exist we require that

$$N_s \geq N_s \frac{N_t}{N_{RF}^t} \quad (2.16)$$

i.e, we need $N_{RF}^t \geq N_t$. This is not practical for hybrid precoding systems, because of which we end up with an overdetermined system with a minimum norm solution.

2.3.2 Phase Sub-Problem

The phase sub-problem can be derived from (2.10) by equating the phase terms on the l.h.s with those on the r.h.s. We obtain the following set of $M_t K N_s$ equations for the q th RF chain as

$$e^{j\psi_{i,j,k}} \approx e^{j\theta_i} e^{j\phi_{q,j,k}},$$

$$\forall 1 \leq i \leq M_t, 1 \leq j \leq N_s, 1 \leq k \leq K, q = \left\lceil \frac{i}{M_t} \right\rceil \quad (2.17)$$

where $-\pi \leq \theta_i \leq \pi, -\pi \leq \phi_{q,j,k} \leq \pi, \forall i, j, k$

Observe from (2.17) and (2.11) that the main difference between the phase and magnitude formulations is that the phase equations (c.f. (2.17)) are obtained upon including the constraint that the analog precoder needs to be shared across the K subcarriers. This is unlike the magnitude formulation, where the equations (c.f. (2.11)) are obtained upon including the constraint that the baseband precoders for the K subcarriers are independent of each other. Applying a logarithmic transformation on the l.h.s and r.h.s of (2.17), we have the following set of $M_t K N_s$ equations

$$\psi_{i,j,k} \approx \theta_i + \phi_{q,j,k},$$

$$\forall 1 \leq i \leq M_t, 1 \leq j \leq N_s, 1 \leq k \leq K, q = \left\lceil \frac{i}{M_t} \right\rceil \quad (2.18)$$

where $-\pi \leq \theta_i \leq \pi, -\pi \leq \phi_{q,j,k} \leq \pi, \forall i, j, k$

2.3. Magnitude-Phase Approach

Following the method of linear least squares to solve for the phases θ_i and $\phi_{q,j,k}$ of the digital and analog precoders respectively, we formulate the phase sub-problem corresponding to the RF chain q in the hybrid precoder as

$$\begin{aligned} & \underset{\theta_i, \phi_{q,j,k}}{\text{minimize}} && \sum_{i=1+(q-1)M_t}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K (\theta_i + \phi_{q,j,k} - \psi_{i,j,k})^2, \\ & \text{subject to} && -\pi \leq \theta_i \leq \pi, \quad -\pi \leq \phi_{q,j,k} \leq \pi, \quad \forall i, j, k \end{aligned} \quad (2.19)$$

The optimization problem in (2.19) is a linear-constrained quadratic programming (QP) problem and can be solved using standard convex optimization techniques [36]. A similar formulation as given in (2.19) can be formed for each RF chain q . Since we have a total of N_{RF}^t RF chains at the transmitter, we should solve a total of N_{RF}^t subproblems to obtain all the phases of the hybrid precoder.

Remark 2 (On the relation between the original formulation in (2.9) and the phase formulation in (2.19)). *The phase formulation given in (2.19) tries to reduce the gap between the original formulation in (2.9) and the magnitude formulation in (2.12), as explained below. Note from (2.12) that in the magnitude formulation, we solve for a lower bound to the original formulation given in (2.9). The $\cos(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k}))$ terms in (2.9) are responsible for the gap between the objective functions in (2.9) and (2.12). To reduce this gap, we can maximize the $\cos(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k}))$ terms w.r.t $(\theta_i$ and $\phi_{q,j,k})$, i.e., we can solve the following optimization problem*

$$\begin{aligned} & \underset{\theta_i, \phi_{q,j,k}}{\text{maximize}} && \sum_{i=1+(q-1)M_t}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K \cos(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k})) \\ & \text{subject to} && -\pi \leq \theta_i \leq \pi, \quad -\pi \leq \phi_{q,j,k} \leq \pi, \end{aligned} \quad (2.20)$$

The problem in (2.20) is non-convex because the cosine function is not convex in $(-\pi, \pi)$. By resorting to a first order Taylor approximation, we can simplify (2.20) as

2.3. Magnitude-Phase Approach

$$\begin{aligned}
 & \underset{\theta, \phi}{\text{maximize}} && \sum_{i=1+(q-1)M_t}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K 1 - \frac{(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k}))^2}{2} \\
 & \text{subject to} && -\pi \leq \theta_i \leq \pi, \quad -\pi \leq \phi_{q,j,k} \leq \pi
 \end{aligned} \tag{2.21}$$

Note that the optimization problem in (2.21) is equivalent to the phase formulation given in (2.19). That is, our phase formulation is trying to minimize the difference between the objective functions in the original formulation and the magnitude formulation.

An important aspect to note from the phase formulation in (2.19) is that, unlike the magnitude formulation in (2.12), (2.19) has to deal with the dependence among the K subcarriers because all the subcarriers share the same analog precoder. Consequently, we cannot solve the phase sub-problems for each subcarrier independently. However, we note from (2.19) that the phase formulation for each RF chain q is independent of the remaining RF chains due to the block-diagonal nature of the analog precoder. Therefore, we can still solve N_{RF}^t phase sub-problems in parallel (one for each RF chain q) to obtain the entire analog precoder.

Similar to the magnitude formulation, the system of linear equations in (2.18) is overdetermined. This is because the available degrees of freedom, i.e, the variables θ_i and $\phi_{q,j,k}$, are lesser in number than the number of constraints on these degrees, i.e., the number of equations involved. Specifically, the overdetermined system of phase formulations has $M_t K N_s$ equations for a set of $M_t + K N_s$ variables. The main cause for the overdetermined nature of the system is that the analog precoder, represented by the phase shifts θ_i , is shared across the K subcarriers. In order to have an optimal solution we need to have

$$\begin{aligned}
 \frac{N_t}{N_{RF}^t} + N_s K &\geq \frac{N_t}{N_{RF}^t} N_s K \\
 N_t + N_s K (N_{RF}^t - N_t) &\geq 0
 \end{aligned}$$

Therefore, for $N_{RF}^t \geq N_t$, we always have an optimal solution and for $N_{RF}^t < N_t$, we have

an optimal solution if

$$N_t \geq N_s K |N_t - N_{RF}^t| \quad (2.22)$$

It can be seen that as N_s or K increase, it becomes increasingly difficult to satisfy (2.22). Thus we end up with an overdetermined system with a minimum norm solution, similar to the magnitude formulation.

2.3.3 Par-ArgMod Algorithm

Algorithm 1 : Par-ArgMod algorithm

- 1: **Input:** \mathbf{H}
 - 2: Compute SVD to get \mathbf{F}_{opt} and set $q=1$;
 - 3: **Solve in parallel** for $q = 1, \dots, N_{RF}^t$
 - 4: For a given subcarrier k , find the magnitude values of the q th row of $\mathbf{F}_{BB}[k]$ using the magnitude formulation (2.12). Solve K such optimization problems, i.e., one for each subcarrier, in parallel.
 - 5: Find the q th block diagonal column entry of \mathbf{F}_{RF} and phases of the q th row of \mathbf{F}_{BB} using the phase formulation (2.19).
 - 6: **end** when the N_{RF}^t parallel problems are solved
 - 7: Rearrange the magnitude and phase values to form \mathbf{F}_{RF} and \mathbf{F}_{BB}
-

In Algorithm 1, also referred to from here on as the Par-ArgMod algorithm, we present a summary of the steps followed in our hybrid precoder design. The input to the proposed algorithm is the channel matrix \mathbf{H} , which is assumed to be perfectly known at the transmitter. We then apply SVD to the channel matrix and obtain the fully digital precoder \mathbf{F}_{opt} . Next, we consider the M_t transmit antennas connected to the RF chain q and obtain the magnitude and phase terms of the hybrid precoder by solving (2.12) and (2.19) respectively. This procedure is repeated in parallel for the N_{RF}^t RF chains, so as to fully obtain the hybrid precoder matrices \mathbf{F}_{RF} and \mathbf{F}_{BB} .

2.4 Joint Formulation and Alt-ArgMod

Introducing the approximation in (2.21) in (2.9) we have

$$F^2 + C^2 B^2 - 2CFB \cos(\psi - (\theta + \phi)) = (F - CB)^2 + CFB \frac{(\psi - (\theta + \phi))^2}{2} \quad (2.23)$$

Alt-ArgMod given in Algorithm 2, alternates between the magnitude and the phase formulations using equation (2.23). A good choice for the initial phase values is arrived at by using the phase formulation in (2.19). Essentially, when the phases are fixed in (2.23), the resulting formulation becomes a QCQP problem which can be written as

$$\begin{aligned} & \underset{B_{q,j,k}}{\text{minimize}} && \sum_{i=(q-1)M_t+1}^{qM_t} \sum_{j=1}^{N_s} (F_{i,j,k} - B_{q,j,k}C)^2 + C_1 F_{i,j,k} B_{q,j,k} \\ & \text{subject to} && \|B_{q,j,k}\|^2 \leq \frac{N_s}{N_t} \end{aligned} \quad (2.24)$$

where $C_1 = C \frac{(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k}))^2}{2}$ is the resulting constant after fixing the phases. Likewise when the resulting magnitudes from (2.24) are fixed in (2.23), the resulting formulation becomes a QP problem with linear constraints which can be written as

$$\begin{aligned} & \underset{\theta, \phi}{\text{minimize}} && \sum_{i=1+(q-1)M_t}^{qM_t} \sum_{j=1}^{N_s} \sum_{k=1}^K C_1 + C_2 \frac{(\psi_{i,j,k} - (\theta_i + \phi_{q,j,k}))^2}{2} \\ & \text{subject to} && -\pi \leq \theta_i \leq \pi, \quad -\pi \leq \phi_{q,j,k} \leq \pi \end{aligned} \quad (2.25)$$

where $C_1 = (F_{i,j,k} - B_{q,j,k}C)^2$ and $C_2 = CF_{i,j,k}B_{q,j,k}$.

In the coming section, we present a detailed analysis on the computational costs incurred by the proposed Par-ArgMod and Alt-ArgMod algorithm and compare them with existing wideband PCS schemes.

Algorithm 2 : Alt-ArgMod algorithm

- 1: **Input:** \mathbf{H}
 - 2: Compute SVD to get \mathbf{F}_{opt} and set $q=1$;
 - 3: **Solve in parallel** for $q = 1, \dots, N_{RF}^t$
 - 4: Initialize the phases in (2.23) using the phase formulation (2.19).
 - 5: Solve for the magnitudes using the resulting formulation as in (2.24).
 - 6: Solve for the phases by replacing the resulting magnitudes from step 5 as in (2.25).
 - 7: Rearrange the magnitude and phase values to form \mathbf{F}_{RF} and \mathbf{F}_{BB} per RF-chain.
 - 8: **Repeat** steps 5,6,7 until convergence criteria is met.
 - 9: **end** when the N_{RF}^t parallel problems are solved
-

2.5 Complexity Analysis

In this section we compare the computational complexity of our proposed algorithm with two existing wideband schemes namely the SDR-AltMin algorithm [20] and the Fixed-Wideband scheme [18]. Other hybrid precoding algorithms also exist, for example, the SIC-based hybrid precoding scheme [16], but the extensions of such schemes to OFDM systems are not currently available and are therefore skipped here.

In Par-ArgMod, we may note that the first step involves the SVD of the $N_r \times N_t$ matrix \mathbf{H} , which can be computed in $\mathcal{O}(N_r^2 N_t)$ operations. The magnitude formulation is a convex QCQP problem, which can be solved using interior point methods in $\mathcal{O}(N_s^3)$ computations. As noted in [37], [38], QCQP and QP problems are given by an iteration complexity of $\mathcal{O}(\sqrt{N_s} \log(N_s/\epsilon))$ with an ϵ -approximate solution using the primal-dual interior-point algorithm for small updates. This should be done once per transmit RF chain and subcarrier, leading to a total of $N_{RF}^t K$ problems for obtaining the magnitude terms of the hybrid precoder. The phase formulation in (18) is a QP problem which can be solved by using interior point methods with a computational complexity of $\mathcal{O}((M_t N_s K)^3)$ and iteration complexity of $\mathcal{O}(\sqrt{M_t N_s K} \log(M_t N_s K/\epsilon))$. This process is followed in N_{RF}^t parallel problems - one per RF chain. Alt-ArgMod also takes similar time along with an additional factor of L which is the number of iterations taken to converge.

The SVD step is common to Par-ArgMod, Alt-ArgMod and SDR-AltMin algorithms be-

2.5. Complexity Analysis

cause it is required to obtain the optimal precoder values. The SDR-AltMin algorithm [20] has a semidefinite programming (SDP) problem inside an alternating minimization block, which makes it increasingly complex. Each iteration of the SDR AltMin algorithm requires $\mathcal{O}((N_s N_{RF}^t)^3)$ computations for the baseband precoder and $\mathcal{O}(N_t N_s K)$ for the analog RF precoder. Assuming the SDR-AltMin algorithm takes L iterations to converge, the total complexity would be $\mathcal{O}(L N_t N_s K + L(N_s N_{RF}^t)^3)$. SVD is also carried out in the Fixed Wideband algorithm [18] to obtain the singular values of the covariance matrix. The covariance matrix itself can be obtained in $\mathcal{O}(M_t^2 N_r K)$ and the baseband precoder can be got in $\mathcal{O}(M_t^2 N_s)$. The complexities of these algorithms are juxtaposed in Table 2.1.

Algorithm	Complexity
SDR-AltMin [20]	$\mathcal{O}(L N_t N_s K + L(N_s N_{RF}^t)^3 \sqrt{N_s N_{RF}^t} \log(N_s N_{RF}^t / \epsilon))$
Fixed Wideband [18]	$\mathcal{O}(M_t^2 N_r K + M_t^2 N_s)$
Par-ArgMod	$\mathcal{O}(\sqrt{M_t N_s K} \log(M_t N_s K / \epsilon) (M_t N_s K)^3 + (N_s)^3 \sqrt{N_s} \log(N_s / \epsilon))$
Alt-ArgMod	$\mathcal{O}(L \sqrt{M_t N_s K} \log(M_t N_s K / \epsilon) (M_t N_s K)^3 + L(N_s)^3 \sqrt{N_s} \log(N_s / \epsilon))$

Table 2.1: Complexity comparison.

Par-ArgMod allows for parallelization; the magnitude and phase problems are solved in parallel and these two problems are in turn solved as N_{RF}^t parallel subproblems. In addition, each of the N_{RF}^t subproblems for the magnitude formulation are solved for the K subcarriers as K parallel subproblems. The Alt-ArgMod also alternates separately for N_{RF}^t RF chains. The SDR-AltMin algorithm is linear in the number of transmit antennas N_t , which is the term we vary in massive MIMO studies. In contrast, thanks to the multiple levels of parallelization involved, our proposed algorithms increase as $\mathcal{O}(M_t^{7/2} \log(M_t / \epsilon))$ and are therefore dependent on M_t rather than N_t , which as we show in section 2.7, needs to be kept relatively constant with increase in N_t .

The proposed algorithms also offer other distinct advantages over existing schemes. Alternating minimization does not have convergence guarantees to a global minimum or even a stationary point, and only converges to a solution where the cost function ceases to decrease [39], [40], [41]. It also requires multiple runs because it imposes the non-trivial task of

choosing good initial values to avoid convergence to bad local optima [20]. These drawbacks are observed in Alt-ArgMod as well, however, convergence is achieved sooner than SDR-AltMin on average because of a good initialization mechanism. This is discussed further in section 2.6.2. Also, the per-RF chain power constraint considered in the proposed algorithms is a safer approach when allocating power to the different RF chains than the sum-power constraint used in Fixed Wideband and SDR-AltMin.

2.6 Simulation Results

In this section, we evaluate the performance of our proposed algorithm. N_s data streams are sent over each of the $K = 12$ subcarriers with $N_t = 36$ and $N_r = 16$ antennas. The antenna elements are separated by half wavelength distance. For the theoretical Saleh-Valenzuela model we set $N_{cl} = 5$ clusters, $N_{ray} = 10$ rays and the average power of each cluster is set to $\sigma_{\alpha,i}^2 = 1$. The azimuth and elevation AoDs and AoAs follow the Laplacian distribution with uniformly distributed mean angles over $[0, 2\pi)$ and angular spread of 10 degrees [20]. All the results are averaged over 1000 independent channel realizations. We compare the spectral efficiency performance achieved by the proposed algorithm with that of the SDR-Alternating Minimization algorithm [20], the Fixed Wideband algorithm proposed in [18], and the codebook-based [42] methods. These three methods serve as our baseline for hybrid precoding with PCS in massive MIMO OFDM systems. We also consider two fully-connected precoding schemes, namely, the OMP [8] and the PE-AltMin [20] for comparison.

2.6.1 Spectral-Efficiency Evaluation

Firstly, Fig. 2.5 shows the variation of spectral efficiency with respect to SNR. Note that all baseline schemes presented in the figure consider a sum-power constraint on the RF chains, whereas, we consider a more-realistic per-RF chain power constraint. We notice that the fully connected architectures follow closely after the optimal precoder, whereas the partially

2.6. Simulation Results

connected structures lag in terms of spectral efficiency performance. The partially connected architectures exhibit near similar performance for low SNR regions which is rather poor when compared to the fully connected architectures and deviate slightly when SNR is high, although the performance gap is not too conspicuous. The DFT-codebook based method is seen to perform poorly in our setting mainly because of the finite size of the codebook. The size of the codebook for a given dimension is fundamentally limited by the number of mutually unbiased bases that are available and the lower bound to the minimum distance that can exist between the codewords.

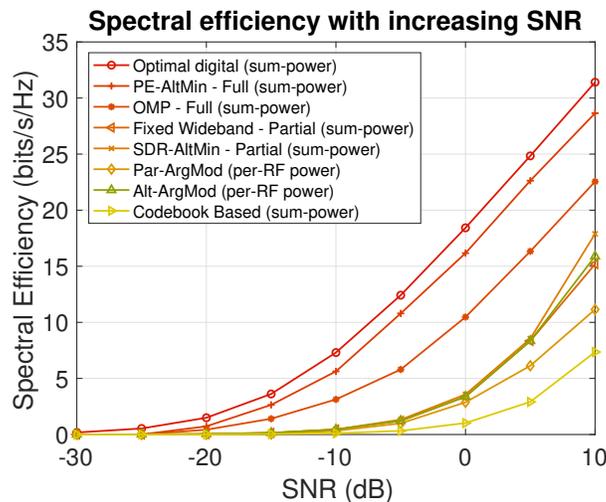


Figure 2.5: Spectral efficiency achieved by precoding algorithms when $N_s = N_{RF}^t = N_{RF}^r = 4$.

Previous works [31],[43] have shown that the achievable capacity with the per-RF chain power constraint is much lower than than the achievable capacity with the sum-power constraint on the RF chains. This difference in the maximum achievable spectral efficiency is a major reason for the suboptimal performance of our scheme at high SNR. Although the Alt-ArgMod outperforms the non-alternating version considerably, further investigation is required on building non-alternating precoding schemes that can respect the per-RF chain power constraint but not degrade in spectral efficiency performance at high SNR. Another reason for the sub-optimality of Par-ArgMod is that we separate the original joint formulation in (2.9) into parallelizable magnitude and phase optimization problems. This separation introduces additional constraints on the degrees of freedom available in the precoder design,

2.6. Simulation Results

thus resulting in a decreased spectral efficiency performance [44]. This argument is validated by the improved performance of Alt-ArgMod which works with the joint formulation given in (2.23). Lastly, due to the partially connected wideband nature of the hybrid precoder, both our magnitude and phase formulations solve overdetermined systems and this adds to the sub-optimality.

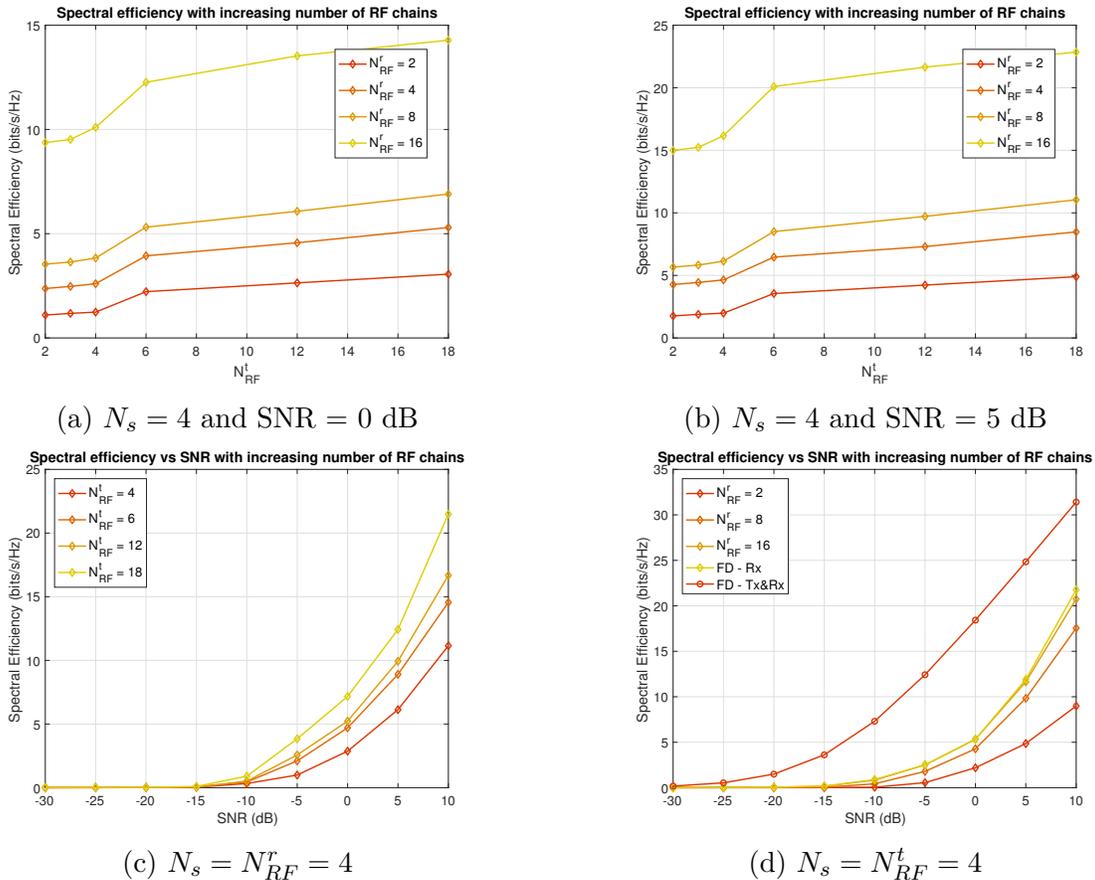


Figure 2.6: Spectral efficiency achieved by Par-ArgMod algorithm (a,b) with varying N_{RF} and (c,d) with varying SNR for different values of N_{RF} .

Next, we fix the number of data streams $N_s = 4$, the signal to noise ratio (SNR) $\frac{\rho}{\sigma_n^2} = 0$ dB and vary the number of RF chains N_{RF}^t and N_{RF}^r at the transmitter and receiver respectively. It can be observed from Fig. 2.6a that increasing number of RF chains at the transmitter can increase the spectral efficiency. In order to achieve further gains, increasing the number of RF chains at the receiver is a reasonable solution. A similar trend is observed when we

increase the SNR to 5 dB and redo the experiment as seen in Fig. 2.6b. There is a jump in spectral efficiency when $N_{RF}^r = N_r$ in both the cases. These figures illustrate the trade-off introduced by hybrid precoders having limited RF chains - we achieve energy savings at the cost of loss in spectral efficiency. It can be observed from Fig. 2.6c that with increasing N_{RF}^t the spectral efficiency increases, while N_s and N_{RF}^r are kept constant. Increasing the number of RF chains at the receiver has a similar effect as seen in Fig. 2.6d. Moreover, using $N_{RF}^r = N_r$ is equivalent to having a fully digital architecture at the receiver.

2.6.2 Power Analysis

The per-RF chain power constraint considered in the proposed algorithm is a safer approach when allocating power to the different RF chains than the sum-power constraint used in Fixed Wideband and SDR-AltMin. This is because unequal power allocation among the RF chains would feed into the non-linearity of the power amplifier and also decrease the overall amplifier efficiency. For further insight, we note that the power expenditure on linear power amplifiers can be expressed as $P_{PA} = P_{in}/\eta$, where P_{in} is the input power to the power amplifiers and η is the power amplifier efficiency. Typically, η depends on the output power P_{out} of the PA and is given by $\eta = \eta_{max} \sqrt{\frac{P_{out}}{P_{max}}}$ [45], where η_{max} is the maximum power amplifier efficiency and P_{max} is the maximum output power of the power amplifier. When P_{max} is higher than the P_{out} (which happens in the case of unequal power allocation), we note that η is smaller and consequently, the power expenditure P_{PA} at the power amplifier is higher. Techniques which minimize the power amplifier losses in massive MIMO systems are still an ongoing topic of investigation [46].

The power achieved by different RF-chains over time is shown in Fig 2.7. The maximum allowed sum power is $N_s N_{RF}^t / N_t$ and per-RF power is N_s / N_t . SDR-AltMin is disparate in the way it assigns power to different RF-chains and breaches the per-RF power limit as seen in Fig. 2.7a, which is not acceptable from a hardware perspective. It also has difficulties converging on some occasions as observed in Fig. 2.7b. On the contrary, Alt-ArgMod is fair

2.6. Simulation Results

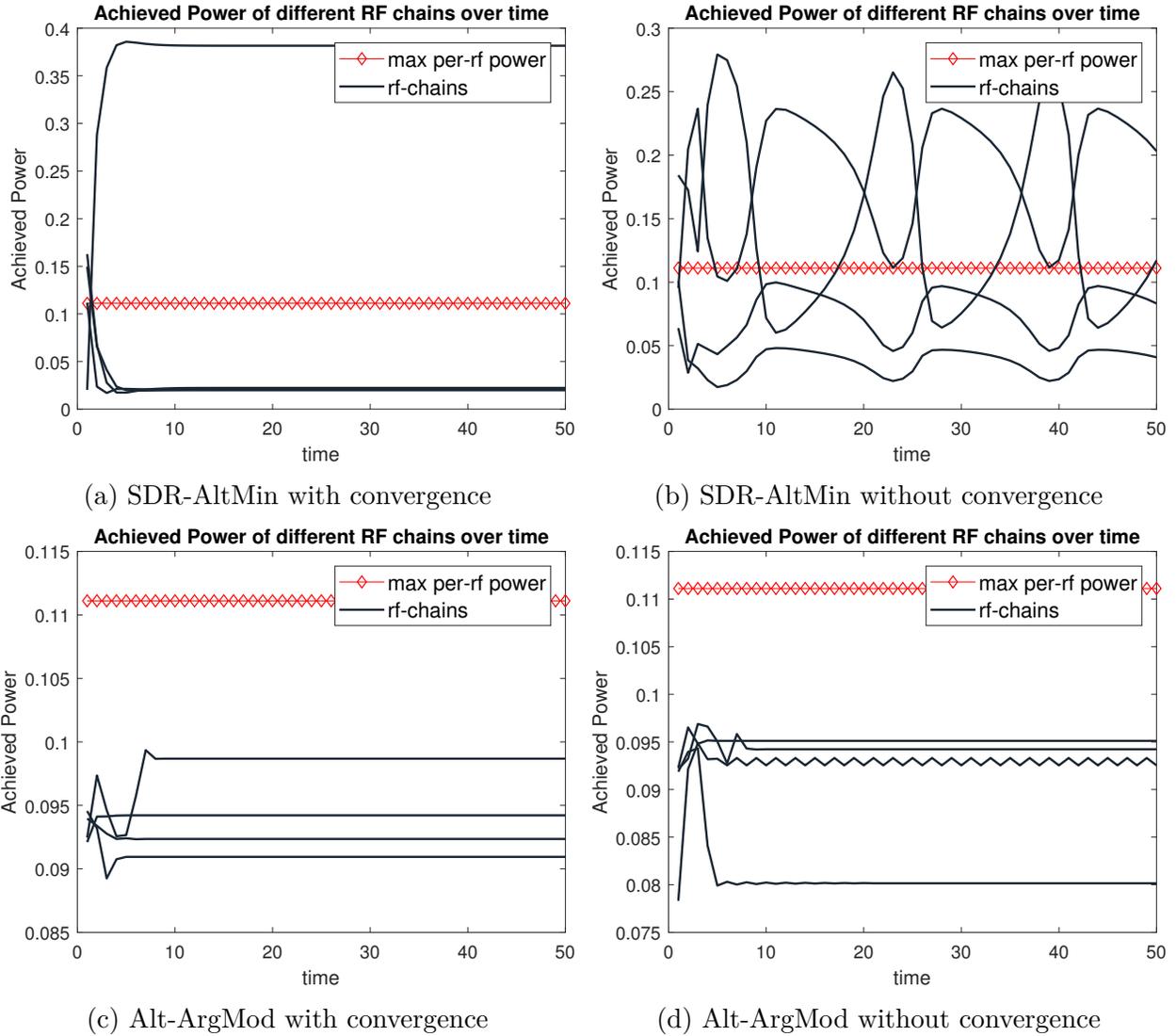


Figure 2.7: Power level achieved by different RF-chains (a,b) with SDR-AltMin and (c,d) with Alt-ArgMod.

in its allocation of power and does not breach the power limit for any RF chain because of the built-in constraints which is shown in Fig 2.7c. Also, as the algorithm alternates per-RF chain, even when convergence may not be achieved for a particular RF-chain, other RF-chains are still able to converge as can be seen in Fig 2.7d. This is in contrast to the SDR-AltMin where all RF chains don't achieve convergence. The above mentioned arguments emphasize that the per-RF chain approach is a welcome departure from most existing schemes which use the sum-power constraint, resulting in unequal power allocation.

2.6.3 Run-Time Evaluation

For practical fast fading systems, it would be of interest to adopt the algorithm that produces faster results by sacrificing some spectral efficiency. Table 2.2 shows the time taken by the SDR Alt-Min, Par-ArgMod and the Alt-ArgMod algorithms, when averaged over 200 runs. The simulations are conducted on a machine with Intel(R) Xeon(R) CPU E5-2630 v2 running at 2.6GHz (24CPUs). The SDR-AltMin algorithm is run on K parallel workers to find \mathbf{F}_{BB} . The Par-ArgMod algorithm is run on $N_{RF}K$ parallel workers for the magnitude problem and N_{RF} parallel workers for the phase problem. Alt-ArgMod is run on similar number of workers as Par-ArgMod.

Algorithm	Mag time	Phase time	Total time
SDR-AltMin [20]	-	-	9.534 s
Par-ArgMod	0.9301 s	1.1241 s	1.1241 s
Alt-ArgMod	-	-	7.2135 s

Table 2.2: Run time comparison

The time taken by the Par-ArgMod algorithm is the maximum of the time taken by magnitude and phase problems. The phase formulation is solved using the primal-dual interior point method for sparse QP problems [38]. We see a nine fold decrease in time taken when using the parallel framework as compared to the alternating approach. This time gain could be significant in fast fading environments where the channel changes rapidly and

precoding has to be sped up in tandem. Moreover, the per RF-chain alternating Alt-ArgMod is much faster than SDR-AltMin.

2.6.4 Realistic Scenarios

In this section, the Saleh-Valenzuela (SV) channel model [6], is extended to realistic scenarios. An Urban Microcell - Street Canyon (UMi-Sc) scenario is considered at a frequency $f_c = 30\text{GHz}$ with a bandwidth of 100MHz and number of subcarriers $K = 12$. Two practical channel models are considered, namely, the 3GPP and NYU channel models. The 3GPP clustered delay line (CDL) channel model is extended to the OFDM case according to [27]. The model assumes that power angular spectrum in azimuth is a wrapped Gaussian and the zenith is Laplacian. The concept of time clusters is maintained with the delay for each cluster generated using an exponential distribution. The delay and the angle are characterized by a joint distribution. The NYU channel model [26], brings in the concept of time clusters and spatial lobes (TCSLs) in addition to the underlying aspects of the 3GPP model. The TC powers are generated using an exponential function of delay and each multipath component is assigned an unique lobe according to a uniform distribution.

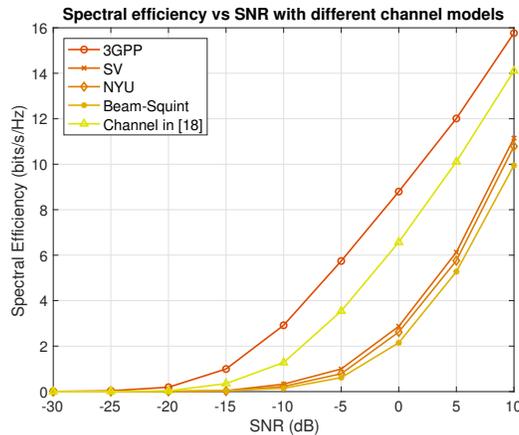


Figure 2.8: Spectral efficiency achieved by Par-ArgMod using different channel models when $N_s = N_{RF}^t = N_{RF}^r = 4$.

In addition to this, we demonstrate the performance of our scheme taking into account

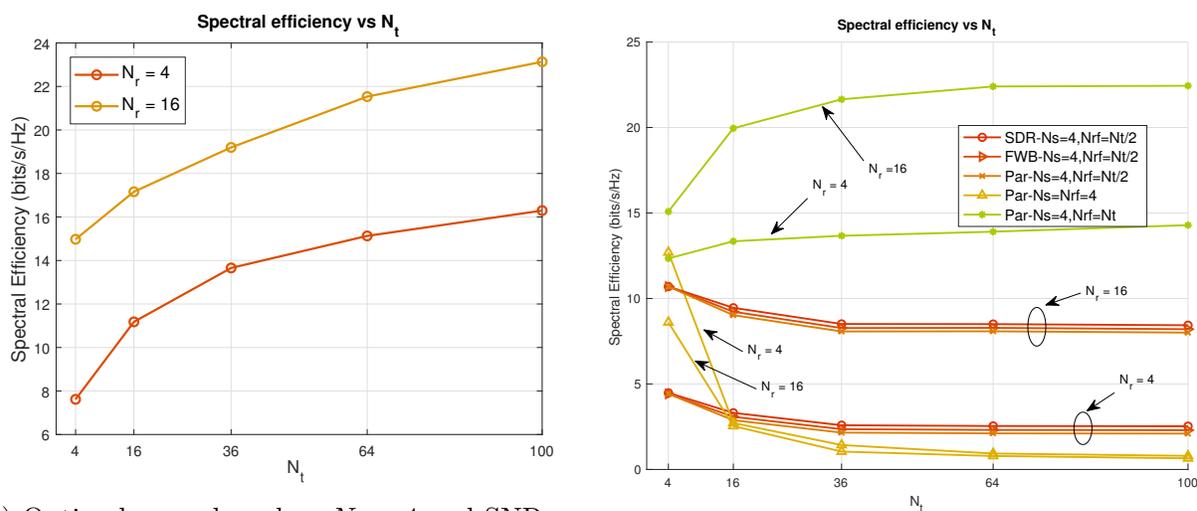
the beam squint effect in wideband systems using large arrays which is investigated in [47], [48]. We model beam squint in USPA following [49] and as noted by the authors observe that the spectral efficiency does suffer. It is also verified that with increasing N_t and K (i.e. bandwidth), the performance worsens due to increased beam-squint. We also show the performance of our scheme with the channel model used in [18], where each ray in a cluster is given a unique phase shift, departing from the channel model used in this work which assigns a unique phase shift to each cluster.

Fig. 2.8 demonstrates the effect of using different channel models to evaluate the performance of the system. The spectral efficiency is estimated optimistically in the case of the 3GPP cluster delay line channel model as the number of clusters and subrays in each cluster are assigned high values for the given UMi-Sc scenario. The number of clusters and subrays used in this work with the SV channel model is lesser when compared to the 3GPP specifications and is closer to the NYU channel model which derives its values from real world channel measurement data. The NYU model claims to be closer to the mmWave setting as the mmWave scattering environment is sparse and won't have large number of multipaths [8]. It is observed that while beam squint affects the spectral efficiency adversely, the use of channel model in [18] does not change the performance to a large extent. Moreover, it is more optimistic in its estimation of spectral efficiency values when compared to the channel model used in our work.

2.7 Critique of Wideband PCS

The variation of spectral efficiency with increasing number of transmit antennas for the optimal precoder in OFDM systems is shown in Fig. 2.9a. The rise in spectral efficiency with increasing number of transmit antennas is expected by the promise of massive MIMO. Increase in number of receive antennas also pumps the spectral efficiency up. This trend is lost in hybrid systems as can be seen in Fig. 2.9b.

2.7. Critique of Wideband PCS



(a) Optimal precoder when $N_s = 4$ and SNR = 0 dB

(b) Partial hybrid precoders at SNR = 0 dB

Figure 2.9: Spectral efficiency achieved by Optimal and hybrid precoders with increasing N_t .

A number of experiments are run to validate this observation. It is seen that having N_s and N_{RF} constant with rising N_t leads to rapid deterioration in performance, with the spectral efficiency dropping to a near zero. In this case we also observe that increasing N_r also results in degradation of performance and could be because increasing number of antennas (both at the transmitter and the receiver) adds to the suboptimality (increasing M leads to highly overdetermined systems in hybrid precoding) but doesn't leverage the positives with fixed N_{RF} . When the number of RF chains is taken to be half the number of antennas (both at the transmitter and the receiver), the spectral efficiency reduces with increasing N_t but at a much lesser pace. This we see as a more realistic experiment as the number of RF chains need to be scaled accordingly with the antennas to take advantage of the increased number of degrees of freedom for spatial multiplexing. However, it is seen that the rising number of antennas makes the hybrid system increasingly suboptimal in all three partially connected architectures compared here. Even though the ratio M of the number of transmitter antennas to the number of RF chains is fixed, i.e., $M = N_t/N_{RF}^t$ is fixed, the spectral efficiency worsens with N_t because the number of subproblems for magnitude and phase formulations is equal to N_{RF} and the suboptimality in each problem adds up with

increasing N_{RF} . Adding to this, each RF chain is limited in power by N_s/N_t , which reduces as N_t increases, hampering the range of values the baseband precoder can take. Lastly, when $N_{RF} = N_t$, it is observed that the spectral efficiency improves with increasing N_t . In this case, there is very little suboptimality in the system and the performance can add up. This confirms our hypothesis that the reduction in the number of RF chains is the major bottleneck in performance scaling with wideband PCS.

The above discussion can be formalized as is done by Proposition 1 in [23]. Extension of Proposition 1 in [23] to wideband systems can be found in the appendix. We follow a similar approach and arrive at conditions for optimality for our system in (2.16) and (2.22). Now, because we use $N_{RF}^t \leq N_t$ (violating (2.16)) and also as N_t is not large enough to satisfy (2.22), the solutions that we obtain are not optimal solutions, but minimum norm solutions in both magnitude and phase. From Fig 2.9b, it can be seen that whenever $N_{RF}^t < N_t$, the spectral efficiency grows worse as N_t increases. This is because (2.16) requires us to have at least $N_{RF} = N_t$, which when satisfied, leads to an increase in spectral efficiency with N_t . Even if we manage to satisfy (2.22), we will violate the more stringent (2.16) and obtain a suboptimal solution. Further investigation is required on how the losses in spectral efficiency with increasing N_t can be mitigated. The current state-of-the-art on hybrid precoding with PCS, including the proposed algorithm suggests that from the spectral efficiency point of view it would be favourable to implement PCS only in systems where the number of RF chains are comparable to the number of antennas.

2.8 Conclusion and Possible Future Work

In this Chapter, we have considered a single-user MIMO system. When OFDM is employed, we have proposed a low-complexity algorithm for the design of a partially connected hybrid precoder which turns out to be much faster than the existing state-of-the-art. The following insights are obtained:

- The run time comparison shows the superiority of the Par-ArgMod algorithm over existing schemes for wideband PCS. In fast fading practical scenarios this would be very desirable. Also, our proposed Alt-ArgMod outperforms other alternating schemes with respect to run-time.
- The per-RF chain power constraint used is more practical in nature than the sum-power constraint because each RF chain is generally equipped with its own power amplifier.
- Departing from the original joint formulation of magnitude and phase in conventional precoding, we solve for the magnitude and phase as two independent subproblems in Par-ArgMod. This allows us to solve the magnitude formulations in parallel for each RF chain and subcarrier. The phase formulations, on the other hand, can be solved in parallel for each RF chain but not for each subcarrier because the analog precoder is shared among the subcarriers. The alternating version (Alt-ArgMod) alternates in parallel for each RF chain but follows a similar parallelization structure.
- Increasing number of antennas without increasing the number of RF chains in tandem is seen to reduce the spectral efficiency performance considerably in wideband PCS due to the overdetermined nature of the hybrid precoding system and the restrictive power constraint on the baseband precoder. From the spectral efficiency point of view, it would therefore be favourable to implement PCS only in systems where the number of RF chains are comparable to the number of antennas.
- Although we propose algorithms that are faster and adhere to practical constraints under the PCS setting, our observations in Section 2.7 tell us that PCS is not a recommended solution for massive MIMO systems where good spectral efficiency is a requirement, especially the ones employing multicarrier transmission like OFDM.

Some possible avenues for future work:

- It is also worthy of investigating further whether the spectral efficiency values in wideband PCS can be improved while respecting the per-RF chain power constraint.
- Mitigate the detrimental effects of increasing the number of transmit antennas with limited RF chains on spectral efficiency in wideband PCS.
- Consider the statistical distribution of phases of the analog precoder in the optimization process and their relationship with the array response vector.
- For massive MIMO systems with high spectral efficiency requirement, investigate FCS design respecting the per-RF chain power constraint. Extend the design to make it workable in real life fast fading scenarios. The resulting design would be the ideal marriage between spectral efficiency, practical implementation and energy savings.
- Extend the proposed hybrid precoding design to the multi-user massive MIMO scenario. The main challenge in this extension would be to integrate the interference cancellation procedure among the multiple users with the proposed procedure of parallelizing the hybrid precoder.
- Extension to distributed massive MIMO systems would also be an interesting avenue to explore as the inherent parallelization in the approach would inspire a distributed setup.

Chapter 3

Signal Detection and Time-Frequency

Localization Using Deep Learning

3.1 Introduction

As wireless devices become pervasive in our day-to-day life, being able to passively detect these devices in the spectrum is an important concern from the security as well as spectrum management perspective. Signal detection techniques have been investigated extensively in literature. A multi-band joint detection technique which jointly detects signal energy levels in multiple frequency bands is introduced in [50] where the spectrum sensing problem is formulated as an optimization problem in an interference limited network. Wavelet edge detection followed by blind source separation is done to separate the signals in the frequency domain in [51]. In both these works, although the signals can be accurately localized and separated in frequency, the joint time-frequency information is lost. In many applications such as detection of the hopping pattern of a wireless device and joint frequency and temporal optimization of the shared spectrum, it becomes a necessity to detect both the time and frequency information of the signals present. In [52], periodic signals are detected using a blind energy detection followed by a cyclostationary detection method where the extracted signals are then classified based on a Chinese restaurant process (CRP). Defining custom features based on RF signatures and cyclostationarity properties may be a viable solution but might not be the best approach to detect various types of heterogeneous signals that deviate from cyclostationary assumptions. This limitation is accompanied by the loss

of temporal information.

Moving away from cyclostationary assumptions requires an agnostic feature extractor network. With advances in deep learning techniques for time-series and image analysis, we can extract rich features out of RF data for downstream tasks such as detection, localization and classification. Audio event detection (AED) is one example where the application of deep learning has been explored in the recent past. The underlying philosophy is that by converting the time series data into spectrograms and then employing deep learning techniques, we can extract certain specific patterns that help detect and localize audio events. In [53], a state-of-the-art object detection framework was adapted to detect monophonic and polyphonic audio events from the spectrograms. A similar approach is proposed in [54], with the added functionality of capturing the long-term temporal context from the extracted features through the use of a convolutional recurrent neural network (CRNN). Both [53] and [54] detect the presence of audio events by converting time series information into time-frequency spectrograms and then learning from the features present in the spectrograms. However, the same philosophy has not been well explored yet for detecting general purpose wireless RF signals present in wideband spectrum.

The idea of using deep learning based frameworks to detect wireless signals has been looked into recently. The work [55] converts the time-frequency information into power spectral density (PSD) based spectrograms. The spectrogram is then fed into a five-layer convolutional neural network (CNN) which is used to perform multi-class classification over different wireless technologies like WiFi, Bluetooth and ZigBee. Although the approach is able to perform classification over heterogeneous devices, it cannot localize them in time and frequency. Localization in time and frequency, if achieved, can be used to study various other properties of these devices like hopping patterns, signal bandwidth and dwell time. This information will be crucial for security purposes because it helps us perform narrow-band jamming to mitigate rogue devices without effecting the other friendly devices on the industrial, scientific and medical (ISM) band. A different time-frequency transformation,

called the Choi-Williams distribution (CWD), is used in [56] to distinguish between different type of coding schemes like polytime codes, Frank code and Costas codes. After image preprocessing, this transformation is fed into a two-layer CNN with pooling and the recorded ratio of successful recognition (RSR) is about 90% for most codes. However, it faces a similar drawback of not being able to localize the signal in time and frequency.

The problem of detecting signals in a spectrogram falls under the more general problem of object detection in images. The state-of-the-art in this regard is to employ CNNs to identify whether an image contains an object(s) and predict the bounding box of the detected object [57]-[63]. Previous state-of-the-art methods, for example [57][58], have employed one-stage object detection using a single CNN to simultaneously obtain the category and location of the objects. Such one-stage methods have recently been outperformed by certain two-stage object detection methods [59]-[63]. In these methods, the first stage generates a set of candidate bounding boxes, commonly referred to as the region proposals. Popular region proposal methods include selective search [62], which is based on greedy superpixel merging, and EdgeBoxes [63], which is based on edge maps and edge groups. The second stage performs a classification task on the region proposals to identify the objects and a refining task on the dominant region proposals to provide the bounding boxes. A major bottleneck of the above mentioned CNN methods is that they perform supervised machine learning and therefore require large amounts of labelled datasets to achieve high accuracy in object detection and localization [57]-[63]. Large labelled datasets are currently available for object detection in day-to-day real-life images containing humans (c.f. PASCAL VOCO [64] and MS COCO [65]) and audio signals (c.f. UrbanSound8k [66] and DCASE [67]). However, there are no standard labelled datasets available online for wireless signals present in the wideband RF spectrum.

In this chapter we introduce a real-time deep learning framework based on the FRCNN [1], for detection and time-frequency localization of narrowband signals present in a wideband RF spectrum. Firstly, we find the most suitable feature extraction network and our experiments

suggest that while weights pretrained on regular images are a good starting point for medium sized networks, making the weights trainable gives much better performance. Following this, we provide design insights with respect to multiple variables such as the STFT parameters, spectrogram and anchor sizes and various thresholds of the model. To evaluate the detection and localization performance of the proposed system, we generate synthetic data as per the recently proposed WiFi-HT protocol, adopt the mAP metric [68] and make the necessary modifications to account for evaluation over varying SNR values. An mAP of 0.9 is recorded when the model is trained and tested on positive SNR values with single-bandwidth signals. In the coming sections we discuss our signal detection framework and introduce the Faster-RCNN architecture.

3.2 Framework for signal detection and time-frequency localization

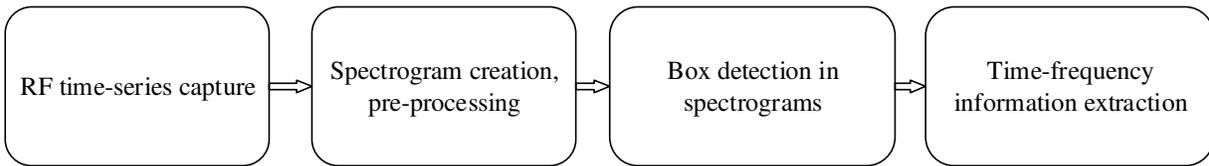
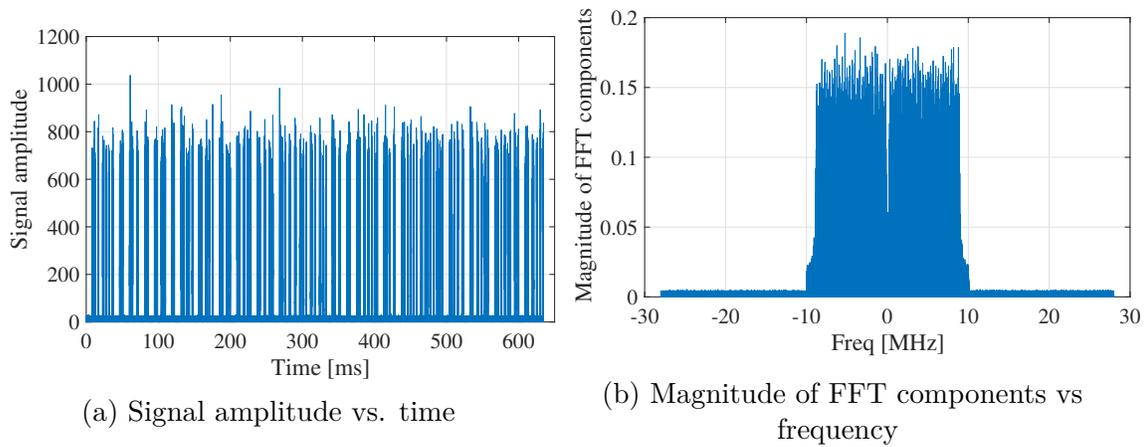


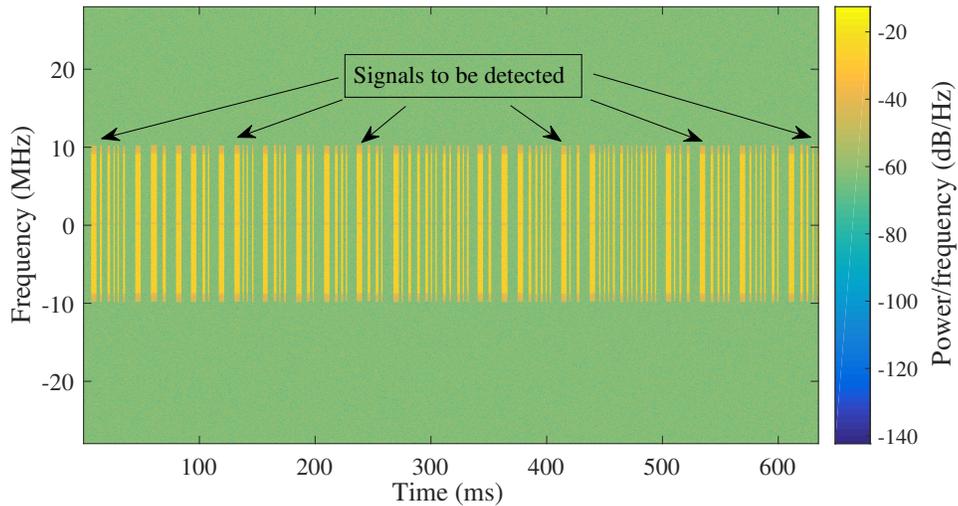
Figure 3.1: Proposed framework for signal detection and time-frequency localization

We propose a deep learning framework to detect and estimate the time-frequency span of all wireless signals present in a wideband RF spectrum. The proposed framework takes the wideband RF time-series data as the input and provides the time and frequency information of each detected signal as the output. An outline of the proposed framework is presented in Fig. 3.1. Details on each stage are presented below.



(a) Signal amplitude vs. time

(b) Magnitude of FFT components vs frequency



(c) Spectrogram image created by plotting the p.s.d values against time and frequency

Figure 3.2: The time content, frequency content, and spectrogram of an example wideband RF capture, when the capture duration is 633 ms, center frequency is 2.4 GHz, wideband bandwidth is 56 MHz, and sampling rate is 56 MHz.

RF time-series capture

In the first stage, we employ a wideband sensor with center frequency f_c and bandwidth W to record time-series RF data in fragments of T ms each. The time and frequency content of an example wideband capture with $f_c = 2.4$ GHz, $W = 56$ MHz, $T = 633$ ms, and a sampling rate of 56MHz is given in Fig. 3.2. While Fig. 3.2a plots the signal amplitude as a function of time, Fig. 3.2b plots the magnitude of the fast Fourier transform (FFT) components as a function of frequency.

Spectrogram creation and pre-processing

For a compact representation of the wideband signal in terms of time and frequency, we apply STFT on the RF time-series captures and obtain the PSD as a function of time and frequency. Three-dimensional spectrogram images are then created by plotting the PSD values along the time and frequency axes. Fig. 3.2c illustrates the spectrogram image created for the RF capture in Fig. 3.2a-3.2b, when the STFT parameters are chosen as follows: number of frequency bins is 4096, number of time bins is 4096, the STFT window is of hann-type, and the window overlap is of 2048 time bins. Few insights on the choice of STFT parameters are given in Section 3.4.1. As may be noted from Fig. 3.2c, the signals to be detected appear in the form of rectangular boxes in the spectrogram image.

From the spectrogram in Fig. 3.2c, we may note that the dimensions of each rectangular box within the spectrogram give us the time and frequency information of the corresponding wireless signal. The problem of signal detection and time-frequency localization therefore boils down to the problem of detecting and estimating the dimensions of rectangular-shaped boxes present in the spectrogram. Before attempting box-detection in the spectrogram image, we may employ some pre-processing steps. For example, we may remove out-of-band transmissions to eliminate unreliable information. We may also employ denoising methods, such as wavelet denoising [69], to improve the SNR of the spectrogram.

Box detection in spectrograms

To detect the rectangular-shaped boxes present in the spectrograms, we take a supervised machine learning approach, wherein, we train a FRCNN model [1] with several labelled spectrogram images. The trained FRCNN model, when input with a test spectrogram image, detects the rectangular-shaped boxes present in the image and reports their dimensions. In Section 3.3, we present details on FRCNN architecture and provide an overview on how the FRCNN model achieves the box detection task at hand.

Time-frequency information extraction

As the final step, we convert the dimensions of each rectangular box reported by the FRCNN model into time and frequency information. For example, using the STFT parameters employed in the spectrogram creation stage, we may scale the x and y dimensions of each box into the time and frequency span of the corresponding signal. The same approach may also be followed to obtain the narrowband center frequency of the signal. In the next section, we present details on the FRCNN architecture and expose several design choices that need to be made to perform the signal detection and time-frequency localization task.

3.3 Faster RCNN Architecture

Faster RCNN is an object detection framework composed of three modules, as illustrated in Fig. 3.3. The first module, which is the base network (BN), takes the image as the input (spectrogram in our case), extracts features that are relevant to the object detection task at hand and outputs a down-scaled feature image. The second module, which is the region proposal network (RPN), takes as input the down-scaled feature image, a set of anchor boxes (ABs) and the ground truths (GTs). The RPN provides as output the region proposals, which are nothing but candidate boxes that are likely to contain the objects of interest. In our case, the objects of interest are the rectangular boxes in the spectrogram. The region

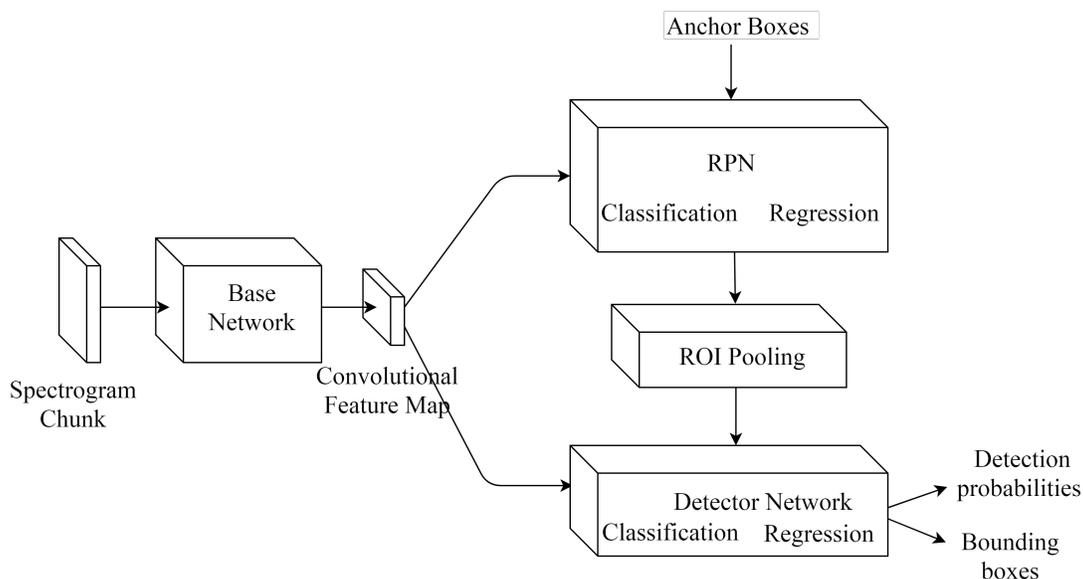


Figure 3.3: Faster RCNN Architecture.

proposals from the RPN, along with the feature image, are fed into the detector network, which is the third module. The detector network assigns object class labels to each region proposal from the RPN, performs a regression task to localize the object within the region proposal, and also provides probabilities with which the assigned labels are true. Essentially the RPN acts as an attention mechanism [70] over the feature image to help the classifier with the object detection and localization task. When spectrograms are provided as input, the entire FRCNN network can be thought of as a single unified framework for detecting and localizing the rectangular boxes (which are nothing but a manifestation of the signals of interest) present in the spectrograms. Details on each module in the FRCNN are presented below.

3.3.1 Base Network

The base network is a CNN that can be either shallow or deep depending on the complexity of features that need to be extracted from the input image. The convolutional layers are interleaved with max pooling layers and the combination of these layers decide the total down-scaling factor. In Section 3.5, we experiment with multiple feature extraction networks,

namely, the VGG-13 [71] with the first 10 convolutional layers, the VGG-13 [71] and the ResNet-50 [72]) networks, and report the impact on the performance of the FRCNN model.

3.3.2 Region Proposal Network

In the RPN, the down-scaled feature map obtained from the base network is passed through an $n \times n$ convolutional layer, where $n = 3$ typically, to obtain a low-dimensional feature vector. Also, a fixed number of user-defined raw region proposals, known as anchors, are created for each pixel in the input feature map to serve as the raw region proposals. The low-dimensional feature vector, along with the raw region proposals created from the anchors, are fed into two fully connected layers to perform a classification and a regression task respectively. The classification task assigns probabilistic labels to each raw region proposal as positive or negative, to denote whether the proposal is likely to contain an object of interest or not. For proposals that are deemed positive by the RPN, the regression task tunes the size of the proposals to suit the dimensions of the object. The role of the anchors is explained next.

Anchors

As briefly mentioned earlier, for each pixel of the down-scaled feature map from the BN, the RPN generates a predefined number N_a of raw region proposals centered at the pixel, where N_a is the number of anchors. Anchors are user-defined raw region proposals whose size and aspect ratio needs to be specified before the training process begins. During the training, if any anchor box has an intersection over union (IoU) greater than a certain threshold (referred to as the *RPN max overlap*) with the ground truth, the RPN treats the anchor box as a positive target. On the other hand, if the IoU is lesser than a certain threshold (referred to as the *RPN min overlap*) the RPN treats the anchor box as a negative target. Any anchor box whose IoU with the ground truth falls between the RPN min and RPN max overlap, is left unutilized and is not acted upon for any further decision making. If the total number of

anchors is N_a , we would therefore have N_a number of proposals per pixel in the input feature map. The regression layer in the RPN will have $4N_a$ outputs per pixel, which correspond to the corner coordinates of the N_a anchor boxes per pixel. Also, the classification layer will have N_a outputs per pixel, to denote the probabilities with which the associated proposals contain the object of interest. Also, for an input feature map of size $W \times H$, we would have a total of WHN_a region proposals in total.

On the WHN_a region proposals thus obtained, we perform a non-max suppression (NMS) operation, as explained below, and obtain a filtered list comprising a small predefined number, say N_r , of region proposals. The filtered list of proposals are fed into the third module, which is the detector network, for further action.

Non-Max Suppression

From the WHN_a region proposals, we choose a small predefined number of region proposals, referred to from here on as the regions of interest (RoIs), and feed them into the detector network. The motivation for NMS is that, by restricting the number of RoIs, improvements are observed in the performance of the detector network and also the overall processing time. During NMS, we firstly sort all the region proposals in decreasing order of their probabilities. Next, we retain the region proposal with the highest probability and suppress all other proposals whose IoU with the retained RoI is greater than a predefined threshold, referred to as the NMS threshold. The same procedure is followed for the RoI with the next highest probability and so on until we have retained a small predefined number N_r of RoIs. This process of suppression is agnostic to the anchor type that the RoI belongs to and only relies on the RPN classification probabilities. After NMS, the RoIs which have high IoU with the ground truth are treated as positive targets and the rest are treated as negative, i.e., background, targets, for the classification and regression tasks executed in the detector network.

3.3.3 Detector Network

The detector network performs an RoI convolutional pooling operation, followed by a classification and a regression operation.

RoI Convolutional Pooling

The ROIs provided by the RPN (after the NMS operation) can be of different sizes, depending on our choice of the anchors and the result of regression task in the RPN. The RoIs need to be converted into fixed size inputs in order to be able to feed them into convolutional layers for classification and regression within the detector network. This action is carried out by a convolutional network known as the RoI pooling network [73]. It takes two inputs, namely, the convolutional feature map from the BN and the filtered RoIs after the NMS operation. For every RoI from the filtered RoI list, the RoI pooling network takes the section of the convolutional feature map that corresponds to the RoI and scales it to some pre-defined output size (e.g., 7×7). The scaling process is carried out by doing the following: (i) dividing the RoI into equal-sized sections of the same dimension as the output, and (ii) finding the maximum value in each section and copying these to the output. The result is that from a list of RoIs of different sizes, we can obtain modified RoIs of a fixed size. The RoI pooling output dimension depends neither on the size of the feature map from the BN nor on the size of the filtered RoIs, but only on the number of sections we divide each RoI into. By yielding fixed-size RoIs as the output, the RoI pooling layer allows us to use densely connected convolutional layers for the ensuing classification and regression tasks.

Classification and Regression in the Detector

After performing RoI convolutional pooling, the fixed size RoIs are fed into a bunch of convolutional layers to convert them into low-dimensional feature vectors. These feature vectors are then input to two densely connected networks to perform a classification and a regression task respectively. The classification task focuses on assigning probabilities that

the RoIs contain an object of interest. The regression task focuses on fine-tuning the size of each positive-labelled RoI to match the dimensions of the object present in it.

When training the detector, if any RoI has an IoU greater than a certain threshold (referred to as the *Detector max overlap*) with the ground truth, the RoI is treated as a positive target. If the IoU is less than the Detector max overlap but is greater than a certain threshold, namely the *Detector min overlap*, the RoI is treated as a negative target. Any RoI whose IoU with the ground truth is less than the Detector min overlap, is left unutilized for any further decision making. During test time, if the total number of RoIs after the NMS operation is N_r , the classification task in the detector yields N_r outputs, each denoting the probability with which the associated RoI contains the signal of interest. For all the RoIs with probability higher than a certain threshold, namely the *Detector positiveness threshold*, the regression task yields 4 outputs, corresponding to the corner coordinates of the regressed RoI.

3.3.4 Loss Functions

For the classification and regression tasks in both the RPN and the detector networks, we need to optimize appropriate loss functions. A standard multi-task loss function, as defined in [1], can be given as:

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*), \quad (3.1)$$

where N_{cls} is the mini-batch size, $N_{reg} = 4 * N_{cls}$ is the total number of coordinates in the mini-batch, i is the anchor index and p_i is the predicted probability of the anchor i containing the object of interest. The p_i^* is the GT label which takes a value of 0 or 1, depending on whether the anchor is negative or positive respectively. The t_i^* is a vector comprising the four parameterized coordinates of the GT box associated with a positive anchor and t_i is that of the predicted bounding box. The p_i and t_i terms are the outputs from the classification and

regression respectively. The classification loss L_{cls} is a simple binary cross-entropy loss [74] over the two object classes of interest, namely the signal and the background. On the other hand, the regression loss L_{reg} is given as

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (3.2)$$

where R is the robust loss function defined in [75]. The term $p_i^* L_{reg}$ indicates that the regression loss is turned on only for positive anchors, i.e., when $p_i^* = 1$, and is quiescent otherwise. Elements of the parameterized coordinate vectors t_i^* and t_i in the regression task are obtained from the corner coordinates of the ground truth, the anchor box and the bounding box as [1]:

$$\begin{aligned} t_x &= (x - x_a)/w_a, & t_y &= (y - y_a)/h_a, \\ t_w &= \log(w/w_a), & t_h &= \log(h/h_a), \\ t_x^* &= (x^* - x_a)/w_a, & t_y^* &= (y^* - y_a)/h_a, \\ t_w^* &= \log(w^*/w_a), & t_h^* &= \log(h^*/h_a), \end{aligned} \quad (3.3)$$

where x , y , w , and h denote the time and frequency coordinates of the center, the time span and the frequency span of the bounding box in the spectrogram respectively. The x_a, y_a, w_a, h_a and x^*, y^*, w^*, h^* are similarly defined for the anchor box and the GT respectively. The regression tasks in both the RPN and the detector networks can be thought of as bounding-box regressions from an anchor box/region proposal to a nearby GT signal, with some transformation of the optimization variables, as given in (3.3), for ease of training. Similar loss functions are used for both the RPN and the classifier with appropriate targets in each case.

3.3.5 Training procedure

We employ the *approximate joint training* method proposed in [1] to train our system. In this method, during each forward pass, the RPN generates proposals which are used for the backpropagation. The updated RPN from the backpropagation generates an additional round of proposals which are then treated as precomputed proposals when training the detector network. After the forward pass of the classifier, both the RPN loss and the detector loss are summed and the resulting loss is used to backpropagate through both the networks. As mentioned in [1], this method ignores the derivative of the sum-loss with respect to the proposal boxes and therefore is an approximation. However, it reduces the training time significantly, is easier to implement, and the loss in performance due to the approximation is minor.

3.4 Design Choices to Adopt FRCNN for Signal Detection and Time-Frequency Localization

As may be noticed from our description of the FRCNN architecture, there are several design choices that we need to make in order to adopt the FRCNN model for the signal detection and time-frequency localization task at hand. Below we provide multiple insights on the major design choices that we have made.

3.4.1 Choice of STFT parameters

To generate the spectrogram images from the raw RF time-series data, we need to apply discrete-time STFT and this in turn requires us to choose a few hyperparameters, namely, the window size, the window type, the window overlap, and the FFT size. We may choose these STFT parameters based on the minimum time and frequency resolutions that we need to achieve. The window size governs the maximum achievable frequency resolution. If the

window is T seconds long, the minimum detectable narrowband bandwidth is $1/T$ Hz. For example, when the sampling rate is $f_s = 56 \times 10^6$ Hz and the window size is 5600, the minimum detectable bandwidth is 10 kHz. While respecting this lower limit, the FFT size allows us to control the number of frequency bins in the spectrogram. For example, an FFT size of 1500 divides the spectrogram into 1500 frequency bins. Therefore, if the wideband bandwidth is 56 MHz, the minimum detectable signal bandwidth would be ≈ 37.3 kHz.

The window size, the window overlap, and the sampling rate govern the maximum achievable time resolution. For example, if the captures signal duration is $T_{sig} = 0.633$ seconds, the sampling rate is $f_s = 56 \times 10^6$ Hz, the window size is $T_{win} = 5600$ and the window overlap is $T_{ov} = 2800$, the maximum achievable time resolution is $T_{sig} * (T_{win} - T_{ov}) / (T_{sig} * f_s - T_{ov}) \approx 50\mu$ seconds. Lastly, the window type governs the amount of discontinuities between successive window segments. For example, windows which are tapered at the ends, such as the Hann and Hamming windows [76],[77], introduce much fewer unnatural discontinuities in the time domain than the ones with non-tapering ends, such as the rectangular window.

3.4.2 Spectrogram size

When feeding the spectrogram into the FRCNN model, care should be taken to ensure that the size of the input image is not too large. This is mainly because of two reasons. Firstly, when we use large images as input to the model, each pass of training takes a long time to complete because of the large input feature map, number of anchors, and the resulting set of computations. This would mean large convergence time. Secondly, the base network that extracts the features has a particular receptive field depending on the base network used, such as the VGG and ResNet. When these receptive fields are large when compared to the size of the signals, it would be hard for the network to make sense of the extracted features as they would contain more of the background than any given signal. In light of these two observations, we restrict the length of our input image to 600 pixels in time. This is achieved by chopping the input spectrogram into chunks of 600 pixels each and recalculating

the relative positions of the ground truths accordingly. Doing so ensures that the size of the signals in the spectrogram is comparable to the full spectrogram size. This approach is also seen to decrease the training time significantly.

3.4.3 Choice of base network

As may be noted from Section 3.3.1, the base network in the FRCNN model handles the crucial task of feature extraction and therefore needs to be chosen based on the type of input data used and the type of object detection task at hand. Standard feature extraction models used in computer vision include VGG-13 [71], ResNet-50 [72], and MobileNet [78], among many other popular ones [79]. These models have been built to detect objects in benchmark datasets for computer vision, such as PASCAL VOC2007 [64] and MC COCO [65]. Example objects that are to be detected from these datasets include humans, animals, and automobiles. It is therefore not clear whether the standard base network models such as VGG-13 and ResNet-50 can perform feature extraction for the signal detection task on RF datasets. We may conduct a few experiments to choose the base network for the task at hand. Firstly, we may consider the publicly available pretrained weights in the VGG-13 and ResNet-50 models and verify if the extracted features are useful for the signal detection task at hand. Secondly, we may set the pretrained weights in these models as initialization points to optimize the weights in the VGG-13 and ResNet-50 models to perform tailor-made feature extraction for the task at hand. Thirdly, we may only consider the architecture of the VGG-13 and ResNet-50 models, randomly initialize the weights and optimize the weights for the feature extraction. We pursue these three experiments in Section 3.5 and provide insights based on the observations we make.

3.4.4 Anchor box sizes and aspect ratios

As we may recall from Fig. 3.3 and Section 3.3.2, anchor boxes aid the RPN in generating region proposals for the detector network. The RPN acts upon the anchor boxes generated

per pixel of the input feature map by performing (i) a classification task which assigns to each anchor the probabilities of it containing the signal of interest and (ii) a regression task which regresses the corner coordinates of the anchor boxes to generate the region proposals. When training the RPN for the classification and regression, anchor boxes whose IoU with the GT are larger than the RPN max overlap are treated as positive targets and the ones having IoU lower than the RPN min overlap are treated as negative targets. Consequently, in order for the training to be successful, we need to carefully choose the anchor sizes and aspect ratios, as well as the RPN min and max overlap values.

Since the anchor boxes serve as raw region proposals for the RPN to act upon, we may choose the anchor sizes and aspect ratios to match the dimensions of the signals that we wish to detect and localize. For example, if the narrowband signals are known follow the IEEE 802.11n HT protocol, we know that the signals have a bandwidth of either 20 MHz or 40 MHz and their duration ranges from about 300 microseconds to the order of 15 milliseconds [2][80]. We may therefore choose multiple anchor boxes, each with a frequency span of either 20 or 40 MHz and with a time-span chosen uniformly randomly in the range [0.3, 15] ms.

3.4.5 RPN max and min overlap

The RPN max overlap needs to be chosen such that at least a few anchor boxes per image have a high enough IoU with the GT to be considered as positive targets for the RPN training. Typically, the RPN max overlap is chosen to be greater than or equal to 0.5 because an IoU of 0.5 or more gives us confidence that the anchor box indeed contains the signal of interest. On a similar note, the RPN min overlap needs to be chosen such that at least a few anchor boxes per image have a low enough IoU with GT to be considered as negative targets for the RPN training. Typically, the RPN min overlap needs to be less than 0.5 because an IoU less than 0.5 denotes that the anchor box may not contain the signal of interest. Higher RPN max overlap and lower RPN min overlap values would deem fewer anchors as positive and negative targets respectively. This would consequently result in the

convergence rates being slower but smoother.

3.4.6 Detector network min and max overlap

The detector max overlap needs to be chosen such that at least a few RoIs per image have a high enough IoU with the GT to be considered as positive targets for training the classification and regression layers present in it. Typically, the detector max overlap is chosen to be greater than or equal to 0.5 because an IoU of 0.5 or more gives us confidence that the RoI indeed contains the signal of interest. On a similar note, the detector min overlap needs to be chosen such that at least a few RoIs per image have an IoU value between the detector max and min overlaps to be considered as negative targets for the training. Typically, the detector min overlap needs to be less than 0.5 because an IoU less than 0.5 denotes that the RoI may not contain the signal of interest. The higher the detector max overlap, the fewer the number of positive target RoIs. Also, the smaller the difference between the detector max and min overlap values, the fewer the number of negative target RoIs. For the training to converge faster, we need to make sure that the pool of positive and negative targets per image is as large as possible and therefore choose the detector max and min overlaps accordingly. In Section 3.5, we conduct simple grid search experiments to select the best possible values for these thresholds.

3.5 Numerical Studies

We now present numerical studies on the performance of the Faster RCNN model for the signal detection and time-frequency localization task under study. Below we provide details on the training and test datasets, the spectrogram generation, the various numerical thresholds chosen for the FRCNN model, and the metric for performance evaluation.

3.5.1 Dataset for training and testing

We consider RF transmissions as per the IEEE 802.11n HT mode protocol [2], popularly known as the WiFi-HT protocol, and generate the time-series data synthetically using MATLAB WLAN toolbox [81]. All the generated RF captures are centered around the 5.8 GHz range, have a time duration of 630ms, a wideband bandwidth of 56 MHz, and an SNR drawn uniformly randomly from the set $\{0, 10, 20, 30\}$ dB. The data sampling rate is 56MHz and the total useful bandwidth, after removing unreliable out-of-band transmissions, is 44.8MHz. On an average, each RF capture contains about 90 WiFi-HT signal packets, each having a narrowband bandwidth of 20MHz. All the signals are randomly subject either to line-of-sight or to non-line-of-sight small-scale fading effects. In total, we generate a dataset of 7 captures per SNR, amounting to around 3780 signals. Out of the 7 captures per SNR, we randomly choose 5 for training and the remaining 2 for test purposes.

3.5.2 Spectrogram generation

For each RF capture, we apply STFT with the following choice of parameters: window size is 5600, window overlap is 2800, window type is Hann, and the FFT size is 1500. The resulting spectrogram, after removal of out-of-band transmissions, contains 1200 frequency bins and 12599 time bins. Each spectrogram image is chopped into fixed chunks of 600 bins in time in order to speed up the training process, as bigger spectrograms take longer time to train and also to maximize the performance of FRCNN (c.f. Section 3.4.2 for details). This allows us to detect signals that span a minimum of 0.05 ms in time and 37.3 kHz in frequency (c.f. Section 3.4.1). Each spectrogram image therefore results in a total of 21 input images to the FRCNN model. Consequently, the FRCNN model encounters a total of 630 input images during training and a total of 252 input images during testing.

3.5.3 Numerical thresholds for the FRCNN model

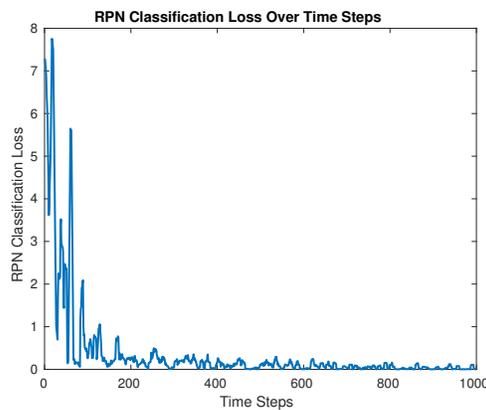
The convolutional layer at the start of the RPN (c.f. Section 3.3.2), which is used to create a low-dimensional feature vector from the feature map generated by the base network, is chosen as per [1] to be of size 3×3 . The anchors are defined such that (i) the time-axis sizes are chosen from the set $\{20, 40, 80, 120\}$, to represent signal time-spans of $\{1, 2, 4, 6\}$ ms respectively, and (ii) the frequency-axis size is chosen to be 17.92, which corresponds to the 90% useful bandwidths for the 20 MHz narrowband transmissions. In total, we therefore have a maximum of $N_a = 4$ anchor boxes. The RPN max overlap and the detector max overlap are chosen from the set $\{0.5, 0.7, 0.9\}$. The RPN min overlap and the detector min overlap are chosen from the set $\{0.1, 0.3\}$. Also, following [1], we set the number of RoIs N_r from the NMS operation to 300, the output RoI size from the RoI pooling network to 7×7 , and the Detector positiveness threshold to 0.5. The λ value in the loss function is set to 1 for all four classification and regression tasks in the RPN and detector. The mini-batch size N_{cls} is set to 256 for the RPN and 32 for the detector.

3.5.4 Training performance evaluation

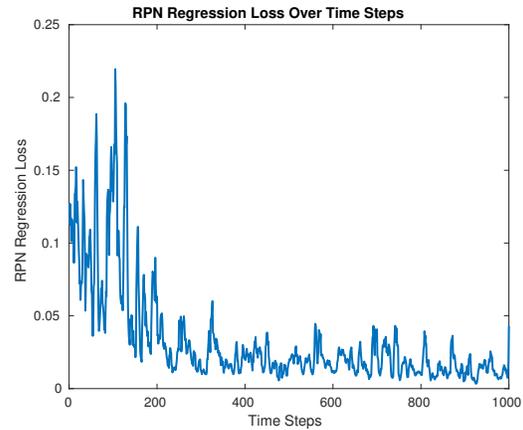
As may be recalled from Sections 3.3.2-3.3.3, there are four main machine learning tasks within the FRCNN model, namely the RPN classification, RPN regression, Detector classification, and Detector regression. To evaluate the training performance, we consider an example experiment with the base network set to VGG-13 (initialized with pre-trained weights and configured as trainable), each training is carried out with one spectrogram chunk, the RPN min overlap set to 0.1, RPN max overlap set to 0.9, Detector min overlap set to 0.1, detector max overlap set to 0.5, the weighted-sum loss function chosen as in Section 3.1 with $\lambda = 1$, and a total of 20 training epochs. In Fig. 3.4, we plot the training loss for the four tasks mentioned above as a function of time. To focus on the trends, moving average was applied over a window of 5 time steps. We notice that all the four training losses converge

3.5. Numerical Studies

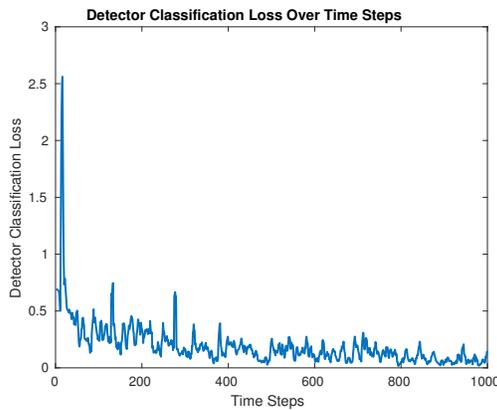
with time to zero. The convergence rates and the fluctuations in the training loss depend on the quantity and quality of positive and negative targets available per mini-batch to the RPN and the Detector. Example positive targets on a spectrogram chunk for the RPN and the positive RoI inputs for the detector networks are shown in Fig. 3.5. As may be noted from Sections 3.3.2 and 3.3.3, the RPN targets are the anchor boxes whose IoU with the ground truth is $\geq RPN \text{ max overlap}$ and the detector inputs are the RoI proposals chosen after non-max suppression.



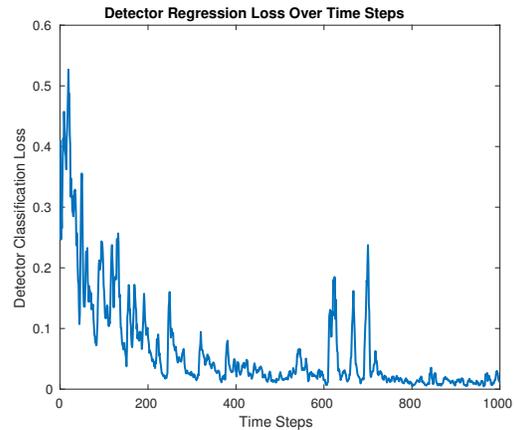
(a) RPN classification loss vs time



(b) RPN regression loss vs time



(c) Detector classification loss vs time



(d) Detector regression loss vs time

Figure 3.4: Training loss convergence for the RPN classification, RPN regression, Detector classification, and Detector regression tasks in the FRCNN model.

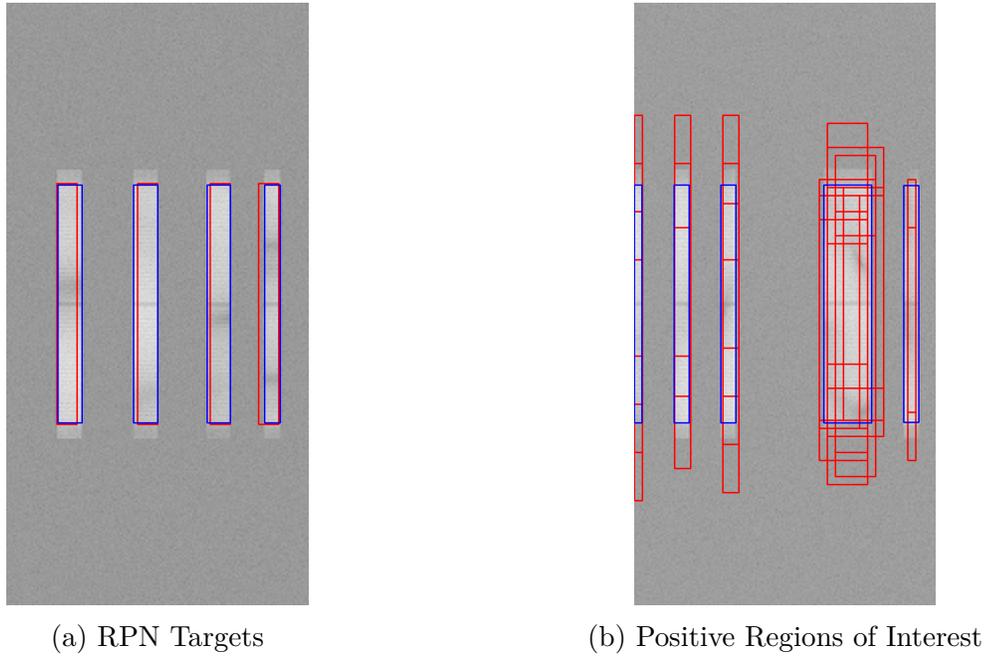


Figure 3.5: Targets for the RPN and the inputs for the Detector. The red boxes are the computed targets and the RoIs where as the blue ones indicate the GT signals.

Algorithm 3 Calculation of the mean average precision (mAP) metric

Require: SNR levels, ground truths, bounding box predictions, detection probabilities

- 1: **for** Each SNR level s **do**
 - 2: Sort the predictions in decreasing order of detection probabilities
 - 3: **for** Each bounding box prediction i **do**
 - 4: Assign a True label if it has $\text{IoU} \geq 0.5$ with any GT (or False if otherwise)
 - 5: Calculate number of true positives (TP), false positives (FP), and false negatives (FN) so far
 - 6: Calculate precision prec_i until the current prediction as

$$\text{prec}_i = \text{TP}/(\text{TP} + \text{FP})$$
 - 7: Calculate recall r_i until the current prediction as

$$r_i = \text{TP}/(\text{TP} + \text{FN})$$
 - 8: **end for**
 - 9: **for** recall levels $r_j = \{0, 0.1, 0.2, \dots, 1\}$ **do**
 - 10: Calculate the maximum achieved precision $\widetilde{\text{prec}}_j$ for recall r_j as

$$\widetilde{\text{prec}}_j = \max_{r_k \geq r_j} \text{prec}(r_k), \text{ where } \text{prec}(r_k) \text{ is the precision at recall } r_k$$
 - 11: **end for**
 - 12: Calculate the average of the maximum precisions for the 11 recall levels as $\text{AP}_s = \sum_{j=1}^{11} \widetilde{\text{prec}}_j$
 - 13: **end for**
 - 14: **return** mean average precision (mAP), calculated as $\frac{1}{S} \sum_{s=1}^S \text{AP}_s$, where S is the number of SNR levels.
-

3.5.5 Prediction Performance evaluation

To evaluate the performance of the FRCNN model in signal detection and time-frequency localization, we consider the mAP metric [68]. The mAP is a standard and widely-used metric for performance evaluation of object detection algorithms in computer vision. An overview of the mAP calculation is given in Algorithm 3. We begin by sorting all the bounding box predictions from the FRCNN model in the decreasing order of the detector classification probabilities. Next, we assign a True or False label for each prediction, depending on whether it has an $\text{IoU} \geq 0.5$ with any ground truth or not. We also calculate the precision and recall values until the current prediction using the formulae given in Steps 5 and 6 of Algorithm 3. We then consider 11 specific recall levels r_j ranging from 0 to 1 in steps of 0.1 and record the maximum achieved precision $\widetilde{\text{prec}}_j$ for each recall level. The mAP is then obtained as the average of the $\widetilde{\text{prec}}_j$ values recorded for the 11 recall levels. Higher mAP values denote better prediction performance.

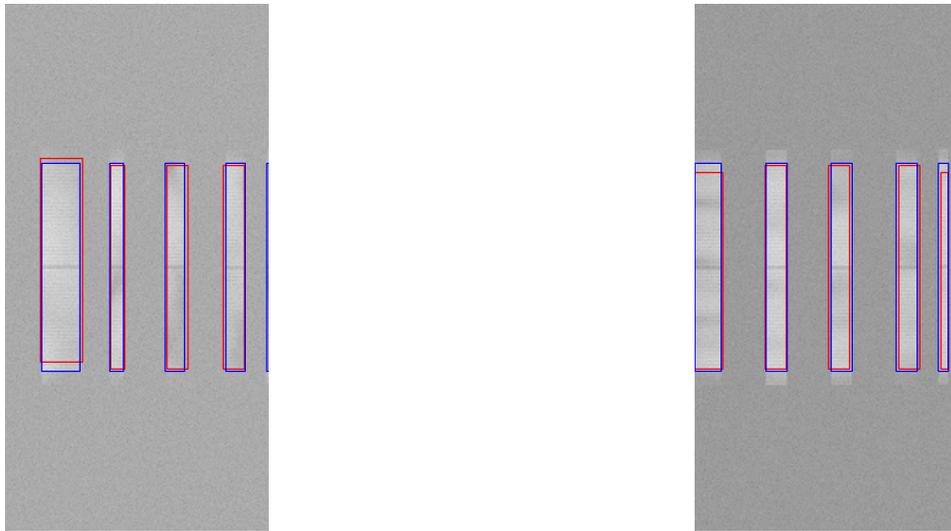


Figure 3.6: Test images of the trained model. The red boxes are the predicted bounding boxes where as the blue ones indicate the GT signals.

Sample prediction images of the model are shown in Fig. 3.6. It should be noted that these images are a result of a particular combination of parameters of the network. In the next few subsections, we analyze the mAP performance of the trained FRCNN model for

various parameters such as different base networks, anchors, RPN min and max overlaps, detector min and max overlaps, and SNR values.

Impact of different base networks

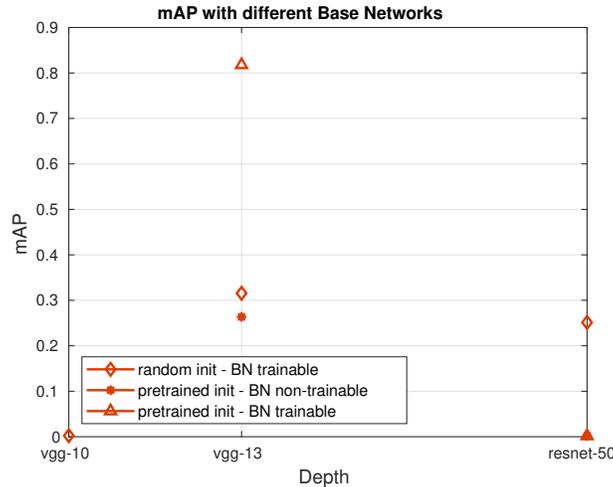


Figure 3.7: mAP for different base networks.

In Fig. 3.7, we plot the mAP values achieved by the FRCNN model when the base network is chosen to be VGG-10, VGG-13, and ResNet-50, where the numbers 10, 13, and 50 denote the number of convolutional layers present. While the architectures for VGG-13 and Resnet-50 are available online [71][72], the VGG-10 architecture is obtained from VGG-13 by removing the last three convolutional layers. We try different three different combinations of feature extraction: (i) use pretrained weights as the initialization for the BN and configure it as non-trainable, (ii) use pretrained weights as the initialization for BN and configure it as trainable, and (iii) use random weights as the initialization for BN and configure it as trainable. With the VGG-10, we only attempt the third method because there are no pretrained weights available for this architecture. We notice that the VGG-13, with the BN set as trainable and the pretrained weights used as the initialization, gives the best mAP performance. The performance drops slightly when the initialization is random and even more when the BN is set as non-trainable. Also, we notice that a very deep network such as the ResNet-50 may

not provide any improvement in mAP over moderately deep network such as the VGG-13 because the training complexity increases with the depth and a moderate depth may actually be sufficient to extract all the features necessary for the task at hand. The VGG-10, which is a slightly shallower network than the VGG-13, achieved poor mAP performance, possibly because it fails to perform feature extraction. For all the experiments presented henceforth, we fix the base network to VGG-13, initialize it with pretrained weights and configure it as trainable.

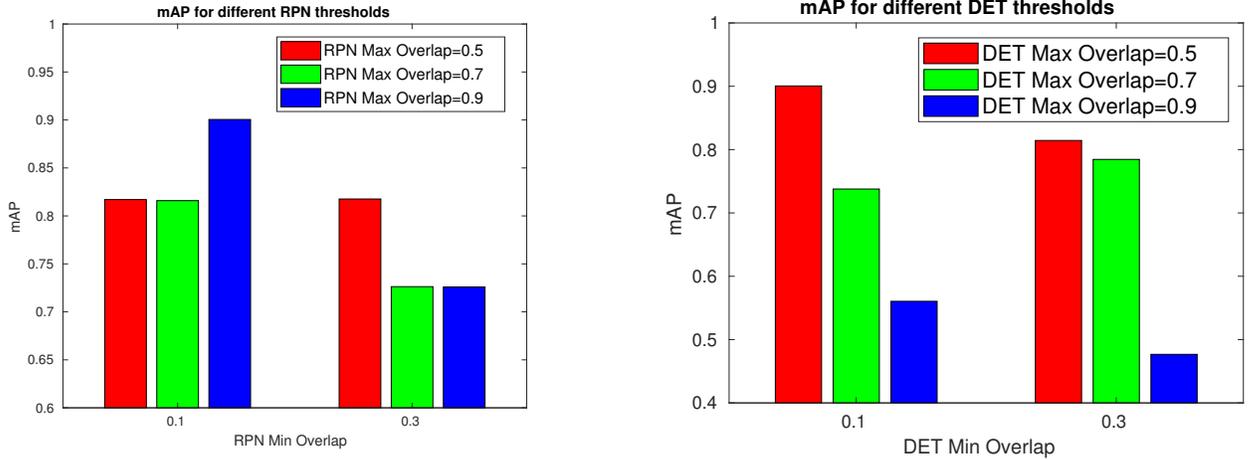
Impact of the number of anchor boxes

We now analyze the effect of using different number of anchors on the mAP performance of the FRCNN model. In Table 3.1, we list out the mAP values achieved with $N_a = 1, 2, 3, 4$. It is observed that the mAP values are fairly constant across different N_a values, with only marginal improvements when N_a is increased from 1 to 4. This reveals that the regression tasks in the RPN and the Detector network are powerful enough to regress from arbitrarily close anchors to the ground truths. Also, the chosen anchors need not have very high IoU with the ground truths for the training to be successful. For all the experiments presented henceforth, we fix the number of anchors to 3, with the anchor dimensions as given in Table 3.1.

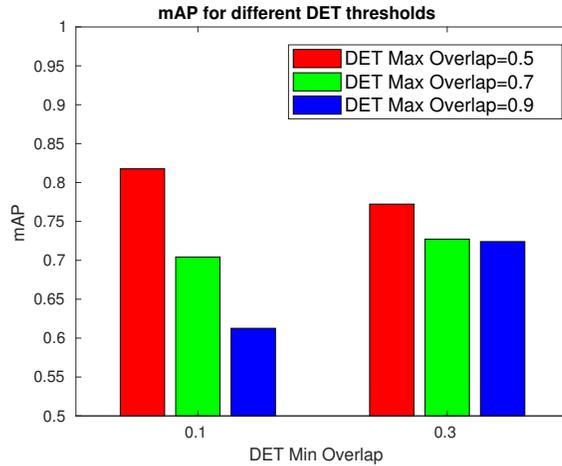
Number of anchors	Anchor dimensions	mAP
1	[1ms, 20MHz]	0.791322
2	[1ms, 20MHz], [2ms, 20 MHz]	0.79436
3	[1ms, 20MHz], [2ms, 20 MHz], [3ms, 20MHz]	0.8176363
4	[1ms, 20MHz], [2ms, 20 MHz], [3ms, 20MHz], [4ms, 20 MHz]	0.8172389

Table 3.1: mAP with different number of anchors.

Impact of RPN and Detector Min and Max Overlaps



(a) DET Min Overlap = 0.1, DET Max Overlap = 0.5 (b) RPN Min Overlap = 0.1, RPN Max Overlap = 0.9



(c) RPN Min Overlap = 0.3, RPN Max Overlap = 0.5

Figure 3.8: mAP with different RPN and Detector thresholds.

We now analyze the impact of the RPN and DET min and max thresholds. Since a grid search can be computationally exhaustive, we employ an alternate-once strategy for finding the optimal threshold values. We first fix the DET Min and Max overlaps to be 0.1 and 0.5 respectively and search over different combinations of RPN min and max overlap values as shown in Fig. 3.8a. Among the different thresholds, we observe that the (RPN Min, RPN Max) combinations (0.1, 0.9) and (0.3, 0.5) achieve better mAP values than the rest. We therefore fix the RPN Min and Max values to these two combinations and search over

the different DET min and max overlap values, as shown in Fig. 3.8b and 3.8c. We notice that the best (DET min, DET max) combination among these runs is (0.1, 0.5) and fix this combination for all ensuing experiments.

The alternate-once strategy presented above is a simplified search approach to obtain the RPN and DET min and max overlap thresholds. Naturally, a more exhaustive grid search method may help us choose better threshold values but would require many more experimental runs. Note that the optimal threshold values obtained either via grid search or via the alternate-once strategy, may not be universal for the FRCNN model because the optimal values can depend on the type and quality of the training data.

Impact of SNR

We now proceed to evaluate the performance of the FRCNN model for different SNR levels. We begin with a training dataset comprising 5 captures per SNR level in the set $\{0, 10, 20, 30\}$ dB. In Fig. 3.9, we plot the mAP values achieved when the test dataset comprises of captures with SNR = $-10, 0, 10, 20,$ and 30 dB respectively. We notice that the mAP values are consistently around 0.9 for positive SNR levels while dropping to a little over 0.5 for -10 dB. To verify whether this trend is universal, we next consider a training dataset comprising 5 captures per SNR level in the set $\{-10, 0, 10, 20, 30\}$ dB, i.e., we have now included a negative SNR level. The mAP values for each SNR level in the test dataset are given in Fig. 3.9. When compared to the case with non-negative SNR levels, we notice a drop in the mAP performance for all the SNR levels except for -10 dB. This is expected as the model should perform well on the data it has seen before, however, the drop in mAP values corresponding to positive SNR when trained with images of negative SNR points to the lack of generalization of the model across positive and negative SNR. It is also observed that the mAP on test captures with SNR less than -10 dB is very poor irrespective of the training strategy. This exposes the need for a denoising mechanism as a preprocessing step on the spectrograms before we feed them into the FRCNN model. We have conducted experiments

with the general-purpose wavelet denoising [69] as a simple pre-processing step and have observed no significant improvement in the mAP values. This observation motivates the need for an advanced custom-made denoising function that can take captures with strongly negative SNR and bring the SNR level at least up to the -10 dB level before feeding them to the FRCNN model. The design of such a custom-made denoising function is an interesting topic of investigation for future work.

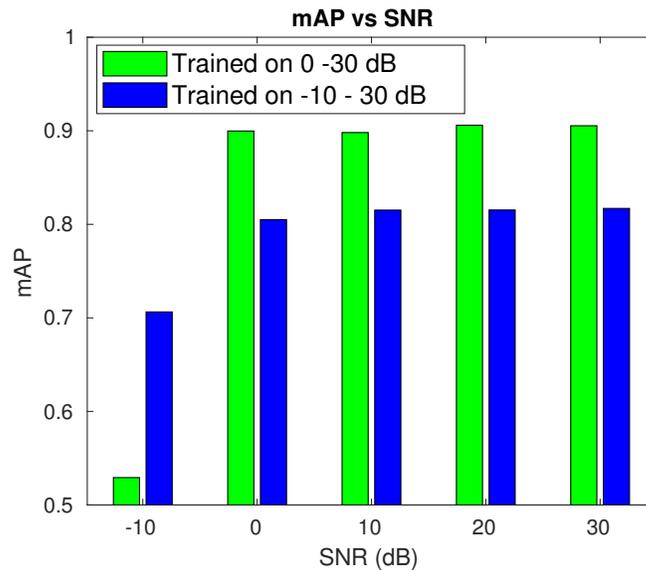


Figure 3.9: mAP vs SNR.

Impact of disparity in signal sizes

In all our experiments so far, we have considered signals with a fixed bandwidth of 20MHz. We now consider an artificially generated WiFi dataset comprising two different signal bandwidths, namely, the 5 MHz and 40 MHz, and SNR levels in the set $\{-10, 0, 10, 20, 30\}$ dB. Due to the nature of the WiFi protocol, both the 5MHz and 40MHz signals have variable dwell time but their bandwidths are different by a factor of 8. We conduct this experiment to verify whether the trained FRCNN model is indifferent to the disparity in the ground truth sizes. The number of anchor boxes N_a is set to 6, with the time spans drawn from $\{1, 2, 3\}$ ms and the frequency spans drawn from $\{5, 40\}$ MHz. We notice that the mAP values drop by a

considerable amount to 0.5320695 from 0.7917968 in the single-bandwidth case. Sample test images showing the prediction performance of the model are given below in Fig. 3.10. The drop in performance could have come from multiple sources: (i) the use of a single non-max suppression threshold may favor region proposals arising from a particular class of ground truth signals, thereby ignoring the other types, (ii) the base network may not be able to simultaneously extract all the important features for both the small and large bandwidth signals, (iii) since the anchor sizes are vastly different, the use of a single RPN min overlap (and RPN max overlap) may result in the RPN assigning higher classification probabilities to region proposals from one particular type of signals. A deeper investigation needs to be conducted on these probable causes and appropriate architectural modifications need to be developed in order to achieve uniform mAP performance across different signal sizes. We are currently exploring this research direction.

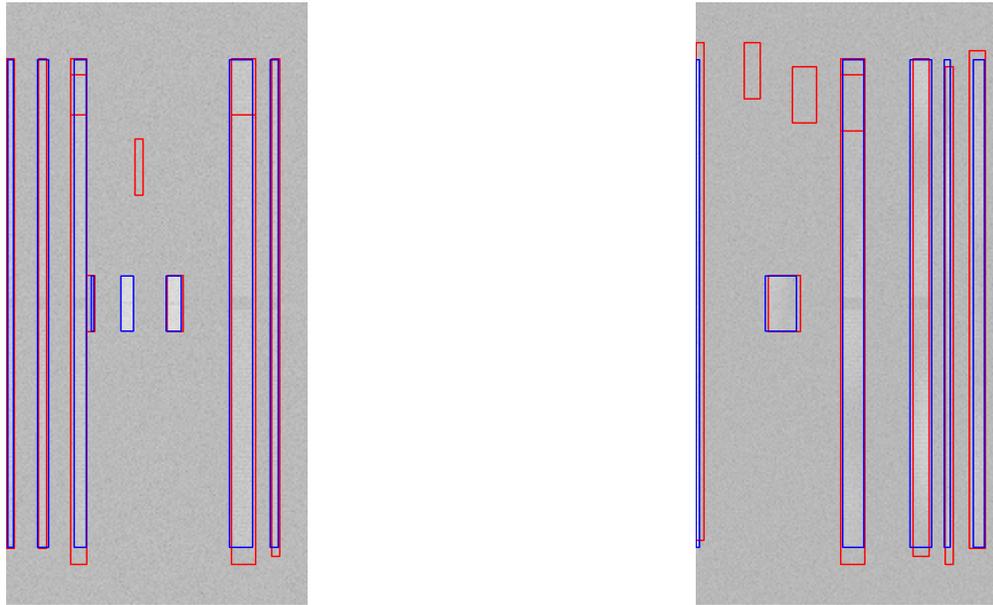


Figure 3.10: Test images of model trained on disparate anchor sizes. The red boxes are the predicted bounding boxes where as the blue ones indicate the GT signals.

3.6 Conclusion and Possible Future Work

In this Chapter, we have proposed a deep learning framework to perform real-time signal detection and time-frequency localization in a wideband RF spectrum of interest. We convert the data capture into a spectrogram and transform the given problem into an object detection problem. The following insights are obtained:

- Our experiments suggest that while weights pretrained on regular images are a good starting point for medium sized networks, making the weights trainable gives much better performance. It is also verified that deeper feature extraction networks don't necessarily improve performance.
- An mAP of up to 0.9 is recorded when the model is trained and tested on positive SNR values with single-bandwidth signals. This is seen to deteriorate as negative SNR signals are included in the training process thus motivating the need for a custom denoiser. The experiments also reveal that the model is agnostic to the number of anchor boxes because of the two-stage regression framework.
- When trained with disparate signal sizes it is noted that the mAP decreases noticeably. We hypothesize that this could be because of the inadequacy of the non-max suppression technique, feature extraction network and the RPN to handle anchors at disparate scales.

Some of the avenues for future work would be:

- The performance of negative SNR captures is observed to be not satisfactory with respect to the mAP. A custom denoising mechanism can be investigated to make sure the SNR level is acceptable before feeding the spectrogram to the signal detection framework.

- It is seen that generalizing over SNR values is not very easy for the model to do. Custom made models corresponding to different SNR ranges can be introduced and their performance improvement over general purpose detectors can be observed.
- It is observed that the employed non-max suppression technique may favour a particular anchor box over another during the training process. A non-max suppression technique could be devised that does not favour any particular anchor box type.
- The performance of the model is seen to deteriorate when disparate signals are introduced in the training process. It would be interesting to change the architecture to accommodate disparity in anchor sizes and come up with custom feature extractors and region proposal networks for disparate sizes.
- The experiments conducted in this work ignore the diversity and distribution in data. It would be interesting to investigate the impact of these two variations on the performance of the model.
- The Faster RCNN model is capable of classifying distinct signal patches. It would be natural to extend the proposed framework to perform modulation classification.
- While models pretrained with regular image datasets such as PASCAL VOC2007 [64] and MC COCO [65] are good starting points for training, it would be favorable to have a model pretrained on RF data to enable enhanced feature extraction and a multitude of diverse downstream tasks.

Chapter 4

Conclusion

In this thesis, we focused on two important directions involving next generation wireless systems. The first chapter of this thesis addressed an energy efficient partially connected transceiver architecture considering practical hardware as well as run-time constraints. While a case could be made about the energy efficiency of partially connected architectures, it should be noted that spectral efficiency does indeed take a hit as a result of the approximations in the design. While the promise of massive MIMO in terms of spectral efficiency is real, this promise it would appear, does not translate very well to the energy efficient hybrid precoding regime. The experiments also do suggest strongly that carefully designed alternating algorithms outperform non-alternating ones, which is accompanied with increased run-time hindering real time implementation of these algorithms. Nevertheless, in applications where energy efficiency is an important requirement and spectral efficiency is not, these architectures could still play a vital role.

The second chapter of this thesis proposed a real-time signal detection and time-frequency localization framework. It is demonstrated that deep learning indeed has the ability to perform high quality signal detection over a range of positive SNR. It would of course be unrealistic to demand that the model perform well across negative SNR as well, as this task would be equivalent to detecting objects in blurred images. A custom denoiser before the signal detection framework would aid in enhancing performance in such a scenario. While the model is very good with single-bandwidth signals, it falls short when multiple signals of disparate sizes are included. We are looking to address this issue in our future research work.

Bibliography

- [1] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in NIPS, 2015.
- [2] E. Perahia, and R. Stacey, *Next Generation Wireless LANs: 802.11n and 802.11ac*. 2nd Ed., United Kingdom, Cambridge University Press, 2013.
- [3] MATLAB 2018a, The MathWorks, Inc., Natick, Massachusetts, United States.
- [4] M. Grant and S. Boyd. (2014). *CVX: Matlab Software for Disciplined Convex Programming (version 2.1)* [Online]. Available: <http://cvxr.com/cvx>.
- [5] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G cellular: It will work!,” *IEEE Access*, vol. 1, pp. 335-349, May 2013.
- [6] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communication*, Englewood Cliffs, NJ, USA:Prentice-Hall, 2014.
- [7] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and Elza Erkip, “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164-1179, Jun. 2014.
- [8] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014.

- [9] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO-A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, Sep. 2017.
- [10] C. Rusu, R. Mendez-Rial, N. Gonzalez-Prelcicy, and R. W. Heath, "Low complexity hybrid sparse precoding and combining in millimeter wave MIMO systems," *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 1340-1345.
- [11] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653-656, Dec. 2014.
- [12] W. Tan, D. Xie, J. Xia, W. Tan, L. Fan, and S. Jin, "Spectral and energy efficiency of massive MIMO for hybrid architectures based on phase shifters," *IEEE Access*, vol. 6, pp. 11751-11759, 2018.
- [13] S. Buzzi, and C. D'Andrea, "Are mmwave low-complexity beamforming structures energy-efficient? Analysis of the downlink MU-MIMO," *Proc. IEEE Globecom Workshops (GC Wkshps)*, pp. 1-6, Dec. 2016.
- [14] J. Singh, and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2670-2683, May 2015.
- [15] C. Kim, T. Kim, and J.-Y. Seol, "Multi-beam transmission diversity with hybrid beamforming for MIMO-OFDM systems," *Proc. IEEE Global Commun. Conf. Workshops (GLOBECOM)*, Atlanta, GA, USA, Dec. 2013, pp. 61-65.
- [16] L. Dai, X. Gao, J. Quan, S. Han, and C.-L. I, "Near-optimal hybrid analog and digital precoding for downlink mmwave massive MIMO systems," *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 1334-1339.

- [17] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186-194, Jan. 2015.
- [18] S. Park, A. Alkhateeb, and R. W. Heath, Jr., "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907-2920, May 2017.
- [19] T. Kim, C. Kim, and J.-Y. Seol, "A low complexity hybrid beamforming algorithm for multi-beam diversity transmission," *IEEE Global Communications Conference (GLOBECOM'14)*, Dec. 2014.
- [20] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485-500, Apr. 2016.
- [21] F. Sohrabi, W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems", *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 2929-2933, Apr. 2015.
- [22] R. Lopez-Valcarce, N. Gonzalez-Prelcic, C. Rusu, R. W. Heath, "Hybrid precoders and combiners for mmWave MIMO systems with per-antenna power constraints", *Proc. GLOBECOM*, 2016.
- [23] N. Li, Z. Wei, H. Yang, X. Zhang, D. Yang, "Hybrid precoding for mmWave massive MIMO systems with partially connected structure", *IEEE Access*, vol. 5, pp. 15142-15151, Jun. 2017.
- [24] J. Lee, and Y. Lee, "AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing," *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, Australia, Jun. 2014, pp. 5838-5842.

- [25] M. Soleimani, R. C. Elliott, W. A. Krzymien, J. Melzer, and P. Mousavi, "Hybrid beamforming and DFT-based channel estimation for millimeter wave MIMO systems," *Proc. Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1-7, Oct. 2017.
- [26] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Transactions on Microwave Theory and Techniques*, vol.64, no.7, pp.2207-2225, July 2016.
- [27] 3GPP, "Study on channel model for frequency spectrum above 6GHz," 3rd Generation Partnership Project (3GPP), TR 38.900 V14.2.0, Jun. 2017
- [28] W. Tan, S. Jin, C. K. Wen, and J. Tao, "Spectral efficiency of multi-user millimeter wave systems under single path with uniform rectangular arrays," *EURASIP Journal on Wireless Communications and Networking*, vol.181, pp.1-13, Oct. 2017.
- [29] A. Cichocki, "Tensors decompositions: New concepts for brain data analysis?," *J. Control Measure. Syst. Integr. (SICE)*, vol. 47, no. 7, pp. 507-517, 2011.
- [30] X. Yu, J. Zhang, and K. B. Letaief, "Hybrid precoding in millimeter wave systems: How many phase shifters are needed?," *Proc. IEEE Global Commun. Conf.*, pp. 1-6, Dec. 2017.
- [31] W. Yu, and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646-2660, June 2007.
- [32] W. Yu, and T. Lam, "Downlink beamforming with per-antenna power constraints," *Proc. IEEE SPAWC*, pp. 1058-1062, Jun. 2005.
- [33] Z. Pi, "Optimal transmitter beamforming with per-antenna power constraints," *Proc. 2012 IEEE International Conf. Commun.*, pp. 3779-3784, Jun. 2012.

- [34] Z. Rezaei, E. Yazdian, and F. S. Tabataba, "Optimal energy beamforming under per-antenna power constraint," [online] Available: <http://arXiv:1702.07545>, Feb. 2017.
- [35] A. C. Jenifer, "Beamforming with per-antenna power constraint and transmit antenna selection using convex optimization technique," *Signal and Image Processing*, 5(3), 59, 2014.
- [36] S. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.:Cambridge Univ. Press, 2004.
- [37] X. Cai, G. Wang, and Z. Zhang, "Complexity analysis and numerical implementation of primal-dual interior point methods for convex quadratic optimization based on a finite barrier," *Numerical Algorithms*, 62(2):289-306, Feb. 2013.
- [38] M. Achache, "A new primal-dual path-following method for convex quadratic programming. Computational and Applied Mathematics," *Computational and Applied Mathematics*, vol. 25, no. 1, pp. 97-110, 2006.
- [39] A. Cichocki, A. Phan, "Fast Local Algorithms for Large Scale Nonnegative Matrix and Tensor Factorizations," *IEICE Trans. Fundamentals of Electronics*, vol. 92, pp. 708-721, Mar. 2009.
- [40] B. M. Browne, A. Langville, P. Pauca, and R. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics and Data Analysis*, vol.52, no.1, pp.155-173, Sep. 2007.
- [41] T.G. Kolda, and B. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455-500, June 2008.
- [42] K. Satyanarayana, M. El-Hajjar, P.-H. Kuo, A. Mourad, and L. Hanzo, "Millimeter wave hybrid beamforming with DFT-MUB aided precoder codebook design," *Proc. Veh. Technol. Conf.*, pp. 1-5, Sep. 2017.

- [43] M. Vu, "MISO capacity with per-antenna power constraint," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1268-1274, May 2011.
- [44] D.P. Palomar, and Y.C. Eldar, *Convex Optimization in Signal Processing and Communications*, Cambridge University Press, 2010.
- [45] M. M. A. Hossain, and R. Jantti, "Impact of efficient power amplifiers in wireless access," *Proc. IEEE GreenCom*, pp. 36-40, Sep. 2011.
- [46] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Energy-efficiency in massive MIMO-based 5G networks: opportunities and challenges," *IEEE Wireless Commun.*, vol. 3, no. 24, pp. 86-94, Jun. 2017.
- [47] M. Cai, K. Gao, D. Nie, B. Hochwald, J.N. Laneman, H. Huang, and K. Liu, "Effect of Wideband Beam Squint on Codebook Design in Phased-Array Wireless Systems," *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1-6, Dec. 2016.
- [48] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and Frequency-Wideband Effects in Millimeter-Wave Massive MIMO Systems," *IEEE Transactions on Signal Processing*, 2018.
- [49] M. Cai, "Modeling and mitigating beam squint in millimeter wave wireless communication," PhD diss., University of Notre Dame, 2018.
- [50] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Wideband spectrum sensing in cognitive radio networks," *Proc. IEEE Int. Conf. Commun.* pp. 901-906 2008-May.
- [51] D. Liu, C. Li, J. Liu, and K. Long, "A novel signal separation algorithm for wideband spectrum sensing in cognitive networks," in *Proc. IEEE Glob. Telecommun. Conf.*, Dec. 2010, pp. 1-6.

- [52] M. Bkassiny, S. Jayaweera, Y. Li, and K. Avery, "Wideband spectrum sensing and non-parametric signal classification for autonomous self-learning cognitive radios," *IEEE Trans. Wireless Commun.* vol. 11 no. 7 pp. 2596-2605 Jul. 2012.
- [53] P. Pham, J. Li, J. Szurley, and S. Das, "Eventness: Object Detection on Spectrograms for Temporal Localization of Audio Events," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*., Apr. 2018, pp. 2491-2495.
- [54] C. Kao, W. Wang, M. Sun, and C. Wang, "R-CRNN: Regionbased convolutional recurrent neural network for audio event detection," in *Proc. of Interspeech'18*, Hyderabad, India, Sep. 2-6 2018.
- [55] N. Bitar, S. Muhammad, H. H. Refai, "Wireless Technology Identification Using Deep Convolutional Neural Networks," Proceedings of *IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017.
- [56] M. Zhang, M. Diao, and L. Guo, "Convolutional neural networks for automatic cognitive radio waveform recognition," *IEEE Access*, vol. 5, pp. 11074-11082, 2017.
- [57] P. Sermanet, *et al.*, "Overfeat: integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, Feb. 2014.
- [58] D. Erhan, *et. al.*, "Scalable object detection using deep neural networks," in *Proc. of IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2014, pp. 2155-2162.
- [59] R. Girshick, *et. al.*, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2014, pp. 580-587.
- [60] K. He, *et. al.*, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *European Conf. Computer Vision (ECCV)*, 2014, pp. 346-361.

- [61] R. Girshick, “Fast R-CNN,” in *IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 1440-1448.
- [62] J. R. Uijlings, *et al.*, “Selective search for object recognition,” *Int. J. Computer Vision (IJCV)*, vol. 104, no. 2, pp 154-171, Apr. 2013.
- [63] C. L. Zitnick and P. Dollar, “Edge boxes: Locating object proposals from edges,” in *European Conference on Computer Vision (ECCV)*, Sept. 2014, pp. 391-405.
- [64] M. Everingham, *et al.*, “The PASCAL visual object classes challenge 2007 (VOC2007) Results,” 2007.
- [65] T.-Y. Lin, *et al.*, “Microsoft COCO: common objects in context,” in *European Conf. Computer Vision (ECCV)*, 2014.
- [66] J. Salamon, C. Jacoby, and J. P. Bello, “A dataset and taxonomy for urban sound research,” in *Proc. of the ACM Int. Conf. on Multimedia (ACM MM)*, 2014, pp. 1041-1044.
- [67] Annamaria Mesaros, *et. al.*, “DCASE 2017 challenge setup: tasks, datasets and baseline system,” in *Proc. of the Detection and Classification of Acoustic Scenes and Events Workshop (DCASE)*, 2017.
- [68] C. D. Manning, P. Raghavan, and H. Schutze, “Chapter 8: Evaluation in information retrieval” in *Introduction to Information Retrieval*, Cambridge University Press. 2008.
- [69] R. Rangarajan, R. Venkataramanan, and S. Shah, “Image denoising using wavelets,” *Wavelet and Time Frequencies*, 2002.
- [70] J. K. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, “Attention-based models for speech recognition,” in *Neural Information Processing Systems (NIPS)*, 2015.
- [71] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *International Conference on Learning Representations (ICLR)*, 2015.

- [72] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv:1512.03385, 2015.
- [73] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *ECCV*, 2014.
- [74] Cross entropy, [Online]. Available: https://en.wikipedia.org/wiki/Cross_entropy
- [75] R. Girshick, "Fast R-CNN," in IEEE International Conference on Computer Vision (ICCV), 2015.
- [76] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," in *Proc. of the IEEE*, vol. 66, no. 1, pp. 51-83, Jan. 1978.
- [77] R. B. Blackman, and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering-Part I," *Bell Syst. Tech. J.*, vol. 37, no. 1, pp. 185-282, 1958.
- [78] A. G. Howard, *et al.*, "Mobilenets: efficient convolutional neural networks for mobile vision applications." [Online] arXiv preprint arXiv:1704.04861, Apr. 2017.
- [79] Pre-Trained Models, [Online]. Available: <https://keras.rstudio.com/articles/applications.html>
- [80] Packet transmission time in 802.11, [Online]. Available: https://sarwiki.informatik.hu-berlin.de/Packet_transmission_time_in_802.11.
- [81] WLAN System Toolbox, [Online]. Available: <https://www.mathworks.com/help/wlan/index.html>

Appendix A

Proof of results in Chapter 2

A.1 Optimality condition for wideband systems PCS

Extending Proposition 1 in [23] to wideband systems, to achieve the maximum data rate in PCS, we have K sets of identical equations, where K is the total number of sub-carriers.

Each set of equations is of the form $Ax = b$ and has an optimal solution only if

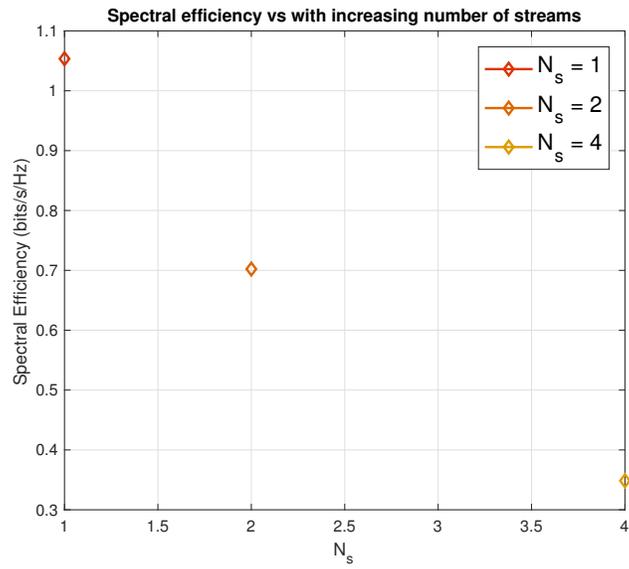
$$N_{RF}^t \geq r_H[k] \tag{A.1}$$

where $r_H[k]$ is the rank of the channel matrix corresponding to the k th subcarrier. Assuming that we have a full column rank channel matrix for all sub-carriers, we have $r_H[k] = N_t, \forall k \in 1, 2, \dots, K$. Therefore, it is required that $N_{RF}^t \geq N_t$ for an optimal solution to exist.

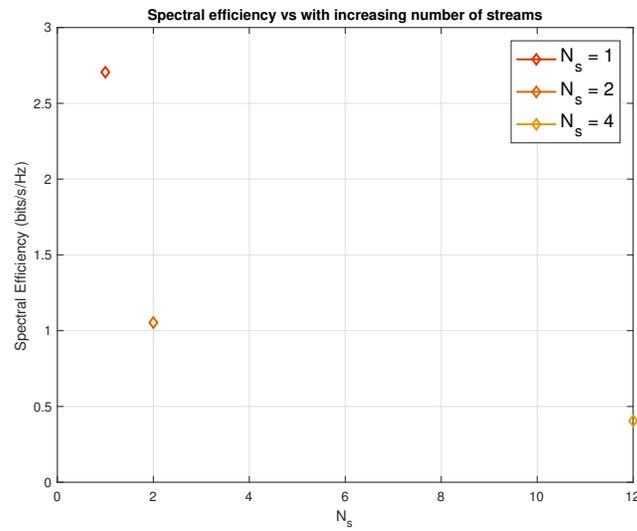
A.2 Results with varying number of streams and subcarriers

We note from equation (2.22) that the system becomes increasingly overdetermined with increase in number of data streams (N_s) and number of sub-carriers (K). We verify this by observing the spectral efficiency performance for different values of N_s and K as shown in Fig. A.1.

A.2. Results with varying number of streams and subcarriers



(a) Spectral efficiency vs N_s with $K = 2$.



(b) Spectral efficiency vs K with $N_s = 1$.

Figure A.1: Spectral efficiency at SNR = 0 dB, with $N_{RF}^t = N_{RF}^r = 4$.